Journal of Banking and Finance 79 (2017) 95–109

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Dynamic portfolio optimization with ambiguity aversion[☆]

Jinqing Zhang[a], Zeyu Jin[a], Yunbi An[b,*]

[a] Institute of Financial Studies, Fudan University, Shanghai, 200433 China
[b] Odette School of Business, University of Windsor, Windsor, Ontario, N9B 3P4 Canada

ARTICLE INFO

Article history:
Received 13 May 2016
Accepted 12 March 2017
Available online 16 March 2017

JEL classification:
G11

Keywords:
Ambiguity aversion
Portfolio optimization
Robust optimization

ABSTRACT

This paper investigates portfolio selection in the presence of transaction costs and ambiguity about return predictability. By distinguishing between ambiguity aversion to returns and to return predictors, we derive the optimal dynamic trading rule in closed form within the framework of Gârleanu and Pedersen (2013), using the robust optimization method. We characterize its properties and the unique mechanism through which ambiguity aversion impacts the optimal robust strategy. In addition to the two trading principles documented in Gârleanu and Pedersen (2013), our model further implies that the robust strategy aims to reduce the expected loss arising from estimation errors. Ambiguity-averse investors trade toward an aim portfolio that gives less weight to highly volatile return-predicting factors, and loads less on the securities that have large and costly positions in the existing portfolio. Using data on various commodity futures, we show that the robust strategy outperforms the corresponding non-robust strategy in out-of-sample tests.

1. Introduction

Portfolio optimization depends crucially on the predicted returns of individual securities; thus, the resulting portfolio may deliver poor out-of-sample performance due to estimation errors in asset returns. For example, if an asset return is overestimated, then the resulting upward-biased position in this asset not only lowers the overall portfolio return but also entails unwarranted transaction costs. Importantly, subsequent adjustments to this biased position may lead to significant turnover and induce further transaction costs. The effect of estimation errors is particularly pronounced for highly frequently-rebalanced dynamic trading strategies, such as momentum and contrarian strategies.

Estimation errors in returns arise because the model used to predict returns might be misspecified, and model parameters have to be estimated based on limited available information. In other words, modelling data-generating processes for returns and their predictors must allow for model and parameter ambiguity or uncertainty, as the complete distributions of returns and return-predicting variables are unknown to investors due to imperfect information (Epstein and Wang, 1994; Hansen and Sargent, 2001). Given the significant impact of estimation errors on portfolio weights and portfolio performance, there is evidence that investors are averse rather than neutral to ambiguity in estimated asset returns (Hansen and Sargent, 2001; Garlappi et al., 2007; Jeong et al., 2015).[1] It is important for active investors to factor ambiguity aversion into the portfolio optimization procedure.

One method adopted in the literature to deal with ambiguity in portfolio optimization is the Bayesian approach (Black and Litterman, 1992; Barberis, 2000; Pástor and Stambaugh, 2012). Under this approach, the predictive distribution of asset returns is recovered by combining the pre-specified prior over the parameters with observations from the data. However, this approach considers only a single prior (Garlappi et al., 2007), and may not be able to produce a stable optimal portfolio when the number of assets is large. The other method to deal with parameter ambiguity is robust optimization, which provides robust decisions in the context of limited distribution information. This approach typically defines a set of distributions that are assumed to include the true distribution of parameters, and then solves for the optimal portfolio based on the worst-case returns that are recovered from the distributions in this set (Epstein and Wang, 1994; Chen and Epstein, 2002; Anderson et al., 2003).

Building on the non-robust framework of Gârleanu and Pedersen (2013) (henceforth, GP), Glasserman and Xu (2013) (henceforth, GX) develop a dynamic portfolio control rule that is robust

☆ We would like to thank Carol Alexander (the editor), two anonymous referees, and the seminar participants at Beihang University for helpful comments. Zhang acknowledges the support from the National Nature Science Funds of China (71471043).

* Corresponding author.
E-mail addresses: zhangjq@fudan.edu.cn (J. Zhang), zeyujin12@fudan.edu.cn (Z. Jin), yunbi@uwindsor.ca (Y. An).

[1] Conceptually, ambiguity aversion is different from risk aversion defined in Markowitz (1952) where the probability of returns is known to investors.

http://dx.doi.org/10.1016/j.jbankfin.2017.03.007
0378-4266/

to model and parameter ambiguity. By acknowledging ambiguity in the estimated model, GX find that the robust trading rules guard against ambiguity by trading less aggressively on signals from return predictors, and can significantly improve performance in out-of-sample tests on historical data. GX distinguish between ambiguity aversion and risk aversion, and document their different effects on the optimal strategy. However, it is unclear how and why ambiguity aversion and risk aversion differ in portfolio selection.

Adopting GX's robust optimization method, this paper investigates portfolio selection in the presence of ambiguity aversion within the framework of GP. We extend GX's model by further distinguishing between ambiguity aversion to returns and ambiguity aversion to return predictors. Based on the Radon–Nikodym theorem, we obtain the explicit forms of the constraints on these two types of ambiguity aversion. Thus, we can separately model ambiguity aversion to returns and to return predictors, and clarify their different effects on the optimal robust strategy. We derive the optimal robust portfolio in closed form, and characterize the properties of the robust trading rules as well as the unique mechanism through which ambiguity aversion impacts the robust strategy.

Our optimal trading strategy gives rise to two principles that are similar to those documented in GP. First, due to transaction costs, the optimal strategy remains to trade partially toward an aim portfolio that minimizes the risk for a given level of return in the whole trading period. However, in our model, the trading rate depends not only on transaction costs and risk aversion, but also on ambiguity aversion. Being ambiguity averse, investors trade faster toward the aim portfolio.

Second, the optimal strategy remains to aim in front of the optimal portfolio derived from Markowitz's (1952) mean-variance model, implying that the aim portfolio is a weighted average of the current Markowitz portfolio and the portfolio with the highest risk-, ambiguity-, and costs-adjusted return in all future periods. However, in the presence of ambiguity aversion, this future portfolio is no longer simply the expected future aim portfolio as documented in GP. Rather, it is a combination of the expected future aim portfolio and the current optimal portfolio.

Importantly, our model gives rise to a third principle that is not present in GP's model, stating that the robust strategy also aims for a low expected loss. Specifically, the aim portfolio loads less on the securities with highly volatile predictors as well as those with large and costly positions in the existing portfolio. Intuitively, highly volatile predictors are more likely to result in large estimation errors in a security's returns, leading to a greatly biased position in the security. There is potential for great losses to arise from these biased positions of the resulting portfolio, particularly when the positions and their associated transaction costs are large. This provides a clear economic interpretation for the role that ambiguity aversion plays in portfolio selection.

Using data on commodity futures, we illustrate that the robust strategy indeed outperforms the non-robust strategy in out-of-sample tests. Additionally, the improvement in performance is particularly pronounced when transaction costs and the level of predictor variability are high. We show that as a result of mitigating the effect of ambiguity, the robust portfolio has smaller and more stable positions than the corresponding non-robust portfolio. Meanwhile, simply scaling down the positions of the non-robust portfolio is not able to achieve the effect of ambiguity aversion on the optimal portfolio choice.

Our research contributes to the literature on portfolio selection and asset pricing with ambiguity aversion in three respects. First, it complements the literature on ambiguity aversion and dynamic asset allocation decisions. Maenhout (2004) considers a dynamic portfolio problem of an investor who is averse to model ambiguity in addition to market risk, and seeks robust decisions within the framework of Anderson et al. (2003). He finds that ambiguity aversion dramatically decreases the optimal share of the portfolio allocated to equities as a result of the high risk premium demanded by ambiguity-averse investors. Garlappi et al. (2007) develop a model with multiple priors and ambiguity aversion, and find that the portfolios delivered by their model tend to over-weight safe assets in the optimal allocations. Branger et al. (2013) analyze the optimal stock-bond portfolio under both learning and ambiguity aversion, and find that both learning and ambiguity aversion impact the size of the stock holdings and also induce some additional hedging demand for the uncertainty due to learning and ambiguity aversion.

Previous studies generally confirm Chen and Epstein's (2002) conjecture that ambiguity aversion and risk aversion are substitutes for each other. By distinguishing between ambiguity aversion to returns and ambiguity aversion to return predictors, our model shows that while return ambiguity aversion and risk aversion impact the robust portfolio in the same fashion, the impacts of predictor ambiguity aversion and risk aversion differ. If there is solely ambiguity aversion to returns, our results are largely similar to those of GP's model with a higher degree of risk aversion. Thus, ambiguity aversion and risk aversion are not substitutes for each other due to the presence of return predictor ambiguity aversion.

Second, this paper contributes to the literature by analyzing the unique mechanisms through which ambiguity aversion helps improve the performance of the robust trading strategy. While previous studies (Garlappi et al., 2007; DeMiguel and Nogales, 2009; Glasserman and Xu, 2013) document the superior performance of robust portfolios, the key drivers of the superior performance are not theoretically analyzed. In contrast, we examine the roles of return predictor variability and transaction costs in shaping ambiguity-averse investors' dynamic trading behavior. In particular, our model shows that investors prefer assets with low predictor variability, as estimation uncertainty is directly associated with the variability of return predictors. We demonstrate that the aim portfolio loads less on the assets with large and costly positions in the existing portfolio, in an effort to reduce the potential loss due to estimation errors. Our research provides economic interpretations for investors' trading behavior with ambiguity aversion, and clearly explains why the robust strategy is able to outperform the non-robust strategy in out-of-sample tests.

Finally, our analysis provides insight into asset pricing with ambiguity aversion. Previous studies show both theoretically and empirically that ambiguity aversion, in addition to risk, affects optimal portfolio choices and, ultimately, equilibrium asset prices (Anderson et al., 2003; Maenhout, 2004; Epstein and Schneider, 2008; Jeong et al., 2015). One underlying assumption of these studies is that risk is the channel through which ambiguity aversion impacts asset pricing. For example, Jeong et al. (2015) consider asset pricing models with stochastic differential utility incorporating ambiguity aversion, and find that models with ambiguity aversion have lower relative risk aversion than models that ignore ambiguity aversion. On the other hand, Anderson et al. (2009) show that ambiguity seems to be different from risk and seems to have a different effect on returns than does risk. Our analysis furthers our understanding of how the ambiguity-return relationship and risk-return relationship differ, and implies that the ambiguity-return relationship hinges on both the variability of return predictors and transaction costs. For a given level of ambiguity aversion, high predictor variability and transaction costs are associated with high ambiguity premium.

The remainder of this paper is organized as follows. Section 2 presents the model and characterizes the optimal robust trading strategy. Section 3 analyzes the properties of the optimal robust strategy and illustrates its trading principles. Section 4 provides a numerical analysis, while Section 5 concludes the paper.

2. Model and optimal trading strategy

In this section, we first review GP's dynamic model. Then, we incorporate ambiguity aversion into the model by allowing estimation errors in both return and return predictor dynamics. Finally, following GX, we derive the optimal robust trading strategy, using the robust optimization approach.

2.1. The model with no ambiguity

We consider an investor who has access to S securities traded at discrete time $t \in \{0, 1, 2, \ldots\}$. The probability space is $(\Omega, F, \{F_t\}, P)$, where Ω is the sample space, $\{F_t\}$ is a filtration, and P is the real probability measure. The securities' rates of return from time t to $t+1$ in excess of the risk-free rate are indicated by an S-column return vector r_{t+1} (henceforth, returns) given by:[2]

$$r_{t+1} = Bf_t + u_{t+1}, \tag{1}$$

where f_t is a K-column vector of return predictors, and B is an $S \times K$ matrix of factor loadings. $u_1, u_2, \ldots, u_t, \ldots$ are i.i.d random vectors, each of which follows a multivariate normal distribution with mean zero and covariance matrix Σ_u.

Eq. (1) implies the investor is able to forecast securities' returns r_{t+1}, based on the available information f_t at time t, whereas B represents the influence of f_t on the predicted returns. Further, f_t is assumed to evolve according to the following mean-reverting process:

$$f_{t+1} = \Phi f_t + v_{t+1}, \tag{2}$$

where Φ is a $K \times K$ matrix of mean-reversion coefficients for the predictors, satisfying $\Phi > 0$ as well as $I - \Phi > 0$ to ensure a stationary process.[3] $v_1, v_2, \ldots$ are the i.i.d shocks affecting the predictors, each of which is assumed to follow a multivariate normal distribution with mean zero and covariance matrix Σ_v. In addition, we assume that v_t is independent of u_t at any time t. Thus, $Cov(u_t, v_t) = \begin{pmatrix} \Sigma_u & 0 \\ 0 & \Sigma_v \end{pmatrix}, \forall t.$

Let x_t denote the vector of shares of the securities invested in a portfolio at time t. Rebalancing the portfolio holdings from x_{t-1} to x_t incurs transaction costs. Following the literature, we assume that the transaction cost associated with this trading is given by:

$$TC_t = \frac{1}{2}\Delta x_t^T \Lambda \Delta x_t, \tag{3}$$

where $\Delta x_t = x_t - x_{t-1}$, and Λ is the cost matrix, which is symmetric and positive definite. It means that a transaction of Δx_t shares leads to a change in the average price by $\Lambda \Delta x_t/2$, yielding a total transaction cost of $\frac{1}{2}\Delta x_t^T \Lambda \Delta x_t$. This cost specification is consistent with that specified in GP and GX, and is motivated partly by tractability.

In GP's model, the investor selects the dynamic trading strategy x_t $(t = 0, 1, 2, \cdots)$ to maximize the present value of all future risk-adjusted returns, net of transaction costs:

$$\max_{x_0, x_1, \cdots} E_0\left\{\sum_t \rho^{t+1}\left(x_t^T r_{t+1} - \frac{\gamma}{2}x_t^T \Sigma_u x_t - \frac{\rho^{-1}}{2}\Delta x_t^T \Lambda \Delta x_t\right)\right\}, \tag{4}$$

subject to $r_{t+1} = Bf_t + u_{t+1}$,

$f_{t+1} = \Phi f_t + v_{t+1}$,

where $\Delta x_0 = 0$, $\rho \in (0, 1)$ is the discount factor, and $\gamma > 0$ measures the investor's risk aversion. Problem (4) is considered a guide to select sensible strategies rather than a precise representation of an investor's preference.

2.2. The model in the presence of ambiguity aversion

Now, we incorporate ambiguity aversion into GP's model. Assume that $\tilde{F}_t$ represents the information that is known by investors at time t. Following GX, we assume that the estimation errors in returns, denoted by $e_{u,t+1}$, and the estimation errors in return predictors, denoted by $e_{v,t+1}$, can be written as follows:

$$e_{u,t+1} = \hat{B}f_t - Bf_t = \tilde{E}_t(r_{t+1}) - E_t(r_{t+1}) = \tilde{E}_t(u_{t+1}), \tag{5}$$

$$e_{v,t+1} = \hat{\Phi}f_t - \Phi f_t = \tilde{E}_t(f_{t+1}) - E_t(f_{t+1}) = \tilde{E}_t(v_{t+1}), \tag{6}$$

where $\hat{B}$ and $\hat{\Phi}$ represent the estimates of B and Φ, respectively. $\tilde{E}_t(\cdot)$ is the conditional expectation operator with respect to the probability measure $\tilde{P}_t$ that is induced by $\tilde{F}_t$.

Eqs. (5) and (6) imply that the underlying reason for estimation errors is the limited information about the distributions driving returns and return predictors. To illustrate this point, we take Eq. (5) as an example. Since a partial sample rather than the full sample is used when regressing the return of a security on its predictors f, the estimated coefficients $\hat{B}$ in the regression are likely biased. Thus, the predicted return $\tilde{E}_t(r_{t+1}) = Bf_t + \tilde{E}_t(u_{t+1}) = \hat{B}f_t$ based on these biased loadings can deviate greatly from its true value $E_t(r_{t+1}) = Bf_t$, and their difference measures the estimation error $\tilde{E}_t(u_{t+1})$. Alternatively, we can interpret return predictors as various indicators for assessing the expected return of a security. Given limited knowledge about the market, investors are not able to accurately evaluate each indicator's ability to predict returns, and thus, their assessment is only an approximation of reality. Note that the estimation errors defined in Eqs. (5) and (6) can arise from misspecification in models or in the data-generating processes.

Next, we constrain the size of estimation errors to reflect investors' aversion to ambiguity in our model. To this end, we constrain the difference between the two probability measures $\tilde{P}_t$ and P_t, as this in turn implies that we constrain the difference between $\tilde{E}_t(r_{t+1})$ and $E_t(r_{t+1})$ as well as the difference between $\tilde{E}_t(f_{t+1})$ and $E_t(f_{t+1})$. By Girsanov's theorem, there exists z_t such that both $\tilde{E}_t(u_{t+1}) = E_t(z_t u_{t+1})$ and $\tilde{E}_t(v_{t+1}) = E_t(z_t v_{t+1})$ hold, where z_t is the Radon–Nikodym derivative of $\tilde{P}_t$ with respect to P_t, or $z_t = d\tilde{P}_t/dP_t$. Following Hansen and Sargent (2001) and Anderson et al. (2003), we constrain the difference between the probability measures by making the relative entropy of the change of measures satisfy $E_t(z_t \log z_t) < \eta$, where η is a constant.

As u_t and v_t are both assumed to be normally distributed, using the change of measure theory, we can prove that the following lemma holds true:[4]

Lemma 1. *At any time t, z_t has a unique form:*

$$z_t = \prod_{i=u,v} \exp\left(e_{i,t+1}^T \Sigma_i^{-1} i_{t+1} - \frac{1}{2}e_{i,t+1}^T \Sigma_i^{-1} e_{i,t+1}\right). \tag{7}$$

Given Eq. (7), the relative entropy constraint $E_t(z_t \log z_t) < \eta$ becomes:

$$\frac{1}{2}e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} \le \eta_1, \tag{8}$$

$$\frac{1}{2}e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} \le \eta_2, \tag{9}$$

[2] A security's return is defined as $r_{t+1} = (p_{t+1} - p_t)/p_t - r_f$, where p_t and p_{t+1} are the security's prices at time t and $t+1$, respectively, and r_f is the risk-free rate for the period from time t to $t+1$. If the security is a share of stock that pays dividends, then p_t is the actual price at time t, and p_{t+1} is the actual price at time $t+1$ plus the total cash dividends from time t to $t+1$.

[3] I represents the identity matrix. $\Phi > 0$ means that matrix Φ is positive definite.

[4] The proofs of this lemma and the following propositions are presented in the Appendix.

where η_1 and η_2 are two constants, satisfying $\eta_1 + \eta_2 = \eta$.

In contrast with GX, we obtain the explicit unique form of the Radon–Nikodym derivative z_t, and split the relative entropy constraint into two ambiguity constraints. This split allows us to specify distinct degrees of ambiguity aversion to returns and to their predictors in our model,[5] and thereby to clarify the different effects of these two types of ambiguity aversion on the optimal trading strategy.

Inequalities (8) and (9) restrict the true values of expected returns and return predictors to lie in an ellipsoid centered at their respective estimated values. Compared with constraints such as $|e_{u,t+1}| < \eta_1$ and $|e_{v,t+1}| < \eta_2$, these ellipsoidal constraints can capture the impacts of securities' correlations on estimation errors. Similarly, Goldfarb and Iyengar (2003), Garlappi et al. (2007), and Delage and Ye (2010) consider $(E_t(r_{t+1}) - \hat{r}_{t+1})^T \Sigma_r^{-1} (E_t(r_{t+1}) - \hat{r}_{t+1}) \le \varepsilon$ as a constraint in robust portfolio optimization, and show that such a constraint is practically relevant. Lemma 1 provides a theoretical explanation for such a specification of constraints in the previous studies.

Finally, being ambiguity averse, investors solve the following robust control problem[6]:

$$\max_{x_0,x_1,\cdots} \min_{e_u,e_v} \tilde{E}_0 \left\{ \sum_{t=0}^{\infty} \rho^{t+1} \tilde{E}_t \left[x_t^T r_{t+1} - \frac{\gamma}{2} x_t^T \Sigma_u x_t - \frac{\rho^{-1}}{2} \Delta x_t^T \Lambda \Delta x_t + \frac{1}{2\theta_1} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} + \frac{1}{2\theta_2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} \right] \right\}, \tag{10}$$

subject to $r_{t+1} = Bf_t + u_{t+1}$,

$f_{t+1} = \Phi f_t + v_{t+1}$,

where $\theta_1 > 0$ and $\theta_2 > 0$ are the coefficients of aversion to ambiguity about returns and their predictors, respectively.[7]

Problem (10) is derived from Problem (4) with robust optimization and two additional constraints Eqs. (8) and (9) as the ambiguity sets, where $1/\theta_1$ and $1/\theta_2$ are the two Lagrange multipliers. As a function of $\eta_1(\eta_2)$, $\theta_1(\theta_2)$ translates the ambiguity set into a penalty term, and a higher value of $\theta_1(\theta_2)$ corresponds to a higher $\eta_1(\eta_2)$. In contrast with Problem (4), Problem (10) incorporates ambiguity aversion to returns and to their predictors, and ensures that even if the worst case in the ambiguity sets occurs, performance is still maximized (Gilboa and Schmeidler, 1989; Hansen and Sargent, 2001; Anderson et al., 2003). The occurrence of any other cases in the ambiguity sets will not make the performance worse, and thus the investor does not need to deviate from the optimal strategy given the ambiguity sets. Such a strategy is actually just sub-optimal, but is robust to estimation uncertainty.

We take return ambiguity as an example to provide an economic interpretation for the model. Since the investor is not completely confident about the estimates of expected returns $\tilde{E}_t(r_{t+1})$, for a given confidence level $c \in (0, 1)$, he/she first specifies an ambiguity set $e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1}/2 \le \eta_1$, such that the true values of expected returns $E_t(r_{t+1})$ lie in this set. If the investor is less confident about the estimated returns (a lower c), he/she considers $E_t(r_{t+1})$ to be likely farther away from $\tilde{E}_t(r_{t+1})$. In this case, he/she tends to set a larger η_1, making the worst case even worse than the worst case with a smaller η_1.[8] Thus, the corresponding optimal robust strategy becomes more conservative than before. As noted by Garlappi et al. (2007) and Cao et al. (2005), this means that the investor is more concerned about the estimation errors, or is more averse to ambiguity in returns.

The above interpretation also helps describe how θ_1 and θ_2 are determined. Take θ_1 as an example. Since a low confidence level c induces a large ambiguity set, the critical value η_1 should satisfy $m(e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1}/2 \le \eta_1) = 1 - c$, where the measuring function $m(\cdot)$ represents the size of the set. Given the facts that c is a positive number below 1 and that $e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1}$ follows a chi-square distribution, we use probability measure $\tilde{P}_t$ as the measuring function m for size consistency. Then, η_1 can be calculated under a chi-square cumulative distribution function. Accordingly, θ_1 can be determined, using Karush-Kuhn-Tucker conditions.[9] This is in sharp contrast with GX, where ambiguity aversion coefficients are arbitrarily set without using any particular method.

2.3. The optimal trading strategy with ambiguity aversion

Using robust dynamic programming (Iyengar, 2005; Hansen et al., 2006; Glasserman and Xu, 2013), we can solve Problem (10). The following proposition characterizes the conditions under which the solution exists, and also characterizes the optimal robust trading strategy and its corresponding value function.

Proposition 1. *If the following two conditions are both satisfied,*

Condition 1. $\Sigma_v^{-1} + 2\rho\theta_2 A_{ff} > 0$,

Condition 2. $J_1 = (\gamma + \theta_1)\Sigma_u + \rho^{-1}\Lambda + \rho A_{xx} + \rho^2\theta_2 A_{xf}(\Sigma_v^{-1} + 2\rho\theta_2 A_{ff})^{-1} A_{xf}^T > 0$, *then Problem (10) has a unique solution, which is given by:*

$$x_t^* = J_1^{-1}\left(Bf_t + \rho A_{xf}(I + 2\rho\theta_2\Sigma_v A_{ff})^{-1}\Phi f_t + \rho^{-1}\Lambda x_{t-1}\right), \tag{11}$$

where the coefficient matrices $A_{xx} > 0$, A_{xf}, and A_{ff} are jointly determined by the following equations:

$$\begin{aligned} A_{xx} &= \bar{\Lambda} - \bar{\Lambda}^T J_1^{-1} \bar{\Lambda}, \\ A_{xf} &= \bar{\Lambda} J_1^{-1}(B + \rho A_{xf}(I + 2\rho\theta_2\Sigma_v A_{ff})^{-1}\Phi), \\ A_{ff} &= \tfrac{1}{2}(\bar{\Lambda}^{-1}A_{xf})^T J_1 (\bar{\Lambda}^{-1}A_{xf}) + \rho\Phi^T A_{ff}(I + 2\rho\theta_2\Sigma_v A_{ff})^{-1}\Phi. \end{aligned} \tag{12}$$

The value function is given by:

$$V(x_{t-1}, f_t) = -\frac{1}{2} x_{t-1}^T A_{xx} x_{t-1} + x_{t-1}^T A_{xf} f_t + f_t^T A_{ff} f_t + A_0. \tag{13}$$

The corresponding conditional expectations of u_{t+1} and v_{t+1} are given by:

$$e_{u,t+1}^* = -\theta_1 \Sigma_u x_t^*, \tag{14}$$

$$e_{v,t+1}^* = -\rho\theta_2\left(\Sigma_v^{-1} + 2\rho\theta_2 A_{ff}\right)^{-1}\left(A_{xf}^T x_t^* + 2A_{ff}\Phi f_t\right). \tag{15}$$

[5] In the empirical analysis, GX consider the cases in which investors are only averse to return ambiguity or predictor ambiguity. For instance, they consider the case in which investors are only averse to predictor ambiguity by setting $\tilde{E}_t(r_{t+1}) = Bf_t$, meaning that there are no estimation errors in returns. This method is not able to truly distinguish between ambiguity aversion to returns and ambiguity aversion to return predictors.

[6] Consistent with GP, we focus on an infinite time horizon, because the optimality equations are easier to solve than they are in the finite horizon case (Merton, 1969). In addition, we can simplify our analysis by avoiding issues such as how to deal with the investment upon termination in the finite horizon case.

[7] These ambiguity coefficients are put in the denominator in Problem (10) to ensure that high values of θ_1 and θ_2 correspond to high values of η_1 and η_2.

[8] We can prove that the worst case in the case of a higher ambiguity coefficient is worse than before. With the penalty term, the objective function includes a quadratic function of $e_{u,t+1}$ ($e_{v,t+1}$). If the investor's degree of ambiguity aversion θ_1 (θ_2) increases, the minimum value of the quadratic function decreases.

[9] It is noteworthy that $e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1}$ follows a chi-square distribution because $e_{u,t+1}$ approximately follows a multivariate normal distribution with mean zero and covariance matrix associated with Σ_u (Wooldridge, 2013). In addition, we postulate a linear relation between the confidence level and the size of the ambiguity set, and let the measuring function be a probability measure. This is just one of the possible methods to calculate the critical value η_1. The issue of how to accurately determine η_1 deserves more research.

From the proof of this proposition, we see that Condition 1 guarantees the existence of the worst performance associated with return predictors, while Condition 2 ensures that the objective function is concave in $x_{t.}$ The economic interpretation is as follows. If Condition 1 does not hold, then for at least one return predictor, investors are not able to find the values that make the performance worst. In other words, performance deteriorates as the estimation errors in these predictors increase. If Condition 2 is not satisfied, then for some assets, as their portfolio weights increase, the performance of the portfolio improves. In this case, there are no optimal holdings for these assets in the trading strategy: the greater the holdings of these assets, the better. However, such a strategy is not truly robust, as the portfolio returns are incorrectly adjusted for ambiguity. This is because $e^*_{v,t+1}$ determined by Eq. (14) is not the worst-case scenario for some predictors, given the fact that Condition 1 is also not satisfied in this case.[10]

Coefficient matrices in Eq. (12) are generally different from those in GP. If $\eta_1 = \eta_2 = 0$, then investors are no longer ambiguity averse. In this case, the last two terms that are used to penalize ambiguity in the objective function in Problem (10) disappear, and the model reduces to GP's model. From Proposition 1, we see that θ_1 solely impacts the first term in the expression of J_1, changing it from $\gamma\Sigma_u$ to $(\gamma+\theta_1)\Sigma_u$. However, θ_2 impacts all coefficient matrices J_1, A_{xf}, and A_{ff}.[11] This shows that if there is solely return ambiguity aversion, our results are largely similar to those of GP's model. The different impact of ambiguity aversion on the optimal trading strategy arises primarily from ambiguity aversion to predictors. If we replace both θ_1 and θ_2 with θ in these coefficient matrices, then they are the same as those in GX. Thus, unlike GX's model, our model clarifies the distinct impacts of the two types of ambiguity aversion on the optimal trading strategy.

The value function $V(x_{t-1}, f_t)$ represents the maximum benefits, adjusted for transaction costs, risk, and ambiguity, of all the future portfolios starting from time t, given the portfolio holdings at time t – 1 and the predicting factors f_t at time t. Note that Eq. (13) and the value function in GP's model are different, although both look like the same. This is because the coefficient matrices with ambiguity aversion differ from those in GP's model.

Eqs. (14) and (15) represent the estimation errors in returns and predictors, respectively, under the worst-case scenario. However, these are not the actual estimation errors. Rather, they are the largest possible estimation errors viewed by investors, as these errors are affected by ambiguity aversion coefficients θ_1 and θ_2. As we can see from Eqs. (14) and (15), these estimation errors are also positively related to the return variability Σ_u and predictor variability Σ_v. This explains why conditional volatilities of asset returns play an important role in measuring premiums driven by ambiguity aversion (Jeong et al., 2015).

3. Properties of the optimal trading strategy with ambiguity aversion

We now investigate the properties of the optimal strategy determined by Problem (10). By examining Eq. (11), we obtain three trading principles, which have important practical implications. The first two principles are similar to those in GP, but are modified to reflect the impact of ambiguity aversion. The third one is a new principle, which is not present in GP.

3.1. The basic trading principles

Proposition 2. *Trade partially toward the aim.*

(i) *The optimal portfolio can be written as:*

$$x_t^* = \left(I - (\kappa+\bar{\Lambda})^{-1}\kappa\right)x_{t-1} + (\kappa+\bar{\Lambda})^{-1}\kappa\cdot aim_t, \tag{16}$$

where $\bar{\Lambda} = \rho^{-1}\Lambda, \kappa = (\gamma+\theta_1)\Sigma_u + \rho A_{xx}$, *and* $aim_t = \kappa^{-1}(Bf_t + \rho A_{xf}\tilde{E}_t^*(f_{t+1}))$. $\tilde{E}_t^*(\cdot)$ *is the conditional expectation operator under the worst-case scenario.*

(ii) *The weight of the aim portfolio is given by:*

$$(\kappa+\bar{\Lambda})^{-1}\kappa = I - \left((\gamma+\theta_1)\Sigma_u + \bar{\Lambda} + \rho A_{xx}\right)^{-1}\bar{\Lambda}. \tag{17}$$

The weight of the aim portfolio measures the speed of trading toward the aim, and is called the trading rate. The trading rate increases with θ_1 and θ_2.

Proposition 2 is in line with Proposition 2 in GP, stating that the optimal portfolio is a weighted average of the existing portfolio x_{t-1} and an aim portfolio aim_t. Thus, transaction costs make it optimal to trade slowly. However, Eq. (17) indicates that ambiguity-averse investors trade faster toward the aim than those investors without ambiguity aversion.

Note that Eq. (16) looks different from GP's Eq. (7). GP's Eq. (7) is obtained by taking the partial derivative of both sides of the valuation function with respective to x_{t-1}. Using this method, based on Eqs. (13) and (B.2) in the appendix, the optimal robust strategy can be rewritten as follows:

$$x_t^* = x_{t-1} + \bar{\Lambda}^{-1}A_{xx}((A_{xx})^{-1}A_{xf}f_t - x_{t-1}). \tag{18}$$

Now, we see that Eq. (18) is similar in format to GP's Eq. (7), although the coefficient matrices A_{xx} and A_{xf} in both equations differ. However, a comparison of Eq. (18) and GP's Eq. (7) does not clearly show the distinction between the robust and non-robust strategies. To better understand the impact of ambiguity aversion on the optimal trading strategy, we plug the expressions of A_{xx}^{nr} and A_{xf}^{nr}, which are obtained from GP's Eqs. (A8) and (A9), into GP's Eq. (7) and simplify the equation. This gives the following equation:

$$(x_t^{nr})^* = x_{t-1}^{nr} + (\gamma\Sigma_u + \rho A_{xx}^{nr} + \bar{\Lambda})^{-1}(\gamma\Sigma_u + \rho A_{xx}^{nr})(aim_t^{nr} - x_{t-1}^{nr}), \tag{19}$$

where $aim_t^{nr} = (\gamma\Sigma_u + \rho A_{xx}^{nr})^{-1}(Bf_t + \rho A_{xf}^{nr}E_t(f_{t+1}))$, and the superscript nr represents the results for the corresponding non-robust strategy. Eq. (16) is obtained using the same method.

A comparison of Eqs. (16) and (19) shows that the effect of aversion to return ambiguity can be interpreted as increasing the risk aversion parameter from γ to $\gamma+\theta_1$. Hence, the presence of ambiguity aversion leads to a more conservative strategy. This is consistent with the findings in previous studies (Chen and Epstein, 2002; Maenhout, 2006; Jeong et al., 2015). Importantly, compared with Eq. (19), Proposition 2 indicates that ambiguity aversion also impacts the aim portfolio itself, with the expectations of return predictors in Eq. (19) being replaced with their values under the worst-case scenario.[12] To better understand how the impact of ambiguity aversion on the optimal trading strategy differs from the

[10] It is easy to prove that if Condition 2 does not hold, then Condition 1 will not hold.

[11] These matrices appear both on the left- and right-hand sides of the expressions of these matrices, as we are not able to provide the explicit expressions of these matrices. However, with the contraction mapping principle, it is easy to prove that the solutions of these equations exist and are unique. In the empirical analysis, we calculate these matrices, using the iterative method of Ljungqvist and Sargent (2004).

[12] Unlike Eq. (19), the aim portfolio in Eq. (16) is expressed in terms of $\tilde{E}_t^*(f_{t+1})$, which in turn depends on the optimal strategy x_t^* as shown by Eq. (15). However, Eq. (18) shows that x_t^* is a function of x_{t-1} and f_t. Plugging Eq. (18) into the aim portfolio would eliminate x_t^* in the expression of the aim portfolio, allowing us to analyze the properties of the aim portfolio.

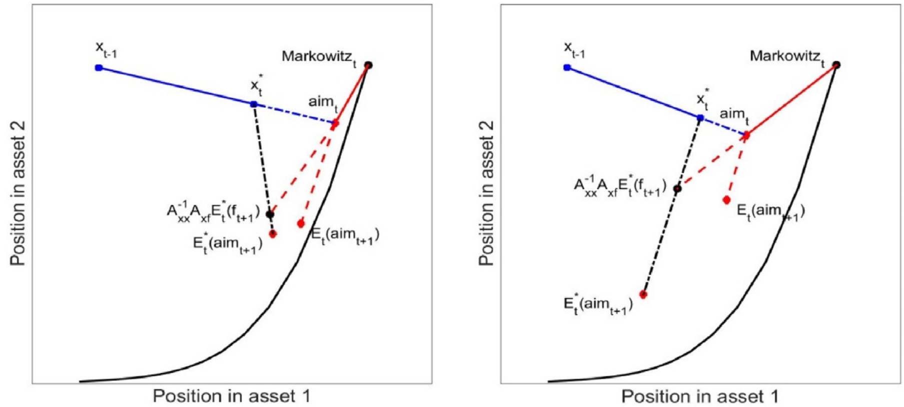


Fig. 1. The aim portfolio in the presence of ambiguity aversion
This figure shows the aim portfolio choice with two securities (assets 1 and 2). The Markowitz portfolio is the current optimal portfolio with ambiguity aversion in the absence of transaction costs. $A_{xx}^{-1}A_{xf}E_t^*(f_{t+1})$ is the portfolio that maximizes risk-, ambiguity-, and costs-adjusted returns at all future dates. $E_t(\cdot)$ is the conditional expectation operator with respect to the probability measure $\tilde{P}_t$ that is induced by $\tilde{F}_t$. $E_t^*(\cdot)$ is the conditional expectation operator under the worst-case scenario. The left panel shows the aim portfolio choice in the case of $\theta_2 = 0.1$, while the right panel shows the aim portfolio choice in the case of $\theta_2 = 1$.

impact of risk aversion, we rearrange the aim portfolio, and obtain the second trading principle.

Proposition 3. *Aim in front of the target.*

(i) *The aim portfolio can be written as:*

$$aim_t = \kappa^{-1}\Big((\kappa - \rho A_{xx})Markowitz_t + \rho A_{xx}\big(A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})\big)\Big), \tag{20}$$

where $Markowitz_t = ((\gamma + \theta_1)\Sigma_u)^{-1} Bf_t$ is the current optimal portfolio with ambiguity aversion in the absence of transaction costs. $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ is the portfolio that maximizes risk-, ambiguity-, and costs-adjusted returns at all future dates under the worst-case scenario.[13]

(ii) *The second component in the aim portfolio can be written as:*

$$A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1}) = \big(\kappa + A_{xx} - \kappa J_1^{-1}\bar{\Lambda}\big)^{-1}\big(\kappa \tilde{E}_t^*(aim_{t+1}) + (A_{xx} - \kappa J_1^{-1}\bar{\Lambda})x_t^*\big), \tag{21}$$

which means that strategy $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ is just a combination of the worst-case expected aim portfolio $\tilde{E}_t^*(aim_{t+1})$ and the optimal portfolio x_t^*. Additionally, $A_{xx} - \kappa J_1^{-1}\bar{\Lambda} = 0$ in the absence of predictor ambiguity aversion.

Proposition 3 reveals that compared with the non-robust case, the robust aim portfolio is still the weighted average of the current Markowitz portfolio and the future portfolio $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$. However, $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ is no longer just the expected future aim portfolio. Rather, it is a combination of the expected future aim portfolio and the current optimal portfolio. This new property can be illustrated in Fig. 1. First, as noted by GP, portfolio aim_t lies in front of portfolio $Markowitz_t$, due to the presence of transaction costs. The weight of the current Markowitz portfolio in the robust aim portfolio increases with return ambiguity aversion (θ_1). Thus, return ambiguity aversion and risk aversion impact the aim portfolio in the same fashion. On the other hand, the weight of the future portfolio $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ increases with predictor ambiguity aversion (θ_2).

Second, the future portfolio $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ is a weighted average of $\tilde{E}_t^*(aim_{t+1})$ and x_t^*, due to the presence of factor ambiguity aversion, while it is equal to $\tilde{E}_t(aim_{t+1})$ in GP's model and is equal to $\tilde{E}_t^*(aim_{t+1})$ with solely return ambiguity aversion. Intuitively, in the presence of predictor ambiguity, the estimated evolution of predictors may deviate from the true evolution. To reduce the impact of the possible estimation biases, the future portfolio $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ moves partially toward the optimal portfolio x_t^*, which is constructed based on both the existing portfolio x_{t-1} and current predictors f_t as shown in Eq. (18). In general, x_t^* has a larger effect on $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ if the degree of predictor ambiguity aversion (θ_2) is higher. As we can see from Fig. 1, the future portfolio $A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})$ is closer to x_t^* in the case of $\theta_2 = 1$ than it is in the case of $\theta_2 = 0.1$.

In summary, Proposition 2 shows the effect of ambiguity aversion on the current robust strategy, while Proposition 3 demonstrates its effect on the robust strategy at future dates. Another reading of Proposition 3 is that ambiguity aversion and risk aversion impact the aim portfolio differently, and this difference arises primarily from aversion to factor ambiguity. To understand why the principle stated in Proposition 3 works, we further investigate the mechanism through which ambiguity aversion impacts the aim portfolio.

3.2. The additional trading principle with ambiguity aversion

Proposition 4. *Aim for a low expected loss.*

(i) *The aim portfolio can be expressed as follows:*

$$aim_t = \kappa^{-1}\Big((\kappa + \bar{\Lambda})J_1^{-1}K_f f_t + (J_1 - \kappa - \bar{\Lambda})J_1^{-1}K_x x_{t-1}\Big), \tag{22}$$

where $K_f = B + \rho A_{xf}(I + 2\rho\theta_2\Sigma_v A_{ff})^{-1}\Phi$ and $K_x = -\bar{\Lambda}$ are the adjustment coefficients for f_t and x_{t-1}, respectively. This suggests that in the aim portfolio, the weight of adjusted predicting factors $K_f f_t$ is $(\kappa + \bar{\Lambda})J_1^{-1}$, while the weight of the adjusted existing portfolio $K_x x_{t-1}$ is $(J_1 - \kappa - \bar{\Lambda})J_1^{-1}$.

(ii) *In the presence of predictor ambiguity aversion, if the variability of predictors Σ_v is large, then $K_f f_t$ is small. This implies that the aim portfolio loads less on the highly volatile predictors. In addition, if the holdings of securities in the existing portfolio are large and costly (a low value of $-\bar{\Lambda}x_{t-1}$), then the aim portfolio loads less on these securities.*

Proposition 4 explains the way in which ambiguity aversion impacts the aim portfolio, and furthers our understanding of how the effects of ambiguity aversion and risk aversion differ. The first implication of Proposition 4 is that with ambiguity aversion to return

[13] Given Eq. (13), it is easy to prove that $A_{xx}^{-1}A_{xf}f_{t+1}$ is the trading strategy that maximizes the value function at time $t+1$.

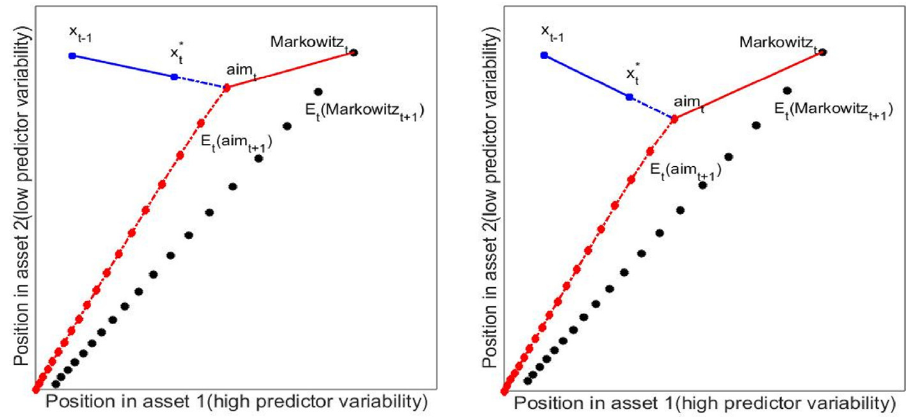


Fig. 2. The effect of factor variability on the aim portfolio
This figure shows the effect of return predictor variability on the aim portfolio choice with two securities (assets 1 and 2). Asset 1's return predictors are more volatile than asset 2's. The Markowitz portfolio is the current optimal portfolio with ambiguity aversion in the absence of transaction costs. $E_t(\cdot)$ is the conditional expectation operator with respect to the probability measure $\tilde{P}_t$ that is induced by $\tilde{F}_t$. The left panel considers the case in which $\theta_2 = 0.1$, while the right panel considers the case in which $\theta_2 = 1$.

predictors, the adjustment coefficient K_f is affected by Σ_v. Without ambiguity aversion, the aim portfolio loads more on securities with more persistent predictors (larger Φ), since such a trading strategy not only generates a high expected return now, but also is expected to generate a high expected return for a longer time in the future (GP, 2013). However, Proposition 4 shows that if the predictors with a slower mean-reversion are highly volatile, the aim portfolio with ambiguity aversion should not give more weights to these predictors. Fig. 2 illustrates the impact of predictor variability on portfolio positions. In the figure, both assets 1 and 2 have the same mean-reverting speed in predictor dynamics, but asset 1′s predictor is more volatile. Since the aim portfolio downweights highly-volatile predictors, it loads more heavily on asset 2. In addition, the difference between these two positions increases with the ambiguity aversion coefficient θ_2.

Intuitively, if return predictors are highly volatile, then the estimated mean-reverting coefficients Φ are more likely to be biased. As a result, these estimated coefficients are not reliable, even if they are large. Investors who follow the strategy with these largely biased estimates may experience huge losses, as asset returns are likely to be seriously misspecified. Hence, investors with ambiguity aversion should trade toward an aim portfolio that is tilted toward the less volatile return predictors.

Another implication of Proposition 4 is that with ambiguity aversion to predictors, the aim portfolio is also affected by transaction costs Λ and existing asset holdings x_{t-1}. To reduce the potential loss due to estimation errors, the aim portfolio should also load less on securities with large and costly existing positions. To illustrate, we consider an optimal strategy with two assets. In the first case, the positions in the two assets in the existing portfolio are the same, but the position in asset 1 has a higher transaction cost. Fig. 3-A shows that the current aim portfolio loads less on asset 1 than on asset 2. In the second case, the transaction costs of the two assets are the same, while asset 1 has a larger position in the existing portfolio. Fig. 3-B shows that the aim portfolio loads less on asset 1, and this remains true until the positions in both assets are identical.

This can be interpreted as follows. If the estimated mean-reverting coefficients Φ are biased, the future returns of corresponding securities may be overestimated. In this case, the resulting upward-biased positions of these securities entail more transaction costs, but are not able to achieve the predicted return levels. Moreover, these biased positions will have to be rebalanced in subsequent periods, which incurs further costs. As a consequence, estimation errors may lead to substantial fluctuations in the security positions over time, driving down the net returns. Even though the estimated coefficients Φ are less likely biased with less volatile predictors, the expected loss can be huge for securities with high transaction costs and large positions in the existing portfolio. Thus, the effects of estimation errors on portfolio performance are more pronounced if the existing security holdings are particularly large and costly.

In summary, this third trading principle suggests two channels through which the aim portfolio minimizes the expected loss. One is to load less on securities with highly volatile predictors to reduce the likelihood of estimation errors occurring. The other is to load less on securities with large and costly existing positions to reduce the size of losses arising from estimation errors. This trading principle is intended to address the ambiguity aversion to return predictors.

4. Performance of the robust trading strategy: an empirical investigation

This section investigates the effectiveness of the robust strategy, using data on various commodity futures, in order to illustrate how the robust strategy can be applied in practice, and to identify the drivers of its superior performance.

4.1. The data

Following GP, we consider 15 commodity futures, including: aluminum, copper, nickel, zinc, lead, and tin from the London Metal Exchange (LME); gasoil from the Intercontinental Exchange (ICE); WTI crude, RBOB unleaded gasoline, and natural gas from the New York Mercantile Exchange (NYMEX); gold and silver from the New York Commodities Exchange (COMEX); and coffee, cocoa, and sugar from the New York Board of Trade (NYBOT). The sample period is from January 1, 1996 to December 31, 2015 for all futures. The data on futures prices is obtained from Bloomberg, the contract multipliers are from the respective exchanges, and the risk-free rate is from Kenneth R. French's website.[14]

For consistency with GP, we use the most liquid futures of all maturities available to construct each futures' data series. Based on these series, we calculate the excess rate of return (henceforth, returns) on each futures at time t.[15] Unlike GP, all data is used

[14] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[15] To calculate the rates of return, we always use the prices of a given contract.

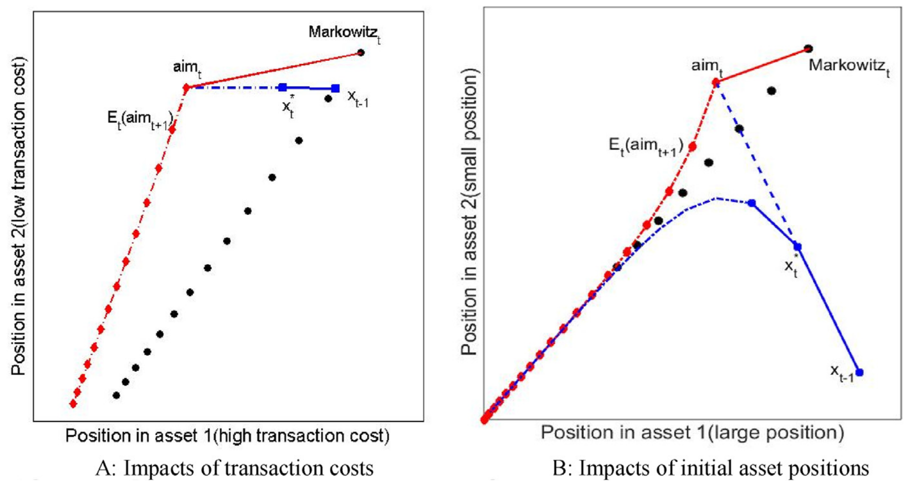


Fig. 3. The effects of transaction costs and asset positions on the aim portfolio
This figure shows the effects of transaction costs (left panel) and initial asset positions (right panel) on the aim portfolio choice with two securities. The Markowitz portfolio is the current optimal portfolio with ambiguity aversion in the absence of transaction costs. $E_t(\cdot)$ is the conditional expectation operator with respect to the probability measure $\tilde{P}_t$ that is induced by $\tilde{F}_t$. The left panel considers the case in which asset 1 has higher average transaction costs than asset 2, while the right panel considers the case in which asset 1 has a larger initial position in the existing portfolio than asset 2.

Table 1
Summary statistics of data.

Commodity	Contract multiplier	June–November 2015 (Rolling window)		December 2015 (Investment window)	
		Average daily returns (%)	Standard deviation of daily returns (%)	Average daily returns (%)	Standard deviation of daily returns (%)
Aluminum	25	−0.14	1.08	0.20	1.24
Cocoa	10	0.02	0.96	−0.18	1.19
Coffee	37,500	−0.07	2.10	0.29	1.66
Copper	25	−0.21	1.59	0.13	1.36
Crude	1000	−0.30	2.72	−0.52	2.84
Gasoil	100	−0.27	2.03	−1.11	2.71
Gold	100	−0.09	0.88	−0.02	1.08
Lead	25	−0.05	1.53	0.42	1.61
NatGas	10,000	−0.20	2.46	0.34	5.24
Nickel	6	−0.26	2.29	−0.03	1.55
Silver	5000	−0.08	1.60	−0.08	1.85
Sugar	112,000	0.20	2.37	0.11	1.64
Tin	5	0.09	1.75	−0.16	0.86
Unleaded	42,000	−0.37	2.91	−0.29	2.98
Zinc	25	−0.22	1.87	0.15	1.66

This table reports summary statistics of daily rates of return and contract's multipliers of various futures considered in our analysis.

for rolling out-of-sample tests to gauge the difference in performance between robust and non-robust strategies. Specifically, for each month from July 1, 1996, we estimate the model parameters using the data on the predictors from the previous six months (Moskowitz et al., 2012; Barroso and Santa-Clara, 2015). For example, we use the data from January to June 1996 (referred to as rolling window) to estimate model parameters for July 1996 (investment window).[16] Table 1 presents the summary statistics of each futures' daily returns for the last rolling window (June to November 2015) and investment window (December 2015).

4.2. Model estimation

We first estimate the factor loadings B and the mean-reversion coefficients Φ for return predictors. Following the literature (Erb and Harvey, 2006; Asness et al., 2013; Gârleanu and Pedersen, 2013), for the futures contract s, we choose predictors $f^{5D,s}$, $f^{1Y,s}$, and $f^{5Y,s}$, which are the moving averages of daily returns over the past five days, one year, and five years, divided by their respective standard deviations. When calculating the average of daily returns over the past year, we skip the most recent month's return to avoid the one-month reversal in returns (Grinblatt and Moskowitz, 2004; Asness et al., 2013). To illustrate the estimation results, we report the results for the last 6-month rolling window (June to November 2015) as follows:

$$\underset{}{r^s_{t+1}} = \underset{(-0.07)}{-0.0011} + \underset{(1.12)}{0.0008} f^{5D,s}_t - \underset{(-1.56)}{0.0174} f^{1Y,s}_t + \underset{(1.48)}{0.0662} f^{5Y,s}_t + u^s_{t+1}, \tag{23}$$

$$f^{5D,s}_{t+1} = 0.8099 f^{5D,s}_t + v^{5D,s}_{t+1}, \quad f^{1Y,s}_{t+1} = 0.9923 f^{1Y,s}_t + v^{1Y,s}_{t+1}, \\ f^{5Y,s}_{t+1} = 0.9963 f^{5Y,s}_t + v^{5Y,s}_{t+1}. \tag{24}$$

The model parameters are estimated using feasible generalized least squares, and the numbers reported in the brackets are the t-statistics. The results in Eq. (23) indicate short- and long-term momentum and medium-term reversals in commodity futures prices. In addition, the results in Eq. (24) confirm that the return predictors are mean reverting with various mean-reversion

[16] Note that the calculation of return predictors for the period from January to June 1996 may use historical data from the period from January 1991 to December 1995.

rates. Based on these estimates, we can write the matrices of factor loadings and the mean-reversion coefficients for the predictors as follows:[17]

$$B = \begin{pmatrix} 0.0008 & -0.0174 & 0.0662 \end{pmatrix} \otimes I_{15\times 15},$$
$$\Phi = \begin{pmatrix} 0.8099 & & \\ & 0.9923 & \\ & & 0.9963 \end{pmatrix} \otimes I_{15\times 15}, \tag{25}$$

where $\otimes$ denotes the Kronecker product of matrices, and $I_{15\times 15}$ is the 15-by-15 identity matrix.

Note that the estimation results in Eq. (23) differ substantially from those in GP and GX, in terms of the magnitude, sign, and significance of the coefficients. The difference in magnitude of the coefficients arises because, in this analysis, rates of return rather than dollar returns (price changes) are used to calculate asset returns and return predictors. The difference in the sign and significance of some predictors arises because we use a 6-month window to estimate Eq. (23) for our rolling out-of-sample tests, while GP's and GX's estimation period is more than 13 years, from January 1, 1996 to January 23, 2009. With a short rolling window for estimation, the significance of the estimates can be reduced, and the sign of these estimates may vary across different estimation periods.[18]

Then, we estimate the variance-covariance matrices Σ_u and Σ_v. Since our model ignores the estimation errors in these two matrices, we use the daily returns and predictors in investment windows to calculate Σ_u and Σ_v to control for the effects of estimation errors in these matrices on the performance of the optimal strategies. Consistent with the model assumptions, both matrices are assumed to be constants in an investment window.

Third, following GP, we assume that transaction costs are proportional to the amount of risk $\Lambda = \lambda\Sigma_u$, and set the multiple $\lambda = 5\times 10^{-7}$. Meanwhile, we set the absolute risk aversion $\gamma = 10^{-9}$, which means that for an agent with \$1 billion under management, the relative risk aversion is one. We also assume that annualized discount rate is 2%, meaning that with approximately 260 trading days in a year, the discount factor is $\rho = 1/(1+0.02/260)$.

Finally, we specify ambiguity aversion coefficients θ_1 and θ_2, based on the analysis in Section 2.2. Take θ_1 as an example. Since $e_{u,t+1}$ follows a multivariate normal distribution with mean zero and covariance matrix $f_t^{\mathrm{T}}(F^{\mathrm{T}}F)^{-1}f_t\Sigma_u$, where $F^{\mathrm{T}} = (f_{-n}, f_{-n+1}, \ldots, f_{-1})$ is the rolling window's predictor matrix and n is the rolling window size (Wooldridge, 2013), $e_{u,t+1}^T(f_t^{\mathrm{T}}(F^{\mathrm{T}}F)^{-1}f_t\Sigma_u)^{-1}e_{u,t+1}$ follows the chi-square distribution $\chi^2(15)$. Thus, for any time t, the constraint η_1 is given by:

$$\eta_1 = \frac{1}{2}f_t^{\mathrm{T}}(F^{\mathrm{T}}F)^{-1}f_t \cdot m^{-1}(1-c), \tag{26}$$

where m is the cumulative distribution function of $\chi^2(15)$, and c is investors' level of confidence in return estimates. According to Karush-Kuhn-Tucker conditions, if θ_1 exists, $e_{u,t+1}^*$ in Eq. (14) should satisfy $(e_{u,t+1}^*)^{\mathrm{T}}\Sigma_u^{-1}e_{u,t+1}^*/2 = \eta_1$. Similarly, we can also obtain an equation that establishes the relationship between θ_2 and η_2. Finally, these two equations are used to solve for θ_1 and θ_2 simultaneously by numerical iteration. For brevity, we assume that θ_1 and θ_2 are constants over time, setting the values equal to their respective estimated medians in the sample period. Using this method, we obtain $\theta_1 = 10^{-10}$ and $\theta_2 = 5\times 10^{-7}$ for a confidence level of 90%. Given that different investors may have different confidence levels, we also consider alternative values of θ_1 and θ_2 in the analysis.

4.3. Performance of robust trading strategy

To evaluate the performance of a trading strategy, we focus on its Sharpe ratio, which is defined as follows:

$$SR = \frac{\text{average daily dollar returns}}{\text{standard deviation of daily dollar returns}} \times \sqrt{260}. \tag{27}$$

Similar to GX, we use dollar returns rather than rates of return to calculate Sharpe ratio. This is because when a portfolio is constructed with futures contracts, it has no principal if ignoring the futures margin, and thus the portfolio's percentage returns cannot be calculated.

Table 2 reports the mean and standard deviation of daily dollar returns before transaction costs, the gross Sharpe ratios, and the net Sharpe ratios (net of transaction costs) for the optimal strategies with various ambiguity-aversion coefficient combinations (θ_1, θ_2). Non-robust in θ_1 or θ_2 represents the strategy with no robustness in returns or predictors, meaning that there is no ambiguity aversion about returns or predictors. In particular, the combination (non-robust in θ_1, non-robust in θ_2) corresponds to the non-robust strategy in GP, and ($\theta_1 = 10^{-10}$,$\theta_2 = 5\times 10^{-7}$) is referred to as the base-case robust strategy.

The results in Table 2 show that in all cases the robust portfolio outperforms the non-robust one, in terms of gross Sharpe ratio. In addition, the gross Sharpe ratio of the robust portfolio improves as the ambiguity aversion coefficients increase. We note that the mean and standard deviation of the robust portfolio returns are lower than those of the non-robust portfolio returns. This suggests that the robust strategy generates a higher gross Sharpe ratio, because the dollar returns on the robust portfolio are less volatile than those on the non-robust portfolio. As the degree of ambiguity aversion increases, the effect of risk reduction becomes more significant, leading to a higher gross Sharpe ratio.

A comparison of the net Sharpe ratios of the robust and non-robust strategies confirms that the robust strategy performs better. Importantly, the improvement in net Sharpe ratio of the robust strategy over the non-robust one is much more pronounced than the improvement in gross Sharpe ratio. Naturally, the Sharpe ratios after transaction costs are reduced in all cases. The percentage reduction in Sharpe ratio after transaction costs is smaller for the robust strategy than for the non-robust strategy, indicating that the robust strategy incurs relatively low transaction costs. Our findings show that by accounting for ambiguity aversion, the robust strategy is able to reduce the transaction costs associated with biased positions in securities with biased estimated returns and predictors, thereby improving net Sharpe ratio in a more significant way than does the non-robust strategy.

To understand why the robust strategy is better able to reduce the volatility of returns and transaction costs, Fig. 4 depicts the positions of gold futures in the base-case robust and the non-robust portfolios.[19] We see from Fig. 4-A that the position of gold futures in the non-robust portfolio is large and fluctuates substantially over time, while the position in the robust portfolio is small and less volatile. To examine the underlying reason for this phenomenon, we focus on the period from January 2006 to December 2007, and depict, in Fig. 4-B, the positions of the two strategies along with the monthly returns on gold futures for the same period. We note that the positions of both strategies change as a result of fluctuations of the futures returns, as past returns are used as predictors to forecast future returns in both strategies. Since the

[17] The vector of the factors is $f_t = (f_t^{5D,1}, \cdots, f_t^{5D,15}, f_t^{1Y,1}, \cdots, f_t^{1Y,15}, f_t^{5Y,1}, \cdots, f_t^{5Y,15})^{\mathrm{T}}$.

[18] In fact, this means that model parameters are estimated with estimation errors, and this can help demonstrate the advantage of the robust trading rules. Similarly, Anderson et al. (2009) also illustrate the contribution of ambiguity premium to asset pricing by allowing model misspecification and insignificance of model parameters.

[19] We depict the gold futures as an example, since its position is the largest one in both the robust and non-robust portfolios.

Table 2
Out-of-sample performance for various strategies.

	Non-robust in θ_2				$\theta_2 = 5 \times 10^{-8}$			
	Mean return	SD of returns	Gross SR	Net SR	Mean return	SD of returns	Gross SR	Net SR
Non-robust in θ_1	8.90×10^6	1.96×10^8	0.73	0.14	4.40×10^6	9.46×10^7	0.75	0.48
$\theta_1 = 10^{-10}$	8.29×10^6	1.78×10^8	0.75	0.17	4.27×10^6	8.94×10^7	0.77	0.50
$\theta_1 = 10^{-9}$	5.32×10^6	9.56×10^7	0.90	0.39	3.37×10^6	6.07×10^7	0.90	0.61
	$\theta_2 = 5 \times 10^{-7}$				$\theta_2 = 5 \times 10^{-6}$			
	Mean return	SD of returns	Gross SR	Net SR	Mean return	SD of returns	Gross SR	Net SR
Non-robust in θ_1	2.51×10^6	4.58×10^7	0.89	0.71	1.27×10^6	1.76×10^7	1.16	1.05
$\theta_1 = 10^{-10}$	2.47×10^6	4.42×10^7	0.90	0.73	1.26×10^6	1.73×10^7	1.17	1.06
$\theta_1 = 10^{-9}$	2.18×10^6	3.40×10^7	1.04	0.85	1.19×10^6	1.50×10^7	1.28	1.16

This table reports out-of-sample Sharpe ratios (SR) for the non-robust strategy and various robust strategies before and after transaction costs, as well as means and standard deviations (SD) of daily dollar returns of these strategies before transaction costs. Non-robust in θ_1 and θ_2 correspond to the strategies with no robustness in returns (u_t) and in return predictors (v_t), respectively. The out-of-sample period is from July 1, 1996 to December 31, 2015.

gold futures returns are high prior to June 2006, the position of the non-robust portfolio is rebalanced from short selling prior to June 2006 to buying long for the period June to October 2006, while the robust strategy is less aggressive in building up a large gold position in the same period. Since the futures returns are actually low from June to October 2006, the non-robust strategy suffers larger losses. Moreover, the non-robust strategy's position decays substantially from November 2006 to March 2007 in response to the low predictors prior to November, which incurs additional transaction costs. This clearly shows the impact of estimation errors of on portfolio positions.

Over the whole sample period, a large percentage of transactions of the non-robust strategy are attributable to over-responses to biased estimated returns and predictors. In contrast, the robust strategy trades less aggressively than the non-robust strategy. This explains why the non-robust strategy entails substantial transaction costs and generates particularly low net Sharpe ratios compared with the robust strategy. Our result is in line with Garlappi et al.'s (2007) finding that the portfolio weights derived from the model with ambiguity aversion are less unbalanced and fluctuate much less over time than do the portfolio weights from the standard mean-variance model.

Proposition 3 indicates that the robust and non-robust strategies differ because of their distinct aim portfolios. To see this, Fig. 4-C depicts the positions of gold futures in the aim portfolios of both strategies. We note that the moving trends and fluctuations in the positions of the aim portfolios for the two strategies are similar to those observed in Fig. 4-A. This indicates that the substantial fluctuations in the position of the non-robust strategy are mainly due to the substantial changes in the position of its aim portfolio.

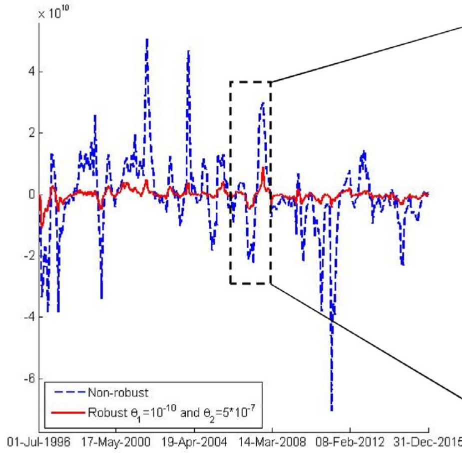
4.4. Impacts of transaction costs and predictor variability on performance

Proposition 4 indicates that the key to the superior performance of the robust strategy is that ambiguity aversion to predictors leads to small positions in the securities with highly volatile predictors and those with large and costly existing portfolio positions. Hence, the improved performance of the robust strategy is related to the levels of transaction costs and predictor variability. To illustrate, we consider three different levels of transaction costs: low costs ($\lambda = 5 \times 10^{-8}$), medium costs ($\lambda = 5 \times 10^{-7}$), and high costs ($\lambda = 5 \times 10^{-6}$), as well as three levels of predictor variability: low variability ($1 \times \Sigma_v$), medium variability ($2 \times \Sigma_v$), and high variability ($4 \times \Sigma_v$).[20] For this exercise, we consider the robust strategies with non-robustness in θ_2 (henceforth, without θ_2) and various values of θ_2. In all the cases considered, θ_1 is set equal to the base-case value. The performance measures include the gross and net Sharpe ratios of these strategies, as well as their means of daily dollar returns (MR) and means of daily dollar transaction costs (MC) expressed as the percentage of those of the corresponding strategies without θ_2.

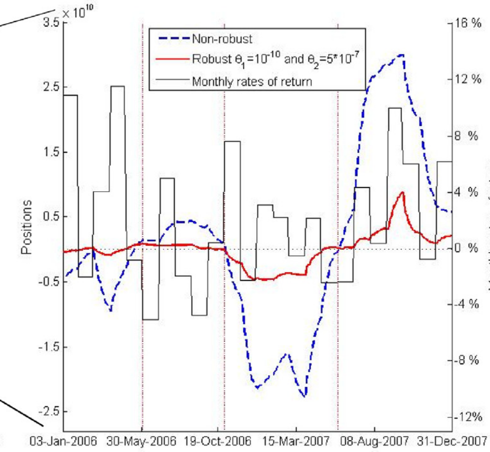
Panel A of Table 3 reports the performance of various portfolios in the cases of low, medium, and high transaction costs. In terms of net Sharpe ratio, the strategy with θ_2 outperforms the strategy without θ_2 for any given level of transaction costs, which is consistent with the findings in Table 2. Moreover, for any given θ_2, the improvement in net Sharpe ratio becomes more pronounced as transaction costs increase. However, in terms of gross Sharpe ratio, the same does not hold true. The gross Sharpe ratios of the strategy with relatively low values of θ_2 are even lower than the gross Sharpe ratio of the strategy without θ_2 when transaction costs are high. This is because, compared with the strategy without θ_2, the strategy with θ_2 is more conservative. As the transaction costs increase, the strategy with θ_2 will lower the positions of all securities in the portfolio, reducing its ability to capture high returns. This is reflected in the values of MR, which are all less than 1 and decline as transaction costs increase for any given θ_2. However, compared with the strategy without θ_2, for any strategy with θ_2, the reduction in TC is greater than the reduction in MR, which is particularly pronounced in the case of high transaction costs. Thus, the strategy with θ_2 is better able to improve its net Sharpe ratio compared with the strategy without θ_2, when transaction costs are particularly high.

Panel B of Table 3 reports the performance of various strategies in the cases of low, medium, and high levels of predictor variability. For any given level of θ_2, we find that the improvement in net Sharpe ratio of the strategy with θ_2 becomes more pronounced, as the predictors become more volatile. While the improvement in gross Sharpe ratio of the strategy with θ_2 decreases slightly with predictor variability in the case of $\theta_2 = 5 \times 10^{-8}$, it increases with predictor variability in the other two cases. For any given θ_2, while both MR and TC of the robust portfolio with θ_2 decline as the level of predictor variability rises, the decline in TC is greater than the decline in MR. Thus, the strategy with θ_2 is better able to reduce

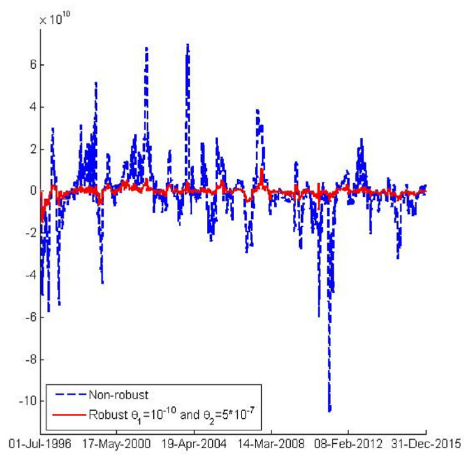
[20] When some or all the predictors have a particularly high level of variability (e.g., $\lambda = 10 \times \Sigma_v$), Condition 1 in Proposition 1 is not satisfied, and thus no estimation error in predictors under the worst-case scenario ($e^*_{v,t+1}$) can be identified. Intuitively, since the predictors are so volatile that investors no longer trust the estimate of Φ, we cannot find the values of the predictors that minimize performance. In our numerical analysis, Conditions 1 and 2 are not binding.



A: Gold positions of the robust and non-robust strategies



B: Gold positions of the robust and non-robust strategies and monthly rates of return for gold futures for the period from January 3, 2006 to December 31, 2007.



C: Gold positions in the robust and non-robust aim portfolios

Fig. 4. The positions of gold futures in the robust and non-robust strategies
A and B display the positions of gold futures in the robust and non-robust strategies over time. C displays gold positions in the robust and non-robust aim portfolios over time.

Table 3
Out-of-sample performance for various robust strategies with different transaction costs and predictor variability.

Panel A: Performance for various strategies with various transaction costs												
	$\lambda = 5 \times 10^{-8}$ (Low costs)				$\lambda = 5 \times 10^{-7}$ (Medium costs)				$\lambda = 5 \times 10^{-6}$ (High costs)			
	Gross SR	Net SR	MR	TC	Gross SR	Net SR	MR	TC	Gross SR	Net SR	MR	TC
Non-robust in θ_2	1.52	1.20	1	1	0.75	0.17	1	1	0.49	−0.43	1	1
$\theta_2 = 5 \times 10^{-8}$	1.65	1.45	0.77	0.45	0.77	0.50	0.51	0.23	0.44	0.06	0.30	0.14
$\theta_2 = 5 \times 10^{-7}$	1.84	1.71	0.56	0.19	0.90	0.73	0.30	0.08	0.48	0.23	0.13	0.04
$\theta_2 = 5 \times 10^{-6}$	2.18	2.09	0.35	0.07	1.17	1.06	0.15	0.02	0.69	0.51	0.06	0.01
Panel B: Performance for various strategies with various levels of factor variability												
	$1 \times \Sigma_v$ (Low variability)				$2 \times \Sigma_v$ (Medium variability)				$4 \times \Sigma_v$ (High variability)			
	Gross SR	Net SR	MR	TC	Gross SR	Net SR	MR	TC	Gross SR	Net SR	MR	TC
Non-robust in θ_2	0.75	0.17	1	1	0.75	0.17	1	1	0.75	0.17	1	1
$\theta_2 = 5 \times 10^{-8}$	0.77	0.50	0.51	0.23	0.76	0.54	0.43	0.16	0.76	0.57	0.35	0.11
$\theta_2 = 5 \times 10^{-7}$	0.90	0.73	0.30	0.08	0.92	0.77	0.23	0.05	0.94	0.82	0.19	0.03
$\theta_2 = 5 \times 10^{-6}$	1.17	1.06	0.15	0.02	1.24	1.14	0.12	0.01	1.322	1.23	0.09	0.01

This table reports out-of-sample Sharpe ratios (SR), means of daily dollar returns (MR), and means of daily dollar transaction costs (TC) for various robust strategies with different levels of transaction costs and predictor variability. MR and TC of each strategy are expressed as the percentage of the corresponding values for the strategy with no robustness in θ_2. For all robust strategies considered, $\theta_1 = 10^{-10}$. We consider three different levels of transaction costs: low costs ($\lambda = 5 \times 10^{-8}$), medium costs ($\lambda = 5 \times 10^{-7}$), and high costs ($\lambda = 5 \times 10^{-6}$), where $\lambda = 5 \times 10^{-7}$ is the base-case transaction cost in our analysis. We consider three levels of predictor variability: $1 \times \Sigma_v$, $2 \times \Sigma_v$, and $4 \times \Sigma_v$. The out-of-sample period is from July 1, 1996 to December 31, 2015.

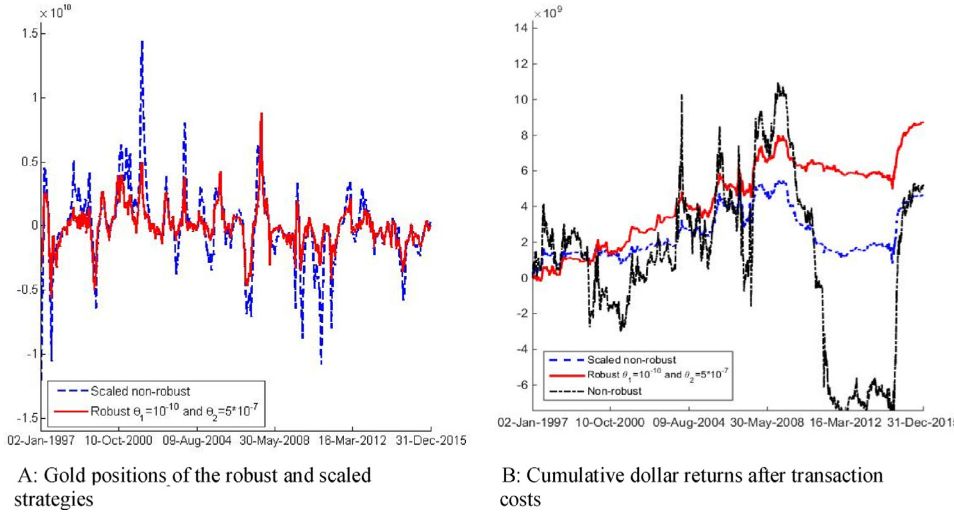


Fig. 5. Gold positions in the robust and scaled portfolios as well as cumulative dollar returns of the robust, non-robust, and scaled non-robust portfolios
This figure depicts the gold positions in the robust and scaled non-robust portfolios as well as cumulative dollar returns after transaction costs of the robust, non-robust, and scaled non-robust portfolios for the period from January 2, 1997 to December 31, 2015.

transaction costs due to estimation errors when predictor variability is high. This is because our model indicates that the holdings of securities in its corresponding portfolio are affected by predictor variability. Collectively, the results suggest that as predictors become more volatile, by loading less on securities with highly volatile predictors, the strategy with θ_2 is better able to deliver superior performance than the strategy without θ_2.

4.5. Robust versus scaled non-robust strategies

Our analysis shows that as a result of controlling for estimation risk, the robust portfolio has smaller positions than the corresponding non-robust portfolio. Interestingly, Barroso and Santa-Clara (2015) document that the position of the momentum strategy can be scaled down to practically eliminate the crash risk of momentum. Kan and Zhou (2007) and DeMiguel et al. (2015) propose to control for estimation risk by scaling down the positions of the Markowitz portfolio in the optimal strategy. In this section, we evaluate the performance of the base-case robust strategy relative to that of a scaled non-robust strategy to better understand the role of ambiguity aversion in portfolio selection.

Inspired by Barroso and Santa-Clara (2015), at each time the model is updated, we scale down the non-robust portfolio in the next investment period as follows:

$$x_t^s = sx_t^{nr} = \frac{\sigma_{-6}^R}{\sigma_{-6}^N} x_t^{nr}, \tag{28}$$

where σ_{-6}^R and σ_{-6}^N are the standard deviations of the previous 6-month realized dollar returns of the robust and non-robust portfolios, respectively. ${x_t}^{nr}$ is the position of the original non-robust portfolio at time t. Clearly, the variance of the scaled non-robust portfolio approximately equals that of the robust portfolio.

Fig. 5A depicts the positions of gold futures in the robust portfolio and the scaled non-robust portfolio, and Fig. 5B displays the cumulative dollar returns on the robust, non-robust, and scaled non-robust strategies after transaction costs. While the gold position in the scaled non-robust portfolio is generally larger than the position in the robust portfolio, it is much smaller than the position in the original non-robust portfolio as illustrated in Fig. 4A. This explains why in Fig. 5B the moving trend of the cumulative dollar returns of the scaled non-robust portfolio and the robust portfolio are similar, and why the volatilities of these portfolios are significantly reduced compared with that of the original non-robust portfolio.

Table 4
Out-of-sample performance of non-robust, scaled non-robust, and robust strategies.

	Mean return	SD of returns	Gross SR	Net SR
Non-robust	8.53×10^6	1.98×10^8	0.70	0.09
Scaled non-robust	1.81×10^6	4.77×10^7	0.61	0.33
Robust	2.36×10^6	4.42×10^7	0.86	0.68

This table reports out-of-sample Sharpe ratios (SR) for the non-robust strategy, scaled non-robust strategy, and base-case robust strategy before and after transaction costs, as well as means and standard deviations (SD) of daily dollar returns of these strategies before transaction costs. The out-of-sample period is from January 2, 1997 to December 31, 2015.

Table 4 reports the Sharpe ratios of the robust, non-robust, and scaled non-robust portfolios, as well as their mean and standard deviation of daily dollar returns before transaction costs. We note that while the gross Sharpe ratio of the scaled non-robust portfolio is slightly lower than that of the original non-robust portfolio, its net Sharpe ratio is much higher. This provides strong evidence that by scaling down positions, the scaled non-robust portfolio can lower transaction costs and portfolio volatility, thereby improving net Sharpe ratio. However, both the gross and net Sharpe ratios of the scaled non-robust portfolio are still lower than those of the robust portfolio. The reason is that scaling down positions of the non-robust portfolio not only lowers its transaction costs, but also reduces its ability to capture high returns. In contrast, while the robust trading strategy is also conservative about building up large positions, it particularly loads less on the assets with highly volatile factors and those with high and costly existing positions. Thus, the robust portfolio is able to generate higher average daily dollar returns than the scaled non-robust portfolio, resulting in a higher Sharpe ratio.

To better understand the superior performance of the robust portfolio relative to the scaled non-robust portfolio, using GP's findings, we rearrange Eq. (28) as follows:

$$\begin{aligned} x_t^s = sx_t^{nr} &= x_{t-1}^s + \bar{\Lambda}^{-1} A_{xx}^{nr} (s \times aim_t^{nr} - x_{t-1}^s) \\ &= x_{t-1}^s + \bar{\Lambda}^{-1} A_{xx}^{nr} \left(\sum_{\tau=t}^{\infty} z(1-z)^{\tau-t} E_t(s \times Markowitz_\tau^{nr}) - x_{t-1}^s \right), \end{aligned} \tag{29}$$

where $z = (\gamma\Sigma_u + \rho A_{xx}^{nr})^{-1}(\gamma\Sigma_u)$ is the weight of the existing Markowitz portfolio in the non-robust aim portfolio. Eq. (29) shows that the scaling factor s is applied solely to the Markowitz portfolio in the non-robust strategy, while the trading rate $\bar{\Lambda}^{-1}A_{xx}^{nr}$ and the weight of the existing Markowitz portfolio z remain unchanged. In contrast, to mitigate the effect of estimation errors, the robust strategy adjusts not only the rate of trading toward the aim portfolio but also the components of the aim portfolio, thereby resulting in reduced portfolio positions. It follows that simply scaling down the position of the non-robust strategy without taking ambiguity aversion into consideration is not able to achieve the superior out-of-sample performance of the robust strategy.

5. Conclusion

Previous work in the finance literature documents that security returns are predictable; however, the predicted returns are just an approximation of reality, due to the presence of model and parameter uncertainties. Estimation errors in security returns could lead to poor portfolio performance, particularly when transaction costs are high. This paper investigates the optimal portfolio choice in the presence of transaction costs and ambiguity aversion. Adopting GX's robust optimization method, we extend GP's model by incorporating ambiguity aversion into the model framework. Unlike GX's model, we allow investors to have different degrees of ambiguity aversion to returns and to return predictors. We not only derive the optimal robust dynamic trading strategy in closed form, but also characterize its properties and clarify the unique mechanism through which the robust strategy improves out-of-sample performance over the non-robust strategy.

Similar to GP's model, our model indicates that the optimal strategy remains to trade partially toward an aim portfolio as well as to aim in front of the optimal portfolio derived from Markowitz's (1952) model. In contrast with GP's non-robust strategy, our robust strategy also aims to reduce the expected loss arising estimation errors. Investors with ambiguity aversion to return predictors trade toward an aim portfolio that loads less on highly volatile predictors. Additionally, the aim portfolio loads less on securities with large and costly existing portfolio holdings. Essentially, the robust strategy is able to minimize the impacts of estimation errors on portfolio performance, by reducing the positions of securities with great parameter uncertainty as well as those with great potential losses associated with parameter ambiguity. This is the key driver of the superior performance of the robust strategy relative to the non-robust strategy.

Using data on commodity futures, we show that the robust strategy outperforms the non-robust strategy in out-of-sample tests. We further find that the robust strategy is better able to improve its performance relative to the non-robust strategy when transaction costs and predictor variability are larger. Simply scaling down the position of the non-robust portfolio is not able to achieve the superior performance of the corresponding robust strategy.

Appendix A. Proof of Lemma 1

Suppose that at any time t, we can specify z_t as follows:

$$z_t = \prod_{j=u,v} \exp\left(e_{j,t+1}^T \Sigma_j^{-1} j_{t+1} - \frac{1}{2} e_{j,t+1}^T \Sigma_j^{-1} e_{j,t+1}\right). \tag{A.1}$$

First, we need to prove that z_t is the Radon–Nikodym derivative. To this end, for $\forall A \in \mathcal{F}_t$, set

$$\tilde{P}_t(A) = \int_A z_t dP_t, \tag{A.2}$$

then $\tilde{P}_t$ satisfies countable additivity, and $\tilde{P}_t(A) \geq 0$ holds true as $z_t \geq 0$. According to Eq. (A.1), given that u_{t+1} and v_{t+1} are mutually independent and are both normally distributed under P_t, we have the following:

$$\tilde{P}_t(\Omega) = \int_\Omega z_t dP_t = E_t(z_t) = 1. \tag{A.3}$$

Thus, $\tilde{P}_t$ is a probability measure. In addition, $\forall A \in \mathcal{F}_t$, if $P_t(A) = 0$, then $\tilde{P}_t(A) = 0$, indicating that $\tilde{P}_t$ is absolutely continuous with respect to P_t. According to the Radon–Nikodym theorem, z_t is the Radon–Nikodym derivative of $\tilde{P}_t$ with respect to P_t.

Next, the characteristic functions of u_{t+1} and v_{t+1} under $\tilde{P}_t$ satisfy:

$$\begin{aligned}\varphi(u_{t+1}, v_{t+1}) &= \tilde{E}_t\left[\exp(ib_u^T u_{t+1} + ib_v^T v_{t+1})\right] \\ &= \int_\Omega \exp(ib_u^T u_{t+1} + ib_v^T v_{t+1}) z_t dP_t,\end{aligned} \tag{A.4}$$

where i is the imaginary unit in a complex number, satisfying $i^2 = -1$. Plugging Eq. (A.1) into Eq. (A.4) yields:

$$\varphi(u_{t+1}, v_{t+1}) = E_t\left[\exp\left(\sum_{j=u,v}(ib_j^T j_{t+1} + e_{j,t+1}^T \Sigma_j^{-1} j_{t+1} - \frac{1}{2} e_{j,t+1}^T \Sigma_j^{-1} e_{j,t+1})\right)\right]. \tag{A.5}$$

Taking u_{t+1} as an example, we simplify the characteristic function as follows:

$$\begin{aligned}&E_t\left[\exp\left(ib_u^T u_{t+1} + e_{u,t+1}^T \Sigma_u^{-1} u_{t+1} - \tfrac{1}{2} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1}\right)\right] \\ &= (2\pi)^{-\frac{S}{2}} |\Sigma_u|^{-\frac{1}{2}} \int_{-\infty}^{\infty} \\ &\exp\left(ib_u^T u_{t+1} + e_{u,t+1}^T \Sigma_u^{-1} u_{t+1} - \tfrac{1}{2} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} - \tfrac{1}{2} u_{t+1}^T \Sigma_u^{-1} u_{t+1}\right) du_{t+1} \\ &= \exp\left(ie_{u,t+1}^T b_u + \tfrac{1}{2} b_u^T \Sigma_u^{-1} b_u\right).\end{aligned} \tag{A.6}$$

Since u_{t+1} and v_{t+1} are mutually independent under P_t, simplifying the characteristic function gives:

$$\varphi(u_{t+1}, v_{t+1}) = \prod_{j=u,v} \exp\left(ie_{j,t+1}^T b_j + \frac{1}{2} b_j^T \Sigma_j^{-1} b_j\right). \tag{A.7}$$

Thus, u_{t+1} and v_{t+1} are mutually independent, and each follows the normal distribution under $\tilde{P}_t$, or $u_{t+1} \sim N(e_{u,t+1}, \Sigma_u)$ and $v_{t+1} \sim N(e_{v,t+1}, \Sigma_v)$. This shows that the specification of z_t in Eq. (A.1) satisfies $E_t(z_t u_{t+1}) = e_{u,t+1}$ and $E_t(z_t v_{t+1}) = e_{v,t+1}$ in our model. Moreover, according to the Radon–Nikodym theorem, this form of the Radon–Nikodym derivative is the unique form that satisfies these requirements in our model.

Finally, given Eq. (A.1), the relative entropy constraint can be simplified as:

$$\begin{aligned}E_t(z_t \log z_t) &= \int_\Omega \log z_t d\tilde{P}_t \\ &= \tfrac{1}{2} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} + \tfrac{1}{2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} \leq \eta.\end{aligned} \tag{A.8}$$

Consequently, we can further divide the relative entropy constraint as follows:

$$\frac{1}{2} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} \leq \eta_1, \quad \frac{1}{2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} \leq \eta_2. \tag{A.9}$$

Apparently, when the two constraints in Eq. (A.9) both hold, the relative entropy constraint in Eq. (A.8) holds and also $\eta = \eta_1 + \eta_2$.

Appendix B. Proof of Proposition 1

For the optimization problem (10), denote:

$$Q_1(x_t, f_t) = x_t^T B f_t - \frac{\gamma}{2} x_t^T \Sigma_u x_t - \frac{1}{2} \Delta x_t^T \bar{\Lambda} \Delta x_t, \tag{B.1}$$

where $\bar{\Lambda} = \rho^{-1}\Lambda$. Then, the value function can be written as:

$$V_t = \max_{x_t} \min_{\binom{e_{u,t+1}}{e_{v,t+1}}} \Big[Q_1(x_t, f_t) + \tilde{E}_t\Big(x_t^T u_{t+1} + \frac{1}{2\theta_1} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} + \frac{1}{2\theta_2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} + \rho V_{t+1}\Big)\Big]. \tag{B.2}$$

Let

$$V_{t+1} = -\frac{1}{2} x_t^T A_{xx} x_t + x_t^T A_{xf} f_{t+1} + f_{t+1}^T A_{ff} f_{t+1} + A_0. \tag{B.3}$$

According to Gârleanu and Pedersen (2013), we obtain:

$$V_t = \max_{x_t} \min_{\binom{e_{u,t+1}}{e_{v,t+1}}} [Q_1(x_t, f_t) + \rho Q_2(x_t, f_t) + f_u(e_{u,t+1}, x_t) + f_v(e_{v,t+1}, x_t, f_t)], \tag{B.4}$$

where $Q_2(x_t, f_t) = -\frac{1}{2} x_t^T A_{xx} x_t + x_t^T A_{xf} \Phi f_t + f_t^T \Phi^T A_{ff} \Phi f_t + A_0 + (\Sigma_v^{1/2} e)^T A_{ff} (\Sigma_v^{1/2} e)$,[21]

$$f_u(e_{u,t+1}, x_t) = \frac{1}{2\theta_1} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} + x_t^T e_{u,t+1},$$

$$\begin{aligned} f_v(e_{v,t+1}, x_t, f_t) &= \frac{1}{2\theta_2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} + \rho e_{v,t+1}^T A_{ff} e_{v,t+1} \\ &\quad + \rho x_t^T A_{xf} e_{v,t+1} + 2\rho (\Phi f_t)^T A_{ff} e_{v,t+1} \\ &= e_{v,t+1}^T \Big(\frac{1}{2\theta_2} \Sigma_v^{-1} + \rho A_{ff}\Big) e_{v,t+1} + \rho \Big(x_t^T A_{xf} + 2(\Phi f_t)^T A_{ff}\Big) e_{v,t+1}. \end{aligned}$$

Based on Eq. (B.4), if $\Sigma_v^{-1} + 2\rho\theta_2 A_{ff} > 0$ holds (Condition 1 in Proposition 1), then the expectations of u_{t+1} and v_{t+1} under the worst-case scenario are:

$$e_{u,t+1}^* = \arg\min_{e_{u,t+1}} f_u(e_{u,t+1}, x_t) = -\theta_1 \Sigma_u x_t, \tag{B.5}$$

$$\begin{aligned} e_{v,t+1}^* &= \arg\min_{e_{v,t+1}} f_v(e_{v,t+1}, x_t, f_t) \\ &= -\rho\theta_2 (\Sigma_v^{-1} + 2\rho\theta_2 A_{ff})^{-1} (A_{xf}^T x_t + 2A_{ff} \Phi f_t). \end{aligned} \tag{B.6}$$

Plugging Eqs. (B.5) and (B.6) into Eq. (B.4) and rearranging it yields:

$$\begin{aligned} V_t(x_{t-1}, f_t) = \max_{x_t} \Big\{ &Q_1(x_t, f_t) + \rho Q_2(x_t, f_t) - \tfrac{\theta_1}{2} x_t^T \Sigma_u x_t \\ &- \tfrac{\rho^2\theta_2}{2} (A_{xf}^T x_t + 2A_{ff}\Phi f_t)^T (\Sigma_v^{-1} + 2\rho\theta_2 A_{ff})^{-1} (A_{xf}^T x_t + 2A_{ff}\Phi f_t) \Big\}. \end{aligned} \tag{B.7}$$

To solve the above optimization problem, we rearrange V_t as follows:

$$V_t = \max_{x_t} \Big(-\frac{1}{2} x_t^T J_1 x_t + x_t^T J_2 + J_3\Big), \tag{B.8}$$

where

$$J_1 = (\gamma + \theta_1)\Sigma_u + \bar{\Lambda} + \rho A_{xx} + \rho^2\theta_2 A_{xf} (\Sigma_v^{-1} + 2\rho\theta_2 A_{ff})^{-1} A_{xf}^T,$$

$$J_2 = B f_t + \bar{\Lambda} x_{t-1} + \rho A_{xf} (I + 2\rho\theta_2 \Sigma_v A_{ff})^{-1} \Phi f_t.$$[22]

If $J_1 > 0$ holds (Condition 2 in Proposition 1), then Problem (10) has a unique solution

$$x_t^* = J_1^{-1} J_2 = J_1^{-1}\Big(B f_t + \bar{\Lambda} x_{t-1} + \rho A_{xf} (I + 2\rho\theta_2 \Sigma_v A_{ff})^{-1} \Phi f_t\Big). \tag{B.9}$$

Plugging Eq. (B.9) into V_t gives the following:

$$A_{xx} = \bar{\Lambda} - \bar{\Lambda}^T J_1^{-1} \bar{\Lambda}, \tag{B.10}$$

$$A_{xf} = \bar{\Lambda} J_1^{-1} (B + \rho A_{xf} (I + 2\rho\theta_2 \Sigma_v A_{ff})^{-1} \Phi), \tag{B.11}$$

$$A_{ff} = \frac{1}{2} (\bar{\Lambda}^{-1} A_{xf})^T J_1 (\bar{\Lambda}^{-1} A_{xf}) + \rho \Phi^T A_{ff} (I + 2\rho\theta_2 \Sigma_v A_{ff})^{-1} \Phi. \tag{B.12}$$

Appendix C. Proof of Proposition 2

Rearranging Eq. (B.9) yields:

$$\begin{aligned} &\big((\gamma + \theta_1)\Sigma_u + \bar{\Lambda} + \rho A_{xx}\big) x_t^* = B f_t + \bar{\Lambda} x_{t-1} \\ &+ \rho A_{xf} (I + 2\rho\theta_2 \Sigma_v A_{ff})^{-1} \Phi f_t - \rho^2\theta_2 A_{xf} (\Sigma_v^{-1} + 2\rho\theta_2 A_{ff})^{-1} A_{xf}^T x_t^* \\ &= B f_t + \bar{\Lambda} x_{t-1} + \rho A_{xf} \tilde{E}_t^*(f_{t+1}), \end{aligned} \tag{C.1}$$

where $\tilde{E}_t^*(\cdot)$ is the conditional expectation operator under the worst-case scenario. Thus, the optimal strategy is

$$x_t^* = \Big(I - (\kappa + \bar{\Lambda})^{-1} \kappa\Big) x_{t-1} + (\kappa + \bar{\Lambda})^{-1} \kappa \cdot aim_t, \tag{C.2}$$

where

$$\kappa = (\gamma + \theta_1)\Sigma_u + \rho A_{xx},$$

$$aim_t = \kappa^{-1} (B f_t + \rho A_{xf} \tilde{E}_t^*(f_{t+1})).$$

Next, we take θ_1 as an example to show that the second part of Proposition 2 holds.

Suppose that at $t+1$, for the ambiguity aversion coefficients $0 < \theta_1^{(1)} < \theta_1^{(2)}$, the value function V_{t+1} satisfies $V_{t+1}^{\theta_1^{(1)}}(x_t, f_{t+1}) \geq V_{t+1}^{\theta_1^{(2)}}(x_t, f_{t+1})$ for all (x_t, f_{t+1}). Then, for any given (x_t, f_{t+1}) and $e_{u,t+1}$, the following is true:

$$\begin{aligned} &\tilde{E}_t\Big(x_t^T u_{t+1} + \tfrac{1}{2\theta_1^{(1)}} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} + \tfrac{1}{2\theta_2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} + \rho V_{t+1}^{\theta_1^{(1)}}\Big) \\ &\geq \tilde{E}_t\Big(x_t^T u_{t+1} + \tfrac{1}{2\theta_1^{(2)}} e_{u,t+1}^T \Sigma_u^{-1} e_{u,t+1} + \tfrac{1}{2\theta_2} e_{v,t+1}^T \Sigma_v^{-1} e_{v,t+1} + \rho V_{t+1}^{\theta_1^{(2)}}\Big). \end{aligned} \tag{C.3}$$

When minimizing both sides with respect to $e_{u,t+1}$, the direction of Inequality (C.3) does not change. This holds true, when maximizing both sides with respect to x_t. Thus, we have $V_t^{\theta_1^{(1)}}(x_{t-1}, f_t) \geq V_t^{\theta_1^{(2)}}(x_{t-1}, f_t)$ for any (x_{t-1}, f_t). Given Eq. (B.3), we can obtain the following inequalities:

$$A_{xx}^{\theta_1^{(1)}} \leq A_{xx}^{\theta_1^{(2)}}, \quad A_{ff}^{\theta_1^{(1)}} \geq A_{ff}^{\theta_1^{(2)}}. \tag{C.4}$$

Similarly, we can prove that the above inequalities hold true for θ_2. Thus, the result (ii) in Proposition 2 follows immediately.

---

[21] In $Q_2(x_t, f_t)$, e is a unit vector, and the last term is obtained from $E_t(z_{t+1} v_{t+1}^T A_{ff} v_{t+1}) = e_{v,t+1}^T A_{ff} e_{v,t+1} + (\Sigma_v^{1/2} e)^T A_{ff} (\Sigma_v^{1/2} e)$.

[22] I is the identity matrix. Since J_3 does not impact the optimal trading strategy, its expression is not provided for brevity.

Appendix D. Proof of Proposition 3

Consider the Markowitz's mean-variance model with ambiguity aversion but without transaction costs:

$$\max_{x_t}\min_{e_{u,t+1}} x_t^T\tilde{E}_t(r_{t+1})-\frac{\gamma}{2}x_t^T\Sigma_u x_t+\frac{1}{2\theta_1}e_{u,t+1}^T\Sigma_u^{-1}e_{u,t+1}.$$

Similar to the proof of Proposition 1, we can prove that $e_{u,t+1}^*=-\theta_1\Sigma_u x_t$. Plugging it into the above optimization problem yields:

$$\max_{x_t} x_t^T Bf_t-\frac{\gamma+\theta_1}{2}x_t^T\Sigma_u x_t.$$

The optimal solution is given by:

$$Markowitz_t=((\gamma+\theta_1)\Sigma_u)^{-1}Bf_t. \tag{D.1}$$

Then, the aim portfolio can be rewritten as:

$$aim_t=\kappa^{-1}\Big((\kappa-\rho A_{xx})Markowitz_t+\rho A_{xx}\big(A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})\big)\Big). \tag{D.2}$$

Further, we note that:

$$\begin{aligned}\kappa\tilde{E}_t^*(aim_{t+1})&=\big(B+\rho A_{xf}\Phi\big)\tilde{E}_t^*(f_{t+1})+\rho A_{xf}\tilde{E}_t^*(e_{v,t+2}^*)\\&=\big(B+\rho A_{xf}(I+2\rho\theta_2\Sigma_v A_{ff})^{-1}\Phi\big)\tilde{E}_t^*(f_{t+1})\\&\quad-\rho^2\theta_2 A_{xf}(\Sigma_v^{-1}+2\rho\theta_2 A_{ff})^{-1}A_{xf}^T\tilde{E}_t^*(x_{t+1}^*),\end{aligned} \tag{D.3}$$

or

$$\kappa\tilde{E}_t^*(aim_{t+1})=J_1\bar{\Lambda}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})-(J_1-\kappa-\bar{\Lambda})\tilde{E}_t^*(x_{t+1}^*). \tag{D.4}$$

Given Eq. (18), we have:

$$\tilde{E}_t^*(x_{t+1}^*)=x_t^*+\bar{\Lambda}^{-1}A_{xx}((A_{xx})^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})-x_t^*). \tag{D.5}$$

Substituting Eq. (D.5) into Eq. (D.4) yields:

$$\kappa\tilde{E}_t^*(aim_{t+1})=(\kappa+\bar{\Lambda})\bar{\Lambda}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})-(J_1-\kappa-\bar{\Lambda})J_1^{-1}\bar{\Lambda}x_t^*. \tag{D.6}$$

Thus, we can obtain the conclusion (ii) in Proposition 3:

$$\begin{aligned}A_{xx}^{-1}A_{xf}\tilde{E}_t^*(f_{t+1})&=A_{xx}^{-1}\bar{\Lambda}(\kappa+\bar{\Lambda})^{-1}\big(\kappa\tilde{E}_t^*(aim_{t+1})\\&\quad+(J_1-\kappa-\bar{\Lambda})J_1^{-1}\bar{\Lambda}x_t^*\big)\\&=(\kappa+A_{xx}-\kappa J_1^{-1}\bar{\Lambda})^{-1}\big(\kappa\tilde{E}_t^*(aim_{t+1})+(A_{xx}-\kappa J_1^{-1}\bar{\Lambda})x_t^*\big).\end{aligned}$$

Meanwhile, it is easy to prove that if there is no ambiguity aversion to return predictors, then

$$A_{xx}-\kappa J_1^{-1}\bar{\Lambda}=0.$$

Appendix E. Proof of Proposition 4

According to Eqs. (16) and (B.6), the aim portfolio can be written as:

$$\begin{aligned}aim_t&=\kappa^{-1}\big(Bf_t+\rho A_{xf}(\Phi f_t+e_{v,t+1}^*)\big)\\&=\kappa^{-1}\big(J_1\bar{\Lambda}^{-1}A_{xf}f_t-(J_1-\kappa-\bar{\Lambda})x_t^*\big).\end{aligned} \tag{E.1}$$

Substituting Eqs. (18) and (B.11) into Eq. (E.1) yields:

$$aim_t=\kappa^{-1}\big((\kappa+\bar{\Lambda})J_1^{-1}K_f f_t+(J_1-\kappa-\bar{\Lambda})J_1^{-1}K_x x_{t-1}\big), \tag{E.2}$$

where

$$K_f=B+\rho A_{xf}(I+2\rho\theta_2\Sigma_v A_{ff})^{-1}\Phi,$$

$$K_x=-\bar{\Lambda}.$$

This is the conclusion (i) in Proposition 4.

References

Anderson, E.W., Ghysels, E., Juergens, J.L., 2009. The impact of risk and uncertainty on expected returns. J. Financ. Econ. 94 (2), 233–263.
Anderson, E.W., Hansen, L.P., Sargent, T.J., 2003. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. J. Eur. Econ. Assoc. 1 (1), 68–123.
Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. J. Finance 68 (3), 929–985.
Barberis, N., 2000. Investing for the long run when returns are predictable. J. Finance 55 (1), 225–264.
Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. J. Financ. Econ. 116 (1), 111–120.
Black, F., Litterman, R., 1992. Global portfolio optimization. Financ. Anal. J. 48 (5), 28–43.
Branger, N., Larsen, L.S., Munk, C., 2013. Robust portfolio choice with ambiguity and learning about return predictability. J. Bank. Finance 37 (5), 1397–1411.
Cao, H.H., Wang, T., Zhang, H.H., 2005. Model uncertainty, limited market participation, and asset prices. Rev. Financ. Stud. 18 (4), 1219–1251.
Chen, Z., Epstein, L., 2002. Ambiguity, risk, and asset returns in continuous time. Econometrica 70 (4), 1403–1443.
Delage, E., Ye, Y., 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper. Res. 58 (3), 595–612.
DeMiguel, V., Martín-Utrera, A., Nogales, F.J., 2015. Parameter uncertainty in multiperiod portfolio optimization with transaction costs. J. Financ. Quant. Anal. 50 (6), 1443–1471.
Demiguel, V., Nogales, F.J., 2009. Portfolio selection with robust estimation. Oper. Res. 57 (3), 560–577.
Epstein, L.G., Schneider, M., 2008. Ambiguity, information quality and asset pricing. J. Finance 63 (1), 197–228.
Epstein, L.G., Wang, T., 1994. Intertemporal asset pricing under Knightian uncertainty. Econometrica 62 (2), 283–322.
Erb, C.B., Harvey, C.R., 2006. The strategic and tactical value of commodity futures. Financ. Anal. J. 62 (2), 69–97.
Garlappi, L., Uppal, R., Wang, T., 2007. Portfolio selection with parameter and model uncertainty: a multi-prior approach. Rev. Financ. Stud. 20 (1), 41–81.
Gârleanu, N., Pedersen, L.H., 2013. Dynamic trading with predictable returns and transaction costs. J. Finance 68 (6), 2309–2340.
Gilboa, I., Schmeidler, D., 1989. Max-min expected utility with non-unique prior. J. Math. Econ. 18 (2), 141–153.
Glasserman, P., Xu, X., 2013. Robust portfolio control with stochastic factor dynamics. Oper. Res. 61 (4), 874–893.
Goldfarb, D., Iyengar, G., 2003. Robust portfolio selection problems. Math. Oper. Res. 28 (1), 1–38.
Grinblatt, M., Moskowitz, T.J., 2004. Predicting stock price movements from past returns: the role of consistency and tax-loss selling. J. Financ. Econ. 71 (3), 541–579.
Hansen, L.P., Sargent, T.J., 2001. Robust control and model uncertainty. Am. Econ. Rev. 91 (2), 60–66.
Hansen, L.P., Sargent, T.J., Turmuhambetova, G., Williams, N., 2006. Robust control and model misspecification. J. Econom. Theory 128 (1), 45–90.
Iyengar, G.N., 2005. Robust dynamic programming. Math. Oper. Res. 30 (2), 257–280.
Jeong, D., Kim, H., Park, J.Y., 2015. Does ambiguity matter? Estimating asset pricing models with a multiple-priors recursive utility. J. Financ. Econ. 115, 361–382.
Kan, R., Zhou, G., 2007. Optimal portfolio choice with parameter uncertainty. J. Financ. Quant. Anal. 42 (3), 621–656.
Ljungqvist, L., Sargent, T., 2004. Recursive Macroeconomic Theory, second ed. MIT Press, Cambridge, MA.
Maenhout, P.J., 2004. Robust portfolio rules and asset pricing. Rev. Financ. Stud. 17 (4), 951–983.
Maenhout, P.J., 2006. Robust portfolio rules and detection-error probabilities for a mean-reverting risk premium. J. Econom. Theory 128 (1), 136–163.
Markowitz, H., 1952. Portfolio selection. J. Finance 7 (1), 77–91.
Merton, R.C., 1969. Lifetime portfolio selection under uncertainty: the continuous–time case. Rev. Econ. Stat. 51 (3), 247–257.
Moskowitz, T.J., Ooi, Y.H., Pedersen, L.H., 2012. Time series momentum. J. Financ. Econ. 104 (2), 228–250.
Pástor, Ľ., Stambaugh, R.F., 2012. Are stocks really less volatile in the long run? J. Finance 67 (2), 431–478.
Wooldridge, J.M., 2013. Introductory Econometrics: A Modern Approach, fifth ed. South-Western, Mason, OH.