Long-term asset allocation under time-varying investment opportunities: Optimal portfolios with parameter and model uncertainty*

Thomas Dangl

Institute of Management Science, TU Wien, Austria
email: thomas.dangl@tuwien.at

Alex Weissensteiner

School of Economics and Management, Free University of Bozen-Bolzano, Italy, email: alex.weissensteiner@unibz.it

December 9, 2017

Abstract

We study the implications of predictability on the optimal asset allocation of ambiguity averse long-term investors. We analyze the term structure of the multivariate risk-return trade-off in a VAR model under full consideration of parameter uncertainty, and we decompose the predictive covariance along different sources of risk/uncertainty. We calibrate the model to real returns of US stocks, US long-term government bonds, cash, real-estate and gold using the term spread and the dividend-price ratio as additional predictive variables. While over short periods the model-implied conditional covariance structure of asset-class returns determines the optimal allocation, we find that over longer horizons the optimal asset allocation is significantly influenced by the covariance structure induced by estimation errors. As a consequence, the ambiguity averse long-term investor tilts her portfolio not simply toward the global minimum-variance portfolio but shrinks portfolio weights toward a seemingly inefficient portfolio which shows maximum robustness against estimation errors. Most interestingly, we find that even though time diversification of stock returns vanishes after consideration of estimation errors, real long-term bond returns are even more affected, making stocks an important asset class for the ambiguity averse long-term investor.

JEL classification: G11; *EFM classification*: 370

Keywords: Portfolio choice; Predictability; Parameter uncertainty; Ambiguity aversion; Strategic asset allocation

*We gratefully acknowledge comments from Nicole Branger, Holger Kraft, Victor DeMiguel, Raman Uppal and Josef Zechner.

1 Introduction

Return predictability has great impact on the optimal long-term asset allocation of risk averse investors, as stressed, e.g., by Campbell and Viceira (2002) or Campbell, Chan, and Viceira (2003). Although there is still an ongoing debate whether out-of-sample predictions are statistically and/or economically significant, a growing body of empirical evidence supports return predictability in different asset classes.[1] The reported autocorrelation in asset returns together with predictive power of explanatory variables lead to time-varying return expectations and have considerable effect on the (co)variance structure of long-horizon returns. Time-varying expectations might give reason for attempts to time the market and for using current predictions to make short-term profits. The long-term investor, however, is primarily interested in the consequences of predictability on the covariance structure of holding-period returns. Since predictability makes the multivariate risk-return trade-off dependent on the investment horizon, investors who consider predictability conclude that part of the unconditional variance of asset returns is actually predictable time variation and thus does not constitute investment risk. This effect – also referred to as time diversification – makes the optimal asset allocation horizon-dependent.[2] The estimated term structure of the risk-return trade-off is, however, susceptible to estimation errors, and these errors tend to increase considerably with an increasing investment horizon, which erodes part of the effect of time diversification.[3]

It is the goal of our study to derive the optimal multivariate asset allocation for a long-term investor who shows aversion against ambiguity in true expected asset returns. Time-varying investment opportunities, i.e., predictability, is modeled in a vector autoregression framework of order one, VAR(1). The portfolio optimization for a mean-variance investor is done under full consideration of parameter uncertainty. We provide closed-form expressions for the decision problem at hand. We treat ambiguity aversion regarding expected holding-period returns in a multi-prior setting as introduced in Garlappi, Uppal, and Wang (2007). With this approach, ambiguity averse investors optimize their mean-variance objective while regarding each possible portfolio selection being paired with the worst possible expected return vector that exceeds a certain critical likelihood threshold.

We calibrate our model to real returns of US bonds, stocks, real estate, T-bills, and gold over the period from 1960-01 to 2015-12 and use the dividend-price ratio and the term

[1] For early papers on predictability in equity and bond returns see Fama and Schwert (1977), Keim and Stambaugh (1986), Fama and French (1988a,b), Campbell and Shiller (1988) and Fama and Bliss (1987). A critical re-examination of studies in equity return predictability is Welch and Goyal (2008). More recent work on equity prediction that addresses the concerns of Welch and Goyal (2008) is, e.g., Campbell and Thompson (2008), Rapach, Strauss, and Zhou (2010), and Dangl and Halling (2012). Recent papers that document predictability in bond premia are Cochrane and Piazzesi (2005), Cieslak and Povala (2015), and Diebold and Li (2006). Predictability in commodity returns is reported, e.g., by Gorton, Hayashi, and Rouwenhorst (2012).

[2] Campbell and Viceira (2005) estimate a VAR model for stocks, bonds and T-bills with dividend-price ratio, term spread, and short-term nominal interest rate as additional predictors and show that the riskiness of stocks decreases for increasing investment horizon, while T-bills and long-term bonds become more risky as the horizon becomes longer.

[3] See, e.g., Pastor and Stambaugh (2012) who find that even with very long time-series of equity-return data, estimation errors in model parameters are large enough to outweigh the variance reduction, or time diversification, originating from predictability.

spread as additional covariates. With this study we fill a gap in the existing literature since it is the first that develops the optimization of a multivariate asset allocation of a risk and ambiguity averse long-term investor in the presence of return predictability. Existing studies like Campbell and Viceira (2002, 2005) and Campbell, Chan, and Viceira (2003) focus on the multivariate term structure of the risk/return trade-off but are silent on estimation uncertainty. Papers that explicitly address parameter uncertainty, like Barberis (2000) or Pastor and Stambaugh (2012) neither analyze multivariate asset allocation nor allow to consider investor's ambiguity aversion with respect to parameter uncertainty. Garlappi, Uppal, and Wang (2007) study multivariate asset allocation with parameter uncertainty and ambiguity aversion, but assume i.i.d. return processes and, thus, do not consider return predictability.

We find that in the presence of predictability the covariance of holding-period returns generally differs from the covariance structure of errors in the estimates of long-term expected returns. Thus, portfolios that are robust against estimation errors differ from portfolios that diversify return volatility. So investors who show ambiguity aversion with respect to errors in estimates of expected returns will prefer portfolios that are seemingly inefficient from the point of view of a pure risk averse investor. While both depend on the investment horizon, estimation errors become disproportionately large compared to return variance when planning for long horizons. Consequently, even moderate ambiguity aversion against misspecification of expected returns, which has only low impact on the asset allocation over short holding periods, eventually becomes the determining factor in the long-term portfolio choice.

In line with Pastor and Stambaugh (2012), we see that time diversification of stock returns effectively vanishes after considering parameter uncertainty. The possible misspecification of expected returns, however, affects other asset classes more severely (most prominently real long-term bond returns), making stocks a relevant asset class for the ambiguity averse long-term investor. With moderate ambiguity aversion, the investor holds a balanced portfolio with considerable long positions in cash. For increasing ambiguity aversion bond holdings are diminished, leaving the investor with an optimal portfolio of stocks, real-estate and cash. Independent of the investment horizon and the model parameterization, gold plays an insignificant role in a long-term portfolio.

This paper focuses on long-term asset allocation, but as an extension we demonstrate that our approach of regarding aversion against ambiguity in estimated risk premia in a VAR(1) setup can actually benefit also the short- and medium-term investor. Optimizing the portfolio under a reasonable level of aversion against model mis-specification protects investors from aggressive attempts to time the market which result in extreme portfolio positions that translate into poor out-of-sample return characteristics and high portfolio turnover. We show empirically that portfolios optimized under ambiguity aversion deliver significant certainty equivalence gains compared to portfolios that are optimized ignoring return predictability (Garlappi, Uppal, and Wang, 2007).

The paper is structured as follows. In Section 2 we discuss existing literature related to our paper. Section 3 introduces the VAR(1) framework we employ to model predictability in asset returns and to decompose total variance-covariance of aggregated long-term returns.

Section 4 describes our data set. Section 5 derives closed-form expressions for our portfolio selection problem under parameter uncertainty and presents empirical results when bonds, stocks, cash, real estate and gold are available. In Section 6 we present empirical evidence that our approach of considering aversion against ambiguity in the model specification can also benefit the short- and medium-term investor. Section 7 concludes. All our proofs can be found in the appendix.

2 Related Literature

Early work of Merton (1969), Samuelson (1969) and Fama (1970) show that when asset prices are generated by a geometric Brownian motion (i.e., returns are i.i.d. and jointly normal) and investors' utility is iso-elastic, the optimal asset allocation is constant and independent of the investment horizon. Deviations from these rather restrictive assumptions generally lead to time-varying investment-opportunity sets and predictability in asset returns, which make optimal investment depend on (i) the current state and (ii) the investment horizon. Investors with a general (non-log) utility function will, thus, exhibit hedging demand against these changes in expected returns and/or covariances over time, see Merton (1971, 1973). Since there is a growing body of empirical evidence for predictability in asset returns (as discussed in Section 1, Footnote 1), we argue that considering predictability is of first-order relevance for long-term portfolio allocation.

Kim and Omberg (1996) propose an analytically tractable model for an investor with HARA utility who trades a risk-free asset and a risky asset. Brennan, Schwartz, and Lagnado (1997) analyze a richer asset allocation problem of a long-term investor who can invest in bonds, stocks and cash. They consider the short rate, the dividend yield as well as the yield on a console bond as additional predictors, and they use a finite difference approximation on a grid to solve the problem numerically. In their setting, an investor with a long horizon places a larger fraction of the portfolio in both stocks and bonds than does a myopic investor. Mean reversion of bonds and stock returns makes these assets less risky in the long run. Furthermore, their out-of-sample results indicate that exploiting predictability in asset returns is also economically significant. Campbell, Chan, and Viceira (2003) use an analytical approximation to solve the investment/consumption problem of an infinitely long-lived investor with Epstein-Zin utility. All mentioned papers assume known parameters in their analysis. In their outline for future research, Brennan, Schwartz, and Lagnado (1997) emphasize that estimation risk should be addressed directly, which would then alleviate the problem of highly leveraged portfolio positions and extreme portfolio turnover.

While Bawa, Brown, and Klein (1979) address estimation risk in portfolio selection problems for i.i.d. returns, Kandel and Stambaugh (1996) are the first to consider parameter uncertainty under return predictability. Their investor uses sample evidence to update his beliefs about parameters in a Bayesian setup. They conclude that considering the predictive variables in portfolio choice can have a substantial influence on investors' utility even though regression evidence for such predictability may be weak. Barberis (2000) considers a long-term investor who can allocate money to Treasury bills and a stock index.

Time-varying investment opportunities are modeled within a VAR framework, which includes asset returns and predictor variables. Parameter uncertainty is considered by the predictive distribution of future returns. He finds that even after incorporating parameter uncertainty, time diversification is still the predominant effect, and a long-horizon investor still allocates more to equities than an investor ignoring predictability, but the effect is smaller than when assuming fixed parameters. Parameter uncertainty might even dominate time diversification, as demonstrated by Pastor and Stambaugh (2012) in a univariate allocation setup, so that the investor who is aware of predictability and parameter uncertainty regards stocks more risky in the long run than in the short run and, thus, allocates less to the risky asset when considering a longer holding period. Xia (2001) examines the optimal portfolio strategy under parameter uncertainty in continuous time. Compared to previous work, in her model the investor is allowed to learn about the predictive relation over time. She illustrates that predictability and stochastic predictive variables introduce a stochastic covariance between the current estimate of the parameters and the stock returns. Learning reduces the sensitivity of the optimal allocation to the predictive variable, and the relationship between both is no longer monotone. Using historical data, she shows that investors who ignore market timing can incur very large opportunity costs.

The standard Bayesian approach in the above mentioned literature treats the unknown parameters as random variables. A pre-specified prior is combined with observations from data to construct a predictive distribution of returns. Optimal portfolios maximize expected utility with respect to this predictive distribution. As emphasized by Garlappi, Uppal, and Wang (2007), the implicit assumption here is that decision-makers are *neutral* to ambiguity in the sense of Knight (1921). However there is substantial evidence that this is not the case, and that especially private investors have a preference for asset allocation rules which are robust with respect to this ambiguity, see, e.g., Li, Tiwari, and Tong (2016).

3 Model

To allow for time-varying investment opportunities, we model the joint dynamics of asset returns and of predictive variables in a VAR(1) framework which we borrow from Barberis (2000)

$$z_t = a + Bz_{t-1} + \epsilon_t, \tag{1}$$

with z_t the $(n \times 1)$ vector of asset returns and covariates. As the asset menu we use long-term US government bonds, US stocks, cash and alternative investments in form of gold plus real estate, since for most homeowners the house is the single most important asset in the portfolio (see e.g. Flavin and Yamashita, 2002). As additional predictors we include the term spread between the yield of long-term US-government bonds and the T-bill rate and the log dividend–price ratio. Therefore, in our setting we have $n = 7$. For a detailed description of the data which we use to calibrate our model, please refer to Section 4. The $(n \times 1)$ vector a consists of intercept coefficients and B is the $(n \times n)$ matrix of slope coefficients. Disturbances are denoted by ϵ_t, an $(n \times 1)$ vector with i.i.d. $N(0, \Sigma_\epsilon)$ distributed elements. A calibration of the model must determine $n^2 + n$ coefficients in a and B plus the $(n + 1)n/2$ elements of Σ_ϵ.

A potentially important drawback of the VAR-based approach is that standard least squares parameter estimates might be contaminated by a finite-sample bias that seriously distorts the asset allocation decision, especially when the model contains variables, as interest rates, dividend-price ratios and term spreads, that are highly persistent, see Bekaert, Hodrick, and Marshall (1997) and Engsted and Pedersen (2012, 2014). More concise, such estimates will generally be biased toward a dynamic system that displays less persistence than the true process. The bias is particularly pronounced when the estimation sample is short and the dynamic process is very persistent (see Bauer, Rudebusch, and Wu, 2012).[4] Therefore, we apply the technique proposed by Kilian (1998); Nicholls and Pope (1988); Pope (1990) to obtain bias-corrected parameter estimators.

We consider a long-term investor in $t = 0$, who is interested in the aggregated returns of the tradable assets bonds, stocks, cash, real estate and gold over an investment horizon of T months. Conditional on a, B and Σ_ϵ, the aggregated log-returns r_T are normally distributed (see Barberis, 2000, eq. (18) and (19)) with mean

$$\mu_T = Ta + (T-1)Ba + (T-2)B^2a + \cdots + B^{T-1}a + (B + B^2 + \cdots + B^T)z_t, \qquad (2)$$

and covariance

$$\begin{aligned}\Sigma_T = \Sigma_\epsilon & \\ &+ (I+B)\,\Sigma_\epsilon\,(I+B)' \\ &+ \left(I+B+B^2\right)\Sigma_\epsilon\left(I+B+B^2\right)' \\ &\vdots \\ &+ \left(I+B+\cdots+B^{T-1}\right)\Sigma_\epsilon\left(I+B+\cdots+B^{T-1}\right)'. \end{aligned} \qquad (3)$$

We address the *uncertainty* in the parameters estimated from a sample with T_{obs} observations within a simulation study by sampling from the predictive distribution. We employ the two-step Bayesian approach described by Zellner (1971). To compute the posterior distribution $p(a, B, \Sigma_\epsilon | z)$ we rewrite the model as

$$\begin{pmatrix} z'_{-(T_{\mathrm{obs}}-1)} \\ \vdots \\ z'_0 \end{pmatrix} = \begin{pmatrix} 1 & z'_{-T_{\mathrm{obs}}} \\ 1 & \vdots \\ 1 & z'_{-1} \end{pmatrix} \begin{pmatrix} a & B \end{pmatrix}' + \begin{pmatrix} \epsilon'_{-(T_{\mathrm{obs}}-1)} \\ \vdots \\ \epsilon'_0 \end{pmatrix}. \qquad (4)$$

With C a $(n \times (n+1))$ matrix which results from joining the vector a with the $(n \times n)$ matrix B,

$$C = \begin{pmatrix} a_1 & B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ a_2 & B_{2,1} & B_{2,2} & \cdots & B_{2,n} \\ \vdots & & & \ddots & \\ a_n & B_{n,1} & B_{n,2} & \cdots & B_{n,n} \end{pmatrix}, \qquad (5)$$

[4] Campbell, Chan, and Viceira (2003) acknowledge the finite-sample bias in their VAR estimates but state that bias corrections are complex in multivariate systems and, hence, they do not attempt to adjust for the bias.

we write the model as

$$Z = XC' + E. \tag{6}$$

When calibrating the model to observed data, estimated coefficients $\hat{a}$ and $\hat{B}$ as well as the covariance structure of the residuals $\hat{\Sigma}_\epsilon$ are affected by estimation errors. From Zellner (1971) we know that true coefficients have a joint inverse Wishart / Normal distribution such that the marginal distribution $p(\Sigma_\epsilon^{-1}|z)$ follows

$$\Sigma_\epsilon^{-1}|z \sim \text{Wishart}(T_{\text{obs}} - n - 2, \hat{\Sigma}^{-1}). \tag{7}$$

Conditional on the covariance matrix Σ_ϵ, model coefficients are normally distributed according to

$$\text{vec}(C')|\Sigma_\epsilon, z \sim N(\text{vec}(\hat{C}'), \Sigma_\epsilon \otimes (X'X)^{-1}). \tag{8}$$

Estimation errors in the model's coefficients propagate in a complex way into estimation errors in the moments of the aggregate T-period return r_T, μ_T and Σ_T stated in Equations (2) and (3). Regarding estimation errors, r_T is generally not normally distributed. We decompose the term structure of the total variance of r_T imposed by time-varying investment opportunities in three parts:

$$\begin{aligned}
\text{var}(r_T|z) &= \mathbb{E}_{\Sigma_\epsilon}(\text{var}(r_T|\Sigma_\epsilon, z)) + \text{var}_{\Sigma_\epsilon}(\mathbb{E}(r_T|\Sigma_\epsilon, z)) \\
&= \mathbb{E}_{\Sigma_\epsilon}[\mathbb{E}_{a,B}(\text{var}(r_T|a, B, \Sigma_\epsilon, z))] + \mathbb{E}_{\Sigma_\epsilon}[\text{var}_{a,B}(\mathbb{E}(r_T|a, B, \Sigma_\epsilon, z))] \\
&\quad + \text{var}_{\Sigma_\epsilon}(\mathbb{E}(r_T|\Sigma_\epsilon, z)) \\
&= \mathbb{E}_{\Sigma_\epsilon}[\mathbb{E}_{a,B}(\Sigma_T)] + \mathbb{E}_{\Sigma_\epsilon}[\text{var}_{a,B}(\mu_T)] + \text{var}_{\Sigma_\epsilon}(\mathbb{E}(r_T|\Sigma_\epsilon, z)) \\
&= \mathbb{E}_{\Sigma_{\epsilon,a,B}}[\Sigma_T] + \mathbb{E}_{\Sigma_\epsilon}[\Omega_T] + \Lambda_T \\
&= \bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T,
\end{aligned} \tag{9}$$

i.e., the expected average covariance, $\bar{\Sigma}_T$, and two additional covariance matrices which reflect the uncertainty in expected asset returns given the ambiguity about the parameters, $\bar{\Omega}_T$, as well as the ambiguity about the residuals' covariance of the VAR model, Λ_T. Hence, a mean-variance optimizer will invest in the classical mean-variance optimal portfolio with $\bar{\mu}_T = \mathbb{E}[\mu_T]$ as the expectations of asset returns and a covariance structure determined by $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T$.

In order to calculate the covariance matrices $\bar{\Sigma}_T$, $\bar{\Omega}_T$ and Λ_T we rely on simulation. We perform 400,000 simulation draws according to the scheme (7) and (8). First, we sample 2,000 covariance matrices from the marginal $p(\Sigma_\epsilon^{-1}|z)$ employing (7) and for each of the sampled Σ_ϵ we simulate 200 sets of model parameters a and B from the conditional distribution described in (8). In line with Stambaugh (1999), we assume that the sampled process is stationary, i.e., when a sampled B exhibits an eigenvalue with modulus above one the observation is dropped and resampled.

4 Data

We use (where necessary calculate) continuously compounded monthly returns for all assets in our investment menu. Returns of stocks, bonds and cash are taken from the updated data set proposed in Welch and Goyal (2008):[5] Equity returns are given by stock returns on the S&P500 index (including dividends) (`CRSP_SPvw`). For bonds and cash we rely on long-term government bond returns (`ltr`) and T-bill rates (`Rfree`). We calculate real estate returns from the Home Price Index, see Shiller (2015).[6] In order to consider the additional rent income for real estate (comparable to a dividend yield of stocks), Davis, Lehnert, and Martin (2008) provide historical rent-to-rice ratios, which we add to the price changes of the Home Price Index.[7] While all other data are available on a monthly frequency, the rent-to-price data are with a quarterly frequency. Therefore, we assume that within the three months of a quarter the indicated rent-to-price ratio is constant. Furthermore, in order to allow for an alternative investment, we add gold to our asset menu.[8] All five time series are deflated by the changes in the Consumer Price Index (`infl`).

As additional predictor variables we use the term spread and the log dividend–price ratio, see Campbell, Chan, and Viceira (2003). We calculate the term spread as the difference between the long-term government bond yield (`lty`) and the T-bill rate (`tbl`). We use 12-months moving sums of dividends paid on the S&P500 index (`D12`). The log dividend–price ratio is the log annual dividends (`D12`) less the log price index (`Index`).

From 1960-01 to 2015-12 we have in total 672 monthly observations to estimate the VAR process. Table 1 shows OLS parameter estimates. Apart from analyzing a different time interval and including the additional assets real estate and gold, our estimates are well in line with those of Campbell, Chan, and Viceira (2003).

[Table 1 about here.]

The most noteworthy observation is the very high and significant persistence of the log dividend-price ratio and the term spread, and the relatively high persistence of cash and real estate, with the consequence that R^2 of the regression of these time series is high. Bond, stock and gold returns are harder to predict. For bonds the returns of stocks and cash as well as the term spread have significant predictive power. In line with previous literature, stocks returns have the lowest R^2. As expected, real estate returns are negatively related to the short interest rate (return of cash). Table 2 describes the correlation structure of the innovations in the VAR system, with annual standard deviations on the main diagonal. Consistent with previous results of Campbell, Chan, and Viceira (2003), unexpected stock returns are highly negatively correlated with shocks to the log dividend-price ratio. Unexpected bond returns are negatively correlated with shocks to the term spread.

[Table 2 about here.]

[5]The data is provided on Amit Goyal's website `www.hec.unil.ch/agoyal/`. Mnemonics are indicated in parenthesis.

[6]The data is provided on Robert Schiller's website `www.econ.yale.edu/~shiller/data.htm`.

[7]The data is provided at `www.lincolninst.edu/subcenters/land-values/rent-price-ratio.asp`.

[8]The data is provided at `www.globalfinancialdata.com`.

Given that OLS parameter estimates for very persistent processes are downward biased toward a stable system, we apply the Pope (1990) correction, which shifts very persistent coefficients upwards (e.g., the OLS parameter for term spread of 0.9544 is corrected to 0.9618, that of the dividend-price ratio of 0.9919 is corrected to 0.9921). That this seemingly "minor" correction is relevant also in terms of risk shows Table 3, in which we compare elements on the main diagonal of the decomposed covariance matrices, see (9), from the OLS estimated VAR model with those of the Pope corrected VAR model. It can be seen that the Pope correction considers the higher estimation risk imposed by very persistent processes, and that the difference to OLS increases with the investment horizon.

[Table 3 about here.]

When comparing the overall variance (i.e. $\bar{\Sigma} + \bar{\Omega} + \Lambda$) of the asset returns in the investment menu, the ratio between Pope-corrected and OLS values given in Table 4 reveals that bonds are affected most from the small-sample adjustment. From our reading of Table 1, this is due to the fact that bond returns are predicted significantly (p-value below 0.01) by the term spread, which is a very persistent process. Thus, estimation errors of the term spread propagate over the considered horizon and translate in estimation errors of bond returns. Although also the log dividend-price ratio is very persistent, its predictive power for stocks is much weaker, which explains why stock returns are less affected by the bias-correction. However, the overall variance of all assets significantly increases when addressing the small-sample bias.

[Table 4 about here.]

[Table 5 about here.]

We use the Pope corrected parameters and sample 2000×200 times to compute the posterior distribution $p(a, B, \Sigma_\epsilon|z)$. Table 5 lists the variance decomposition (9) for the five asset classes on a horizon of 120 months as well as on a horizon of 600 months. The contribution of Λ_T, i.e., the effect of estimation uncertainty in the covariance of VAR residuals, plays a minor role for all asset classes independent of the horizon. This observation is consistent with Kan and Zhou (2007) who show that the covariance of residuals can be estimated with good precision when the number of asset returns is low compared to the number of observations. In fact the main contribution to long-term return variance originates from $\bar{\Sigma}_T$, the expected covariance of returns, and from $\bar{\Omega}_T$, the variance of expected returns driven by estimation errors in VAR coefficients a and B. The contribution of $\bar{\Omega}_T$ increases with the time horizon. While the contribution of errors in expected returns is roughly 15% on a horizon of 120 months, it increases to 43%–48% on a horizon of 600 months. Please note that $\bar{\Omega}_T$ implies a covariance structure that differs from $\bar{\Sigma}_T$, hence, the asset allocation is more severely influenced by estimation errors in expected returns on longer horizons, as we will see in Section 5.

In Figure 1 we illustrate the effect of these different sources of variance on asset-class volatility. The upper-right panel of Figure 1 shows that the overall p.a. risk of stocks is slightly decreasing with the investment horizon, even though considering parameter uncertainty leads to a substantial upward correction in the long-term riskiness of stocks (for a discussion see, e.g., Barberis, 2000; Pastor and Stambaugh, 2012). Compared to stock returns, due to their persistence the overall p.a. risk of the returns of bonds, cash, gold and real estate clearly increases in time.

[Figure 1 about here.]

The corresponding term structure of correlations for the different asset pairs are illustrated in Figure 2. In general, the 95% confidence intervals of the correlation between all asset classes indicated with dashed-dotted lines results to be huge, with values between -0.8 and $+0.8$. For the pairwise correlation between bond returns, stock returns and real estate returns we observe a small positive short-term and a significantly higher long-term average dependence. Furthermore, the long-term average pairwise correlation between each of these assets and cash is negative, i.e., bonds, stocks and real estate benefit in the long run from a decline in the short interest rate, and vice verse.[9] While the long-term average correlation between gold and stocks/real estate is negative, the pairwise correlation between gold and bonds is positive. Our tentative explanation is that during times of financial turmoil stock and real estate prices decline, while safe-haven instruments as governmental bonds and gold are on high demand. All confidence bounds are wide, so ignoring estimation errors and giving asset-allocation advice based on the expected term structure of return correlations is not recommended.

[Figure 2 about here.]

Given the well-known challenge to identify *proper* expected returns for portfolio selection problems, we follow the idea of Sharpe (1974) and compute implicit T-returns conveyed by the asset weights of anchor portfolios. Therefore, we propose to use the 2013 Survey of Consumer Finances (SCF) conducted by the Federal Reserve. The main purpose of the survey is to analyze the financial condition of families in the United States and to study the effects of changes in the economy. The Survey of Consumer Finance is updated every three years and, among others, categorizes the asset allocation of US families according to different discrimination criteria as, e.g., income, age of head, family structure, education of head, race, work status of head, region, wealth etc. For our investigation, we identify *age of head* as the most relevant characteristic in revealing information about the investment horizon. We distinguish: (a) young families with an age of head between 35 and 44, and (b) older families with an age of head between 65 and 74, arguing that younger families to have a longer investment horizon (young families 50 years versus older families 10 years). We aggregate and categorize asset classes to five groups: bonds (saving bonds plus bonds minus loans secured by residential property), stocks (stocks, business equity), cash (transaction accounts plus certificates of deposits minus short term loans as installment loans, lines of credit and credit card balances), real estate (primary and other residential property, equity in non residential property) and gold. Other asset classes with mixed portfolios (pooled investment funds, retirement accounts, cash life insurance and other managed assets) are assigned to bonds, stocks and cash according to the relative weight in the composition of the financial wealth of a family. Since we lack detailed information, the implicit assumption

[9] The cash returns are given by the monthly T-bill rate.

here is that *on average* the asset allocation in those mixed products corresponds to the weights directly held in the classical financial assets. Furthermore, we assign the investment class "Other" to our category gold. We denote these market-capitalization implied assets returns by $\hat{m}_T$. Figure 3 compares the calculated asset allocation of the two groups.

[Figure 3 about here.]

The most significant difference can be identified in the weights of bonds and real estate. While for young households the weight of real estate in the overall portfolio is considerably higher than for older households, the bond investments (netted for loans secured by residential property) is lower. We consider this finding as an empirical evidence for older households having a shorter planning horizon and investing therefore a higher fraction of their wealth in bonds.

Let us mention two alternative approaches to specifying expected returns. First, expected returns can be determined endogenously in a general equilibrium setup. If assets are in limited supply and different types of investors form their optimized portfolios, aggregate demand drives prices and consequently expected returns. We have worked out the optimization analysis in general equilibrium for two types of investors, an investor that shows only risk aversion but no ambiguity aversion and an investor with both risk and ambiguity aversion. Such a model leads to interesting cross-dependencies since changes in the parametrization of one investor type influences optimal portfolios of the other through the demand channel. While these effects are undoubtedly very interesting, we think that the estimation and the comprehensive interpretation of such a model is beyond the scope of this paper. Thus, we will discuss optimal asset allocation and the treatment of ambiguity aversion in a partial equilibrium with market implied expected returns as described above. The second alternative to implicit return expectations is to resort directly to the VAR model. In Section 6 we discuss portfolio properties when applying the VAR estimate of risk premia for a variety of different investment horizons and different levels of ambiguity aversion. It can be seen that under reasonable levels of ambiguity aversion optimized long-term portfolios based on expected premia directly from the VAR estimates are very similar to optimal portfolios based on premia implied by SCF data.

5 Portfolio Optimization

We consider a long-term investor who decides on the optimal asset allocation according a mean-variance criterion and shows a certain degree of ambiguity aversion against misspecification in expected long-term returns. The investor uses the VAR(1) framework discussed in the previous sections to exploit predictability in returns and employs the multi-prior approach of Garlappi, Uppal, and Wang (2007) to deal with ambiguity in expected returns. This framework is a classical *max-min* approach to consider ambiguity aversion within a restricted set of specifications for the long-term expected returns – the set of priors considered. The choice of priors is done with respect to the likelihood of the model specification, and we will discuss this approach in more detail in this section.

Starting point of the mean-variance optimization is the decomposition of the predictive variance derived in Equation (9),

$$\begin{aligned}
\operatorname{var}(r_T) &= \bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T \\
&= \mathbb{E}(\Sigma_T) + \mathbb{E}\left[\operatorname{var}_{a,B}(\mu_T)\right] + \operatorname{var}_{\Sigma_\epsilon}(\mu_T).
\end{aligned}$$

The ambiguity averse mean-variance investor will consider the first part of this decomposition $\bar{\Sigma}_T$ as return variation affecting her mean-variance objective and treat the estimation uncertainty in expected returns, conveyed by $\bar{\Omega}_T + \Lambda_T$, within the multi-prior approach.

In contrast to this complex approach, we also discuss/analyze four less sophisticated investors and interpret similarities and differences in their asset allocations: (a) We consider a mean-variance investor who ignores predictability and uses the sample covariance matrix to determine the optimal asset allocation. (b) We regard a mean-variance investor who is aware of predictability but ignores parameter uncertainty, i.e., who uses a mean-variance criterion that considers $\hat{\Sigma}_T$ from (3) as the covariance structure of returns, where $\hat{\Sigma}_T$ refers to the matrix $\Sigma_T|z$, the matrix which is estimated from observed data (see, e.g., Campbell, Chan, and Viceira, 2003). (c) We look at a mean-variance investor who additionally considers the effect of estimation uncertainty on Σ_T but takes expected long-term returns as given. This investor uses only the first part, $\bar{\Sigma}_T$, from the decomposed predictive variance. (d) Finally, we refer to a mean-variance investor without ambiguity aversion who fully recognizes that estimation errors in the expected returns increase the predictive variance. This investor acknowledges the full predictive variance $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T$ in her mean-variance optimization (see, e.g., Barberis, 2000).

The VAR(1) model (1) is written in log returns, which have a clear advantage in time aggregation of μ_T in (2). For portfolio optimization we use the discrete counterpart, denoted by m_T. Since we are not interested in market timing strategies, but only in the long-term implications of predictability, we estimate expected discrete holding-period returns from representative portfolios, as discussed in the previous section, and denote these implied estimates $\hat{m}_t$.

We formulate the optimization problem of an ambiguity averse mean-variance investor, who addresses aversion against ambiguity in expected returns within a multi-prior framework, in the following way

$$\max_{w} \min_{\bar{m}_T} w'\bar{m}_T - \frac{\gamma}{2} w'\bar{\Sigma}_T w, \tag{10}$$

$$\text{s.t. } (\hat{m}_T - \bar{m}_T)' \left(\bar{\Omega}_T + \Lambda_T\right)^{-1} (\hat{m}_T - \bar{m}_T) \leq \varepsilon, \tag{11}$$

$$w'\mathbf{1} = 1. \tag{12}$$

We assume an investor with relative risk aversion of γ who seeks to optimize a mean-variance criterion, i.e., the risk/return trade-off implied by $\bar{\Sigma}_T$ and $\bar{m}_T$, with $\bar{\Sigma}_T$ the expected covariance of returns under consideration of parameter uncertainty and $\bar{m}_T$ the unknown vector of expected holding-period returns, see Equation (10). Equation (12) is the usual portfolio constraint. To ensure robustness with respect to a misspecification of expected returns $\bar{m}_T$, the objective is simultaneously minimized with respect to $\bar{m}_T$, chosen from a set of available priors, see Equation (11). This so-called multi-prior approach follows the idea of Garlappi, Uppal, and Wang (2007) which states that under consideration of the first two moments of estimation errors in $\bar{m}_T$, robustness is ensured by doing the *max-min* optimization (10) over all models whose normal-likelihood exceeds a given threshold ε (such a constraint results in a hyper-ellipsoid, see, e.g., Meucci, 2009). Ignoring higher moments of $\bar{m}_T$, the left-hand side of (11), also known as Mahalanobis distance, is distributed according to Hotelling's T^2 distribution, see, e.g., Marida, Kent, and Bibby (1979). If the number of observations is large, the distribution converges to a χ^2 distribution with n degrees of freedom. Hence, interpreting ε as the quantile of the $\chi^2_{(n)}$ distribution allows to state a confidence level for the robustness of the optimization. While we acknowledge that predictability introduces higher moments in $\bar{m}_T$, we stick to the assumption that investors regard all $\bar{m}_T$ contained in a critical hyper-ellipsoid defined by (11). We do this because empirical quantiles of the left-hand-side of (11) are almost identical to quantiles of the $\chi^2_{(n)}$ distribution for short investment horizons. For longer horizons we consider the deviations from joint normality by determining ε quantiles of the empirical distribution obtained from simulation.

Figure 4 shows the empirical distribution of the Mahalanobis distance of $\hat{m}$ and $\bar{m}$ for $T = 120$ months and for $T = 600$ months together with a χ^2 distribution with five degrees of freedom. Vertical solid and dashed lines indicate the 90% and 95% quantiles of the $\chi^2_{(5)}$ distribution (solid red) as well as empirical quantiles at $T = 120$ (solid black) and $T = 600$ (dotted black). We see that for investment periods up to 10 years, the $\chi^2_{(5)}$ distribution approximates the empirical distribution quite well. For longer horizons non-normality in $\bar{m}$ leads to higher probability for large deviations. A reasonable choice of ε for 90% confidence is between 9.2 ($T = 120$) and 9.5 ($T = 600$) and for 95% confidence between 11.1 and 14.7.

[Figure 4 about here.]

Proposition 1. *Let* $A = \mathbf{1}'\bar{\Sigma}_T^{-1}\mathbf{1}$, $B = \hat{m}_T'\bar{\Sigma}_T^{-1}\mathbf{1}$ *and* $D = \bar{\Sigma}_T^{-1}(\Phi_T^{-1} + \bar{\Sigma}_T^{-1})^{-1}\bar{\Sigma}_T^{-1}$ *with* $\Phi_T(w) = \frac{\sqrt{\varepsilon}}{\gamma\sqrt{w'(\bar{\Omega}_T+\Lambda_T)w}}(\bar{\Omega}_T + \Lambda_T)$, *and given the min-max problem defined by* (10)-(12), *the optimal asset allocation* w^* *is given by*

$$\begin{aligned} w^* &= \frac{1}{\gamma}\left(\bar{\Sigma}_T^{-1} - D\right)\left(\hat{m}_T + \frac{\gamma + \hat{m}_T' D\mathbf{1} - B}{A - \mathbf{1}'D\mathbf{1}}\mathbf{1}\right) \\ &= \underbrace{\frac{1}{\gamma}\bar{\Sigma}_T^{-1}\left(\hat{m}_T + \frac{\gamma - B}{A}\mathbf{1}\right)}_{\text{mean-variance portfolio ignoring uncertainty in } \bar{m}_T} + \text{ambiguity hedge.} \end{aligned} \tag{13}$$

All our proofs can be found in the appendix. From (13) we see that optimal portfolio weights can be expressed as the sum of the mean-variance optimal portfolio of an investor who ignores estimation uncertainty in $\bar{m}_T$ (while considering the effect of estimation uncertainty in the covariance structure) plus an ambiguity hedge portfolio (zero-sum investment) that ensures the desired protection against ambiguity in $\bar{m}_T$. The overall hedging demand against ambiguity depends on the magnitude of ε.

Proposition 2. *The first-order condition, see Appendix A.1 (24), reveals that there exists* *a coefficient of ambiguity aversion ε^* with*

$$\varepsilon^*(\gamma) = \gamma^2 w'(\bar{\Omega}_T + \Lambda_T)w, \tag{14}$$

such that the optimal portfolios of the ambiguity averse investor and that of a pure risk averse investor coincide. With $\varepsilon > \varepsilon^*(\gamma)$, the ambiguity averse investor overweights uncertainty from estimation errors relative to the pure return variance they produce. With ε below the critical level of $\varepsilon^*(\gamma)$, ambiguity aversion is so low that investors underweight estimation errors relative to the pure risk averse investor.[10]

For $\varepsilon \searrow 0$, i.e., if ambiguity aversion vanishes, the hedging demand vanishes and w^* converges to the optimal portfolio of a pure mean-variance optimizer who ignores estimation uncertainty in $\bar{m}_T$,

$$w^* = \frac{1}{\gamma}\bar{\Sigma}_T^{-1}\left(\hat{m}_T - \frac{B-\gamma}{A}\mathbf{1}\right). \tag{15}$$

This means that for vanishing ambiguity aversion there is no smooth transition to the portfolio of the pure risk averse investor, it is rather a transition to the portfolio of the investor who ignores errors in the estimate of expected returns. Thus, in interpreting results, the critical level of ambiguity aversion ε^* serves as an implicit threshold that indicates whether the ambiguity averse investor treats estimation errors more cautiously than the pure risk averse investor.

For $\varepsilon \nearrow \infty$, the optimal portfolio converges to

$$w^* = \frac{1}{\mathbf{1}'(\bar{\Omega}_T + \Lambda_T)^{-1}\mathbf{1}}(\bar{\Omega}_T + \Lambda_T)^{-1}\mathbf{1}, \tag{16}$$

such that the hedge against estimation errors in expected returns completely dominates the asset allocation. Such a portfolio optimally diversifies the joint errors in expected returns. For an investor who is not ambiguity averse this portfolio is essentially inefficient.

Our result differs from Garlappi, Uppal, and Wang (2007). In the absence of predictability, the covariance structure introduced by parameter uncertainty is the same as the covariance structure of return volatility (i.e., the covariance of estimation errors is a scaled version of the sample covariance of returns). As a consequence, a growing ambiguity aversion pushes the optimal portfolio toward the global minimum-variance portfolio along the invariant efficient portfolio frontier, i.e., increasing risk aversion and increasing ambiguity aversion both move the optimal portfolio along the existing frontier toward the minimum-variance portfolio. In our setting, due to predictability, parameter uncertainty induces a covariance structure that differs from the covariance structure of return volatility. As a consequence, an increasing ambiguity aversion does not "only" shrink the optimal portfolio toward the minimum-variance portfolio, but pushes optimal portfolios from the efficient frontier into the seemingly inefficient-portfolio area.

In order to determine market implied expected asset returns for different investment horizons, we reverse the fix-point problem in (13) for a representative agent with $\gamma = 4$ and

[10] The dotted vertical lines in Figure 4 show the location of the corresponding ε^* for a relative risk aversion of $\gamma = 4$ and investment horizons of $T = 120$ and $T = 600$.

set ε to the critical value ε^* defined in Proposition 2 such that $\hat{m}_T$ gives the asset weights of older households with a planning horizon of $T = 120$, and younger household with a planning horizon of $T = 600$, respectively. We use data given by the Survey of Consumer Finance 2013 sorted according to *age of head* to identify their asset allocation, see our discussion in Section 4. The choice of the representative agent using ε^* is motivated by the empirical fact that institutional investors became the dominant market player over time, while ambiguity aversion seems to be mainly found for retail investors, see Li, Tiwari, and Tong (2016).[11] Institutional investors are typically viewed as being relatively sophisticated with a better understanding of the market. Consequently, such investors face less ambiguity, but they account for the full predictive variance $\bar{\Omega}_T + \bar{\Sigma}_T + \Lambda$ in their optimization.

Given the portfolio constraint (weights must sum to one) is binding in the anchor portfolio of the representative investor, reverse optimization does not yield unique return expectations, see Zagst and Pöschik (2008). Since the binding constraint influences the portfolio choice, portfolio weights reveal the investors expectations only up to one degree of freedom, i.e., relative to some reference return that must be specified exogenously. We choose cash as the reference asset and assume a real long-term return of zero. Implicit returns backed out by reverse optimization of the anchor portfolio specify the remaining asset class returns, see Figure 5.

[Figure 5 about here.]

[Figure 6 about here.]

Figure 6 illustrates efficient frontiers of investors with different sophistication levels (as discussed above) from the viewpoint of an investor who only considers the sample covariance matrix in his optimization. His efficient frontier is indicated with the solid green line. In addition, we show also efficient frontiers of more sophisticated agents: First, we consider investors who are aware of predictability from (3) but ignore parameter uncertainty, such that they base their portfolio decision on $\hat{\Sigma}_T$ (black-dotted line). Second, we consider investors who are aware of predictability and parameter uncertainty but take expected long-term returns as given, i.e., they base their decision on $\bar{\Sigma}_T$ (black-dashed line). Third, we show the optimal risk-return trade-off of investors who consider the full predictive variance $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T$ but are not ambiguity averse (black-solid line). Finally, we consider investors who are aware of the covariance of returns $\bar{\Sigma}_T$ and show aversion against ambiguity in expected returns, i.e., $\bar{\Omega}_T + \Lambda_T$ is treated in the multi-prior approach (dependent on the degree of ambiguity aversion, their efficient frontier is indicated with different colors). For $\varepsilon > \varepsilon^*$ the ambiguity averse investor overweights uncertainty from estimation errors in the mean (drawn by colored-solid lines), while below this critical value the investor underweights estimation errors (drawn by colored-dashed lines), see Proposition 2. The diamond and the bullet points indicate the global minimum-variance portfolios for the above mentioned (more sophisticated) investors. Of course, from the naive investor's viewpoint all choices of the others are seemingly inefficient. The left panel shows results for a planning

[11] Blume and Keim (2012) illustrate, e.g., that he proportion of U.S. public equities managed by institutions has risen steadily over the past six decades, from about 7 or 8% of market capitalization in 1950, to about 67% in 2010.

horizon of $T = 120$ months, the right panel shows results for a planning horizon of $T = 600$, respectively. The comparison between both panels confirms that parameter uncertainty becomes much more relevant in the long run.

[Figure 7 about here.]

[Figure 8 about here.]

In Figure 7 and 8 we analyze the impact of an increasing ambiguity aversion on the optimal portfolio composition given holding periods of $T = 120$ and $T = 600$ months. The risk aversion γ is set equal to 4. In each figure, the upper-left panel shows the optimal portfolio of an investor who considers only $\bar{\Sigma}_T$, i.e., neglects uncertainty in the expected returns. The upper-right panel gives the optimal portfolio of an investor who considers the full predictive variance $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T$ without being ambiguity averse. Given our choice of the representative agent, these portfolios correspond to those in Figure 3. The lower panels show optimal asset allocations of investors who consider the full predictive variance and are ambiguity averse, i.e., have $\varepsilon > \varepsilon^*$. A higher ambiguity aversion tilts the optimal asset allocation toward cash. By comparing Figure 7 and 8, it can be seen that for longer investment horizons stocks and real estate (bonds) become more (less) attractive. In this sense our results are in line with Barberis (2000), who also finds that stock allocation increases with horizon. We attribute this result to the fact that with our data set the predictive volatility of stock returns is relatively independent of the time horizon, while those of the other assets is increasing. In all cases, gold plays only a minor role in the optimal composition of a portfolio.

In order to better distinguish our contribution from earlier papers in the literature, in Figure 9 we compare the optimal asset allocations for $T = 120$ (upper panels) and $T = 600$ (lower panels) of investors with different sophistication level. While in the left panels we show the optimal asset allocation for a relative risk aversion of $\gamma = 4$, the right panels show the corresponding global minimum-variance portfolios of extreme risk/ambiguity averse investors. For all four combinations we consider again (a) an investor who completely neglects predictability and parameter uncertainty in asset returns, i.e., bases his investment decision on the sample covariance matrix, (b) an investor who considers predictability but ignores parameter uncertainty, i.e., uses $\hat{\Sigma}_T$, (c) an investor who optimizes his portfolio based on $\bar{\Sigma}_T$, i.e., takes expected returns as given, (d) an investor who fully recognizes the entire predictive variance without ambiguity aversion, which – due to our choices for the representative agent – gives the market capitalization from the SCF, (e) an ambiguity averse investor who considers the full predictive variance. In the left panels, all investors use the same market-implied expected returns of the representative agent, and ambiguity averse investors have to consider the constraint $\varepsilon = 10$ (which corresponds to a confidence level of roughly 91% for both investment horizons). The left panels show, in line with well-known previous results of the literature, that erroneously neglecting parameter uncertainty, i.e., underestimating the overall risk, leads to extreme asset weights. The investor (a), who bases his decision on the sample covariance matrix, is tempted to borrow heavily short-term money in order to invest in real estate. For an investment horizon of $T = 600$

his debt-to-equity ratio is up to 28.[12] Considering parameter uncertainty leads to more balanced portfolios. The right panels show that the difference in the optimal asset allocation of minimum-variance portfolios can be substantial. In our data set, predictability makes stocks more attractive compared to cash. For both considered investment horizons, an increasing ambiguity aversion tilts the optimal portfolio toward cash. However, compared to an extremely risk averse investor, extremely ambiguity averse investors allocate a significantly higher amount of their wealth to stocks. For both, the extremely risk averse investor and the extremely ambiguity averse investor, bonds and gold play no role in the optimal asset allocation.

[Figure 9 about here.]

6 Extension: Short- and Medium-term Investment

We use the data set of Section 4 to show empirically that a short- and medium-term investor, who aims for exploiting predictable variation in expected risk premia, can benefit from our approach of introducing ambiguity aversion in a VAR(1) setup. This is an extension to the analysis of this paper, and an in-depth study of timing strategies under ambiguity aversion shall be left for future research. While data availability limits the application of informative backtest studies of long-term asset allocation strategies, short- and medium-term investment (up to a horizon of a few years) can be investigated using historical data. We therefore compare out-of-sample results of our VAR approach that considers predictability to the approach *without* predictability (GBM) proposed by Garlappi, Uppal, and Wang (2007). In both cases we assume a risk-aversion of $\gamma = 4$. We use an expanding window to estimate the parameters, starting with 25 years of monthly observations (from 1960 to 1984), and calculate out-of-sample returns for the period from 1985 onwards. In line with the corresponding model assumptions, we allow the GBM approach (as a one-period model) to re-allocate the portfolio every month, and take non-overlapping investment periods of length T for the VAR approach. In Table 6 we present results for holding periods of one, two and five years.

[Table 6 about here.]

For different levels of ambiguity aversion and different holding periods we calculate the certainty equivalent (CE) as well as the first four moments of the out-of-sample returns for both strategies.[13] To avoid that our results are mainly driven by the specific starting month of the investigation, we shift the starting point of the analysis up to eleven months and report mean and median of the twelve calculations.

[12] This extreme leverage is in line with ballooned average loan-to-value ratios of mortgaged homeowners (up to 95%) during and after the Great Recession, see, e.g., data presented by Bullard (2012).

[13] Garlappi, Uppal, and Wang (2007, p. 51) show the relationship between ε, used in their equation (16), and ϵ, used in their numerical exhibition. Given that compared to them we use directly the covariance matrix of *expected* returns (instead of the covariance matrix of the returns), we can omit the term $1/T$, and a value of $\epsilon = 3$ in their setting corresponds to $\varepsilon = 15.2$ for 25 years of monthly observations, i.e., has to be scaled by $5.07 \approx 299 \times 5/(300 - 5)$.

The VAR approach is able to exploit predictability, which results in a higher mean out-of-sample return compared to the GBM approach. The increase in mean is accompanied by an increase in the standard deviation of returns. Both effects are higher for low levels of ε and short investment horizons T. Further evidence for the successful exploitation of predictability is the generation of right-skewed out-of-sample returns for low levels of ambiguity aversion and for short investment horizons. Furthermore, Table 6 reports the certainty equivalent provided by the two approaches for an investor with CRRA utility with relative risk aversion of $\gamma = 4$. For short investment horizons the VAR approach delivers higher CEs than the GBM approach, and higher levels of ambiguity aversion reduce the tendency to exploit time variation in expected returns and, thus, reduce the CE advantages of the VAR strategy. For longer holding periods, however, the VAR model faces a higher exposure to potential model mis-specification than models which ignore predictability. Having high tolerance against mis-specifications (low ε) over a 60 months holding period results in lower CEs than using the GBM model. Using high ε, i.e., investing in a portfolio that focuses on diversifying estimation errors, the VAR model has again a CE advantage over a model that does not regard predictability. Consequently, the introduction of the ambiguity-aversion approach proposed by Garlappi, Uppal, and Wang (2007) into the VAR setup apparently improves its applicability to short- and medium-term investment.

Figure 10 compares the asset allocation of the VAR approach for $T = 12$, $T = 60$ and $T = 600$ to the asset allocation of the one-period GBM approach with $\varepsilon = 25$. For longer investment horizons the impact of predictability on the covariance becomes more important than short-term predictability in expected returns, and, as a consequence, the asset allocation decisions are more balanced. It can be seen that while the optimal asset allocation of the GBM approach is almost entirely composed of cash and real estate, the VAR approach, in order to exploit predictability in the data, trades actively also the other assets.

[Figure 10 about here.]

Without ambiguity-aversion, over-confidence in the estimated parameters often leads to extreme portfolio weights in a VAR model (even if the analysis is conducted from a steady state perspective with very long/infinite investment horizons, as done in e.g. Campbell, Chan, and Viceira, 2003). With the consideration of potential mis-specifications in expected returns, the asset allocations becomes more stable – even when starting from the practically relevant vector of realized state variables as done here. The proposed asset allocations of the VAR approach in Figure 10 seem to be reasonable, e.g. for $T = 60$ with asset weights between -0.60 and 0.04 for bonds, between -0.18 and 0.59 for stocks, between -5.97 and 0.40 for cash, between 0.81 and 6.83 for real estate and between -0.10 and 0.21 for gold. In line with empirical observations, for both the VAR and GBM approach real estate is a very attractive investment class.

7 Conclusion

In this paper we investigate the optimal asset allocation of long-term investors who consider return predictability but show aversion against model mis-specification. Predictability induces a term structure in the covariance of asset returns, with the consequence that the optimal asset allocation depends on the investment horizon. Return predictability is captured by a VAR(1) model that considers ambiguity aversion in the way proposed by Garlappi, Uppal, and Wang (2007), where investors perform a *max-min* optimization over a set of priors about expected returns. We demonstrate how to decompose the term-structure of return correlation along three sources of risk/uncertainty: The expected covariance of returns, the covariance of expected returns imposed by estimation errors of the VAR parameters, and the covariance of expected returns imposed by estimation errors of the VAR model's residual covariance. We derive a closed-form expression for the optimal portfolio decision, and we show that ambiguity aversion under predictability of asset returns does no longer correspond to a "shrinkage" toward the global minimum-variance portfolio. Thus, aversion against the ambiguity in expected returns leads to optimal investment in a portfolio that is not mean-variance efficient but optimally diversifies estimation errors.

The model is calibrated to a data set that covers more than 50 years of real returns of US stocks, US long-term government bonds, cash, real-estate and gold, and uses the term spread and the dividend-price ratio as additional predictors. Expected returns are estimated from market-implied expectations using portfolio weights of young and older US households as references.

Our main findings from an inspection of the calibrated model can be summarized as follows: The risk of model mis-specification significantly contributes to the overall predictive volatility and is highly relevant for the asset allocation decision over long investment horizons. For stocks the term structure of overall annual volatility is slightly decreasing in the investment horizon, for the other assets it increases significantly. We show that the 95% confidence intervals for the annual volatility as well as for the correlation pairs are wide. Thus, long-term portfolio advice needs to be done under consideration of estimation errors. Neglecting parameter uncertainty leads to overconfidence and extreme portfolio compositions, where cash is used to leverage the investments in real estate and bonds. Ambiguity against errors in the model parametrization reduces leverage and even turns cash into an interesting investment vehicle. Stocks and real estate are further relevant asset classes for ambiguity averse long-term investors, while bonds and gold play only a minor role.

A Appendix

A.1 Proof of Proposition 1

Equation (10)-(12) can be written as Lagrangian

$$\mathcal{L}(\bar{m}_T, \lambda, \delta) = w'\bar{m}_T - \frac{\gamma}{2} w'\bar{\Sigma}_T w - \lambda \left[\varepsilon - (\hat{m}_T - \bar{m}_T)' \left(\bar{\Omega}_T + \Lambda_T\right)^{-1} (\hat{m}_T - \bar{m}_T)\right] + \delta(1 - w'\mathbf{1}). \tag{17}$$

It is well known that $\bar{m}_T^*$ is a solution of the constraint problem only if there exists a scalar $\lambda^* \geq 0$, such that $(\bar{m}_T^*, \lambda^*)$ is a solution of the following unconstrained problem

$$\min_{\bar{m}_T} \max_{\lambda} \mathcal{L}(\mu, \lambda). \tag{18}$$

From the first-order conditions with respect to $\bar{m}_T$ in (17) we obtain

$$\bar{m}_T^* = \hat{m}_T - \frac{(\bar{\Omega}_T + \Lambda_T)w}{2\lambda}. \tag{19}$$

Substituting this in (17) we get

$$\mathcal{L}(\bar{m}_T, \lambda, \delta) = w'\hat{m}_T - \frac{\gamma}{2}w'\bar{\Sigma}_T w - \frac{1}{4\lambda}w'(\bar{\Omega}_T + \Lambda_T)w - \lambda\varepsilon + \delta(1 - w'\mathbf{1}). \tag{20}$$

Therefore, the max-min problem is equivalent to the following maximization problem,

$$\max_{w,\lambda} w'\hat{m}_T - \frac{\gamma}{2}w'\bar{\Sigma}_T w - \frac{1}{4\lambda}w'(\bar{\Omega}_T + \Lambda_T)w - \lambda\varepsilon + \delta(1 - w'\mathbf{1}). \tag{21}$$

Solving for λ we obtain

$$\lambda = \frac{1}{2}\sqrt{\frac{w'(\bar{\Omega}_T + \Lambda_T)w}{\varepsilon}} > 0, \tag{22}$$

which results in

$$\mathcal{L}(\bar{m}_T, \delta) = w'\hat{m}_T - \frac{\gamma}{2}w'\bar{\Sigma}_T w - \sqrt{\varepsilon w'(\bar{\Omega}_T + \Lambda_T)w} + \delta(1 - w'\mathbf{1}). \tag{23}$$

The first-order conditions with respect to w gives

$$\hat{m}_T - \delta\mathbf{1} = \gamma\left(\bar{\Sigma}_T + \frac{\sqrt{\varepsilon}}{\gamma\sqrt{w'(\bar{\Omega}_T + \Lambda_T)w}}(\bar{\Omega}_T + \Lambda_T)\right)w. \tag{24}$$

Let $\Phi_T(w) = \frac{\sqrt{\varepsilon}}{\gamma\sqrt{w'(\bar{\Omega}_T + \Lambda_T)w}}(\bar{\Omega}_T + \Lambda_T)$ and given the so-called Woodbury matrix identity $(\bar{\Sigma}_T + \Phi_T)^{-1} = \bar{\Sigma}_T^{-1} - \bar{\Sigma}_T^{-1}(\Phi_T^{-1} + \bar{\Sigma}_T^{-1})^{-1}\bar{\Sigma}_T^{-1}$ we get

$$w = \frac{1}{\gamma}\left(\bar{\Sigma}_T^{-1} - \bar{\Sigma}_T^{-1}(\Phi_T^{-1} + \bar{\Sigma}_T^{-1})^{-1}\bar{\Sigma}_T^{-1}\right)(\hat{m}_T - \delta\mathbf{1}). \tag{25}$$

Using $w'\mathbf{1} = 1$, we can write

$$w'\mathbf{1} = \frac{1}{\gamma}(\hat{m}_T - \delta\mathbf{1})'\left(\bar{\Sigma}_T^{-1} - \bar{\Sigma}_T^{-1}(\Phi_T^{-1} + \bar{\Sigma}_T^{-1})^{-1}\bar{\Sigma}_T^{-1}\right)\mathbf{1} = 1. \tag{26}$$

Let's define $A = \mathbf{1}'\bar{\Sigma}_T^{-1}\mathbf{1}$, $B = \hat{m}_T'\bar{\Sigma}_T^{-1}\mathbf{1}$ and $D = \bar{\Sigma}_T^{-1}(\Phi_T^{-1} + \bar{\Sigma}_T^{-1})^{-1}\bar{\Sigma}_T^{-1}$ we can rewrite (26) as

$$\frac{1}{\gamma}(B - \hat{m}_T'D\mathbf{1} - \delta(A - \mathbf{1}'D\mathbf{1})) = 1 \tag{27}$$

$$-\frac{\gamma + \hat{m}_T'D\mathbf{1} - B}{A - \mathbf{1}'D\mathbf{1}} = \delta, \tag{28}$$

which in (25) gives

$$w^* = \frac{1}{\gamma}\left(\bar{\Sigma}_T^{-1} - D\right)\left(\hat{m}_T + \frac{\gamma + \hat{m}_T' D\mathbf{1} - B}{A - \mathbf{1}' D\mathbf{1}}\mathbf{1}\right).$$

Thus, the optimal weights w are implicitly given as a fixed-point of the above equation, since the matrix D on the right hand side of the equation depends on w. For a proof of the existence and uniqueness of these weights see Appendix A.2. □

A.2 Proof of Existence and Uniqueness of the Optimal Portfolio

Optimal portfolio weights are implicitly given by (13) which we rewrite as

$$w(s) = \frac{1}{\gamma}\left(\bar{\Sigma}_T + \frac{\sqrt{\varepsilon}}{\gamma s}\left(\bar{\Omega}_T + \Lambda_T\right)\right)^{-1}\left(\hat{m}_T + \frac{\gamma - \hat{m}_T'\left(\bar{\Sigma}_T + \frac{\sqrt{\varepsilon}}{\gamma s}\left(\bar{\Omega}_T + \Lambda_T\right)\right)^{-1}\mathbf{1}}{\mathbf{1}'\left(\bar{\Sigma}_T + \frac{\sqrt{\varepsilon}}{\gamma s}\left(\bar{\Omega}_T + \Lambda_T\right)\right)^{-1}\mathbf{1}}\mathbf{1}\right), \tag{29}$$

which implies that determining optimal portfolio weights resembles to finding a positive root of the characteristic equation

$$f(s) = w(s)'\left(\bar{\Omega}_T + \Lambda_T\right)w(s) - s^2. \tag{30}$$

To prove the existence of a positive root of (30) we examine the first part of the characteristic equation and define

$$w_k = \frac{1}{\gamma}\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)^{-1}\left(\hat{m}_T + \frac{\gamma - \hat{m}_T'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)^{-1}\mathbf{1}}{\mathbf{1}'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)^{-1}\mathbf{1}}\mathbf{1}\right), \tag{31}$$

with

$$\begin{aligned}
w_0 &= \frac{1}{\gamma}\bar{\Sigma}_T^{-1}\left(\hat{m}_T + \frac{\gamma - \hat{m}_T'\bar{\Sigma}_T^{-1}\mathbf{1}}{\mathbf{1}'\bar{\Sigma}_T^{-1}\mathbf{1}}\mathbf{1}\right), \\
w_\infty = \lim_{k\to\infty} x(k) &= \frac{1}{k\gamma}\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\left(\hat{m}_T + \frac{\gamma - \hat{m}_T'\frac{1}{k}\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}}{\mathbf{1}'\frac{1}{k}\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}}\mathbf{1}\right) \\
&= \frac{1}{k\gamma}\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\left(\hat{m}_T + \frac{k\gamma - \hat{m}_T'\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}}{\mathbf{1}'\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}}\mathbf{1}\right) \\
&= \frac{1}{\mathbf{1}'\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}}\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}.
\end{aligned}$$

Note that $w_{k=\frac{\sqrt{\varepsilon}}{\gamma s}}$ coincides with $w(s)$. Since $\lim_{k\to\infty} w_k$ is the minimum-variance portfolio corresponding to the covariance structure $\left(\bar{\Omega}_T + \Lambda_T\right)$, we conclude that

$$\inf\left\{w_k'\left(\bar{\Omega}_T + \Lambda_T\right)w_k\right\} = w_\infty'\left(\bar{\Omega}_T + \Lambda_T\right)w_\infty = \frac{1}{\mathbf{1}'\left(\bar{\Omega}_T + \Lambda_T\right)^{-1}\mathbf{1}} > 0.$$

The portfolio x_0 is the mean variance portfolio corresponding to $\bar{\Sigma}_T$ and risk aversion γ. Thus, we have

$$0 < w_\infty' \left(\bar{\Omega}_T + \Lambda_T\right) w_\infty \leq w_0' \left(\bar{\Omega}_T + \Lambda_T\right) w_0 < \infty$$

We now set $k = \frac{\sqrt{\varepsilon}}{\gamma s}$ and conclude that

$$\begin{aligned} \lim_{s\to 0} f(s) &= \frac{1}{\mathbf{1}' \left(\bar{\Omega}_T + \Lambda_T\right)^{-1} \mathbf{1}} > 0, \\ \lim_{s\to\infty} f(s) &= w_0' \left(\bar{\Omega}_T + \Lambda_T\right) w_0 - \lim_{s\to\infty} s^2 = -\infty. \end{aligned}$$

Since f is continuous, we conclude that f has a positive root, independent of the parametrization of the model.

The proof of the uniqueness of the positive root is done in three steps:

- First, we restrict the considerations on portfolios w with $w'\hat{m}_T = \text{const}$ and show that if w_0, the minimum-variance portfolio with respect to $\bar{\Sigma}_T$, differs from w_∞, the minimum-variance portfolio with respect to $\left(\bar{\Omega}_T + \Lambda_T\right)$, then $w_k'\bar{\Sigma}_T w_k$ is strictly increasing and $w_k \left(\bar{\Omega}_T + \Lambda_T\right) w_k$ is strictly decreasing in k with $k > 0$.
- Second, we show that within $\hat{m}_T = \text{const}$, non-uniqueness of the root of f implies $\frac{d}{dk}\{w_k' \left(\bar{\Omega}_T + \Lambda_T\right) w_k\} \geq 0$. Thus, excluding the non-interesting stable case we conclude that for $w'\hat{m}_T = \text{const}$, $f(s)$ has a unique root.
- Third, we prove that the frontier in the $\sigma_\Sigma = \sqrt{w'\bar{\Sigma}_T w}$ vs. $\bar{m}_T$ space is concave which implies the uniqueness of the optimal portfolio stated in Proposition 1.

We start by defining $M(m) = \{w \in \mathbb{R}^n | w'\mathbf{1} = 1 \wedge w'\hat{m}_T = m\}$, which is the space of all portfolios that have a given expected return (measured with respect to $\hat{m}_T$, not $\bar{m}_T$). Restricting our considerations first to $M(m)$, w_k is the minimum-variance portfolio on $M(m)$ with respect to the covariance matrix $\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)$.

The characteristic function f from Equation (30) for w restricted to M has a positive root, since $w'(s)\left(\bar{\Omega}_T + \Lambda_T\right) w(s)$ equals $0 < w_\infty' \left(\bar{\Omega}_T + \Lambda_T\right) w_\infty < \infty$ at $s = 0$ and $0 < w_0' \left(\bar{\Omega}_T + \Lambda_T\right) w_0 < \infty$ for $s \to \infty$. Also under the portfolio restriction M, $f(0) > 0$ and $f(s \to \infty) \to -\infty$, hence, there exists at least one positive root.

Any portfolio transaction between two elements of $M(m)$ is orthogonal to w_k. In particular, $\frac{dw_k}{d_k}$ is such a transaction from which we conclude that

$$\begin{aligned} 0 &= \frac{dw_k}{d_k}' \left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k \\ &= \frac{dw_k}{d_k}' \bar{\Sigma}_T w_k + k\frac{dw_k}{d_k}' \left(\bar{\Omega}_T + \Lambda_T\right) w_k. \\ \Rightarrow \quad k\frac{d}{d_k}\left(w_k' \left(\bar{\Omega}_T + \Lambda_T\right) w_k\right) &= -\frac{d}{d_k}\left(w_k'\bar{\Sigma}_T w_k\right). \end{aligned} \tag{32}$$

Thus, for $k > 0$ the sensitivity of the variances of the minimum-variance portfolio w_k measured with respect to $\bar{\Sigma}_T$ and $\left(\bar{\Omega}_T + \Lambda_T\right)$ show opposite sign. Since w_0 is the minimum-variance portfolio with respect to $\bar{\Sigma}_T$ and w_∞ is the minimum-variance portfolio with respect to $\left(\bar{\Omega}_T + \Lambda_T\right)$ (both restricted to $M(m)$) and the variance of w_k is continuous and differentiable in k, $\frac{d}{d_k}(w_k' \left(\bar{\Omega}_T + \Lambda_T\right) w_k)$, we must have at least some k with $\frac{d}{d_k}(w_k' \left(\bar{\Omega}_T + \Lambda_T\right) w_k) < 0$. Now assume that $\frac{d}{d_k}\left(w_k' \left(\bar{\Omega}_T + \Lambda_T\right) w_k\right)$ changes sign for some $k > 0$. Then $\exists k > 0$ with

$$\begin{aligned} \frac{d}{d_k}\left(w_k' \left(\bar{\Omega}_T + \Lambda_T\right) w_k\right) &= \frac{d}{d_k}(w_k'\bar{\Sigma}_T w_k) = 0, \\ \frac{dw_k}{d_k}' \left(\bar{\Omega}_T + \Lambda_T\right) w_k &= \frac{dw_k}{d_k}' \bar{\Sigma}_T w_k = 0, \end{aligned}$$

which implies that $\frac{dw_k}{d_k}$ is a portfolio transaction within the iso-variance ellipsoid of both $\bar{\Sigma}_T$ and $\left(\bar{\Omega}_T + \Lambda_T\right)$. The Lagrangian for the constraint minimization under given expected target return m is

$$L(w, \lambda_1, \lambda_2) = w_k' \left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k + \lambda_1(1 - w_k'\mathbf{1}) + \lambda_2(m - w_k'\hat{m}_T).$$

The first derivative of the the Lagrangian with respect to w_k must be constant equal to zero independent of k, i.e.,

$$\begin{aligned} \frac{d}{dk}\left[\frac{d}{dw_k}[w_k' \left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k]\right] &= -\frac{d}{dk}\left[\frac{d}{dw_k}[\lambda_1(1 - w_k'\mathbf{1}) + \lambda_2(m - w_k'\hat{m}_T)]\right], \\ \frac{d}{dk}\left[2\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k\right] &= \frac{d}{dk}\left[\lambda_1\mathbf{1} + \lambda_2\hat{m}_T\right], \\ 2\left[\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)\frac{dw_k}{dk}' + \left(\bar{\Omega}_T + \Lambda_T\right) w_k\right] &= \frac{d\lambda_1}{dk}\mathbf{1} + \frac{d\lambda_2}{dk}\hat{m}_T. \end{aligned}$$

Let x be a portfolio transaction within $M(m)$, then it follows

$$\begin{aligned} 2\left[x'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)\frac{dw_k}{dk} + x'\left(\bar{\Omega}_T + \Lambda_T\right) w_k\right] &= \frac{d\lambda_1}{dk}\underbrace{x'\mathbf{1}}_{=0} + \frac{d\lambda_2}{dk}\underbrace{x'\hat{m}_T}_{=0}, \\ \Rightarrow \quad x'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)\frac{dw_k}{dk} &= -x'\left(\bar{\Omega}_T + \Lambda_T\right) w_k. \end{aligned} \tag{33}$$

This is true for all portfolio transactions within M and since $\frac{dw_k}{dk}$ is itself a portfolio transaction within M, it follows

$$\frac{dw_k}{dk}'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)\frac{dw_k}{dk} = 0, \quad \Rightarrow \quad \frac{dw_k}{dk} = 0.$$

But Equation (33) is valid for all transactions within M, so we get

$$x'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right)\frac{dw_k}{dk} = -x'\left(\bar{\Omega}_T + \Lambda_T\right) w_k = 0, \forall x \in M,$$

which means that w_k equals w_∞, the minimum-variance portfolio with respect to $\left(\bar{\Omega}_T + \Lambda_T\right)$ in M.

Now show that if at some $k_1 > 0$ we have $\frac{dw_k}{dk}'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k = 0$ (and, hence, $w_{k_1} = w_\infty$), then $w_k = w_\infty$ for all $k \in [k_1, \infty)$. If we assume the contrary, i.e., there is some nonempty range (k_2, k_3) in $[k_1, \infty)$ such that w_k strictly deviates from w_∞ for all $k \in (k_2, k_3)$. Then $w_k'\left(\bar{\Omega}_T + \Lambda_T\right) w_k$ must exceed $w_\infty\left(\bar{\Omega}_T + \Lambda_T\right) w_\infty$, but must return to this value (at least asymptotically) for growing k. So $\frac{dw_k}{dk}' \left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k$ must first be positive and then turn negative, implying (by continuity) that it must change signs which in turn requires some $k \in (k_2, k_3)$ with $w_k = w_\infty$. This is a contradiction to the assumption that w_k strictly deviates from w_∞ in (k_2, k_3). So we know if $\frac{dw_k}{dk}' \left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k = 0$ at some k_1, then w_k is equal to the minimum-variance portfolio with respect to $\left(\bar{\Omega}_T + \Lambda_T\right)$ for all $k \in [k_1, \infty)$.

Next we show that w_k must also equal w_0, i.e., the minimum-variance portfolio with respect to $\bar{\Sigma}_T$, for all $k \in [k_1, \infty)$. Assume the contrary, then $w_k = w_\infty$ differs from w_0 at some $k \in [k_1, \infty)$. Then, $\exists x$ a portfolio transaction within $M(m)$, such that $x'\bar{\Sigma}_T w_\infty \neq 0$. But w_k is the minimum-variance portfolio with respect to $\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)$, so it follows that

$$0 = x'\left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k = x'\bar{\Sigma}_T w_\infty + k \underbrace{x'\left(\bar{\Omega}_T + \Lambda_T\right) w_\infty}_{=0} \neq 0.$$

In a last step, we show that $w_k = w_0$ for all $k > 0$. This is done again by contradiction. Assume there is some nonempty range (k_2, k_3) in $(0, k_1]$ where w_k differs from w_0, thus, we argue that $\frac{dw_k}{dk}'\bar{\Sigma}_T w_k$ must change sign inside (k_2, k_3). Also $\frac{dw_k}{dk}' \left(\bar{\Sigma}_T + k\left(\bar{\Omega}_T + \Lambda_T\right)\right) w_k$ changes sign at the same location, say at $k_4 \in (k_2, k_3)$. But then it follows that $w_k = w_0 = w_\infty \in [k_4, \infty)$, which contradicts our assumption. Finally we know that if $w_0 \neq w_\infty$, $w_k'\bar{\Sigma}_T w_k$ is strictly increasing for $k > 0$ and $w_k\left(\bar{\Omega}_T + \Lambda_T\right) w_k$ is strictly decreasing for $k \geq 0$.

To prove the uniqueness of the root of the characteristic function f restricted to M, we assume that there exist three roots $s_1 < s_2 < s_3$, where f cuts the abscissa from above at s_1 (since it is positive for small s), cuts the abscissa from below at s_2 and again from above at s_3.

Consider roots s_2 and s_3 and define $\sigma^2(s) = w_k'\left(\bar{\Omega}_T + \Lambda_T\right) w_k|_{k=\frac{\sqrt{\varepsilon}}{\gamma s}}$. Then

$$\begin{aligned}
s_2 &< s_3, \\
f(s_2) &= f(s_3) = 0, \\
f(s) &\geq 0, \ \forall s \in [s_2, s_3], \\
\sigma(s)/s &\geq 1, \ \forall s \in [s_2, s_3], \\
k_2 = \frac{\sqrt{\varepsilon}}{\gamma s_2} &> k_3 = \frac{\sqrt{\varepsilon}}{\gamma s_3}, \\
w_{k_3}'\bar{\Sigma}_T w_{k_3} &> w_{k_2}'\bar{\Sigma}_T w_{k_2}.
\end{aligned}$$

Using Equation (32),

$$\begin{aligned}
\underbrace{w_{k_3}'\bar{\Sigma}_T w_{k_3} - w_{k_2}'\bar{\Sigma}_T w_{k_2}}_{>0} &= \int_{k_2}^{k_3} \frac{d}{dk}(w_k'\bar{\Sigma}_T w_k) dk \\
&= 2\int_{k_2}^{k_3} \frac{dw_k}{dk}' \bar{\Sigma}_T w_k dk \\
&= -2\int_{k_2}^{k_3} k\frac{dw_k}{dk}' \left(\bar{\Omega}_T + \Lambda_T\right) w_k dk.
\end{aligned}$$

To integrate the last term by parts, we use

$$\frac{d}{dk}(w_k'(k\left(\bar{\Omega}_T+\Lambda_T\right))w_k) \quad = \quad 2k\frac{dw_k}{dk}'\left(\bar{\Omega}_T+\Lambda_T\right)w_k + w_k'\left(\bar{\Omega}_T+\Lambda_T\right)w_k,$$

and conclude

$$\begin{aligned}
-2\int_{k_2}^{k_3} k\frac{dw_k}{dk}'\left(\bar{\Omega}_T+\Lambda_T\right)w_k dk &= -kw_k'\left(\bar{\Omega}_T+\Lambda_T\right)w_k|_{k_2}^{k_3} + \int_{k_2}^{k_3} w_k'\left(\bar{\Omega}_T+\Lambda_T\right)w_k dk \\
&= k_2 w_{k_2}'\left(\bar{\Omega}_T+\Lambda_T\right)w_{k_2} - k_3 w_{k_3}'\left(\bar{\Omega}_T+\Lambda_T\right)w_{k_3} \\
&\quad + \int_{k_2}^{k_3} w_k'\left(\bar{\Omega}_T+\Lambda_T\right)w_k dk \\
&= \frac{\sqrt{\varepsilon}}{\gamma s_2}s_2^2 - \frac{\sqrt{\varepsilon}}{\gamma s_3}s_3^2 + \int_{s_2}^{s_3}\sigma^2(s)(-\frac{\sqrt{\varepsilon}}{\gamma s^2})ds \\
&= \frac{\sqrt{\varepsilon}}{\gamma}(s_2-s_3) + \frac{\sqrt{\varepsilon}}{\gamma}\int_{s_2}^{s_3}\frac{\sigma^2(s)}{s^2}ds \\
&\leq \frac{\sqrt{\varepsilon}}{\gamma}(s_2-s_3) + \frac{\sqrt{\varepsilon}}{\gamma}\int_{s_2}^{s_3} 1 ds \\
&= \frac{\sqrt{\varepsilon}}{\gamma}(s_2-s_3) + \frac{\sqrt{\varepsilon}}{\gamma}(s_3-s_2) = 0.
\end{aligned}$$

Two distinct roots at s_2 and s_3 with $f(s) \geq 0$ in $[s_2, s_3]$ is not consistent with $\frac{d}{dk}(w_k'\bar{\Sigma}_T w_k)$ strictly positive everywhere. In the case where f touches the abscissa from below at s_2, equivalent considerations lead to $\frac{d}{dk}(w_k'\bar{\Sigma}_T w_k) \leq 0$ at s_2. So we conclude that on $M(m)$, $f(s)$ has a unique positive root, which we denote $k^* = k^*(m)$ and the corresponding portfolio $w_{k^*} = w_{k^*}(m)$.

As a side result we get the derivative of the root of $f(s)$ with respect to ε by applying the implicit function theorem. If s^* is the root of $f(s)$ and $k = \frac{\sqrt{\varepsilon}}{\gamma s}$ then according to the definition (30) it must hold that

$$\begin{aligned}
0 &= \underbrace{\frac{d}{dk}\left(w_k'\left(\bar{\Omega}_T+\Lambda_T\right)w_k\right)}_{<0}\underbrace{\frac{dk}{d\varepsilon}}_{>0}\Bigg|_{s=s^*} + \underbrace{\frac{df}{ds}}_{<0}\Bigg|_{s=s^*}\frac{ds^*}{d\varepsilon} \\
\Rightarrow \frac{ds^*}{d\varepsilon} &= -\frac{\frac{d}{dk}\left(w_k'\left(\bar{\Omega}_T+\Lambda_T\right)w_k\right)}{\frac{df}{ds}} < 0. \qquad (34)
\end{aligned}$$

By construction s^* equals $\sqrt{w_{k^*}'\left(\bar{\Omega}_T+\Lambda_T\right)w_{k^*}}$, hence, an increase in ε leads to a decrease in $\sqrt{w_{k^*}'\left(\bar{\Omega}_T+\Lambda_T\right)w_{k^*}}$.

The final step of our proof is to show that in the $\sigma_\Sigma = \sqrt{w'\bar{\Sigma}_T w}$ vs. $\bar{m}_T$ space the set of all available portfolios is convex. Referring to Equations (19) and (22) we write

$$\bar{m}_T = \hat{m}_T - \frac{\sqrt{\varepsilon}\left(\bar{\Omega}_T+\Lambda_T\right)w}{\sqrt{w'\left(\bar{\Omega}_T+\Lambda_T\right)w}}, \qquad (35)$$

i.e., according to the max-min approach to ambiguity aversion, the expected return $w'\bar{m}_T$

is the minimum among all available priors, for given w. Transferring the results for portfolios restricted to $M(m)$ into the σ_Σ / $\bar{m}_T$ space we find that w_{k^*} maximizes mean-variance utility for all w with $w'\mathbf{1} = 1$. The globally optimal portfolio for given γ is then characterized by one particular w_{k^*}. Assume mean-variance utility is maximized at w, then $w = w_{k^*}(w'\hat{m}_T)$, otherwise we can argue that $w_{k^*}(w'\hat{m}_T)$ has higher mean-variance utility than w.

As a final step of the proof, we show that the portfolio frontier is concave in the σ_Σ / $\bar{m}_T$ space. Take two portfolios, w_1 and w_2 and consider $w(h) = hw_1 + (1-h)w_2$ for $h \in [0,1]$, then due to diversification with respect to $\bar{\Sigma}_T$ and $\left(\bar{\Omega}_T + \Lambda_T\right)$, respectively, we have

$$\begin{aligned}
\sigma_\Sigma(w(h)) &= \sqrt{w(h)'\bar{\Sigma}_T w(h)} \leq h\sigma_\Sigma(w_1) + (1-h)\sigma_\Sigma(w_2), \\
\bar{m}_T(w_h) &= hw_1'\hat{m}_T + (1-h)w_2'\hat{m}_T - \sqrt{\varepsilon}\sqrt{w(h)'\left(\bar{\Omega}_T + \Lambda_T\right)w(h)} \\
&\geq h[w_1'\hat{m}_T - \sqrt{\varepsilon}\sqrt{w_1'\left(\bar{\Omega}_T + \Lambda_T\right)w_1}] \\
&\quad +(1-h)[w_2'\hat{m}_T - w_1'\hat{m}_T - \sqrt{\varepsilon}\sqrt{w_2'\left(\bar{\Omega}_T + \Lambda_T\right)w_2}] \\
&= h\bar{m}_T(w_1) + (1-h)\bar{m}_T(w_2).
\end{aligned}$$

The whole linear section w_1, w_2 in the σ_Σ / $\bar{m}_T$ space is (weakly) dominated by $w(h)$, $h \in [0,1]$. Hence, we conclude that the efficient frontier is concave and consequently the portfolio which satisfies the first-order condition of Proposition 1 is the unique maximum. □

A.3 Proof of Proposition 2

As a side result of the proof of Proposition 1, (34) in Appendix A.2 shows that $w'(\bar{\Omega}_T + \Lambda_T)w$ is monotonically decreasing in ε. Consequently,

$$\frac{\sqrt{\varepsilon}}{\gamma\sqrt{w'(\bar{\Omega}_T + \Lambda_T)w}}$$

is monotonically increasing in ε. Therefore, $\varepsilon > \varepsilon^*(\gamma)$ implies that the ambiguity averse investor overweights uncertainty in expected returns from estimation errors. □

References

Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *The Journal of Finance* 55, 225–264.

Bauer, Michael D, Glenn D Rudebusch, and Jing Cynthia Wu, 2012, Correcting estimation bias in dynamic term structure models, *Journal of Business & Economic Statistics* 30, 454–467.

Bawa, Vijay S, Stephen J Brown, and Roger W Klein, 1979, *Estimation risk and optimal portfolio choice* (North Holland, Amsterdam).

Bekaert, Geert, Robert J Hodrick, and David A Marshall, 1997, On biases in tests of the expectations hypothesis of the term structure of interest rates, *Journal of Financial Economics* 44, 309–348.

Blume, Marshall E., and Donald B. Keim, 2012, Institutional investors and stock market liquidity: Trends and relationships, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2147757 (Accessed on 10/26/2016).

Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado, 1997, Strategic asset allocation, *Journal of Economic Dynamics and Control* 21, 1377 – 1403.

Bullard, James, 2012, The afthermath of the housing bubble, https://www.stlouisfed.org/~/media/Files/PDFs/Bullard/remarks/BullardBipartisanPolicyCenter5June2012Final.pdf (Accessed on 10/26/2016).

Campbell, John Y, Yeung Lewis Chan, and Luis M Viceira, 2003, A multivariate model of strategic asset allocation, *Journal of Financial Economics* 67, 41–80.

Campbell, John Y, and Robert J Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Campbell, John Y, and Samuel B Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *Review of Financial Studies* 21, 1509–1531.

Campbell, John Y, and Luis M Viceira, 2002, *Strategic Asset Allocation: Portfolio Choice for Long-term Investors* (Oxford University Press, USA).

———, 2005, The term structure of the risk-return trade-off, *Financial Analysts Journal* 61, 34–44.

Cieslak, Anna, and Pavol Povala, 2015, Expected returns in Treasury bonds, *The Review of Financial Studies* 28, 2859–2901.

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *The American Economic Review* 95, 138–160.

Dangl, Thomas, and Michael Halling, 2012, Predictive regressions with time-varying coefficients, *Journal of Financial Economics* 106, 157 – 181.

Davis, Morris A, Andreas Lehnert, and Robert F Martin, 2008, The rent-price ratio for the aggregate stock of owner-occupied housing, *Review of Income and Wealth* 54, 279–284.

Diebold, Francis X., and Canlin Li, 2006, Forecasting the term structure of government bond yields, *Journal of Econometrics* 130, 337–364.

Engsted, Tom, and Thomas Q Pedersen, 2012, Return predictability and intertemporal asset allocation: Evidence from a bias-adjusted var model, *Journal of Empirical Finance* 19, 241–253.

——— , 2014, Bias-correction in vector autoregressive models: A simulation study, *Econometrics* 2, 45–71.

Fama, Eugene F., 1970, Multiperiod consumption-investment decisions, *The American Economic Review* 60, 163–174.

——— , and Robert R. Bliss, 1987, The information in long-maturity forward rates, *The American Economic Review* 77, 680–692.

Fama, Eugene F, and Kenneth R French, 1988a, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Fama, Eugene F., and G. William Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115–146.

Flavin, Marjorie, and Takashi Yamashita, 2002, Owner-occupied housing and the composition of the household portfolio, *The American Economic Review* 92, 345–362.

Garlappi, Lorenzo, Raman Uppal, and Tan Wang, 2007, Portfolio selection with parameter and model uncertainty: A multi-prior approach, *Review of Financial Studies* 20, 41–81.

Gorton, Gary B., Fumio Hayashi, and K. Geert Rouwenhorst, 2012, The fundamentals of commodity futures returns, *Review of Finance* 17, 35–105.

Kan, Raymond, and Guofu Zhou, 2007, Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42, 621–656.

Kandel, Shmuel, and Robert F. Stambaugh, 1996, On the predictability of stock returns: An asset-allocation perspective, *The Journal of Finance* 51, 385–424.

Keim, Donald B, and Robert F Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357–390.

Kilian, Lutz, 1998, Small-sample confidence intervals for impulse response functions, *Review of Economics and Statistics* 80, 218–230.

Kim, Tong Suk, and Edward Omberg, 1996, Dynamic nonmyopic portfolio behavior, *Review of Financial Studies* 9, 141–161.

Knight, Frank, 1921, *Risk, Uncertainty, and Profit* (Houghton Miffin, New York).

Li, C Wei, Ashish Tiwari, and Lin Tong, 2016, Investment decisions under ambiguity: Evidence from mutual fund investor behavior, *Management Science* forthcoming.

Marida, Kantilal V., John T. Kent, and John M. Bibby, 1979, *Multivariate Analysis* (Academic Press, London).

Merton, Robert C., 1969, Lifetime portfolio selection under uncertainty: The continuous-time case, *The Review of Economics and Statistics* 51, 247–257.

——— , 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373–413.

——— , 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Meucci, Attilio, 2009, *Risk and Asset Allocation* (Springer Science & Business Media).

Nicholls, D.F., and Alun L. Pope, 1988, Bias in the estimation of autoregressions, *Australian Journal of Statistics* 30, 296–309.

Pastor, Lubos, and Robert F. Stambaugh, 2012, Are stocks really less volatile in the long run?, *The Journal of Finance* 67, 431–477.

Pope, Alun L., 1990, Biases of estimators in multivariate non-Gaussian autoregressions, *Journal of Time Series Analysis* 11, 249–258.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *The Review of Financial Studies* 23, 821–862.

Samuelson, Paul A., 1969, Lifetime portfolio selection by dynamic stochastic programming, *The Review of Economics and Statistics* 51, 239–246.

Sharpe, William F., 1974, Imputing expected security returns from portfolio composition, *The Journal of Financial and Quantitative Analysis* 9, 463–472.

Shiller, Robert J, 2015, *Irrational Exuberance* (Princeton university press).

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Xia, Yihong, 2001, Learning about predictability: The effects of parameter uncertainty on dynamic assetallocation, *The Journal of Finance* 56, 205–246.

Zagst, Rudi, and Michaela Pöschik, 2008, Inverse portfolio optimisation under constraints, *Journal of Asset Management* 9, 239–253.

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley & Sons, New York).

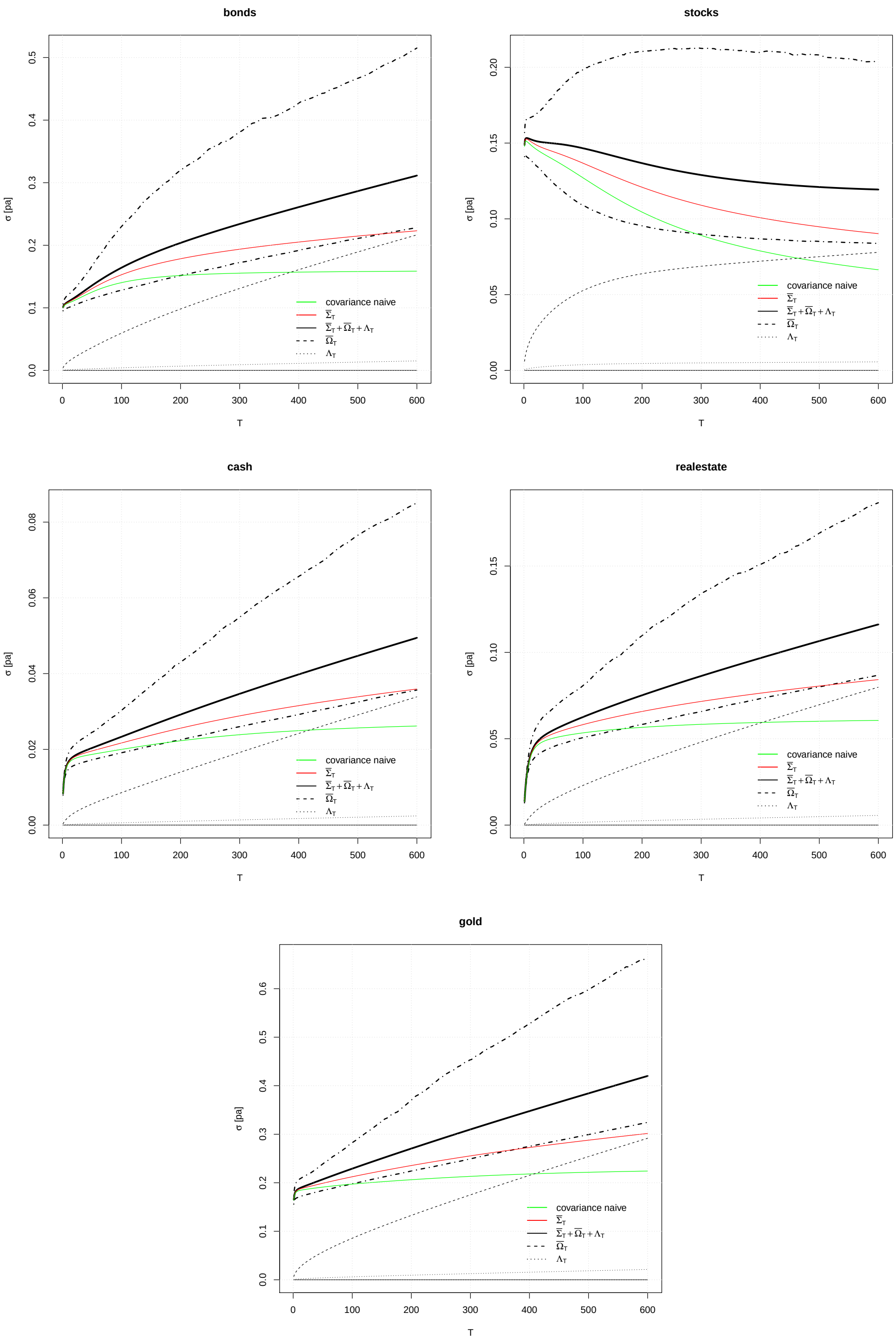

Figure 1: Contribution of overall volatility as proposed in (9) for bonds, stocks, cash, real estate and gold (values indicated in standard deviation p.a.). The solid lines show the term structure of volatility for an investor: (a) who is aware of predictability but neglects parameter uncertainty (green line), (b) who is aware of predictability and parameter uncertainty but takes expected returns as given (red line), and (c) who uses the full predictive variance (black line with black dashed-dotted lines indicating the 95% confidence interval). The dashed (dotted) line indicates the volatility from purely considering $\bar{\Omega}_T$ (Λ_T).

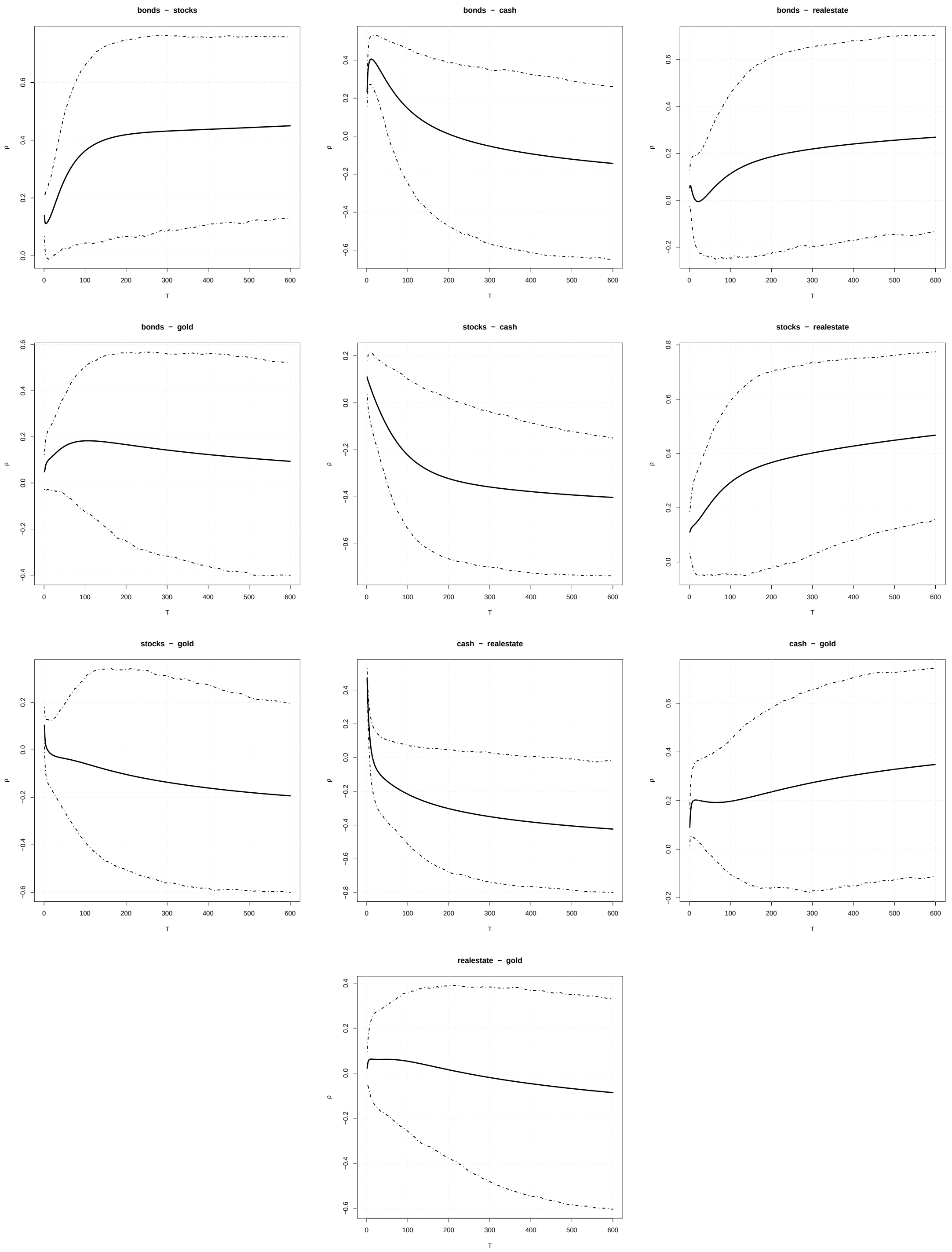

Figure 2: Term structure of return correlations of pairs of asset classes based on the full predictive covariance $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda_T$ (95% confidence bounds as dashed-dotted lines).

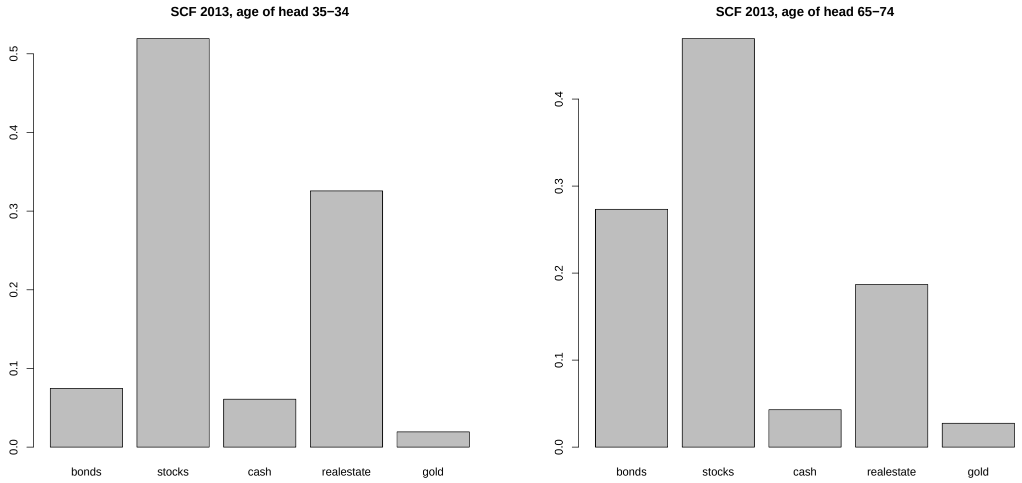

Figure 3: Asset allocation of US families with respect to age of head (Survey of Consumer Finances, 2013).

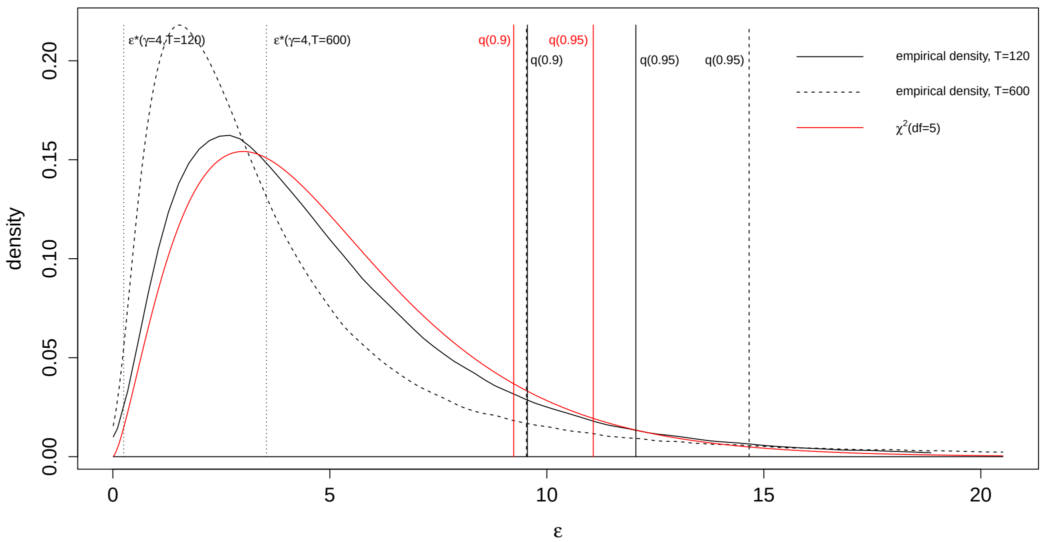

Figure 4: Empirical density of the Mahalanobis distance $(\hat{m}_T - \bar{m}_T)' (\bar{\Omega}_T + \Lambda_T)^{-1} (\hat{m}_T - \bar{m}_T)$ for $T = 120$ months and $T = 600$ months together with the density of a χ^2- distribution with $n = 5$ degrees of freedom. Vertical solid and dashed lines indicate quantiles at the 90% and 95% level, respectively. Vertical dotted lines indicate critical confidence levels ε^* such that the full predictive variance is considered without additional ambiguity aversion.

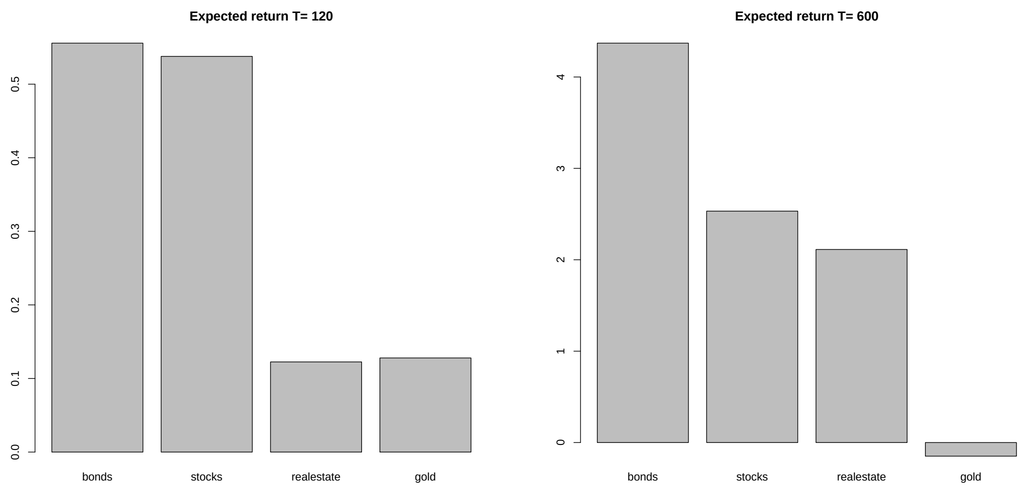

Figure 5: Market implied expected (excess) returns $\hat{m}_T$ relative to the expected return of cash (set equal to zero) given the market capitalizations of the two identified groups from the Survey of Consumer Finance, 2013. The left panel shows results for families with an *age of head* between 65 and 74 and an assumed planning horizon of $T = 120$ months, while the right panel gives corresponding results for families with an *age of head* between 35 and 44 and $T = 600$ months.

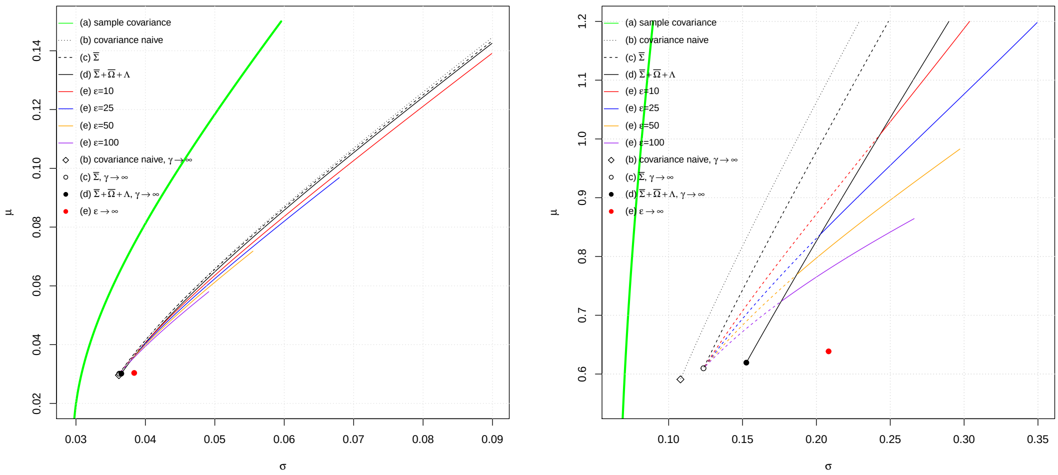

Figure 6: Efficient risk-return combinations of investors with different sophistication level from the viewpoint (a) of an investors who considers only the sample covariance matrix, i.e., neglects predictability and parameter uncertainty. His efficient frontier is indicated with a green line. The other agents consider: (b) only predictability but not parameter uncertainty (black-dotted line), (c) predictability and parameter uncertainty with long-term means as given (black-dashed line), (d) the full predictive variance without ambiguity aversion (black solid line), and (e) the full predictive variance with different degrees of ambiguity aversion (colored-solid line for $\varepsilon > \varepsilon^*$, coloured-dashed line for $\varepsilon < \varepsilon^*$). The diamond and the bullet points indicate the global minimum-variance portfolios of investors (b)-(e). The left panel shows results for $T = 120$, the right panel for $T = 600$.

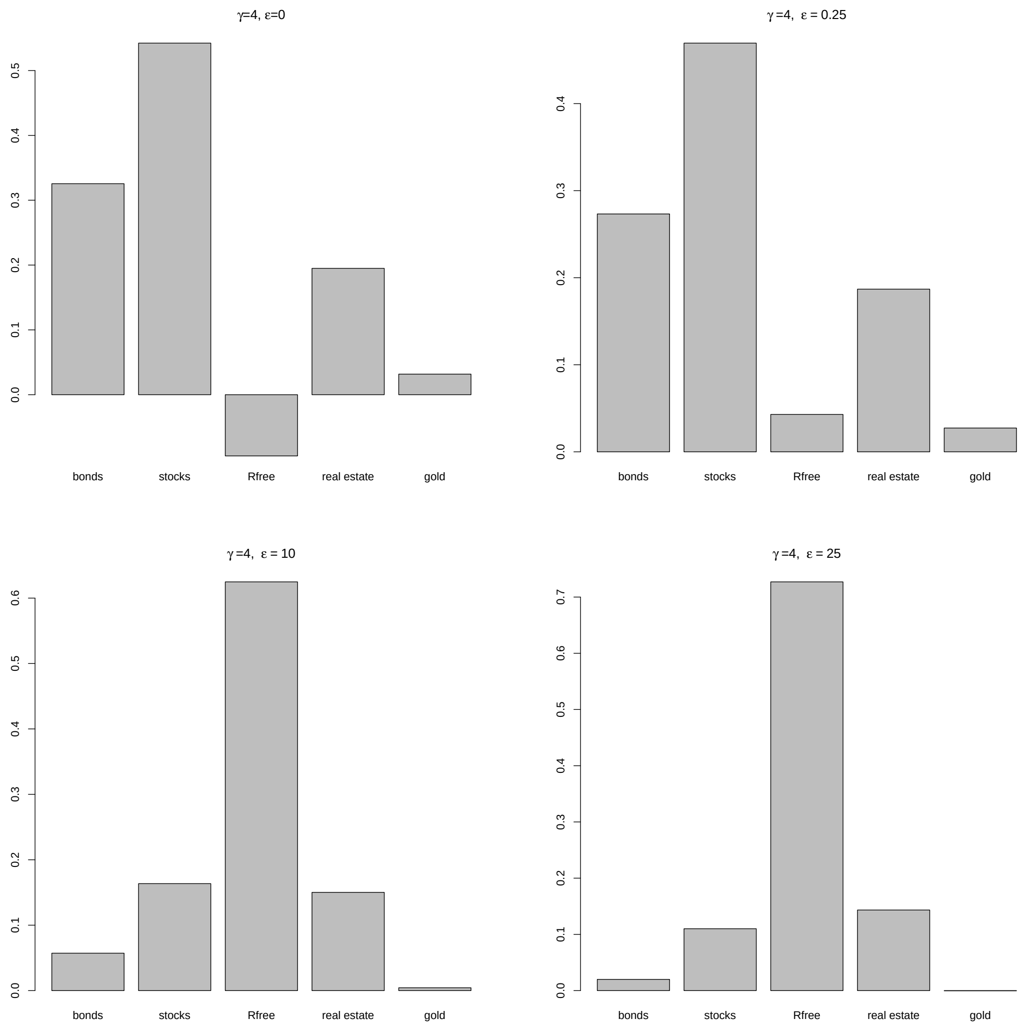

Figure 7: Optimal asset allocation with a risk aversion $\gamma = 4$, and different levels of ambiguity aversion $\varepsilon \in \{0, 0.25, 10, 25\}$ with $T = 120$. The upper-left panel shows the optimal portfolio of an investor who considers only $\bar{\Sigma}_T$, the upper-right panel gives the optimal asset allocation of an investor who considers the full predictive variance $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda$, i.e., $\varepsilon = \varepsilon^*$. The lower panels illustrate the optimal asset allocation of investors who recognize the full predictive variance but are ambiguity averse with respect to uncertainty in expected returns, i.e., $\varepsilon > \varepsilon^*$.

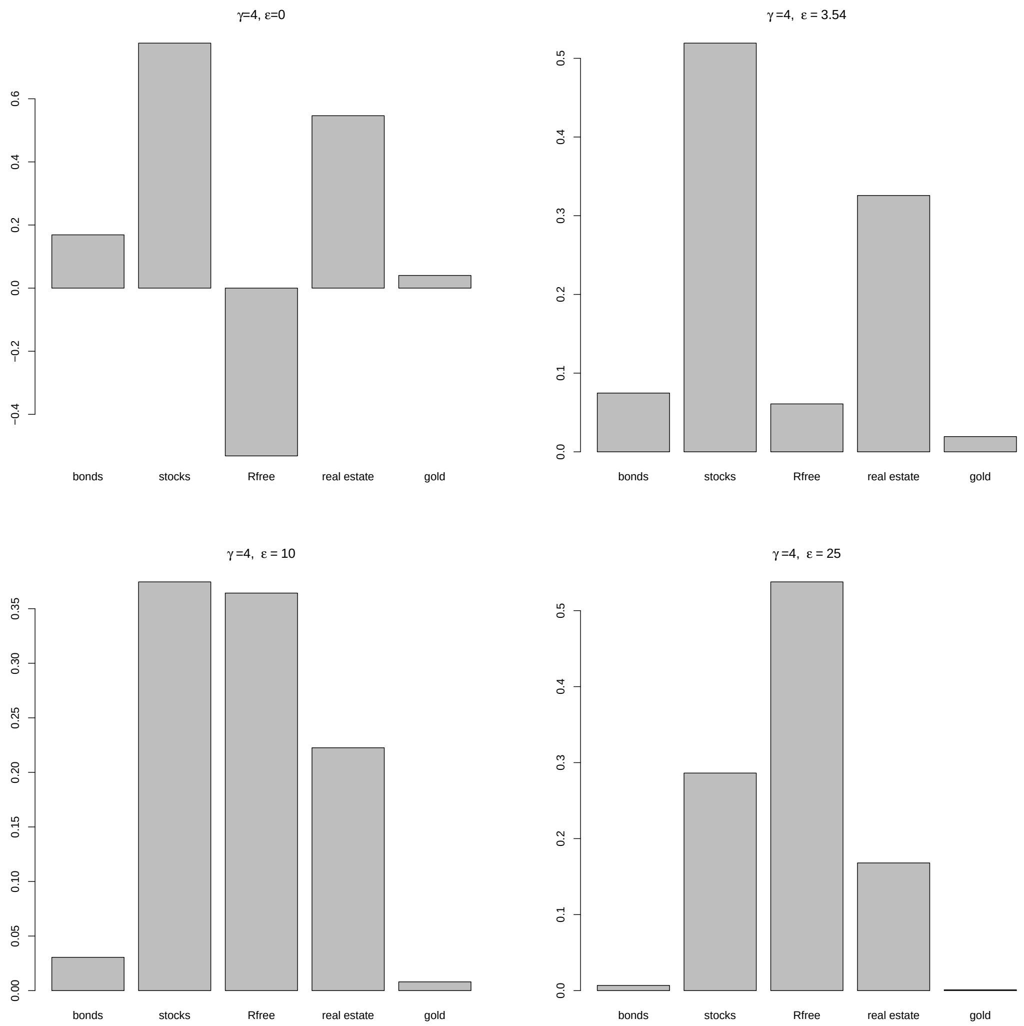

Figure 8: Optimal asset allocation with a risk aversion $\gamma = 4$, and different levels of ambiguity aversion $\varepsilon \in \{0, 3.54, 10, 25\}$ with $T = 600$. The upper-left panel shows the optimal portfolio of an investor who considers only $\bar{\Sigma}_T$, the upper-right panel give the optimal asset allocation of an investor who considers the full predictive variance $\bar{\Sigma}_T + \bar{\Omega}_T + \Lambda$, i.e., $\varepsilon = \varepsilon^*$. The lower panels illustrate the optimal asset allocation of investors who recognize the full predictive variance but are ambiguity averse with respect to uncertainty in expected returns, i.e., $\varepsilon > \varepsilon^*$.

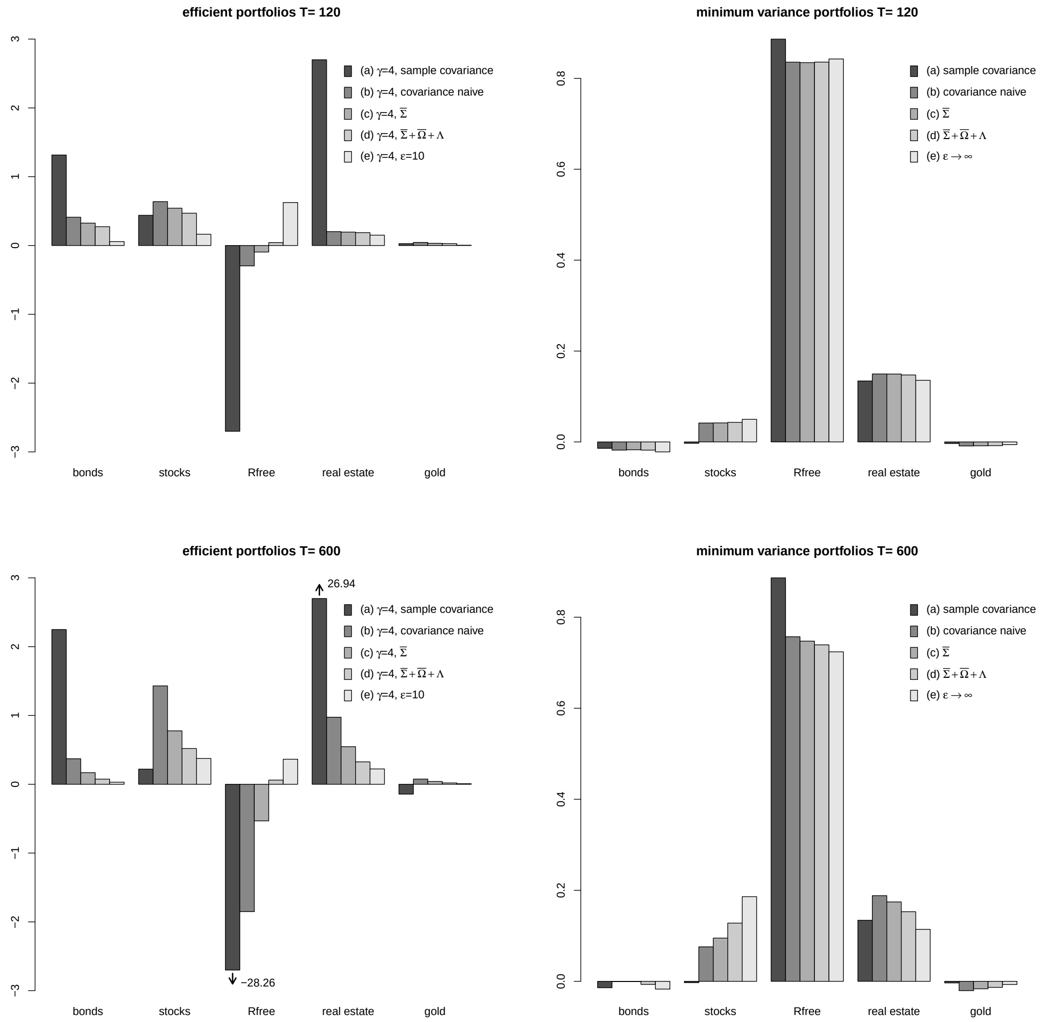

Figure 9: Comparison of optimal portfolios for $\gamma = 4$ (left panels) and global minimum-variance portfolios (right panels) of investors with a planning horizon of $T = 120$ (upper panels) and $T = 600$ (lower panels), who (a) base their investment decision on the sample covariance and neglect predictability and parameter uncertainty, (b) consider predictability in asset returns but neglect parameter uncertainty, (c) optimize their portfolio based on $\bar{\Sigma}_T$, i.e., neglect uncertainty in expected returns, (d) recognize the full predictive variance without ambiguity aversion, and (e) recognize the full predictive variance with ambiguity aversion.

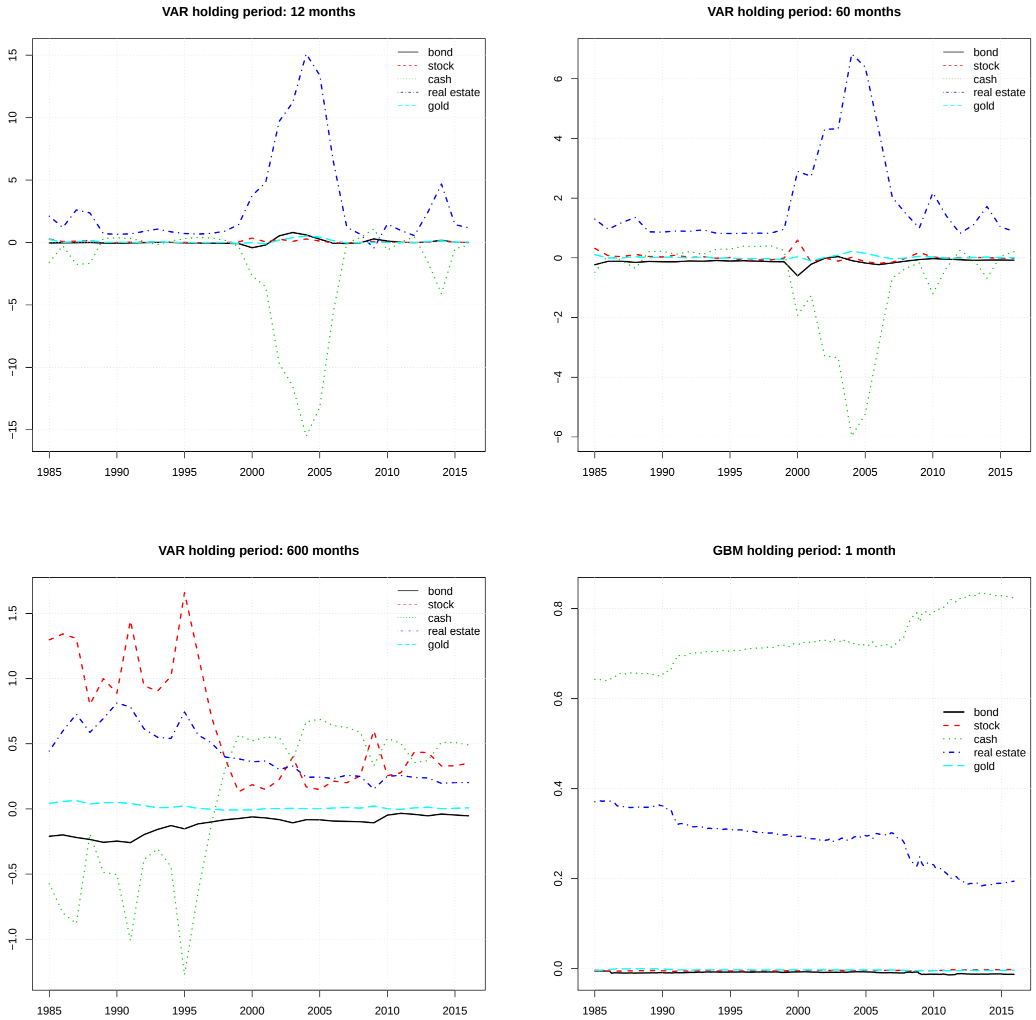

Figure 10: Optimal portfolio weights ($\varepsilon = 25$) for the VAR approach with $T = 12$ (upper-left panel), $T = 60$ (upper-right panel), and $T = 600$ (lower-left panel), and for the GBM approach with $T = 1$ (lower-right panel).

Table 1: OLS based VAR estimation parameters (standard errors in parenthesis)

	Dependent variable:						
	bonds(t)	stocks(t)	cash(t)	real estate(t)	gold(t)	spr(t)	d(t)-p(t)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
bond(t-1)	0.036	0.147^{**}	0.004	−0.001	0.052	0.043^{***}	-0.132^{**}
	(0.039)	(0.058)	(0.003)	(0.005)	(0.064)	(0.006)	(0.059)
stocks(t-1)	-0.105^{***}	0.028	-0.005^{**}	−0.001	−0.058	0.001	−0.042
	(0.027)	(0.039)	(0.002)	(0.004)	(0.043)	(0.004)	(0.040)
cash(t-1)	1.331^{***}	0.588	0.541^{***}	-0.255^{***}	1.219^{*}	−0.033	−0.504
	(0.426)	(0.630)	(0.035)	(0.058)	(0.698)	(0.061)	(0.636)
real estate(t-1)	0.179	0.131	-0.035^{**}	0.743^{***}	0.487	-0.061^{**}	−0.004
	(0.205)	(0.303)	(0.017)	(0.028)	(0.336)	(0.029)	(0.306)
gold(t-1)	−0.015	-0.059^{*}	−0.001	−0.003	0.095^{**}	−0.0004	0.064^{*}
	(0.024)	(0.035)	(0.002)	(0.003)	(0.039)	(0.003)	(0.035)
spr(t-1)	0.306^{***}	0.240^{**}	−0.001	0.015	0.164	0.954^{***}	-0.238^{**}
	(0.079)	(0.116)	(0.007)	(0.011)	(0.129)	(0.011)	(0.117)
d(t-1)-p(t-1)	−0.0001	0.008^{*}	0.0004	−0.00000	0.006	−0.00001	0.992^{***}
	(0.003)	(0.004)	(0.0002)	(0.0004)	(0.005)	(0.0004)	(0.004)
const	−0.005	0.027^{*}	0.002^{**}	0.001	0.018	0.001	−0.024
	(0.010)	(0.015)	(0.001)	(0.001)	(0.017)	(0.001)	(0.015)
Observations	671	671	671	671	671	671	671
R^2	0.056	0.027	0.301	0.533	0.029	0.922	0.988
Adjusted R^2	0.046	0.017	0.293	0.528	0.019	0.922	0.988
Residual Std. Error (df = 663)	0.029	0.043	0.002	0.004	0.047	0.004	0.043
F Statistic (df = 7; 663)	5.659^{***}	2.664^{**}	40.705^{***}	108.172^{***}	2.871^{***}	$1{,}125.882^{***}$	$8{,}015.214^{***}$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: Cross-correlation of residuals. Standard deviation (p.a.) on the main diagonal.

	bond	stock	riskfree	real estate	gold	spr	d-p
bond	0.100	0.137	0.233	0.052	0.048	-0.365	-0.119
stock	–	0.148	0.110	0.111	0.104	0.001	-0.989
riskfree	–	–	0.008	0.470	0.091	-0.147	-0.065
real estate	–	–	–	0.014	0.022	-0.023	-0.087
gold	–	–	–	–	0.164	-0.032	-0.091
spr	–	–	–	–	–	0.014	-0.011
d-p	–	–	–	–	–	–	0.149

Table 3: Elements on the main diagonal of OLS versus Pope estimated variance.

	OLS			Pope		
	$\bar{\Sigma}_{120}$	$\bar{\Omega}_{120}$	Λ_{120}	$\bar{\Sigma}_{120}$	$\bar{\Omega}_{120}$	Λ_{120}
bond	0.213	0.038	0.0002	0.255	0.046	0.0002
stock	0.161	0.028	0.0001	0.178	0.031	0.0002
riskfree	0.005	0.001	0.00000	0.005	0.001	0.00000
realestate	0.031	0.006	0.00003	0.036	0.007	0.00003
gold	0.445	0.082	0.0004	0.472	0.092	0.0005
	$\bar{\Sigma}_{600}$	$\bar{\Omega}_{600}$	Λ_{600}	$\bar{\Sigma}_{600}$	$\bar{\Omega}_{600}$	Λ_{600}
bond	1.809	1.616	0.008	2.490	2.345	0.012
stock	0.332	0.248	0.001	0.407	0.304	0.002
riskfree	0.051	0.044	0.0002	0.065	0.057	0.0003
realestate	0.264	0.245	0.001	0.355	0.319	0.002
gold	3.939	3.551	0.019	4.547	4.255	0.023

Table 4: Ratio of the overall variance between Pope corrected and OLS variance.

	bond	stock	riskfree	realestate	gold
T=120	1.199	1.104	1.131	1.153	1.069
T=600	1.412	1.225	1.282	1.324	1.175

Table 5: Variance contribution for different time horizons of the considered asset classes according to (9). The contribution of estimation errors in expected returns to the total return variance increases with the investment horizon. Estimation errors in expected returns are mainly induced via uncertainty in VAR coefficients a and B ($\bar{\Omega}_T$). Errors in expected returns resulting from estimation errors in the covariance matrix of residuals Σ_ϵ (Λ_T) are of minor importance.

	bonds	stocks	cash	real estate	gold
$\bar{\Sigma}_{120}$	84.56%	84.99%	84.31%	84.27%	83.67%
$\bar{\Omega}_{120}$	15.36%	14.93%	15.61%	15.66%	16.24%
Λ_{120}	0.07%	0.08%	0.08%	0.08%	0.08%
$\bar{\Sigma}_{600}$	51.37%	57.17%	52.84%	52.55%	51.53%
$\bar{\Omega}_{600}$	48.39%	42.61%	46.92%	47.22%	48.21%
Λ_{600}	0.24%	0.22%	0.24%	0.23%	0.26%

Table 6: Certainty equivalent (CE) and moments of out-of-sample returns for the model *with* predictability (VAR) versus the model *without* predictability (GBM) (values are not annualized).

$\varepsilon = 10$						
T		CE	mean	sd	skew	kurt
VAR 12	mean	0.1611	0.4727	0.6908	1.7995	5.7779
	median	0.1645	0.4518	0.6818	1.8403	5.8259
GBM 12	mean	0.0225	0.0239	0.0267	-0.4982	3.2005
	median	0.0226	0.0240	0.0268	-0.4631	2.9703
VAR 24	mean	0.1226	0.5503	0.8393	1.8570	6.2522
	median	0.1304	0.5481	0.8534	1.8624	6.2740
GBM 24	mean	0.0439	0.0489	0.0520	-0.3377	2.5139
	median	0.0442	0.0490	0.0517	-0.3344	2.4991
VAR 60	mean	-0.2473	0.6020	0.9071	0.6276	2.9104
	median	-0.1961	0.5885	0.8976	0.6228	2.9346
GBM 60	mean	0.1060	0.1293	0.1247	-0.1434	1.3858
	median	0.1053	0.1301	0.1258	-0.1052	1.3775

$\varepsilon = 25$						
T		CE	mean	sd	skew	kurt
VAR 12	mean	0.1291	0.2851	0.4714	2.1853	7.2205
	median	0.1241	0.2707	0.4647	2.2356	7.2278
GBM 12	mean	0.0203	0.0216	0.0251	-0.4278	3.2463
	median	0.0204	0.0216	0.0251	-0.3955	3.0774
VAR 24	mean	0.1326	0.2729	0.4571	2.3667	8.1671
	median	0.1308	0.2675	0.4618	2.4134	8.4503
GBM 24	mean	0.0397	0.0440	0.0486	-0.2764	2.5226
	median	0.0399	0.0441	0.0482	-0.2673	2.5132
VAR 60	mean	0.0216	0.2594	0.3998	0.2376	2.7570
	median	0.0329	0.2572	0.4011	0.2060	2.7493
GBM 60	mean	0.0955	0.1161	0.1166	-0.1108	1.4554
	median	0.0948	0.1168	0.1178	-0.0802	1.4490

$\varepsilon = 100$						
T		CE	mean	sd	skew	kurt
VAR 12	mean	0.0415	0.0463	0.0525	1.1598	5.7344
	median	0.0398	0.0445	0.0538	1.1033	5.5811
GBM 12	mean	0.0186	0.0197	0.0240	-0.3634	3.2509
	median	0.0186	0.0197	0.0239	-0.3382	3.1301
VAR 24	mean	0.0586	0.0665	0.0666	-0.1247	2.3938
	median	0.0588	0.0669	0.0669	-0.0668	2.3813
GBM 24	mean	0.0363	0.0402	0.0462	-0.2245	2.5141
	median	0.0364	0.0403	0.0458	-0.2193	2.5273
VAR 60	mean	0.1158	0.1486	0.1486	-0.1417	1.4557
	median	0.1161	0.1488	0.1489	-0.0789	1.3957
GBM 60	mean	0.0868	0.1056	0.1110	-0.0812	1.5334
	median	0.0862	0.1062	0.1124	-0.0607	1.5073