Why stocks may disappoint ☆

Andrew Ang[a,b,c], Geert Bekaert[a,c],*, Jun Liu[d]

[a]*Columbia University, Columbia Business School, 3022 Broadway Uris Hall, New York NY 10027, USA*
[b]*University of Southern California*
[c]*NBER*
[d]*C509 Anderson School, UCLA CA 90095, USA*

Abstract

We provide a formal treatment of both static and dynamic portfolio choice using the Disappointment Aversion preferences of Gul (1991. Econometrica 59(3), 667–686), which imply asymmetric aversion to gains versus losses. Our dynamic formulation nests the standard CRRA asset allocation problem as a special case. Using realistic data generating processes, we find reasonable equity portfolio allocations for disappointment averse investors with utility functions exhibiting low curvature. Moderate variation in parameters can robustly generate

substantial cross-sectional variation in portfolio holdings, including optimal non-participation in the stock market.

1. Introduction

The U.S. population displays a surprisingly large variation in equity holdings, including a majority of households that hold no stocks at all (see, among many others, Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995; Heaton and Lucas, 1997; Vissing-Jørgensen, 2002). Driven in part by a large equity premium, standard portfolio choice models often predict large equity positions for most investors and fail to generate the observed cross-sectional variation in portfolio choice (see, for example, Campbell and Viceira, 1999). In an effort to explain these portfolio puzzles, one approach is to combine transactions costs, such as a fixed cost to entering the stock market, with various sources of background risk. Another approach considers heterogeneous preferences. Standard constant relative risk aversion (CRRA) preferences cannot resolve these puzzles, since they cannot generate non-participation at any level of risk aversion, except in the presence of large transactions costs (see Liu and Loewenstein, 2002). However, a rapidly growing literature builds on the framework of Kahneman and Tversky (1979) and investigates asset allocation in the presence of loss aversion—that is, investors are assumed to maintain an asymmetric attitude towards gains versus losses (see Benartzi and Thaler, 1995; Berkelaar and Kouwenberg, 2000; Aït-Sahalia and Brandt, 2001; Gomes, 2003).[1] Portfolio choice problems with loss aversion generate more realistic (that is, lower) equity holdings than standard models.

We provide a formal treatment of portfolio choice in the presence of loss aversion; however, rather than relying on Kahneman and Tversky (1979)'s behavioral prospect theory, we use the axiomatic Disappointment Aversion (DA) framework of Gul (1991). Gul's preferences are a one-parameter extension of the expected utility framework and have the characteristic that good outcomes, i.e., outcomes above the certainty equivalent, are downweighted relative to bad outcomes. The larger weight given to outcomes which are bad in a relative sense gives rise to the name "disappointment-averse" preferences, but they also imply an aversion to losses.

In the literature, DA preferences have only appeared in equilibrium models with consumption, not in portfolio choice problems. For instance, Epstein and Zin (1990,

---

[1]Roy (1952), Maenhout (2004), Stutzer (2000) and Epstein and Schneider (2004) provide an alternative treatment of an investor's asymmetric response to gains and losses, by modelling agents who first minimize the possibility of undesirable outcomes.

2001) embed a number of alternative preferences, including DA preferences, into an infinite horizon consumption-based asset pricing model with recursive preferences. Bekaert et al. (1997) consider asset return predictability in the context of an international consumption-based asset pricing model with DA preferences. Both these models endogenously generate more realistic equity premiums than models with standard preferences. When we investigate portfolio choice under DA preferences, we show that investors with a sufficient degree of disappointment aversion do not participate in the equity market.

Loss aversion is not only introspectively an attractive feature of preferences, but as we demonstrate, it also circumvents the problem posed by Rabin (2000): within the expected utility framework, anything but near risk neutrality over modest stakes implies manifestly unrealistic risk aversion over large stakes. Whereas both behavioral Kahneman and Tversky (1979) loss aversion (LA) preferences and DA preferences share this advantage, DA preferences are a useful alternative to LA preferences for three main reasons.

First, DA utility is axiomatic and normative. Although DA utility is non-expected utility, it is firmly grounded in formal decision theory. Gul (1991) replaces the independence axiom underlying expected utility by a slightly weaker axiom that accommodates the violation of the independence axiom commonly observed in experiments (the Allais paradox), but retains all the other assumptions and axioms underlying expected utility. The similarity between the DA utility and expected utility frameworks yields a number of benefits. For example, DA preferences embed CRRA preferences as a special case. Thus, the portfolio implications of loss aversion are directly comparable to a large body of empirical work in standard preference settings; moreover, they allow us to retain as much of the insight offered by expected utility theory as possible.

Second, we demonstrate that with LA utility, finite optimal solutions do not always exist, particularly with empirically relevant data generating processes (DGPs). Third, DA preferences eliminate the arbitrary choices required by LA. In particular, Kahneman and Tversky (1979)'s prospect theory offers no guidance with regard to choosing and updating the reference point against which gains and losses are compared. With DA utility, on the other hand, the reference point is the certainty equivalent and hence is endogenous. Moreover, we propose a tractable and natural dynamic DA setting that nests a dynamic CRRA problem and endogenously updates the reference point.

Our paper proceeds in four steps. Sections 2 and 3 develop a portfolio choice framework under DA preferences. Specifically, in Section 2, we focus on a static setting and show that DA preferences can generate stock non-participation. Section 3 generalizes the set-up to a dynamic long-horizon framework that preserves CRRA preferences as a special case of DA preferences. Given the popularity of LA preferences, we consider them as an alternative in Section 4. Section 5 explores the empirical implications of portfolio choice under DA preferences. We calibrate two DGPs (one with and one without predictability) to U.S. data on Treasury bills and stock returns and we then examine static and dynamic asset allocation for a wide set of parameters. Finally, Section 6 concludes.

2. Static asset allocation under DA preferences

We use the case of standard CRRA utility to set up the basic asset allocation framework in Section 2.1. Section 2.2 extends the framework to DA preferences and derives a stock market non-participation result.

2.1. CRRA utility

The investment opportunity set of an investor with initial wealth W_0 consists of a risky asset and a riskless bond. The bond yields a certain return of r and the risky asset yields an uncertain return of y, both continuously compounded. The investor chooses the proportion of her initial wealth to invest in the risky asset α to maximize the expected utility of end-of-period wealth W, which is uncertain. The terminal wealth problem avoids the computational complexities of allowing for consumption decisions and makes our work comparable to both the standard portfolio choice literature (e.g. Kim and Omberg, 1996; Brennan et al., 1997; Liu, 1999; Barberis, 2000), and the asset allocation with loss aversion literature (e.g. Benartzi and Thaler, 1995; Berkelaar and Kouwenberg, 2000; Aït-Sahalia and Brandt, 2001; Gomes, 2003).

Formally, the problem is

$$\max_{\alpha} \ \mathrm{E}[U(W)], \tag{1}$$

where W is given by

$$W = \alpha W_0(\exp(y) - \exp(r)) + W_0 \exp(r). \tag{2}$$

Denoting risk aversion by γ, under CRRA preferences the utility function $U(W)$ takes the form

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma}. \tag{3}$$

Since CRRA utility is homogenous in wealth, we set $W_0 = 1$.

The first-order condition (FOC) of Eq. (1) is solved by choosing α such that

$$\int_{-\infty}^{\infty} W^{-\gamma}(\exp(y) - \exp(r)) \, \mathrm{d}F(y) = 0, \tag{4}$$

where $F(\cdot)$ is the cumulative density function of the risky asset's return. This expectation can be computed by numerical quadrature as described in Tauchen and Hussey (1991). This procedure involves replacing the integral with a probability-weighted sum, i.e.,

$$\sum_{s=1}^{N} p_s W_s^{-\gamma}(\exp(y_s) - \exp(r)) = 0. \tag{5}$$

The N values of the risky asset return, $\{y_s\}_{s=1}^{N}$, and the associated probabilities, $\{p_s\}_{s=1}^{N}$, are chosen by a Gaussian quadrature rule, where W_s represents the investor's

terminal wealth when the risky asset return is y_s. Quadrature approaches to solving asset allocation problems have been taken by Balduzzi and Lynch (1999), Campbell and Viceira (1999), and Ang and Bekaert (2002), among others. For future reference, we denote the excess return $\exp(y) - \exp(r)$ by x_e.

2.2. Disappointment aversion

2.2.1. Definition

DA utility μ_W is implicitly defined by

$$U(\mu_W) = \frac{1}{K} \left(\int_{-\infty}^{\mu_W} U(W)\, dF(W) + A \int_{\mu_W}^{\infty} U(W)\, dF(W) \right), \tag{6}$$

where $U(\cdot)$ is the felicity function that we choose to be power utility (i.e., of the form $U(W) = W^{(1-\gamma)}/(1 - \gamma)$), $A \leqslant 1$ is the coefficient of disappointment aversion, $F(\cdot)$ is the cumulative distribution function for wealth, μ_W is the certainty equivalent (the certain level of wealth that generates the same utility as the portfolio allocation determining W), and K is a scalar given by

$$K = \Pr(W \leqslant \mu_W) + A\, \Pr(W > \mu_W). \tag{7}$$

If $0 \leqslant A < 1$, the outcomes below the certainty equivalent are weighted more heavily than the outcomes above the certainty equivalent. These preferences are outside the standard expected utility framework because the level of utility at the optimum (or the certainty equivalent of wealth) appears on the right-hand side. Routledge and Zin (2003) provide an extension of the Gul (1991) framework where the reference point can be below the certainty equivalent, but we restrict our analysis to the case where outcomes are compared to the certainty equivalent. Although this is a non-expected utility function, CRRA preferences are a special case for $A = 1$. When $A < 1$ individuals are averse to losses, or disappointment averse.

For DA preferences, the optimization problem becomes

$$\max_{\alpha}\ U(\mu_W), \tag{8}$$

where the certainty equivalent is defined in Eq. (6) and end-of-period wealth W is given by Eq. (2). For $U(\cdot)$ given by power utility, optimal utility remains homogenous in wealth and we set $W_0 = 1$. The implicit definition of μ_W makes the optimization problem non-trivial (see Epstein and Zin, 1989, 2001), so we relegate a rigorous treatment to an appendix, available upon request.

The FOC for the DA investor is

$$\frac{1}{A}\, E\left[\frac{\partial U(W)}{\partial W}(\exp(y) - \exp(r)) \mathbf{1}_{\{W \leqslant \mu_W\}} \right]$$
$$+ E\left[\frac{\partial U(W)}{\partial W}(\exp(y) - \exp(r)) \mathbf{1}_{\{W > \mu_W\}} \right] = 0, \tag{9}$$

where $\mathbf{1}$ is an indicator function. If μ_W were known, we could solve Eq. (9) for α in the same way as in the case of expected utility. The only difference is that for states

below μ_W, the original utilities have to be scaled up by $1/A$. However, μ_W is itself a function of the outcome of optimization (that is, μ_W is a function of α). Hence, Eq. (9) must be solved simultaneously with Eq. (6) which defines μ_W.

The similarity between expected utility and DA preferences allows us to derive a new algorithm to solve the DA asset allocation problem (Eqs. (6) and (9)). Specifically, the DA problem can be viewed as a CRRA maximization problem with a changed probability distribution such that the probabilities above the certainty equivalent are downweighted by A and the new probabilities are then re-normalized. We present the details of this new approach in Appendix A.

2.2.2. Non-participation

In an expected utility framework investors always hold a positive amount of equity if the risk premium is positive. However, with DA preferences it may be optimal to not participate in the stock market.[2] This immediately implies that CRRA preferences cannot deliver the same empirically relevant dispersion in stock holdings that we can obtain with DA preferences.

Proposition 2.1 (*Non-participation under disappointment aversion*). *Suppose the expected excess return* $E(x_e)$ *is positive. Then under DA preferences, there exists a level of* A, $A = A^*$, *such that for* $A < A^*$, *investors hold no equity. This non-participation level* A^* *is independent of risk aversion* γ.

Proof. See Appendix B.

Appendix B shows that A^* is given by

$$A^* = -\frac{E[x_e | x_e \leqslant 0] \Pr(x_e \leqslant 0)}{E[x_e | x_e > 0] \Pr(x_e > 0)}.$$

The intuition for the non-participation result is straightforward. As A decreases, a DA investor becomes more averse to losses. Consequently, her optimal allocation to equities decreases. At a particular A, say A^*, the optimal portfolio weight becomes zero. If x_e were to have a discrete distribution, then the more dramatic the negative excess return states and the higher their probabilities, the less disappointment aversion it takes for α^* to reach zero and the higher A^* will be. This critical point does not depend on the curvature of the utility function since as α^* approaches zero, the certainty equivalent approaches $R_f = \exp(r)$ and the marginal utility terms cancel out in the FOCs. For $A < A^*$, the optimal allocation remains zero. Shorting is not optimal, since the certainty equivalent is increasing in α for $\alpha < 0$. This occurs because for $\alpha < 0$, negative excess return states have higher wealth than R_f and hence are downweighted.

The fact that A^* only depends on the excess return distribution generalizes to the multiple risky asset case when asset returns are jointly normally distributed. In that case, two fund separation applies and the excess return distribution of the tangency

---

[2]Dow and Werlang (1992), Epstein and Schneider (2004), and Liu (2002) show that a similar non-participation result can be obtained with ambiguity-averse preferences.

portfolio determines A^*. However, under alternative distributions, non-participation in one asset need not imply non-participation in another asset.

To illustrate non-participation, consider a binomial model to approximate the excess return $x_e \equiv \exp(y) - \exp(r)$. In this simplest case, the excess return can be u with probability p and d with probability $1 - p$. Note that in standard notation with binomial trees, u and d refer to gross returns of the risky asset, but here we use them to indicate the excess return states. Under this setting, the critical level of A which results in non-participation is

$$A^* = -\frac{(1-p)d}{pu}.$$

To calibrate the binomial tree, we assume that U.S. equity returns are log-normally distributed. For quarterly stock return data from 1926 to 1998, we find that the mean continuously compounded equity return is 10.63% and the volatility is 21.93% (see Table 1). The mean continuously compounded short rate is 4.08%, so that the continuously compounded equity premium is 6.55%. Denote the implied average simple gross return and volatility by m and s, respectively. We match these two moments by setting $u = m + s - \exp(r) = 0.3504$ and $d = m - s - \exp(r) = -0.1553$, with $p = 0.5$. The implied simple excess return premium from the binomial approximation is 9.76%. For this model, we find that $A^* = 0.44$. That is, if an investor's utility in the loss region, relative to the utility in the gain region, is scaled up by $1/0.44 = 2.27$, she chooses to not participate in the market. To appreciate the importance of this result, suppose we would like to generate low stock holdings using a CRRA utility function, assuming the binomial stock model as the DGP. For comparison, to obtain an optimal equity allocation of 5%, γ must be set equal to 33.7.

However, the extreme states inherent in the two-date approximation exaggerate the non-participation region. We can determine the correct answer for a log-normal distribution by numerical integration, and determine that $A^* = 0.36$. A two-period/three-date binomial tree is sufficient to approximate the log-normal solution much more closely. At time 0, there are two possible states to be realized at time 1, and at time 1 (after six months), there are again two possible states for time 2 from each of the two branches of time 1, giving a total of three possible states at the end of the

Table 1
Summary statistics of the data

	Stock	T-bill	Excess
Mean	0.1063	0.0408	0.0655
Std	0.2193	0.0173	0.2197
Autocorrelation	−0.0575	0.9273	−0.0532

All data are quarterly. Stock data represent S&P 500 returns, with dividends. The T-bill data are three-month T-bill returns from CRSP. Excess returns refer to stock returns in excess of T-bill returns. All returns are continuously compounded. The mean and standard deviation are annualized by multiplying by four and two, respectively. The data sample is 1926–1998.

year with a recombining tree. At each branch, the probability of an upward move is p. Calibrating the tree, the three states are $uu = 0.4808$, $ud = du = 0.0699$, and $dd = -0.2301$. Note that only the lowest state is disappointing. In this case,

$$A^* = -\frac{(1-p)^2 dd}{p^2 uu + 2p(1-p)ud} = 0.37. \tag{10}$$

Because the historical equity premium we use is high and its estimation is subject to substantial sampling error, Fig. 1 shows the region of stock non-participation as a function of different expected equity returns. To produce the plot, we vary the expected equity return in the binomial gamble from 0% to 20% and plot A^* on the vertical axis. The circle shows the empirical expected total equity return of 10.63%, or the empirical risk premium of 6.55%, which corresponds to $A^* = 0.37$. For an
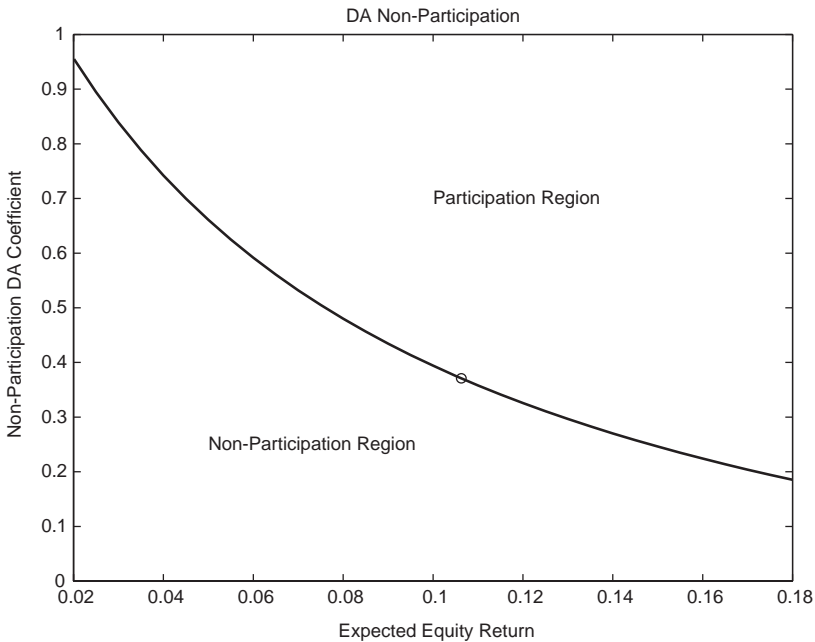

Fig. 1. Stock market participation under disappointment aversion. The plot shows the critical level A^* required for an investor to hold equity as a function of the total excess expected return (on the x-axis). For any A higher than the solid line denoting A^*, investors hold a positive amount of equity ("participation region"). For any A lower than the line investors hold zero equity ("non-participation region"). To produce the plot, we use a binomial tree with three states (two periods) for excess returns and graph A^* given by Eq. (10). To calibrate the binomial tree, let $\bar{\mu}$ denote the mean and $\bar{\sigma}$ denote the standard deviation of continuously compounded returns (10.63% and 21.93%, respectively) and let the continuously compounded risk-free rate be $r = 4.08\%$. Define the log-normal mean for one period as $m = \exp(\frac{1}{2}\bar{\mu} + \frac{1}{2}\bar{\sigma}^2)$ and standard deviation $s = \exp(\frac{1}{2}\bar{\mu} + \frac{1}{2}\bar{\sigma}^2)\sqrt{(\exp(\bar{\sigma}^2) - 1)}$ for a half-year node. The three states uu, ud, and dd at the end of the year are given by $uu = \tilde{u}^2 - \exp(r)$, $ud = \tilde{u} \times \tilde{d} - \exp(r)$, and $dd = \tilde{d}^2 - \exp(r)$, where $\tilde{u} = m + s$ and $\tilde{d} = m - s$. The circle shows $A^* = 0.37$, corresponding to the empirical total expected return 10.63%, or an expected excess return of 6.55%.

expected excess return of 16%, A^* drops to just over 0.20. Values of A above the line in Fig. 1 induce investors to participate in the market. This is the "participation" region. Values of A below A^* define the "non-participation" region, where investors hold no equity. While this is an illustration of non-participation with a simple binomial model of equity returns, we compute optimal non-participation regions for more realistic DGPs in Section 5.

3. Dynamic asset allocation under DA preferences

We embed DA preferences in a dynamic asset allocation setting, which nests dynamic CRRA asset allocation as a special case. The dynamic setting is important for several reasons. First, the recent empirical portfolio choice literature has devoted much attention to the dynamic effects of asset allocation (see Brennan et al., 1997, among many others). Second, our dynamic extension of DA utility has a number of desirable mathematical and rational properties that are hard to replicate with LA preferences. Finally, our dynamic extension enables the standard technical tools, in particular dynamic programming, to be used with portfolio choice problems with DA utility. Section 3.1 discusses how we solve the dynamic asset allocation problem under CRRA utility before we present our formulation of dynamic portfolio choice under DA preferences in Section 3.2.

3.1. Dynamic CRRA utility

Our problem for dynamic CRRA utility is to find a series of portfolio weights $\alpha = \{\alpha_t\}_{t=0}^{T-1}$ to maximize

$$\max_{\alpha_0,\ldots,\alpha_{T-1}} \mathrm{E}_0[U(W_T)], \tag{11}$$

where $\alpha_0,\ldots,\alpha_{T-1}$ are the portfolio weights at time 0 (with T periods remaining), ..., to time $T-1$ (with one period remaining), and $U(W) = W^{1-\gamma}/(1-\gamma)$. Wealth W_t at time t is given by $W_t = R_t(\alpha_{t-1})W_{t-1}$, with

$$R_t(\alpha_{t-1}) = \alpha_{t-1}(\exp(y_t) - \exp(r_{t-1})) + \exp(r_{t-1}).$$

Since CRRA utility is homogenous in wealth, we set $W_0 = 1$ as in the static case.

Using dynamic programming, we obtain the portfolio weights at each horizon t by using the investor's (scaled) indirect utility, $Q_{t+1,T}$:

$$\alpha_t^* = \arg\max_{\alpha_t} \mathrm{E}_t[Q_{t+1,T} W_{t+1}^{1-\gamma}], \tag{12}$$

where $Q_{t+1,T} = \mathrm{E}_{t+1}[(R_T(\alpha_{T-1}^*)\ldots R_{t+2}(\alpha_{t+1}^*))^{1-\gamma}]$ and $Q_{T,T} = 1$. The FOCs of the investor's problem are, for all t,

$$\mathrm{E}_t[Q_{t+1,T} R_{t+1}^{-\gamma}(\alpha_t) x_{e,t+1}] = 0, \tag{13}$$

where $x_{e,t+1} = (\exp(y_{t+1}) - \exp(r_t))$ is the excess return at time $t+1$. This expectation can be solved using quadrature in a manner similar to that for the

static problem. For N states, we must track N values of $Q_{t+1,T}$ at each horizon. There are also N portfolio weights, one corresponding to each state, at each horizon. Hence α_t^* represents one of N portfolio weights at horizon t, depending on which state is prevailing at that time in the conditional expectation of Eq. (12).

In Eq. (12), if (y_{t+1}, r_{t+1}) is independent of (y_t, r_t) for all t, then $Q_{t+1,T}$ is independent of $W_{t+1} \equiv R_{t+1}^{1-\gamma}(\alpha_t)$, so the indirect utility in Eq. (12) becomes

$$\mathrm{E}_t[Q_{t+1,T} W_{t+1}^{1-\gamma}] = \mathrm{E}_t[Q_{t+1,T}]\mathrm{E}_t[R_{t+1}^{1-\gamma}(\alpha_t)]. \tag{14}$$

Since $\mathrm{E}_t[Q_{t+1,T}]$ does not depend on α_t, the objective function for the optimization problem at time t is equivalently $\mathrm{E}_t[R_{t+1}^{1-\gamma}(\alpha_t)]$. Thus, the problem reduces to a single-period problem and there is no horizon effect.

3.2. Dynamic DA utility

The generalization of DA utility to multiple periods is non-trivial. Therefore, we first explore a two-period example in Section 3.2.1, which highlights various considerations we must address to generalize DA to a dynamic, long horizon set-up. Section 3.2.2 presents our dynamic programming algorithm for the full-fledged multi-period case.

3.2.1. Two period example

Suppose there are three dates $t = 0, 1, 2$ and two states u, d for the excess equity return at dates $t = 1, 2$. Hence, this is the two-period binomial tree example of Section 2.2.2, but we allow for rebalancing after each period. Without loss of generality we specify the risk-free rate to be zero. The distribution of returns is independent across time. In this special setting, $R_t(\alpha_{t-1})$ is given by $1 + \alpha_{t-1}u$ in state u and $1 + \alpha_{t-1}d$ in state d. The agent chooses optimal portfolios at dates $t = 0$ and 1.

At $t = 1$ for each state u and d, the investor chooses α_1 to maximize μ_1 given by

$$K_1 \mu_1^{1-\gamma} = \mathrm{E}[R_2^{1-\gamma}(\alpha_1)\mathbf{1}_{\{R_2(\alpha_1) \leqslant \mu_1\}}] + A\mathrm{E}[R_2^{1-\gamma}(\alpha_1)\mathbf{1}_{\{R_2(\alpha_1) > \mu_1\}}], \tag{15}$$

where $K_1 = \Pr(R_2(\alpha_1) \leqslant \mu_1) + A\Pr(R_2(\alpha_1) > \mu_1)$. Since the distribution is IID, the optimal utility μ_1^* is the same across states, that is $\mu_1^*(u) = \mu_1^*(d)$.

Suppose at $t = 0$ the investor defines the DA utility function as

$$K_0 \mu_0^{1-\gamma} = \mathrm{E}_0[(R_1(\alpha_0)R_2(\alpha_1^*))^{1-\gamma}\mathbf{1}_{\{R_1(\alpha_0)R_2(\alpha_1^*) \leqslant \mu_0\}}]$$
$$+ A\mathrm{E}_0[(R_1(\alpha_0)R_2(\alpha_1^*))^{1-\gamma}\mathbf{1}_{\{R_1(\alpha_0)R_2(\alpha_1^*) > \mu_0\}}], \tag{16}$$

where $K_0 = \Pr(R_1(\alpha_0)R_2(\alpha_1^*) \leqslant \mu_0) + A\Pr(R_1(\alpha_0)R_2(\alpha_1^*) > \mu_0)$. That is, she computes the certainty equivalent of end-of-period wealth, given her current information. There are four states $\{uu, ud, du, dd\}$ with portfolio returns $\{(1 + \alpha_0 u)(1 + \alpha_1^* u), (1 + \alpha_0 u)(1 + \alpha_1^* d), (1 + \alpha_0 d)(1 + \alpha_1^* u), (1 + \alpha_0 d)(1 + \alpha_1^* d)\}$. Since we cannot a priori assume $\alpha_0^* = \alpha_1^*$, returns are not necessarily recombining (the *ud* return can be different from the *du* return) and we must track all the return states both at $t = 1$ and $t = 0$. Hence, the number of states increases exponentially with the

number of periods. Moreover, the optimization is time-dependent, so portfolio weights may depend on the horizon even when returns are IID.

This example highlights two related difficulties in extending DA utility to a dynamic case, in contrast to the computationally convenient, recursive, dynamic programming approach presented in Section 3.1 for CRRA utility. First, the number of states increases exponentially with the horizon while the computational advantage of dynamic programming relies on the dimension of the state-space being kept the same at each horizon. Second, while the reference point is endogenously determined each period, it depends on all possible future return paths. Furthermore, an interesting feature of the set-up of Eq. (16) is that even with IID returns there are horizon effects. These complexities make solving dynamic DA problems not only more challenging, but they also make extending DA portfolio choice problems to a context with DGPs that require extra state variables (to accommodate predictability, for example) next to impossible. Therefore, we develop a dynamic extension of DA which does not suffer from these problems. Most importantly, our approach is tractable enough to apply to realistic DGPs.

We present a dynamically consistent way to compute the certainty equivalent which does not increase the state-space with each horizon and which endogenously updates the reference point. The key assumption is that future uncertainty, as far as the choice of the future endogenous reference point is concerned, is captured in the certainty equivalent. This assumption is similar to the way the recursive formulation of Kreps and Porteus (1979) and Epstein and Zin (1989) captures future uncertainty. We illustrate this dynamic DA formulation with the simple two-period example. Instead of using actual future returns to compute the certainty equivalent at $t = 0$, we use the certainty equivalent at $t = 1$

$$K_0\mu_0^{1-\gamma} = \mathrm{E}_0[(R_1(\alpha_0)\mu_1^*)^{1-\gamma}\mathbf{1}_{\{R_1(\alpha_0)\mu_1^* \leqslant \mu_0\}}] + A\mathrm{E}_0[(R_1(\alpha_0)\mu_1^*)^{1-\gamma}\mathbf{1}_{\{R_1(\alpha_0)\mu_1^* > \mu_0\}}],$$

(17)

where K_0 is now defined as $K_0 = \Pr(R_1(\alpha_0) \leqslant \mu_0) + A \Pr(R_1(\alpha_0) > \mu_0)$. In this formulation, there are only two states $\{u, d\}$ and we only need to track $\{(1 + \alpha_0 u)\mu_1^*, (1 + \alpha_0 d)\mu_1^*\}$. Hence, the state-space remains at two states each period.

This investor uses the next period's indirect utility μ_1^* to form the DA utility this period, so (17) is a dynamic programming problem. Notice that the endogenous reference point also updates itself and depends on the future optimal return. Finally, this generalization of DA utility to a dynamic setting also preserves the property that the CRRA dynamic program (using the CRRA indirect utility) is a special case for $A = 1$. Like CRRA utility, the DA portfolio weights in this generalization of DA utility to a dynamic setting do not exhibit horizon effects if the return DGP is IID.

Although the non-recombining utility specification in (16) has many undesirable features, it remains a valid theoretical preference specification. We illustrate the differences between the optimal asset allocations resulting from solving the problem in (16) versus the specification in (17) in the case of the two-period binomial tree. At

date 1, both problems look the same. Assuming a positive risk premium, we know that $\alpha_1^* \geqslant 1$. Hence, the up state has a lower weight. (If the risk premium is negative, the down state would have a lower weight.) Therefore, the DA utility can be written as

$$\mu_1^{1-\gamma} = \frac{(1-p)(1+\alpha_1 d)^{1-\gamma} + Ap(1+\alpha_1 u)^{1-\gamma}}{(1-p) + Ap}. \tag{18}$$

The corresponding FOC for α_1 is

$$(1-p)(1+\alpha_1 d)^{-\gamma} d + Ap(1+\alpha_1 u)^{-\gamma} u = 0,$$

so the optimal portfolio weight α_1^* is given by

$$\alpha_1^* = \frac{\kappa - 1}{u - \kappa d},$$

where

$$\kappa = \left(-\frac{Apu}{(1-p)d} \right)^{1/\gamma}.$$

At date $t = 0$, the problem in Eq. (16) remains unchanged because returns are IID (see Eq. (13)). For the non-recombining case, the actual utility specification depends on the magnitude of μ_0 relative to the four states. The ordering of the states depends on whether α_0 is smaller or larger than α_1. In an appendix available upon request, we fully analyze this case and show that the utility lies in one of four different regions. For each region, the FOCs can be derived, and we must check whether the resulting optimal utility is indeed in the assumed region. Corner solutions are also possible, where the optimal solution lies at the border of a region—this may happen for low levels of A.

Table 2 summarizes our findings for two different calibrations of the binomial tree. On the left, we report α_0 and α_1 assuming one period is equal to six months and the total horizon is one year. On the right, we assume one period is equal to one quarter and the total horizon is six months. The α_1 weight is also the optimal solution for the specification in (15). It is clear that there are indeed horizon effects with the specification using Eq. (16), which can be quantified by looking at the difference between α_0 and α_1. Interestingly, $\alpha_0 > \alpha_1$, so longer horizons mitigate the disappointment aversion. The differences are small for high A but become larger for low A. Nevertheless, we still obtain non-participation for $A < 0.65$.

3.2.2. Dynamic DA algorithm

Building on the DA utility defined in Eq. (17), we present an algorithm for solving the dynamic asset allocation problem under DA preferences. Our problem is similar to the problem described in Eq. (11), but the utility function is now DA utility. We start the dynamic program at horizon $t = T - 1$, and solve

$$\max_{\alpha_{T-1}} \mu_{T-1}(\alpha_{T-1}), \tag{19}$$

Table 2
Optimal portfolio weights for the two-period binomial tree

A	Rebalancing frequency is one half-year		Rebalancing frequency is one-quarter	
	α_0	α_1	α_0	α_1
1.00	0.8901	0.8901	0.9136	0.9136
0.95	0.8943	0.8144	0.8496	0.7974
0.90	0.8009	0.7286	0.7763	0.6752
0.85	0.7483	0.6383	0.6386	0.5456
0.80	0.6825	0.5431	0.4881	0.4106
0.75	0.5667	0.4423	0.3228	0.2666
0.70	0.4384	0.3354	0.1402	0.1135
0.65	0.2957	0.2214	0.0000	0.0000
0.60	0.1359	0.0995	0.0000	0.0000

The table lists optimal disappointment aversion portfolio weights in equity for the two-period (three dates) recombining binomial tree as described in Section 3.2.1. The curvature coefficient is $\gamma = 2.00$ for all cases. The binomial tree is calibrated to U.S. stock return data. In the left-hand columns labelled "rebalancing frequency is one-half year," $u = 0.2132$ and $d = -0.1198$, with the base period being six months, and the horizon being one year. In the right-hand columns labelled "rebalancing frequency is one-quarter," $u = 0.1365$ and $d = -0.0908$, with the base period being one quarter, and the horizon being six months.

where μ_{T-1} is defined by

$$K_{T-1}\mu_{T-1}^{1-\gamma} \equiv \mathrm{E}_{T-1}[R_T^{1-\gamma}(\alpha_{T-1})\mathbf{1}_{\{R_T(\alpha_{T-1}) \leqslant \mu_{T-1}\}}]$$
$$+ A\mathrm{E}_{T-1}[R_T^{1-\gamma}(\alpha_{T-1})\mathbf{1}_{\{R_T(\alpha_{T-1}) > \mu_{T-1}\}}] \qquad (20)$$

with $K_{T-1} = \Pr(R_T(\alpha_{T-1}) \leqslant \mu_{T-1}) + A\Pr(R_T(\alpha_{T-1}) > \mu_{T-1})$. We solve for the optimal portfolio weight α_{T-1}^*, with the corresponding optimal utility μ_{T-1}^*, as in the one-period problem. At this horizon, the allocation problem is equivalent to the static problem, but we solve for each quadrature state, which yields N optimal state-dependent portfolio weights and utilities.

At horizon $t = T - 2$ we solve

$$\max_{\alpha_{T-2}} \mu_{T-2}(\alpha_{T-2}), \qquad (21)$$

where μ_{T-2} is defined by

$$K_{T-2}\mu_{T-2}^{1-\gamma} \equiv \mathrm{E}_{T-2}[R_{T-1}^{1-\gamma}(\alpha_{T-2})(\mu_{T-1}^*)^{1-\gamma}\mathbf{1}_{\{\{R_{T-1}(\alpha_{T-2})\mu_{T-1}^* \leqslant \mu_{T-2}\}\}}]$$
$$+ A\mathrm{E}_{T-2}[R_{T-1}^{1-\gamma}(\alpha_{T-2})(\mu_{T-1}^*)^{1-\gamma}\mathbf{1}_{\{R_{T-1}(\alpha_{T-2})\mu_{T-1}^* > \mu_{T-2}\}}] \qquad (22)$$

with $K_{T-2} = \Pr(R_{T-1}(\alpha_{T-2}) \leqslant \mu_{T-2}) + A\Pr(R_{T-1}(\alpha_{T-2}) > \mu_{T-2})$. To solve for α_{T-2}^* and μ_{T-2}^* at a particular state at $t = T - 2$, we need only track the N states for μ_{T-1}^* at $T - 1$. We continue this process for $t = T - 3$ until $t = 0$.

If $A = 1$, then at horizon $t = T - 2$ the DA utility reduces to

$$\mu_{T-2}^{1-\gamma} = \mathrm{E}_{T-2}[R_{T-1}^{1-\gamma}(\alpha_{T-2})(\mu_{T-1}^*)^{1-\gamma}] = \mathrm{E}_{T-2}[R_{T-1}^{1-\gamma}(\alpha_{T-2})Q_{T-1,T}], \qquad (23)$$

which is the standard CRRA problem. Note that if returns are IID, then at each horizon, exactly the same DA problem applies and the portfolio weights are independent of the horizon. More generally, to solve the DA problem at each horizon, we simultaneously use the FOC and the definition of the certainty equivalent, which also occurs in the static case (see Eqs. (6) and (9)).

4. Disappointment aversion versus loss aversion

A large literature documents how the risk attitudes of individuals differ from the predictions of expected utility theory. The behavioral work of Kahneman and Tversky (1979) has been very influential in this area. In Section 4.1, we define LA utility following Kahneman and Tversky (1979). Both the LA and DA preferences capture similar features of human behavior, and we comment on how LA and DA preferences imply risk aversion with respect to both small and large stakes, a feature not shared by CRRA utility.

LA preferences have been far more popular than DA preferences in applied finance work. This is surprising for a number of reasons. First, the grounding of DA preferences in decision theory makes them more attractive to economists using rational dynamic programming tools. Second, whereas we are able to formulate a mathematically well-defined static and dynamic asset allocation framework under DA preferences, this is considerably harder under LA preferences. Sections 4.2–4.4 briefly illustrate some of the problems encountered when applying the original Kahneman–Tversky formulation to an asset allocation framework. These problems include the real possibility of infinite optimal asset allocations and the sensitivity of the asset allocation to the choice of the reference point. The shortcomings of LA have led researchers employing this behavioral utility function to modify the original Kahneman–Tversky specification and we discuss various implementations in Section 4.5. The attraction of the DA framework is precisely that it accommodates loss aversion without other behavioral implications.

4.1. Loss aversion and Rabin (2000) gambles

4.1.1. Kahneman and Tversky (1979) loss aversion
With χ representing a gain or loss relative to a reference point B_0, the LA utility of Kahneman and Tversky (1979) is given by

$$U(\chi) = -\lambda E[(-\chi)^{(1-\gamma_1)}\mathbf{1}_{\{\chi \leqslant 0\}}] + E[\chi^{(1-\gamma_2)}\mathbf{1}_{\{\chi > 0\}}], \tag{24}$$

where $\mathbf{1}$ is an indicator variable, $\chi = W - B_0 = R_f + \alpha x_e - B_0$ is the gain or loss of final wealth W relative to the benchmark B_0, $R_f = \exp(r)$ is the gross risk-free rate and $x_e = \exp(y) - \exp(r)$ is the excess stock return where y is the equity return. Kahneman and Tversky (1979) argue that the expectation in Eq. (24) should be taken under a subjective measure, but for now we assume that the objective (real) measure holds.

The parameter λ governs the additional weight on losses. According to Kahneman and Tversky, $\lambda = 2.25$, so losses are weighted 2.25 times as much as gains, and $\gamma_1 = \gamma_2 = 0.12$, which implies the same amount of curvature across gains and losses. Following the behavioral literature, we consider only the case of $0 \leqslant \gamma_1 < 1$ and $0 \leqslant \gamma_2 < 1$ since the felicity function $(-\lambda(-\chi)^{1-\gamma_1}\mathbf{1}_{\{\chi \leqslant 0\}} + \chi^{(1-\gamma_2)}\mathbf{1}_{\{\chi > 0\}})$ is monotone in wealth only if $0 \leqslant \gamma_1 < 1$ and $0 \leqslant \gamma_2 < 1$.[3] Hence, both LA and DA preferences incorporate an asymmetric treatment of good and bad outcomes that is not present in standard expected utility. LA utility is different from DA utility because the LA felicity function is not globally concave in wealth. When expressed in wealth levels, the LA utility function is S-shaped, which implies risk-seeking behavior in the loss region and risk aversion in the gain region. The non-concavity has important consequences for optimal portfolio choice under LA utility.

4.1.2. Rabin (2000) gambles

Rabin (2000) demonstrates a striking problem arising in the expected utility framework. His "calibration theorem" is best illustrated with an example. Suppose that for some ranges of wealth (or for all wealth levels), a person turns down gambles where she loses $100 or gains $110, each with equal probability. Then she will turn down 50%–50% bets of losing $1000 or gaining ANY sum of money. We call such a gamble a "Rabin gamble." Since DA preferences do not fall into the expected utility category, they do not necessarily suffer from the Rabin-gamble problem.

Fig. 2 illustrates this. Imagine an investor with $10,000 wealth. If he has CRRA preferences, a risk aversion level of $\gamma = 10$ makes him reject the initial $-100/+110$ gamble. The graphs in the left-most column show both his utility and willingness-to-pay relative to the Rabin gamble of losing $1000 and gaining the amount on the x-axis. The willingness-to-pay to avoid the gamble is the difference between the certain wealth the investor has available by not participating in the gamble minus the certainty equivalent of the gamble. If the willingness-to-pay is negative, rational agents would accept the gamble. The last amount on the right-hand side of the x-axis represents $25,000. It is apparent from the top graph that the marginal utility of additional wealth becomes virtually zero very fast. The willingness-to-pay to avoid the gamble asymptotes to about $280, even if the potential gain is over $1,000,000. The extreme curvature in the utility function drives the continued rejection of the second gamble even as the possible amount of money to be gained increases to infinity.

With DA preferences, an investor need not display an extremely concave utility function to dislike the original $-100/+110$ gamble, because he hates to lose $100. The middle column of Fig. 2 shows utility levels and willingness-to-pay for DA utility. An investor with $\gamma = 2$ and $A = 0.85$ rejects the original gamble, but this investor loves the Rabin gambles. In fact, the willingness-to-pay decreases rapidly and quickly becomes negative. As an example, our DA investor would be willing to pay $3664 to enter a bet where she can gain $25,000 but may lose $1000 with equal

---

[3]Note that the utility function implies different preference orderings if χ is expressed in different units unless $\gamma_1 = \gamma_2$ or the difference between γ_1 and γ_2 is very small. Expressing χ in returns (so χ has no dimension) circumvents this problem. Note that LA utility is not defined at $W = B_0$ for $\gamma > 1$.
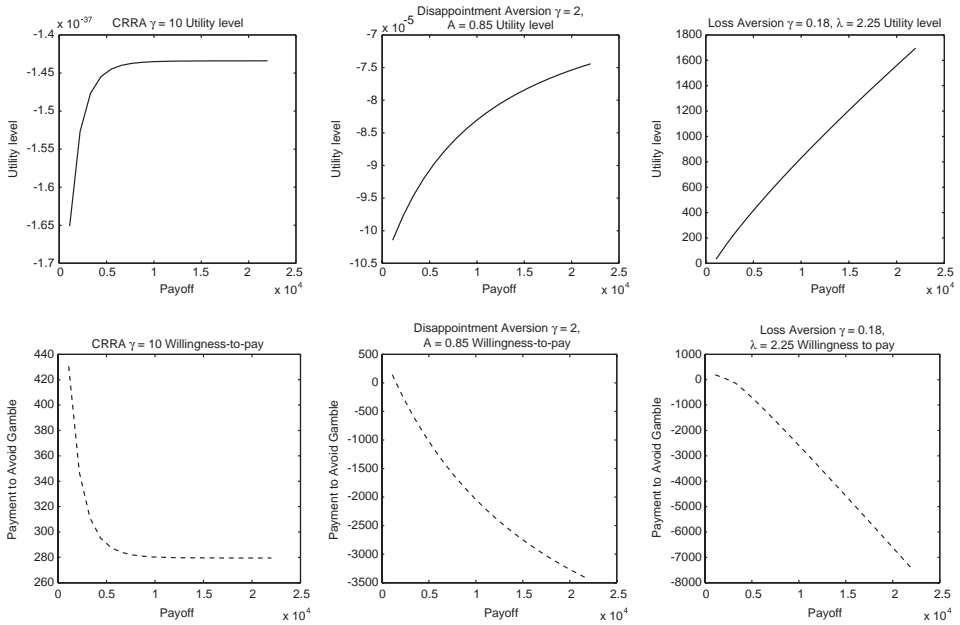
Fig. 2. Rabin gambles. Utility levels (top row) and willingness to pay (bottom row) for a gamble with the following characteristics: an initial wealth of $10,000, a 50% probability of a loss of $1000 and a 50% probability of a gain represented on the *x*-axis. The left column considers CRRA utility with $\gamma = 10$, the middle column Disappointment Aversion utility with curvature parameter $\gamma = 2$ and disappointment level $A = 0.85$, and the right column loss aversion utility with curvature parameter $\gamma = 0.18$ and loss parameter $\lambda = 2.25$.

probability. For lower γ, or higher A, this amount increases. For example, if $\gamma = 0$ and $A = 0.85$, the DA investor would be willing to pay $10,946 to take on this $-1000/+25,000$ gamble.

For LA preferences, we must first introduce a notion of willingness-to-pay because the LA utility in Eq. (24) is defined over gains and losses. However, since gains and losses are always evaluated relative to a benchmark, wealth is implicitly given as the gain or loss plus the reference point. Denoting the LA utility in Eq. (24) as U_{LA}, we define the certainty equivalent of LA, μ_W^{LA}, as

$$\mu_W^{\mathrm{LA}} = \begin{cases} U_{\mathrm{LA}}^{1/(1-\gamma_2)} + B_0 & \text{if } U_{\mathrm{LA}} > 0, \\ -\left(-\dfrac{U_{\mathrm{LA}}}{\lambda}\right)^{1/(1-\gamma_1)} + B_0 & \text{if } U_{\mathrm{LA}} \leqslant 0, \end{cases} \tag{25}$$

where B_0 is the benchmark of the gamble, which is also our chosen initial wealth.

The last column of Fig. 2 shows utility levels and willingness-to-pay for a LA investor with benchmark parameters $\gamma = 0.18$ and $\lambda = 2.25$ from Kahneman and Tversky (1979). This investor also rejects the initial $-100/+110$ gamble but likes the Rabin gambles. For example, the LA investor would be willing to pay $8671 to enter

a bet to gain \$25,000 with probability one half, and lose \$1000 with probability one half. Hence, both the DA and LA preference functions can resolve the Rabin puzzle. From introspection, over-weighting losses relative to gains seems to yield much more reasonable attitudes towards risk.

4.2. Characterizing optimal LA portfolio weights

When the portfolio weight in equities α is very large in absolute magnitude (so that $\chi \to \alpha x_e$ as $\alpha \to \pm\infty$), the utility function approaches

$$-\lambda \alpha^{1-\gamma_1} E[(-x_e)^{1-\gamma_1} \mathbf{1}_{\{x_e \leqslant 0\}}] + \alpha^{1-\gamma_2} E[x_e^{1-\gamma_2} \mathbf{1}_{\{x_e > 0\}}], \quad \text{for } \alpha \to +\infty$$

and

$$|\alpha|^{1-\gamma_1} E[(-x_e)^{1-\gamma_1} \mathbf{1}_{\{x_e \leqslant 0\}}] - \lambda |\alpha|^{1-\gamma_2} E[x_e^{1-\gamma_2} \mathbf{1}_{\{x_e > 0\}}], \quad \text{for } \alpha \to -\infty, \quad (26)$$

where x_e is the excess return on equity. Hence, the term with the higher exponent on α dominates. In particular, for $\gamma_1 > \gamma_2$ the second term dominates so there is no finite optimal portfolio weight. The behavioral literature has mostly considered only the case of $\gamma_1 = \gamma_2 = \gamma$ following Kahneman and Tversky (1979) (see Benartzi and Thaler, 1995; Berkelaar and Kouwenberg, 2000; Barberis et al., 2001). Even in this restricted case, extreme LA portfolio weights are likely. Sharpe (1998) analyzes a closely related bilinear utility function, which can be represented as $\lambda x_e \mathbf{1}_{\{x_e \leqslant 0\}} + x_e \mathbf{1}_{\{x_e > 0\}}$. Sharpe shows that this bilinear utility function implies extreme portfolio weights under empirically relevant circumstances (see also Aït-Sahalia and Brandt, 2001). Appendix C outlines general conditions under which finite portfolio solutions with LA preferences are possible. Interestingly, we find that LA may produce local maxima (see Benartzi and Thaler, 1995), even though the global maximum is either $-\infty$ or $+\infty$. Such local maxima exist for realistic parameter values.

4.3. Choice of the LA reference point

Kahneman and Tversky (1979)'s prospect theory gives no guidance with regard to the choice of the reference point B_0, which must be set exogenously. As noted by Benartzi and Thaler (1995), different horizons can turn the LA optimization into a totally different problem. For example, for very long rebalancing periods, the effect of the benchmark is negligible if the benchmark is current wealth or current wealth times the risk-free rate. This is because the benchmark in accumulated wealth $W = R_f + \alpha x_e - B_0$ is swamped by the equity returns over long horizons. Moreover, in a dynamic setting there is no clear guidance about how the loss aversion reference point should be updated. If the choice of reference point is current wealth times the risk-free rate, as specified by Barberis et al. (2001), then the LA optimal portfolio weight, if finite, is zero:

Proposition 4.1. *If the benchmark B_0 is equal to current wealth times the risk-free rate then the optimal portfolio weight $\alpha^* = 0$ or it is unbounded.*

Proof. See Appendix C.

Hence, for this particular benchmark, the only possible portfolio weights are $-\infty$, 0, or $+\infty$. In contrast, in DA utility, the reference point defining elating outcomes ("gains"), versus disappointing outcomes ("losses") is endogenous and we show that DA portfolio weights are finite.

4.4. Subjective probability transformations

When the risk premium is zero, CRRA or DA investors hold zero equity. This is not always the case for LA investors. Kahneman and Tversky (1979) propose to use a subjective rather than an objective probability distribution to take the expectations in Eq. (24). Kahneman and Tversky call the transformed objective probabilities "decision weights." The transformation involves over-weighting small probability events and under-weighting large probability events. Whereas the literature that empirically applies loss aversion (Benartzi and Thaler, 1995; Berkelaar and Kouwenberg, 2000; Barberis et al., 2001; Gomes, 2003) has not used these probability transformations, it is useful to point out one of their undesirable properties. In particular,

Proposition 4.2. *If the risk premium is zero, then the optimal portfolio weight in LA could be less than zero ($\alpha^* < 0$).*

Proof. See Appendix C.

Short positions may result with LA even with zero risk premiums since the LA probability transformation downweights probabilities, rather than events. Similar to LA, DA also involves an implied probability transformation. However, Appendix A shows that DA uses a CRRA maximization problem with transformed probabilities such that the probabilities for wealth above the certainty equivalent are down-weighted. The major difference between the probability transformation of DA and LA is that DA's probability transformation is endogenous, while LA's is arbitrary. The subjective probability transformation of LA also violates first-order stochastic dominance and transitivity. In contrast, these properties are maintained under DA preferences (see Machina, 1982; Gul, 1991).

4.5. Loss aversion in the literature

The possible non-finite optimal portfolio weight under LA preferences is due to the global non-concavity of the LA utility function. Specifically, LA utility, as defined by Kahneman and Tversky (1979) and used by Benartzi and Thaler (1995), is finite for negative wealth. Proposition 4.2 shows that LA investors may be more risk-seeking than CRRA investors. An important consequence is that there is no guarantee that non-corner solutions can be found. Various approaches have been taken in the literature to practically implement LA. Most of these approaches rely on imposing additional restrictions on the original specification so that the utility function is sufficiently negative, or negative infinity, at zero wealth. This allows the

original LA utility to be "pseudo-concavified" (Berkelaar and Kouwenberg, 2000), but changes the fundamental nature of the original specification.[4]

For example, to avoid corner solutions, Gomes (2003) adds a term to the LA utility function, which for sufficiently large losses makes the utility function again concave:

$$U(W) = V_{\text{BL}} - \lambda \text{E}[((-\chi))^{(1-\gamma_1)} \mathbf{1}_{\{\underline{W} < W \leqslant B_0\}}] + \text{E}[(\chi)^{(1-\gamma_2)} \mathbf{1}_{\{W > B_0\}}],$$

where V_{BL} is defined by Gomes as

$$V_{\text{BL}} = \text{E}[W^{1-\gamma} \mathbf{1}_{\{W \leqslant \underline{W}\}}] - c$$

and c is a constant set to make the utility function continuous at $W = \underline{W}$. By using another CRRA term, negative wealth can be assigned negative infinite utility, forcing wealth to be positive. Another approach is taken by Berkelaar and Kouwenberg (2000), who do not modify Kahneman and Tversky (1979)'s original specification, but instead explicitly restrict wealth to be positive.

Barberis et al. (2001)'s utility function has two components. The first component is a log-utility function defined over consumption. The second component is defined over wealth and embeds loss aversion. The loss aversion component is piece-wise linear ($\gamma_1 = 0$ and $\gamma_2 = 0$), following Benartzi and Thaler (1995). In an asset allocation framework, their utility function can be written as

$$U(W) = \text{E}[\log(W)] - \lambda \text{E}[(-\chi) \mathbf{1}_{\{W \leqslant B_0\}}] + \text{E}[\chi \mathbf{1}_{\{W > B_0\}}]. \tag{27}$$

The log utility function endogenously enforces a positive wealth constraint, since wealth at zero yields negative infinite utility. Barberis et al. choose the reference point as current wealth times the risk-free rate. In the original Kahneman and Tversky (1979) formulation without the Barberis–Huang–Santos log-utility term, we know from Proposition 4.1 that the only finite optimal equity portfolio weight for this choice of reference point is zero. These practical implementation problems inherent in LA preferences make DA preferences a very viable alternative to model loss aversion.

5. Disappointment aversion and stock holdings

5.1. Data and data generating processes

To examine portfolio choice under realistic DGPs, we use quarterly U.S. data from 1926 to 1998 on nominal stock returns and Treasury bill interest rates. We use two main DGPs in this paper that largely conform to the DGP's prevalent in the extensive literature on dynamic asset allocation (e.g. Kandel and Stambaugh, 1996;

---

[4]Benartzi and Thaler (1995), although working with the original Kahneman and Tversky (1979) LA specification, do not perform an optimization problem. They evaluate LA utility as a function of the equity portfolio holding, but only between zero and one. Calibrating a binomial tree to the U.S. stock market, we find that this is likely a local optimum, with utility being maximized at $\alpha = +\infty$.

Balduzzi and Lynch, 1999; Campbell and Viceira, 1999; Barberis, 2000). In our first model, stock returns are IID over time and the interest rate follows a first-order autoregressive system. In our second model, we accommodate predictability. Following most of the dynamic asset allocation literature, we consider only one possible predictor of stock returns and consider a system in which an instrument linearly predicts stock returns in the conditional mean of equity returns. Whereas many authors have focused on yield variables, we use the interest rate itself. This has the advantage of reducing the state space and introduces an interesting dynamic since the predictor itself is the return on an investable asset. We are also unlikely to lose much predictive power, since Ang and Bekaert (2003) find that the short rate is the most robust predictor of international stock returns. Ang and Bekaert (2003) and Goyal and Welch (2003) demonstrate that the dividend yield, which has been previously used by many authors to forecast returns, has no forecasting power when data of the late 1990s are added to the sample.

Our two DGPs for nominal data are special cases of a bivariate vector autoregression (VAR) on stock returns and interest rates

$$X_t = c + \Phi X_{t-1} + \Sigma^{1/2}\varepsilon_t, \tag{28}$$

where $X_t = (\tilde{y}_t\, r_t)'$, $\tilde{y}_t = y_t - r_{t-1}$ is the continuously compounded excess equity return and r_t is the risk-free rate, measured by the quarterly T-bill interest rate, and $\varepsilon_t \sim N(0, I)$.

The "No Predictability" model imposes the condition that all elements of Φ equal zero except Φ_{22}, and the "Predictability" model constrains all elements of Φ except Φ_{12} and Φ_{22} to be zero. Estimates for these DGPs are reported in Table 3. In both systems, there is negative contemporaneous correlation between shocks to short rates and stock returns. The predictability system reveals that the short rate is not a significant predictor of stock returns over the full sample period. In fact, predictability is much stronger in the post-1940 period. Although we do not report results for this alternative sample explicitly, we investigate a DGP estimated on post-1940 data. With this DGP, equity is relatively more attractive, but our main results are unchanged.

We now derive optimal asset allocations for various parameter configurations under the two DGPs. Since the DGPs are first-order Markov processes, they lend themselves easily to discretization, which we detail in Appendix D.

5.2. No predictability case

In this system, the excess premium is constant and IID, while short rates are autoregressive and negatively correlated with equity returns. For a given risk aversion, portfolio allocations in this system depend on the horizon, but they do not depend on the level of the short rate (as we show later). This is not surprising given that our set-up is similar to that of Liu (1999). Liu proves this result analytically in a continuous-time problem with the short rate following a Vasicek (1977) model. Under the Vasicek term structure model, excess returns of bonds have a constant risk premium, have constant volatilities, and are perfectly correlated with the short rate.

Table 3
Parameter estimates for the data generating processes

	No predictability system	Predictability system
c_1	0.0161	0.0223
	(0.0064)	(0.0099)
c_2	0.0008	0.0008
	(0.0003)	(0.0003)
Φ_{12}	—	−0.6049
		(0.7416)
Φ_{22}	0.9273	0.9273
	(0.0219)	(0.0219)
σ_1	0.1095	0.1094
	(0.0015)	(0.0015)
σ_2	0.0032	0.0032
	(0.0000)	(0.0000)
ρ	−0.0474	−0.0475
	(0.0585)	(0.0585)

The model is

$$X_t = \mu + \Phi X_{t-1} + \Sigma^{1/2}\varepsilon_t,$$

with $X_t = (\tilde{y}_t\, r_t)'$, $\tilde{y}_t$ the excess one-period stock return, and r_t the short rate. Stock returns are S&P500 returns and the short rate is the three-month T-bill interest rate. All returns are continuously compounded. All elements of Φ are constrained to be zero except for Φ_{22} in the no predictability system. In the predictability system, Φ_{11} and Φ_{21} are constrained to be zero. The correlation between the errors of $\tilde{y}_t$ and r_t is denoted ρ. The data sample is quarterly from 1926 to 1998.

Similarly, in our no predictability system, excess stock returns have a constant risk premium and a constant volatility. Although in our setting the correlation between equities and the short rate is not unity, Liu's result obtains. Given there is no short rate dependence, we only discuss general patterns in optimal equity portfolio weights.

Because there is little guidance on the choice of parameter values, we characterize portfolio choice for DA preferences (which include CRRA as a special case) across a wide set of parameter values. Fig. 3 establishes benchmark asset allocations for CRRA preferences restricting the curvature parameter γ to the interval [2, 10], a range suggested by decades of empirical research, with $\gamma = 2$ as the most popular choice (see Friend and Blume, 1975). Moderately risk-averse CRRA agents ($\gamma = 2$) should put close to 100% of their portfolio in equities. Equity allocations of 50% to 60% start to appear at γ's between 3 and 4, but CRRA utility never produces a non-participation result. Fig. 3 also shows that the equity proportion is slightly larger for longer horizons and hence agents gradually decrease their equity proportions as they age.

Table 4 reports the asset allocation results for $\gamma = 2$ and 5 (the middle of the [2, 10] range) and two horizons (three months and ten years), and also reports standard errors for the weights, computed using the delta method (see Ang and Bekaert, 2002). We check the accuracy of these standard errors using a small-scale Monte
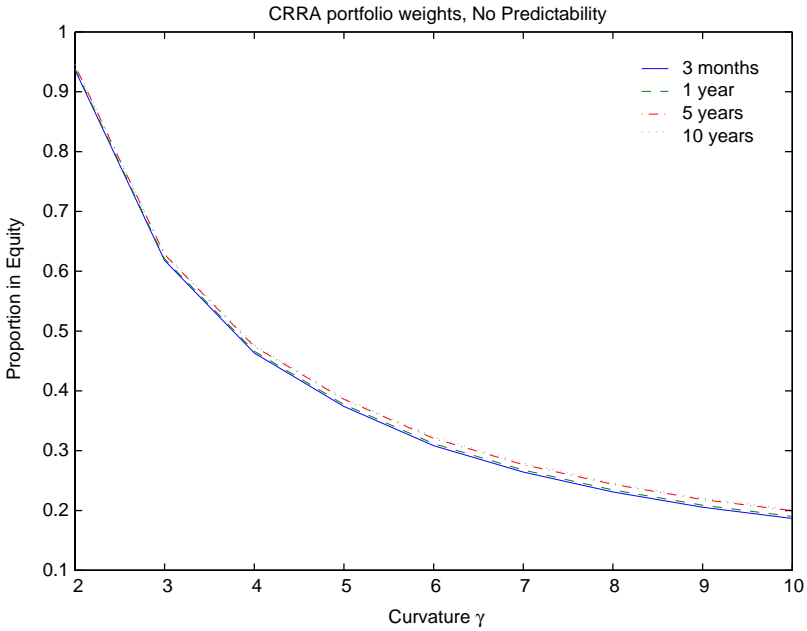
Fig. 3. Equity portfolio weights for CRRA utility in the no predictability system. Optimal equity portfolio weights for CRRA utility for the no predictability system. The curvature parameter γ is on the horizontal axis. We show the portfolio weight for a horizon of three months, one year, five years and ten years (dynamically rebalancing each quarter) on the vertical axis. Portfolio weights are the same for all short rate states.

Carlo (with 400 parameter draws); this is extremely time-consuming because of the long computation time required to solve the DA problem for every parameter draw. The Monte Carlo standard errors are very similar to the standard errors computed using the delta-method. For low levels of A, the delta-method tends to overstate the standard errors because it fails to fully account for the non-linearity induced by non-participation.

Because the equity return mean is measured with large sampling error, these standard errors are quite large. Nevertheless, the weights at $\gamma = 2$ (0.927) and $\gamma = 5$ (0.370) are both significantly different from zero at the 1% level and, using a Wald test, significantly different from each other (p-value = 0.0006). However, for both $\gamma = 2$ and 5, the equity weight is no longer statistically significant from zero once A reaches 0.80. The Wald test for the hypothesis that the one-quarter horizon portfolio weights corresponding to $A = 1.00$ and 0.65 for $\gamma = 2$ are the same rejects the null with p-value less than 0.0001; similarly, a test that the one-quarter horizon portfolio weights for $A = 1.00$ and 0.65 are equal for $\gamma = 5$ also rejects with p-value less than 0.0001.

Table 4 shows one of our main results. For the no predictability system, the critical A^* required to induce investors to participate in the market is $A^* = 0.6030$. That is,

Table 4
Portfolio weights for the no predictability system

A	Curvature parameter $\gamma = 2$			Curvature parameter $\gamma = 5$		
	1 qtr	10 yr	χ^2 test p-value	1 qtr	10 yr	χ^2 test p-value
1.00	0.9270	0.9359	0.9870	0.3703	0.3835	0.9524
	(0.2728)	(0.2730)		(0.1098)	(0.1110)	
0.95	0.8331	0.8455	0.9815	0.3324	0.3473	0.9459
	(0.2729)	(0.2638)		(0.1094)	(0.1106)	
0.90	0.7337	0.7499	0.9764	0.2925	0.3091	0.9395
	(0.2728)	(0.2748)		(0.1090)	(0.1113)	
0.85	0.6283	0.6485	0.9716	0.2503	0.2688	0.9319
	(0.2743)	(0.2921)		(0.1085)	(0.1089)	
0.80	0.5165	0.5408	0.9645	0.2057	0.2262	0.9229
	(0.2717)	(0.2741)		(0.1080)	(0.1044)	
0.75	0.3976	0.4324	0.9537	0.1584	0.1809	0.9126
	(0.2698)	(0.3281)		(0.1075)	(0.0985)	
0.70	0.2710	0.3116	0.9399	0.1080	0.1327	0.9142
	(0.2689)	(0.2701)		(0.1069)	(0.1224)	
0.65	0.1357	0.1824	0.9240	0.0542	0.0844	0.8906
	(0.2667)	(0.2233)		(0.1062)	(0.1138)	

Critical A^* to induce participation

	1 qtr	10 yr
A^*	0.6030	0.6001

Optimal portfolio weights for disappointment aversion utility for various horizons for curvature parameter $\gamma = 2$ and 5 for the system without predictability. Portfolios are rebalanced quarterly. The critical level of A required to participate in the equity market is given by A^*. Standard errors are given in parentheses. The χ^2 test reports a p-value that the one-quarter horizon portfolio weight is the same as the ten-year horizon portfolio weight.

scaling up the utility of disappointing outcomes by 1.66 (1/0.6030) produces non-participation. Note that all investors hold zero equity at A^* irrespective of γ (see Proposition 2.1). Hence, variation in A (from 1 to 0.6030) for a CRRA investor with the "normal" curvature in the utility function of $\gamma = 2$ leads to variation in equity holdings from close to 100% to 0%. For $\gamma = 2$, dropping A to 0.85 is sufficient to bring the equity allocation close to 60%. The effect on asset allocation of lower A is less dramatic for higher γ, which is apparent from the column with $\gamma = 5$.

Turning now to horizon effects, Table 4 shows that the portfolio weights for one-quarter and ten-year horizons are very similar. The columns labeled χ^2 p-value in Table 4 report the p-value of a Wald test that the portfolio weights for the one-quarter and ten-year horizons are the same. This tests overwhelmingly fails to reject for all γ and A. Looking at the point estimates of the portfolio weights across horizons, the horizon effect becomes larger when A is decreased, reaching a 4.67% difference for $A = 0.65$ and $\gamma = 2$.

While the statistical significance and economic magnitude of the horizon effects are small, their existence in a system without predictability warrants some further

elaboration. In our estimated system, portfolio weights do not depend on the short rate but there is a weak positive horizon effect: agents with longer horizons hold more equity. From Samuelson (1991) and others, processes with positive persistence induce negative horizon effects (they are "riskier" over longer periods), whereas negatively correlated processes induce positive horizon effects. In our empirical estimates, shocks to stock returns and short rates are slightly negatively correlated (-0.0474) (see Table 3), which induces weak positive hedging demands.

The size of hedging demands is primarily determined by the rebalancing horizon, the predictor variable used to forecast equity returns and the correlation between predictor innovations and returns. Brandt (1999) and Ang and Bekaert (2002) find that frequent rebalancing reduces the size of hedging demands and Aït-Sahalia and Brandt (2001) and Lynch (2001) also find that the magnitude of hedging demands depends very much on the choice of predictor variable.

We also consider a system with heteroskedasticity, changing the interest rate process in Eq. (28) to a simple square root model $r_{t+1} = c_r + \rho r_t + \sigma_r \sqrt{r_t} \varepsilon_{t+1,r}$ with the equity return given by $y_{t+1} = c_y + \sigma_{y,r} \sqrt{r_t} \varepsilon_{t+1,r} + \sigma_y \varepsilon_{t+1,y}$ to match the same unconditional moments implied by the VAR (28). Because σ_r and the conditional correlation between interest rates and excess equity returns are small, the results are rather uninteresting. The portfolio weights are invariably slightly smaller than what we obtain for the homoskedastic case, with the differences becoming slightly larger with the horizon. However, the differences are very small, never exceeding 0.008, and we do not report the results to conserve space. We also incorporate heteroskedasticity into the predictability system in Section 5.4 with a square root process for interest rates, and find it totally dominated by the conditional mean effects.

5.3. The impact of the rebalancing frequency

With LA preferences, the rebalancing frequency is very important. For example, if the benchmark is current wealth, the longer the rebalancing frequency, the more irrelevant the benchmark becomes. This observation is critical to the argument of Benartzi and Thaler (1995), who claim that it is myopic loss aversion which accounts for the puzzling lack of equity holdings among investors. Of course, in our framework, the rebalancing frequency is likely to be less important, since the reference point is endogenous and changes with the rebalancing frequency in an internally consistent fashion.

Table 5 confirms this conjecture. The first line of the table simply expands on the results of Table 4, showing that longer horizons induce slightly higher equity allocations. The table then displays results for three other rebalancing frequencies: two quarters, one year, and two years. These portfolio weights are computed by temporally aggregating the one-quarter VAR (see Appendix D) and then discretizing the resulting dynamic system as a first-order VAR over the myopic frequency. Formally, temporal aggregation leads to a VARMA(1,1) system in the new frequency, but taking the MA component into account in the optimal asset allocation is infeasible. For a myopic horizon, the VAR approximation should be very accurate.

Table 5
Portfolio weights with different rebalancing frequencies for the no predictability system

Rebalancing frequency	Horizon			
	1 qtr	2 qtrs	1 yr	2 yrs
Portfolio weights for different rebalancing frequencies				
1 qtr	0.6283	0.6290	0.6332	0.6403
	(0.2743)	(0.2724)	(0.2743)	(0.2991)
2 qtrs		0.6348		
		(0.2717)		
1 year			0.6683	
			(0.2714)	
2 years				0.6462
				(0.2741)
Critical A^* for different rebalancing frequencies				
1 qtr	0.6030	0.6028	0.6024	0.6017
2 qtrs		0.5965		
1 year			0.5656	
2 years				0.5927

The table lists optimal portfolio weights for various horizons for disappointment aversion utility with curvature parameter $\gamma = 2$ and disappointment level $A = 0.85$ for the system without predictability. For the portfolio weights in the row labelled ''1 qtr,'' we report optimal portfolio weights for the quarterly VAR rebalanced each quarter, for different horizons from one quarter to two years. For the rows labelled ''2 qtrs'' to ''2 years,'' the DGP is a time-aggregated VAR. For example, for the two-quarter rebalancing frequency, we use a VAR time-aggregated over two quarters. These portfolio weights are myopic. The critical level of A required to participate in the equity market is given by A^*. Standard errors are given in parentheses.

Table 5 shows that the effect of changing the rebalancing frequency is very small. Because there is no predictability, the main effect comes from changes in the annualized volatility of interest rates with different rebalancing frequencies, which first decreases from $0.0173 \times \sqrt{4} = 0.0346$ at the one-quarter horizon to 0.0286 at the one-year horizon (making equities relatively more attractive) and then increases back to $0.0560 \div 2 = 0.0396$ for the two-year frequency. The mechanism here is very different from the drastic change in the benchmark level that drives the results in Benartzi and Thaler (1995). Our results are driven by changes in the DGP at the different frequencies.

5.4. Predictability case

Table 6 reports myopic portfolio weights corresponding to three annualized interest rates levels 0.0392, 0.0816, and 0.1208 for $\gamma = 2$, 5 and various disappointment levels A, for the system with predictability of excess returns. The special case of CRRA utility is given by $A = 1$. The interest rates represent a state close to the unconditional mean ($r = 0.0392$), an extremely high interest rate ($r = 0.1208$), and one in the middle of the range. As interest rates increase, the equity holding decreases. The effect is quite pronounced. For example, a $\gamma = 2$ investor

Table 6
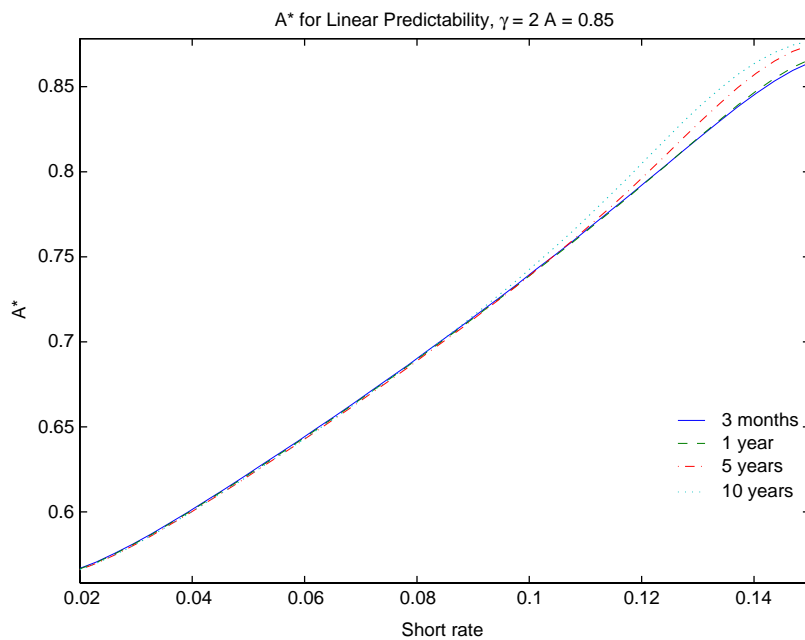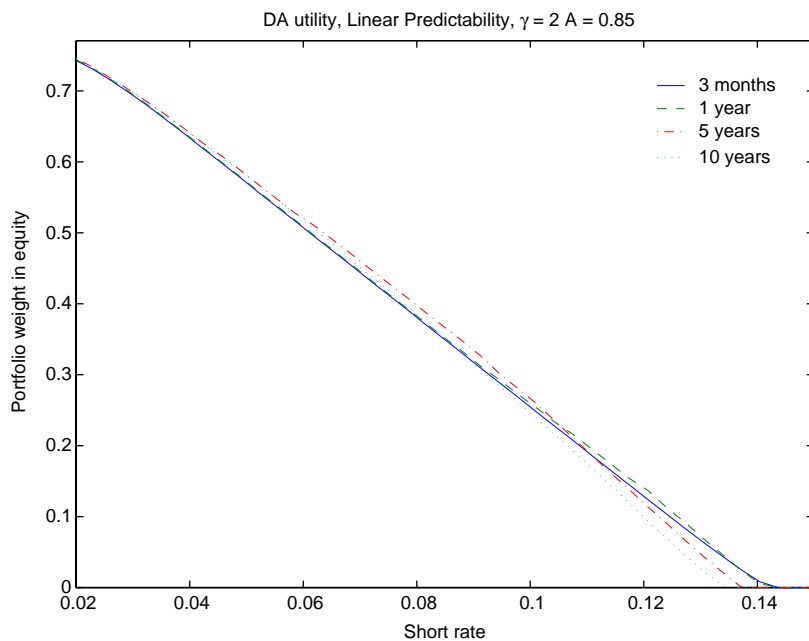Myopic portfolio weights for the predictability system

A	$r = 0.0392$		$r = 0.0816$		$r = 0.1208$	
	$\gamma = 2$	$\gamma = 5$	$\gamma = 2$	$\gamma = 5$	$\gamma = 2$	$\gamma = 5$
1.00	0.9379	0.3747	0.6704	0.2675	0.4224	0.1685
	(0.2735)	(0.1101)	(0.4172)	(0.1668)	(0.6769)	(0.2701)
0.95	0.8438	0.3367	0.5758	0.2296	0.3278	0.1308
	(0.2736)	(0.1097)	(0.4171)	(0.1664)	(0.6768)	(0.2695)
0.90	0.7443	0.2967	0.4760	0.1897	0.2283	0.0911
	(0.2735)	(0.1093)	(0.4170)	(0.1660)	(0.6768)	(0.2690)
0.85	0.6389	0.2545	0.3705	0.1477	0.1235	0.0493
	(0.2737)	(0.1088)	(0.4187)	(0.1656)	(0.6665)	(0.2682)
0.80	0.5269	0.2098	0.2590	0.1033	0.0130	0.0052
	(0.2724)	(0.1083)	(0.4118)	(0.1652)	(0.6687)	(0.2677)
0.75	0.4079	0.1625	0.1408	0.0562	0.0000	0.0000
	(0.2715)	(0.1078)	(0.4127)	(0.1645)		
0.70	0.2810	0.1120	0.0153	0.0061	0.0000	0.0000
	(0.2697)	(0.1072)	(0.4097)	(0.1634)		
0.65	0.1455	0.0581	0.0000	0.0000	0.0000	0.0000
	(0.2675)	(0.1065)				
	$A^* = 0.5998$		$A^* = 0.6941$		$A^* = 0.7943$	

The table lists myopic (three-month horizon) portfolio weights for the system with short rate predictability. We list weights corresponding to three (annualized) interest rate states 0.0392, 0.0816, and 0.1208 for $\gamma = 2$ and 5 and various disappointment levels A. The critical disappointment level required to participate in the equity market is given as A^*.

holds 64% in the market portfolio for $A = 0.85$ at $r = 0.0392$. When r increases to 0.0816 this investor's portfolio weight decreases to 37%. This is not surprising. In the system with predictability, higher interest rates lower the conditional equity premium (Table 3 shows that a 1% increase in the short rate decreases the equity premium by 60 basis points). Of course, the standard errors on the weights remain large.

The critical A^* required for investors to participate in the equity market now depends on the interest rate and rises from 0.60, over 0.69 to 0.79 for the three interest rates reported in Table 6. The critical level A^* increases with the interest rate because higher short rates lower the equity premium, giving stocks more room to disappoint. Stock non-participation now occurs for smaller degrees of

Fig. 4. Equity portfolio weights and A^* under disappointment aversion utility in the predictability system. The top plot shows equity portfolio weights under DA utility for the system with short rate predictability for curvature parameter $\gamma = 2$ and disappointment level $A = 0.85$. The short rate is on the horizontal axis, and the portfolio weights are on the vertical axis. We show portfolio weights corresponding to four different horizons (three months, one year, five years, ten years); rebalancing is always quarterly. The bottom plot shows the corresponding minimum A^* required to induce equity participation corresponding to the same horizons as the top plot.

DA utility, Linear Predictability, $\gamma = 2$ A = 0.85

A* for Linear Predictability, $\gamma = 2$ A = 0.85

disappointment aversion (higher A). At high interest rate levels, since A^* is high, modest cross-sectional variation in A produces substantial variation in equity participation. The naïve reaction of some retail investors to pull out of the stock market when money market rates are high is thus optimal in this framework.

In Fig. 4, we graph optimal equity weights under DA preferences with $\gamma = 2$ and $A = 0.85$ for various horizons as a function of the interest rate. As is also the case for CRRA utility, the portfolio weight curve is downward sloping and hedging demands are small. At first, hedging demands are positive and increase with higher interest rates. For very high interest rates, they become negative. At interest rates above approximately 12%, as the horizon increases, the interest rate at which the investor switches entirely to risk-free bonds decreases. To help gain intuition on this result, the bottom panel graphs A^* as a function of the interest rate for various horizons. At very high interest rates, A^* increases with horizon which causes equity holdings to decrease and produces the negative hedging demands. This is because for longer horizons, A^* is determined not only by the one-period-ahead distribution of excess returns, but also depends on future certainty equivalents of wealth. For very high interest rates (above 14.7%), the equity premium is negative. At high interest rate levels, for long horizons the probability of landing in the negative equity premium region is larger than for short horizons. This effect increases A^* for longer horizons at high interest rates. For low interest rates, around 2–4%, the probability of ending in the negative equity premium region is almost zero, so at low interest rates we find small positive hedging demands, as in the no predictability case.

In Table 6, the drop in equity holdings going from $A = 1.0$ to 0.85 is about 30%, and more generally, the portfolio weights decline almost linearly with the interest rate. This prompts the question whether the state dependence of DA utility is different from CRRA utility. If this is the case, we may find DA outcomes using CRRA utility with a higher risk aversion coefficient. Fig. 5 vividly illustrates that CRRA utility cannot replicate DA asset allocations. For each short rate, we start from the optimal equity weight at a horizon of one quarter for a DA investor with $\gamma = 5$ or $\gamma = 2$ and with $A = 0.85$. We then find a CRRA investor characterized by γ who chooses the same portfolio. If the above claim were true, we should find a horizontal line. In contrast, the line starts out relatively flat but then rapidly ratchets upward non-linearly for higher short rates, so the aversion of the DA investor to stocks increases non-linearly with higher interest rates. The implied CRRA risk aversion increases as a function of the short rate because the higher the short rate, the lower the equity premium, and thus the more stocks can disappoint. At very high interest rates, a DA investor with $A = 0.85$ holds zero equity, which can only be captured by infinite CRRA risk aversion.

6. Conclusions

In this article, we use the disappointment aversion preference framework developed by Gul (1991) to study the dynamic asset allocation problem. DA preferences incorporate loss aversion in that they treat gains and losses
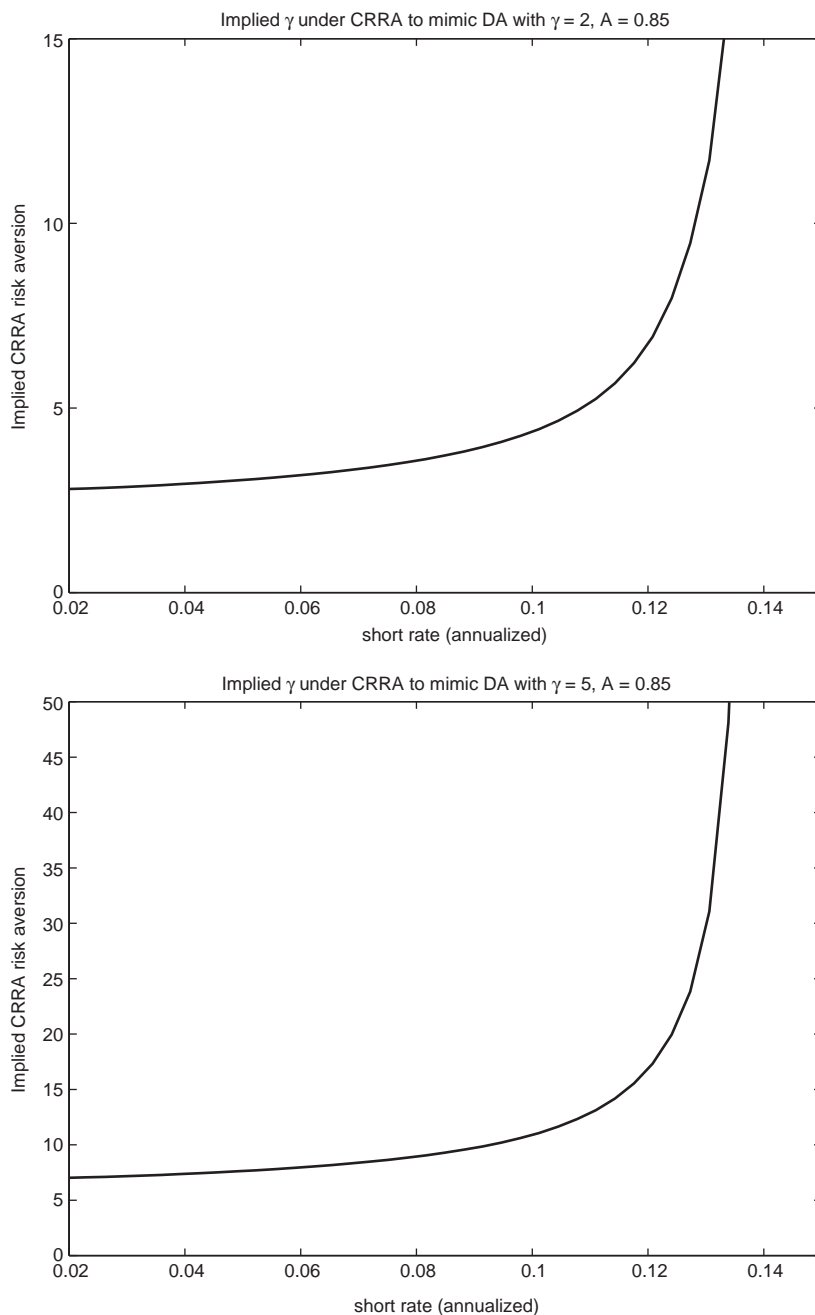
Fig. 5. Implied CRRA risk aversion for disappointment aversion utility in the predictability system. We plot the implied risk aversion γ under CRRA utility which produces the same portfolio weight as under DA utility for curvature parameter $\gamma = 2$ and disappointment level $A = 0.85$ (top plot) and $\gamma = 5$ and $A = 0.85$ (bottom plot). We show (annualized) short rates on the horizontal axis and the implied CRRA coefficient on the vertical axis.

asymmetrically, but are fully axiomatically motivated and admit easy comparison with standard expected utility. From the perspective of the smooth concave nature of constant relative risk averse (CRRA) preferences, the behavior of many investors often appears puzzling: investors often do not invest in the stock market, and a portfolio choice model with predictable equity returns often leads to substantially levered equity positions. Investors who are averse to disappointing outcomes should hold significantly less equity even with moderate curvature in the utility function. Moreover, we show that for high enough disappointment aversion, an investor's optimal equity position is zero.

By calibrating a number of data generating processes to U.S. data on stock and bond returns, we find very reasonable portfolios for disappointment-averse investors with utility functions exhibiting quite low curvature. DA preferences induce horizon effects and state dependence of asset allocation in such a way as to not be replicable by a CRRA utility function with higher curvature. Despite the large equity premium, stocks may disappoint! Whereas the primary focus of the recent literature has been on the effects of predictability or background risk on portfolio choice, our results suggest the importance of understanding the investor's attitude towards risk. The proper specification of an investor's utility function matters as much as, if not more than, the proper specification of the stochastic environment. Consequently, it is encouraging to see related work such as Barberis et al. (2001) who embed prospect theory in a dynamic portfolio choice model with consumption.

Whether heterogeneity in preferences or heterogeneity in circumstances is the more fruitful direction to pursue to explain the portfolio choice evidence remains to be seen. There is a scarcity of experimental work on risk preferences, and almost none on the kind of preferences we examine in this paper. Loomes and Segal (1994) focus on the implications of different utilities for the order of risk aversion. Standard CRRA preferences exhibit second-order risk aversion (the insurance premium the investor is willing to pay to avoid a gamble is proportional to the variance of the gamble), while DA preferences exhibit first-order risk aversion (the insurance premium is proportional to volatility). They observe both first- and second-order risk aversion in their subjects. Although they note that the first-order risk aversion embedded in DA preferences may not be strong enough relative to their experimental evidence, their results coupled with ours definitely suggest to take heterogeneity in preferences as a potentially important determinant of portfolio choice.

There are a number of interesting avenues for future work. Disappointment-averse agents dislike negative skewness much more than standard CRRA agents. Hence, the regular occurrence of equity market crashes inducing such skewness may further scare investors away from equity investments or it may induce them to buy (costly) insurance against such crashes. This may account for the recent popularity of put-protected products which seem to have lured many investors into the stock market. In an international context, the occurrence of correlated bear markets (see, e.g., Ang and Bekaert, 2002; Das and Uppal, 2004) may induce home bias in asset preferences for disappointment-averse investors. Although DA preferences yield portfolio allocations promisingly close to actual holdings in partial equilibrium settings, we

must ultimately investigate whether DA preferences can be accommodated in an equilibrium model of risk.

Appendix A. Solving the DA portfolio allocation problem

To solve Eqs. (6) and (9) numerically, we use quadrature to approximate the definition of μ_W in Eq. (6) by

$$\mu_W^{1-\gamma} = \frac{1}{K} \left(\sum_{s:W_s \leqslant \mu_W} p_s W_s^{1-\gamma} + \sum_{s:W_s > \mu_W} A p_s W_s^{1-\gamma} \right), \qquad (A.1)$$

and the FOC in Eq. (9) by

$$\sum_{s:W_s \leqslant \mu_W} p_s W_s^{-\gamma}(\exp(y_s) - \exp(r)) + \sum_{s:W_s > \mu_W} A p_s W_s^{-\gamma}(\exp(y_s) - \exp(r)) = 0. \qquad (A.2)$$

We solve Eqs. (A.1) and (A.2) simultaneously to yield the portfolio weight α that maximizes the utility of this disappointment-averse investor. Appendix D discusses the discretization procedure.

Let $x_e = (\exp(y) - \exp(r))$ denote the excess stock return. With N quadrature points there are N outcomes for x_e, $\{x_{es}\}_{s=1}^N$, with probability weights $\{p_s\}_{s=1}^N$. Without loss of generality, we can order x_e from low to high across states s. The utility equivalent μ_W^* corresponding to the optimal portfolio weight α^* can be in any of N intervals:

$$[\exp(r) + \alpha^* x_{e1}, \ \exp(r) + \alpha^* x_{e2}),$$
$$[\exp(r) + \alpha^* x_{e2}, \ \exp(r) + \alpha^* x_{e3}),$$
$$\vdots$$
$$[\exp(r) + \alpha^* x_{e,N-1}, \ \exp(r) + \alpha^* x_{eN}).$$

Suppose μ_W^* lies in $[\exp(r) + \alpha^* x_{ei}, \ \exp(r) + \alpha^* x_{e,i+1})$ for some state i. Then α^* solves

$$\sum_{s:W_s \leqslant \exp(r)+\alpha^* x_{e,i}} p_s(W_s^*)^{-\gamma} x_{es} + \sum_{s:W_s > \exp(r)+\alpha^* x_{e,i+1}} A p_s(W_s^*)^{-\gamma} x_{es} = 0, \qquad (A.3)$$

where $W_s^* = \exp(r) + \alpha^* x_{es}$. Eq. (A.3) is a CRRA maximization problem with a changed probability distribution $\pi_i = \{\pi_{is}\}_{s=1}^N$, where the probabilities for wealth above the certainty equivalent are downweighted, i.e., the probabilities π_{is} are transformed from the original quadrature probabilities p_s by the relation

$$\pi_i \equiv \frac{(p_1, \ldots, p_i, A p_{i+1}, \ldots, A p_N)'}{(p_1 + \cdots + p_i) + A(p_{i+1} + \cdots + p_N)}. \qquad (A.4)$$

Our algorithm is as follows. We start with a state i and solve the CRRA problem with probability distribution π_i. Then we compute the certainty equivalent, μ_{Wi}^*,

given by

$$\mu^*_{Wi} = \left(\sum_{s=1}^{N} (W^*_s)^{1-\gamma} \pi_{is} \right)^{1/(1-\gamma)}. \tag{A.5}$$

Then, we check if in this state the following is true:

$$\mu_{Wi} \in [\exp(r) + \alpha^*_i x_{ei}, \ \exp(r) + \alpha^*_i x_{e,i+1}). \tag{A.6}$$

If this is true for $i = i^*$, then $\alpha^* = \alpha^*_i$ and $\mu^*_W = \mu^*_{Wi}$. As the states are ordered in increasing wealth across states for a given portfolio weight, it is easy to do a bisection search algorithm (with intermediate CRRA optimizations) to obtain the DA portfolios. If we start our search for i^* at the midpoint of the N states and find that $\mu_{Wi} > (<) \exp(r) + \alpha^*_i x_{e,i+1}$, then we begin a search in the upper (lower) half of the state space.

Gul (1991)'s appendix describes a similar algorithm. Both our algorithm and Gul's require the solution of an optimization problem in each discrete state. The difference is that in our algorithm we solve a simple smooth CRRA problem, whereas Gul requires a non-linear maximization involving an indicator function. For his optimization problem, gradient-based search algorithms cannot be used, and thus our algorithm is numerically more tractable.

We can extend this solution to the dynamic DA problem in Section 3.2. Specifically, if wealth W_s is increasing across states s for a given portfolio weight, and the certainty equivalent for horizon t, $\mu_{s,t}$, is also increasing across states for a given portfolio weight, then $\tilde{W}_s = R_s \mu_{s,t}$ is also increasing across states s.

Appendix B. Proof of Proposition 2.1

Define

$$A^* = -\frac{\mathrm{E}[x_e | x_e \leqslant 0] \Pr(x_e \leqslant 0)}{\mathrm{E}[x_e | x_e > 0] \Pr(x_e > 0)}. \tag{B.1}$$

As we formally show, A^* is the level of disappointment such that for $A \leqslant A^*$, $\alpha^* = 0$ and for $A > A^*$, $\alpha^* > 0$. Note that this definition of A is independent of risk aversion γ.

Considering optimality at $\alpha = 0$ is a special case since the certainty equivalent equals the gross risk-free rate $R_f = \exp(r)$ and since the definition of disappointing or elating states switches when α changes from negative to positive (if $x_e > 0$, $R_f + \alpha x_e > R_f$ only for positive α). Therefore, we must consider left- and right-hand side derivatives to determine optimality.

Consider first $A < A^*$. We show that the optimal asset allocation at A is $\alpha^* = 0$. We start by denoting the certainty equivalent $v(A, \alpha)$ as a function of the disappointment level A and the portfolio weight α

$$v(A, \alpha)^{1-\gamma} = \frac{1}{K} \{ \mathrm{E}[U(W) \mathbf{1}_{\{W \leqslant v(A,\alpha)\}}] + A \mathrm{E}[U(W) \mathbf{1}_{\{W > v(A,\alpha)\}}] \}, \tag{B.2}$$

with $K = \Pr(W \leqslant v(A, \alpha)) + A \Pr(W > v(A, \alpha))$. Recall that $W = R_f + \alpha x_e$ in our setting.

The derivative of $v(A, \alpha)$ with respect to α is given by

$$v(A, \alpha)^{-\gamma} \cdot \frac{\partial v(A, \alpha)}{\partial \alpha} = \frac{1}{K} \{ E[W^{-\gamma} x_e \mathbf{1}_{\{W \leqslant v(A, \alpha)\}}] + A E[W^{-\gamma} x_e \mathbf{1}_{\{W < v(A, \alpha)\}}] \}. \quad \text{(B.3)}$$

This is the well-known first-order condition, derived for instance in Epstein and Zin (2001) and Bekaert et al. (1997). This expression is the same as the derivative of the terms in the integrands in (B.2). However, taking the derivative of $v(\cdot)$ with respect to α also involves taking the derivatives of K with respect to α and the derivatives of the certainty equivalent in the integration limits, both with respect to α. In the NBER working version of this paper, we explicitly show that the latter two derivatives of the indicator functions sum to zero.

When α approaches zero, we have $W = R_f$ and $v(A, 0) = R_f$. Hence, we can equivalently express $\mathbf{1}_{\{W \leqslant v(A, \alpha)\}}$ as $\mathbf{1}_{\{\alpha x_e \leqslant 0\}}$ and, analogously, $\mathbf{1}_{\{W > v(A, \alpha)\}}$ as $\mathbf{1}_{\{\alpha x_e > 0\}}$. Clearly, the value of these indicator functions depends on whether we approach zero from the left or the right. Let us first take the LHS derivative of $v(\cdot)$ at $\alpha = 0$. First, note that because $\alpha < 0$,

$$\mathbf{1}_{\{\alpha x_e \leqslant 0\}} = \mathbf{1}_{\{x_e \geqslant 0\}} \quad \text{and} \quad \mathbf{1}_{\{\alpha x_e > 0\}} = \mathbf{1}_{\{x_e < 0\}}. \quad \text{(B.4)}$$

Second, the terms $v(A, 0)^{-\gamma}$ and $W^{-\gamma}$ cancel on each side of the equation. Consequently, we obtain

$$\left. \frac{\partial v}{\partial \alpha} \right|_{\alpha = 0^-} = \frac{1}{K} \{ E[x_e \mathbf{1}_{\{x_e \geqslant 0\}}] + A E[x_e \mathbf{1}_{\{x_e < 0\}}] \}, \quad \text{(B.5)}$$

where $K = \Pr(x_e \geqslant 0) + A \Pr(x_e < 0)$. Since $\alpha < 0$, states in which $x_e < 0$ have higher wealth than the certainty equivalent and these are now downweighted by A, since $A \leqslant 1$. But then

$$\left. \frac{\partial v}{\partial \alpha} \right|_{\alpha = 0^-} \geqslant \frac{E[x_e]}{K} > 0 \quad \text{(B.6)}$$

by the assumption of a positive risk premium $E[x_e] > 0$ and because $K > 0$. Hence, we conclude that $\partial v / \partial \alpha > 0$ and it must be that $\alpha^* \geqslant 0$ because the utility function is globally concave in α.

Now let us consider the case of the RHS derivative and $\alpha^* > 0$. In this case, we have

$$\mathbf{1}_{\{\alpha x_e \leqslant 0\}} = \mathbf{1}_{\{x_e \leqslant 0\}} \quad \text{and} \quad \mathbf{1}_{\{\alpha x_e > 0\}} = \mathbf{1}_{\{x_e > 0\}}. \quad \text{(B.7)}$$

Consequently, we obtain

$$\left. \frac{\partial v}{\partial \alpha} \right|_{\alpha = 0^+} = \frac{1}{K} \{ E[x_e \mathbf{1}_{\{x_e \leqslant 0\}}] + A E[x_e \mathbf{1}_{\{x_e > 0\}}] \}. \quad \text{(B.8)}$$

Here, as is usual, the good states are positive excess return states, since $\alpha > 0$ and they are downweighted by A. By assumption, $A \leqslant A^*$, so

$$\left. \frac{\partial v}{\partial \alpha} \right|_{\alpha = 0^+} < \frac{1}{K} \{ E[x_e \mathbf{1}_{\{x_e \leqslant 0\}}] + A^* E[x_e \mathbf{1}_{\{x_e > 0\}}] \} = 0, \quad \text{(B.9)}$$

where the equality follows by definition of A^*. Hence, it must be the case that $\alpha^* \leqslant 0$.

Combining the two cases above, we have $\alpha^* = 0$. Note that in the above argument for the utility function increasing in α for $\alpha < 0$, we only used the fact that $E[x_e] > 0$ and $A \leqslant 1$. We use the extra assumption $A < A^*$ to show the utility function is decreasing in α for $\alpha > 0$. When $A > A^*$, the utility function is increasing at $\alpha = 0^+$ as well as at $\alpha = 0^-$, therefore $\alpha^* > 0$. Note that the RHS of Eq. (B.9) also constitutes the FOC at A^*.

Appendix C. Optimal portfolio solutions under LA utility

The following proposition gives conditions under which a finite portfolio choice solution with LA preferences is possible.

Proposition C.1 (*Existence of optimal LA portfolio weights*). *Consider the LA utility function in Eq. (24), with $\gamma_1 = \gamma_2 = \gamma$ and $0 \leqslant \gamma < 1$. Then there exists a finite solution for the optimal portfolio weight α^* only when both $B_1 < 0$ and $B_2 < 0$, where B_1 and B_2 are given by*

$$B_1 = -\lambda E[(-x_e)^{1-\gamma}\mathbf{1}_{\{x_e \leqslant 0\}}] + E[x_e^{1-\gamma}\mathbf{1}_{\{x_e > 0\}}]$$

and

$$B_2 = E[(-x_e)^{1-\gamma}\mathbf{1}_{\{x_e \leqslant 0\}}] - \lambda E[x_e^{1-\gamma}\mathbf{1}_{\{x_e > 0\}}]. \tag{C.1}$$

Under these conditions, the optimal weight α^* depends on the benchmark B_0 but is independent of λ.

Proof. When the portfolio weight $\alpha \to +\infty$ then

$$U \to (\alpha)^{1-\gamma}B_1,$$

so $U \to +\infty$ if $B_1 > 0$ and there is no optimal weight. Similarly,

$$U \to (|\alpha|)^{1-\gamma}B_2,$$

when $\alpha \to -\infty$, so $U \to +\infty$ if $B_2 > 0$ and there is no optimal weight. Therefore, the optimal portfolio weight can only exist if $B_1 < 0$ and $B_2 < 0$.

If both $B_1 < 0$ and $B_2 < 0$, then as $\alpha \to \infty$, $U \to -\infty$ and as $\alpha \to -\infty$, $U \to -\infty$. Since U is monotonic in wealth for $0 \leqslant \gamma < 1$ there must exist an optimal solution α^*. □

Proof of Proposition 4.1. If $B_0 = W \exp(r)$, then LA utility becomes

$$U = (|\alpha|)^{1-\gamma}B_2\mathbf{1}_{\alpha \leqslant 0} + \alpha^{1-\gamma}B_1\mathbf{1}_{\alpha > 0}, \tag{C.2}$$

where B_1 and B_2 are defined in Eq. (C.1). If $B_1 < 0$ and $B_2 < 0$, the utility U is maximized at $\alpha = 0$. If $B_2 > 0$, then $\alpha \to -\infty$ and $U \to +\infty$, so $\alpha^* = -\infty$. Similarly, if $B_1 > 0$, then $U \to +\infty$ as $\alpha \to +\infty$, so $\alpha^* = +\infty$. □

Proof of Proposition 4.2. Suppose the risk premium is zero, and the probability of a negative equity return occurring is smaller than the probability of a positive equity return. Then the probability transformation of prospect theory assigns a higher

probability weight to the negative return, which makes the risk premium negative under the subjective measure. Hence an agent with these preferences shorts the stock. $\square$

Appendix D. Data generating processes

We estimate the following VAR:

$$X_t = c + \Phi X_{t-1} + u_t, \tag{D.1}$$

where $u_t \sim$ IID N$(0, \Sigma)$. For our system $X_t = (\tilde{y}_t \, r_t)'$, where $\tilde{y}_t = y_t - r_{t-1}$ is the excess equity return and r_t is the short rate. The optimal lag choice by the Bayesian Information Criteria (BIC) is one lag.

The system without predictability has

$$\Phi = \begin{pmatrix} 0 & 0 \\ 0 & \rho \end{pmatrix}$$

and in the system with predictability,

$$\Phi = \begin{pmatrix} 0 & b \\ 0 & \rho \end{pmatrix}.$$

Eq. (D.1) can be written in compact form as

$$X = BZ + U, \tag{D.2}$$

where $X = (X_1 \ldots X_T)$ $(2 \times T)$, $B = [c\Phi]$ (2×3), $U = (u_1 \ldots u_T)$ $(2 \times T)$, $Z = (z_0 \ldots z_{T-1})$ $(3 \times T)$ with $z_t = [1 X_t']'$ (3×1). The restrictions are written as $R\beta = r$ with $\beta = \text{vec}(B)$. The unrestricted maximum likelihood estimator, where Φ is unconstrained is given by

$$\hat{\beta} = ((ZZ')^{-1}Z \otimes I)Y,$$

where $Y = \text{vec}(X)$. The restricted maximum likelihood estimator is given by

$$\hat{\beta}^c = \hat{\beta} + ((ZZ')^{-1} \otimes I)R'(R((ZZ')^{-1} \otimes I)R')^{-1}(r - R\hat{\beta}) \tag{D.3}$$

and $\hat{B} = \text{devec}(\hat{\beta}^c)$.

The estimate of Σ is given by $\hat{\Sigma} = 1/T \, (\hat{U}'\hat{U})$, where $\hat{U} = X - \hat{B}Z$. The estimated covariance of $\hat{\beta}_c$ is given by

$$\widehat{\text{cov}}(\hat{\beta}_c) = \Gamma \otimes \hat{\Sigma} - (\Gamma \otimes \hat{\Sigma})R'(R(\Gamma \otimes \hat{\Sigma})R')^{-1}R(\Gamma \otimes \hat{\Sigma}), \tag{D.4}$$

where $\Gamma = (ZZ')^{-1}$. The estimated covariance of $\text{vech}(\hat{\Sigma})$ is given by

$$\widehat{\text{cov}}(\text{vech}(\hat{\Sigma})) = \frac{2}{T} \, D^{-1}(\hat{\Sigma} \otimes \hat{\Sigma})(D^{-1})', \tag{D.5}$$

where D^{-1} is the Moore-Penrose inverse of D, the duplication matrix which makes $\text{vec}(C) = D \text{vech}(C)$ for a symmetric matrix C.

D.1. Time aggregation of VARs

Define the time-aggregated process $\bar{X}_{t+k,k} = X_{t+1} + \cdots + X_{t+k}$ over k horizons. If X_t follows the VAR given by $X_{t+1} = \mu + \Phi X_t + \varepsilon_{t+1}$, with $\varepsilon_{t+1} \sim$ IID N$(0, \Sigma)$, then we can define a time-aggregated VAR as

$$\bar{X}_{t+k,k} = \bar{\mu} + \bar{\Phi}\bar{X}_{t,k} + u_{t+k,k}. \tag{D.6}$$

The companion form of the time-aggregated VAR $\bar{\Phi}$ is simply $\bar{\Phi} = \Phi^k$ and $\bar{\mu}$ is given by

$$\bar{\mu} = (I + \Phi + \cdots + \Phi^k)\mu. \tag{D.7}$$

The conditional covariance $E_t(u_{t+k,k}u'_{t+k,k}) = \bar{\Sigma}$ is given by

$$
\begin{aligned}
\bar{\Sigma} = {} & \Sigma + (I + \Phi)\Sigma(I + \Phi)' + \cdots + (I + \Phi + \cdots + \Phi^k)\Sigma(I + \Phi + \cdots + \Phi^k)' \\
& + (\Phi + \Phi^2 + \cdots + \Phi^k)\Sigma(\Phi + \Phi^2 + \cdots + \Phi^k)' \\
& + (\Phi^2 + \cdots + \Phi^k)\Sigma(\Phi^2 + \cdots + \Phi^k)' + \cdots + \Phi^k\Sigma(\Phi^k)'.
\end{aligned}
\tag{D.8}
$$

D.2. Discretization of VARs

We construct an approximate discrete Markov chain to the VAR in Eq. (D.1) using the quadrature-based methods of Tauchen and Hussey (1991). For the system for $X_t = (\tilde{y}_t r_t)$, with $\tilde{y}_t = y_t - r_{t-1}$ the excess equity return and r_t the short rate, $\tilde{y}_t$ may be dependent on lagged r_t but not vice versa, so r_t is the driving variable in the system. We choose $N = 50$ points for the short rate over a uniform grid and denote these as $\{r_i\}$. The short rate is very persistent, so many points are necessary for an accurate approximation (see Tauchen and Hussey, 1991). We use a uniform grid because points chosen by Gaussian-Hermite quadrature perform poorly in optimization as they are too widely spaced. We construct the transition probabilities Π_r $(N \times N)$ for going from state r_i to r_j, $1 \leqslant i,j \leqslant N$ by evaluating the conditional density of r_j (which is conditionally Normal) and then normalizing the densities so that they sum to unity. This is the driving process of the discretized system.

We choose $M = 30$ discrete states for $\tilde{y}_t$. These states are chosen using Gaussian-Hermite points approximating the unconditional distribution of $\tilde{y}_t$ implied by Eq. (D.1). To include $\tilde{y}_t$ in the discretization we note that for each state r_i, an $N \times M$ vector π_i can be constructed giving the transition probabilities going from state r_i $(1 \leqslant j \leqslant N)$ to $(r_j, \tilde{y}_j)$ $(1 \leqslant j \leqslant N \times M)$. The distribution of $\tilde{y}_t$ conditional on r_i is normal, and is discretized by evaluating the distribution of $\tilde{y}_t$ conditional on r_i for going from state r_i to state $(r_j, \tilde{y}_j)$. A Choleski decomposition is used to take account of the contemporaneously correlated error terms u_t in Eq. (D.1). The vectors π_i can be stacked to give a $N \times NM$ probability transition matrix Π_{ry} giving the probabilities from $\{r_i\}$, $1 \leqslant i \leqslant N$ to $\{r_j, \tilde{y}_j\}$, $1 \leqslant j \leqslant NM$. The Markov chain constructed in this way matches first and second moments of the VAR in Eq. (D.1) to three-to-four significant figures. It is possible to also construct a square Π matrix, but this matrix will have repeated rows.

References

Aït-Sahalia, Y., Brandt, M., 2001. Variable selection for portfolio choice. Journal of Finance 54, 1297–1351.

Ang, A., Bekaert, G., 2002. International asset allocation with time-varying correlations. Review of Financial Studies 15 (4), 1137–1187.

Ang, A., Bekaert, G., 2003. Stock return predictability—is it there? Working paper, Columbia University.

Balduzzi, P., Lynch, A.W., 1999. Transaction costs and predictability: some utility cost calculations. Journal of Financial Economics 52, 47–78.

Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance 55 (1), 225–264.

Barberis, N., Huang, M., Santos, J., 2001. Prospect theory and asset prices. Quarterly Journal of Economics 116 (1), 1–53.

Bekaert, G., Hodrick, R.J., Marshall, D.A., 1997. The implications of first-order risk aversion for asset market risk premiums. Journal of Monetary Economics 40, 3–39.

Benartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110, 73–92.

Berkelaar, A., Kouwenberg, R., 2000. Optimal portfolio choice under loss aversion. Working paper.

Brandt, M., 1999. Estimating portfolio and consumption choice: a conditional Euler equations approach. Journal of Finance 64 (5), 1609–1645.

Brennan, M.J., Schwartz, E.S., Lagnado, R., 1997. Strategic asset allocation. Journal of Economic Dynamics and Control 21, 1377–1403.

Campbell, J.Y., Viceira, L.M., 1999. Consumption and portfolio decisions when expected returns are time varying. Quarterly Journal of Economics 114, 433–495.

Das, S.J., Uppal, R., 2004. International portfolio choice with systemic risk. Journal of Finance 59, 2809–2834.

Dow, J., Werlang, S., 1992. Uncertainty aversion, risk aversion, and the optimal choice of portfolio. Econometrica 60 (1), 197–204.

Epstein, L.G., Schneider, M., 2004. Learning under ambiguity. Working paper, University of Rochester.

Epstein, L.G., Zin, S.E., 1989. Substitution, risk aversion and the temporal behavior of consumption and asset returns. Econometrica 57, 937–969.

Epstein, L.G., Zin, S.E., 1990. 'First-order' risk aversion and the equity premium puzzle. Journal of Monetary Economics 387–407.

Epstein, L.G., Zin, S.E., 2001. The independence axiom and asset returns. Journal of Empirical Finance 8, 537–572.

Friend, I., Blume, M., 1975. The demand for risky assets. American Economic Review 65 (5), 900–922.

Gomes, F.J., 2003. Portfolio choice and trading volume with loss-averse investors. Journal of Business, forthcoming.

Goyal, A., Welch, I., 2003. The myth of predictability: does the dividend yield forecast the equity premium? Management Science 49 (5), 639–654.

Gul, F., 1991. A theory of disappointment aversion. Econometrica 59 (3), 667–686.

Haliassos, M., Bertaut, C.C., 1995. Why do so few hold stocks? Economic Journal 110, 1–35.

Heaton, J., Lucas, D.J., 1997. Market frictions, savings behavior and portfolio choice. Macroeconomic Dynamics 1, 76–101.

Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–291.

Kandel, S., Stambaugh, R.F., 1996. On the predictability of stock returns: an asset-allocation perspective. Journal of Finance 51 (2), 385–424.

Kim, T.S., Omberg, E., 1996. Dynamic nonmyopic portfolio behavior. Review of Financial Studies 9 (1), 141–161.

Kreps, D.M., Porteus, E.L., 1979. Dynamic choice theory and dynamic programming. Econometrica 47 (1), 91–100.

Liu, H., Loewenstein, M., 2002. Optimal portfolio selection with transactions costs and finite horizons. Review of Financial Studies 15 (3), 805–835.

Liu, J., 1999. Portfolio selection in stochastic environments. Working paper, UCLA.

Liu, W.F., 2002. Savings and portfolio decision: observable implications with Knightian uncertainty. Working paper, University of Washington.

Loomes, G., Segal, U., 1994. Observing different orders of risk aversion. Journal of Risk and Uncertainty 9 (3), 239–256.

Lynch, A.W., 2001. Portfolio choice and equity characteristics: characterizing the hedging demand induced by return predictability. Journal of Financial Economics 62 (1), 67–130.

Machina, M., 1982. 'Expected utility' analysis without the independence axiom. Econometrica 50 (2), 277–323.

Maenhout, P.J., 2004. Robust portfolio rules and asset pricing. Review of Financial Studies 17 (4), 951–983.

Mankiw, N.G., Zeldes, S.P., 1991. The consumption of stockholders and nonstockholders. Journal of Financial Economics 29 (1), 97–112.

Rabin, M., 2000. Risk aversion and expected utility theory: a calibration theorem. Econometrica 68 (5), 1281–1292.

Routledge, B.R., Zin, S.E., 2003. Generalized disappointment aversion and asset prices. NBER working paper 10107.

Roy, A.D., 1952. Safety first and the holding of assets. Econometrica 20, 431–449.

Samuelson, P.A., 1991. Long-run risk tolerance when equity returns are mean regressing: pseudoparadoxes and vindication of businessmen's risk. In: Brainard, W.C., Nordhaus, W.D., Watts, H.W. (Eds.), Money, Macroeconomics and Economic Policy. MIT Press, Cambridge, MA, pp. 181–200.

Sharpe, W.F., 1998. Morningstar's risk-adjusted ratings. Financial Analysts Journal July–August 21–33.

Stutzer, M.A., 2000. A portfolio performance index. Financial Analysts Journal 56 (3), 52–61.

Tauchen, G., Hussey, R., 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. Econometrica 59 (2), 371–396.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.

Vissing-Jørgensen, A., 2002. Limited asset market participation and the elasticity of intertemporal substitution. Journal of Political Economy 110 (4), 825–853.