Supplementary Appendix to

LEARNING UNDER AMBIGUITY*

Larry G. Epstein   Martin Schneider

January 30, 2007

Abstract

This appendix provides a detailed treatment of the portfolio choice problem studied in our paper "Learning under Ambiguity", filling in details of calculations that were omitted in the text of the paper.

1 Investor Problem

Section 1 of this appendix restates the portfolio problem studied in Section 5 of "Learning under Ambiguity". Section 2 defines beliefs and derives the dynamics of beliefs under ambiguity. Section 3 computes optimal portfolio weights for a benchmark Bayesian investor and a myopic ambiguity-averse investor.

Time is measured in months; there are k trading dates per month. The state space is $S = \{1, 0\}$.[1] The return on stocks $R(s_t) = e^{r(s_t)}$ realized in period t is either high or low: we fix (log) return realizations $r(1) = \sigma/\sqrt{k}$ and $r(0) = -\sigma/\sqrt{k}$. In addition to stocks, the investor also has access to a riskless asset with constant per period interest rate $R^f = e^{r^f/k}$, where $r^f < \sigma$. We consider investors who plan for T months starting in month t and who care about terminal wealth according to the utility function $V_T(W_{t+T}) = \log W_{t+T}$. Investors may rebalance their portfolio at all $k(T-t)$ trading dates between t and T. Let $\omega^*_{t,T,k}$ denote the optimal fraction of wealth invested in stocks at date t for an investor who plans for T months, when there are k trading dates per month.

Consider the investor's problem when beliefs are given by a general process of one-step-ahead conditionals $\{\mathcal{P}_\tau(s^\tau)\}$. The history s^τ of state realizations up to trading date

*Epstein: Department of Economics, U. Rochester, lepn@troi.cc.rochester.edu; Schneider: Department of Economics, NYU, and Federal Reserve Bank of Minneapolis, ms1927@nyu.edu.

[1] In this appendix, we call the binary states 1 and 0, rather than *hi* and *lo* as in the text of "Learning under Ambiguity", because this allows simpler notation below.

$\tau = t + j/k$ – the jth trading date in month $t+1$ – can be summarized by the fraction ϕ_τ of the state $s_t = 1$ observed up to τ. The value function of the log investor takes the form $V_\tau(W_\tau, s^\tau) = h_\tau(\phi_\tau) + \log W_\tau$. The process $\{h_\tau\}$ satisfies $h_{t+T} = 0$ and

$$\begin{aligned} h_\tau(\phi_\tau) &= \max_{\omega_\tau} \min_{p_\tau \in \mathcal{P}_\tau(s^\tau)} E^{p_\tau}\left[\log\left(R^f + \left(R\left(s_{\tau+1/k}\right) - R^f\right)\omega_\tau\right) + h_{\tau+1/k}\left(\phi_{\tau+1/k}\right)\right] \\ &= \min_{p_\tau \in \mathcal{P}_\tau(s^\tau)} \max_{\omega_\tau} E^{p_\tau}\left[\log\left(R^f + \left(R\left(s_{\tau+1/k}\right) - R^f\right)\omega_\tau\right) + h_{t+1/k}\left(\phi_{\tau+1/k}\right)\right], \quad (1) \end{aligned}$$

where we have used the minimax theorem to reverse the order of optimization.

We are interested in the optimal portfolio of an investor who has seen a monthly sample of log real US stock returns $\{r_\tau\}_{\tau=1}^t$. However, agents in the model are assumed to observe not only monthly returns, but actually binary returns $R(s_\tau)$ at every trading date. To model an agent's history up to date t, we thus construct a sample of tk realizations of $R(s_\tau)$ such that the implied empirical distribution of monthly log returns is the same as that in the data. Let σ denote the monthly standard deviation of log returns. The sample of binary returns up to some integer date τ is summarized by the fraction of states $s_t = 1$, defined by

$$\phi_\tau = \hat{\phi}_k(\bar{r}_\tau) := \frac{1}{2} + \frac{1}{2}\frac{\bar{r}_\tau}{\sigma\sqrt{k}}. \quad (2)$$

where $\bar{r}_\tau := \frac{1}{\tau}\Sigma_{j=1}^\tau r_j$ is the mean of the monthly return sample. For given k, the sequence $\{\phi_\tau\}$ pins down a sequence of monthly log returns $\left\{\Sigma_{j=1}^k \log R\left(s_{\tau+j/k}\right)\right\}$ that is identical to the data sample $\{r_\tau\}$.

2 Beliefs

As a Bayesian benchmark, we assume that the investor has an improper beta prior over the probability p of the high state, so that the posterior mean of p after t months (or tk state realizations) is equal to ϕ_τ, the maximum likelihood estimator of p. The Bayesian's probability of a high state next period is then also given by ϕ_τ. The optimal portfolio follows from solving (1) when $\mathcal{P}_\tau(s^\tau)$ is a singleton set containing only the measure that puts probability ϕ_τ on the high state.

For an ambiguity-averse investor, beliefs are defined as in Section 3 of "Learning under Ambiguity". We briefly review the general model here. Beliefs are represented by

$$(\Theta, \mathcal{M}_0, \mathcal{L}_k, \alpha),$$

where Θ is a parameter space, $\mathcal{M}_0$ is a set or priors on Θ, $\mathcal{L}_k$ is a set of likelihoods and α is a number between zero and one. A *theory* is a pair (μ_0, ℓ^t), where μ_0 is a prior belief on Θ and $\ell^t = \left(\ell_{1/k}, \ell_{2/k}.., \ell_t\right) \in \mathcal{L}_k^{tk}$ is a *sequence* of likelihoods. Let $\mu_t(\cdot\,; s^t, \mu_0, \ell^t)$ denote the posterior derived from the theory (μ_0, ℓ^t) by Bayes' Rule, given the data s^t.

The set of posteriors contains posteriors that are based on theories not rejected by a likelihood ratio test:

$$\mathcal{M}^{\alpha}_{t,k}(s^t) = \{\mu_t\left(s^t;\mu_0,\ell^t\right):\ \mu_0 \in \mathcal{M},\ \ell^t \in \mathcal{L}^t_k,$$
$$\int \Pi^t_{j=1}\ell_j\left(s_j|\theta\right)d\mu_0(\theta) \geq \alpha \max_{\substack{\tilde{\mu}_0\in\mathcal{M}_0 \\ \tilde{\ell}^t\in\mathcal{L}^t_k}} \int \Pi^t_{j=1}\tilde{\ell}_j\left(s_j|\theta\right)d\tilde{\mu}_0\}. \tag{3}$$

The set of one-step-ahead conditional belief is defined by

$$\mathcal{P}_t(s^t) = \left\{p_t\left(\cdot\right) = \int_{\Theta}\ell(\cdot\mid\theta)\,d\mu_t(\theta) : \mu_t \in \mathcal{M}^{\alpha}_t(s^t),\ \ell \in \mathcal{L}_k\right\}, \tag{4}$$

This process enters the specification of recursive multiple priors preferences in (1).

The particular representation $(\Theta, \mathcal{M}_0, \mathcal{L}_k, \alpha)$ assumed for the portfolio choice problem is defined as follows. The ambiguity averse investor perceives the mean monthly log return as $\theta+\lambda_t$, where $\theta \in \Theta := \mathbb{R}$ is fixed and can be learned, while λ_t is driven by many poorly understood factors affecting returns and can *never* be learned. The set $\mathcal{L}_k$ consists of all $\ell\left(\cdot\mid\theta\right)$ such that

$$\ell^k\left(hi|\theta\right) = \frac{1}{2} + \frac{1}{2}\frac{\theta+\lambda}{\sigma\sqrt{k}},\quad \text{for some } \lambda \text{ with } |\lambda| < \bar{\lambda}. \tag{5}$$

The set of priors $\mathcal{M}_0$ on Θ consists of Dirac measures. For simplicity, we write $\theta \in \mathcal{M}_0$ if the Dirac measure on θ is included in the set of priors, and we define

$$\mathcal{M}_0 = \left\{\theta : |\ \theta\ | \ \leq\ \bar{\lambda} + 1/\sigma\right\}.$$

The condition ensures that the probability (5) remains between zero and one for all $k \geq 1$.

Beliefs depend on history only via the fraction of high returns $\phi_t = \hat{\phi}_k\left(\bar{r}_t\right)$. We write $\theta \in \mathcal{M}^{\alpha}_{t,k}\left(\bar{r}_t\right)$ if the Dirac measure on θ is included in the posterior set at the end of month t after history $\hat{\phi}_k\left(\bar{r}_t\right)$. The posterior set for fixed k can be characterized as follows:

Proposition S1. *The posterior set is a subinterval* $\left[\underline{\theta}_k\left(\bar{r}_t\right), \bar{\theta}_k\left(\bar{r}_t\right)\right]$ *of* $\mathcal{M}_0$ *with both bounds strictly increasing in* $\bar{r}_t$*.*

Proof. The history s^t consists of tk realizations of the state. Write the likelihood of a sample s^t under some theory, here identified with a pair $\left(\theta, \lambda^t\right)$, as

$$L_k\left(s^t,\theta,\lambda^t\right) = \prod_{\tau=0}^{t-1}\prod_{j=1}^{k}\left(\frac{1}{2}+\frac{1}{2}\frac{\theta+\lambda_{\tau+j/k}}{\sigma\sqrt{k}}\right)^{s_{\tau+j/k}}\left(\frac{1}{2}-\frac{1}{2}\frac{\theta+\lambda_{\tau+j/k}}{\sigma\sqrt{k}}\right)^{1-s_{\tau+j/k}}. \tag{6}$$

Let $\tilde{\lambda}^t$ denote the sequence that maximizes (6) for fixed θ. This sequence is independent of θ and has $\tilde{\lambda}_i = \overline{\lambda}$ if $s_i = 1$ and $\tilde{\lambda}_i = -\overline{\lambda}$ if $s_i = 0$, for all $i \leq t$. It follows that

$L_k\left(s^t,\theta,\tilde{\lambda}^t\right)$ depends on the sample only through the fraction ϕ_t of high returns observed. The posterior set can be written

$$\mathcal{M}^{\alpha}_{t,k}\left(s^t\right)=\left\{\theta:\frac{1}{tk}\log L\left(s^t,\theta,\tilde{\lambda}^t\right)\geq\max_{\tilde{\theta}}\frac{1}{tk}\log L\left(s^t,\tilde{\theta},\tilde{\lambda}^t\right)-\frac{1}{tk}\log\left(\alpha\right)\right\}\tag{7}$$

Indeed, if $\theta\in\mathcal{M}^{\alpha}_{t,k}$, then there exists *some* λ^t such that the theory $\left(\theta,\lambda^t\right)$ passes the criterion for an admissible theory put forward in (3). Thus the theory $\left(\theta,\tilde{\lambda}^t\right)$ must also pass that criterion, since its likelihood is at least as high. In contrast, if $\theta\notin\mathcal{M}^{\alpha}_{t,k}$, then there is *no* λ^t such that the theory $\left(\theta,\lambda^t\right)$ passes the criterion.

Using the fraction of high states $\phi_t=\frac{1}{tk}\sum_i s_i$, rewrite the log data density as

$$\log L\left(s^t,\theta,\tilde{\lambda}^t\right)=f\left(\theta;\phi_t\right):=\phi_t\log\left(\frac{1}{2}+\frac{1}{2}\frac{\theta+\bar{\lambda}}{\sigma\sqrt{k}}\right)+\left(1-\phi_t\right)\log\left(\frac{1}{2}-\frac{1}{2}\frac{\theta-\bar{\lambda}}{\sigma\sqrt{k}}\right).$$

The likelihood ratio criterion becomes

$$f\left(\theta;\phi_t\right)\geq\max_{\tilde{\theta}}\left\{f\left(\theta;\phi_t\right)\right\}-\frac{1}{tk}\log\left(\alpha\right).\tag{8}$$

For $\phi_t\in(0,1)$, the function f is strictly concave and achieves a unique maximum at the MLE

$$\theta^*\left(\phi_t\right)=\left(2\phi_t-1\right)\left(\sigma\sqrt{k}+\bar{\lambda}\right)=\left(2\phi_k\left(\bar{r}_t\right)-1\right)\left(\sigma\sqrt{k}+\bar{\lambda}\right)=\bar{r}_t\left(1+\bar{\lambda}/\sigma\sqrt{k}\right).$$

Since f is strictly concave and $\lim_{\theta\to\bar{\lambda}+1/\sigma}f\left(\theta;\phi_t\right)=\lim_{\theta\to-\bar{\lambda}-1/\sigma}f\left(\theta;\phi_t\right)=-\infty$, the set of θs that pass the likelihood ratio test is a subinterval of $\mathcal{M}_0$, with bounds that satisfy (8) with equality. Using $\phi_t=\phi_k\left(\bar{r}_t\right)$, the posterior set can be written as an interval $\left[\underline{\theta}_k\left(\bar{r}_t\right),\bar{\theta}_k\left(\bar{r}_t\right)\right]$.

To see why both bounds are strictly increasing in $\bar{r}_t$, suppose that $\tilde{\theta}$ satisfies (8) with equality. Apply the implicit function theorem to obtain

$$\left.\frac{d\theta}{d\phi}\right|_{\theta=\tilde{\theta}}=\frac{-f_2\left(\tilde{\theta};\phi_t\right)+f_2\left(\theta^*\left(\phi_t\right);\phi_t\right)}{f_1\left(\tilde{\theta};\phi_t\right)},$$

where f_i is the derivative of f with respect to its ith argument. Since f is strictly concave, $f_1\left(\tilde{\theta};\phi_t\right)>0$ if $\tilde{\theta}<\theta^*\left(\phi_t\right)$ and $f_1\left(\tilde{\theta};\phi_t\right)<0$ if $\tilde{\theta}>\theta^*\left(\phi_t\right)$. In addition, it can be verified that $f_{21}\left(\tilde{\theta};\phi_t\right)>0$, so that $f_2\left(\tilde{\theta};\phi_t\right)-f_2\left(\theta^*\left(\phi_t\right);\phi_t\right)<0$ if $\tilde{\theta}<\theta^*\left(\phi_t\right)$, but $f_2\left(\tilde{\theta};\phi_t\right)-f_2\left(\theta^*\left(\phi_t\right);\phi_t\right)>0$ if $\tilde{\theta}>\theta^*$. Taken together, these facts imply that $d\theta/d\phi>0$. Since $\phi_k'\left(\bar{r}_t\right)>0$, it follows that the bounds are strictly increasing in $\bar{r}_t$. ■

A simple formula for the posterior set is obtained by taking the limit as $k\to\infty$:

Proposition S2. *The limit of the posterior set is given by*

$$\left[\lim_{k\to\infty}\underline{\theta}_k\left(\bar{r}_t\right),\lim_{k\to\infty}\bar{\theta}_k\left(\bar{r}_t\right)\right]=\left[\bar{r}_t-t^{-\frac{1}{2}}\sigma b_\alpha,\bar{r}_t+t^{-\frac{1}{2}}\sigma b_\alpha\right], \tag{9}$$

where $b_\alpha=\sqrt{-2\log\alpha}$.

Proof. Substitute for θ^* and $\phi_t=\phi_k\left(\bar{r}_t\right)$ defined in (2) to obtain

$$\begin{aligned}
&\left(1+\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\log\left(1+\frac{\theta+\bar{\lambda}}{\sigma\sqrt{k}}\right)+\left(1-\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\log\left(1-\frac{\theta-\bar{\lambda}}{\sigma\sqrt{k}}\right)\\
\geq\;&\left(1+\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\log\left(1+\frac{\bar{r}_t+\bar{\lambda}}{\sigma\sqrt{k}}+\frac{\bar{r}_t\bar{\lambda}}{\sigma^2k}\right)\\
&+\left(1-\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\log\left(1-\frac{\bar{r}_t-\bar{\lambda}}{\sigma\sqrt{k}}-\frac{\bar{r}_t\bar{\lambda}}{\sigma^2k}\right)-\frac{2}{tk}\log\left(\alpha\right)
\end{aligned}$$

For each term of the form $\log\left(1+x\right)$, perform a Taylor expansion around $x=0$:

$$\begin{aligned}
&\left(1+\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\left(\frac{\theta+\bar{\lambda}}{\sigma\sqrt{k}}-\frac{1}{2}\frac{(\theta+\bar{\lambda})^2}{\sigma^2k}+O\left(k^{-\frac{3}{2}}\right)\right)\\
&+\left(1-\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\left(-\frac{\theta-\bar{\lambda}}{\sigma\sqrt{k}}-\frac{1}{2}\frac{(\theta-\bar{\lambda})^2}{\sigma^2k}+O\left(k^{-\frac{3}{2}}\right)\right)\\
\geq\;&\left(1+\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\left(\frac{\bar{r}_t+\bar{\lambda}}{\sigma\sqrt{k}}-\frac{1}{2}\frac{(\bar{r}_t+\bar{\lambda})^2}{\sigma^2k}+\frac{\bar{r}_t\bar{\lambda}}{\sigma^2k}+O\left(k^{-\frac{3}{2}}\right)\right)\\
&+\left(1-\frac{\bar{r}_t}{\sigma\sqrt{k}}\right)\log\left(-\frac{\bar{r}_t-\bar{\lambda}}{\sigma\sqrt{k}}-\frac{1}{2}\frac{(\bar{r}_t-\bar{\lambda})^2}{\sigma^2k}-\frac{\bar{r}_t\bar{\lambda}}{\sigma^2k}+O\left(k^{-\frac{3}{2}}\right)\right)-\frac{2}{tk}\log\left(\alpha\right)
\end{aligned}$$

Multiply out to obtain

$$\begin{aligned}
&\frac{\bar{\lambda}}{\sigma\sqrt{k}}+\frac{2\bar{r}_t\theta}{\sigma^2k}-\frac{1}{2}\frac{(\theta+\bar{\lambda})^2}{\sigma^2k}-\frac{1}{2}\frac{(\theta-\bar{\lambda})^2}{\sigma^2k}+O\left(k^{-\frac{3}{2}}\right)\\
\geq\;&\frac{\bar{\lambda}}{\sigma\sqrt{k}}+\frac{2\bar{r}_t^2}{\sigma^2k}-\frac{1}{2}\frac{(\bar{r}_t+\bar{\lambda})^2}{\sigma^2k}-\frac{1}{2}\frac{(\bar{r}_t-\bar{\lambda})^2}{\sigma^2k}+O\left(k^{-\frac{3}{2}}\right)-\frac{2}{tk}\log\left(\alpha\right)
\end{aligned} \tag{10}$$

All the terms involving $\bar{\lambda}$ cancel from this inequality. Multiplying by σ^2k thus yields

$$\left(\bar{r}_t-\theta\right)^2\leq\sigma^2\frac{-2\log\alpha}{t}+O\left(k^{-\frac{1}{2}}\right), \tag{11}$$

which implies (9). ■

Finally, consider the set of one-step-ahead beliefs $\mathcal{P}_{t,k}\left(\bar{r}_t\right)$. Following (4), it contains all likelihoods of the type (5) for some $\theta\in\mathcal{M}_{t,k}^{\alpha}\left(s^t\right)$ and $|\lambda|<\bar{\lambda}$. It can thus be summarized by an interval of probabilities for the high state next period, denoted $\left[\underline{p}_k(\bar{r}_t), \bar{p}_k(\bar{r}_t)\right]$, with both bounds strictly increasing in the sample mean $\bar{r}_t$:

$$\begin{aligned} \underline{p}_k(\bar{r}_t) &= \frac{1}{2} + \frac{1}{2}\frac{\underline{\theta}_k(\bar{r}_t) - \bar{\lambda}}{\sigma\sqrt{k}} \\ \bar{p}_k(\bar{r}_t) &= \frac{1}{2} + \frac{1}{2}\frac{\bar{\theta}_k(\bar{r}_t) + \lambda}{\sigma\sqrt{k}}. \end{aligned} \tag{12}$$

To get an idea about the shrinkage of the interval of possible equity premia, consider the lowest and highest mean log returns per month. In the limit as $k \to \infty$,

$$\left[\lim_{k\to\infty}\left(\underline{\theta}_k(\bar{r}_t) - \bar{\lambda}\right), \lim_{k\to\infty}\left(\bar{\theta}_k(\bar{r}_t) + \bar{\lambda}\right)\right] = [\bar{r}_t - \bar{\lambda} - t^{-\frac{1}{2}}\sigma b_\alpha, \bar{r}_t + \bar{\lambda} + t^{-\frac{1}{2}}\sigma b_\alpha].$$

3 Optimal portfolio weights

It is helpful to begin with the maximization step in (1), given some arbitrary probability $0 < p < 1$ for the high state. The optimal weight on stocks is

$$w(p) = \frac{e^{r^f/k}\left(p\left(e^{\sigma/\sqrt{k}} - e^{r^f/k}\right) + (1-p)\left(e^{-\sigma/\sqrt{k}} - e^{r^f/k}\right)\right)}{\left(e^{\sigma/\sqrt{k}} - e^{r^f/k}\right)\left(e^{r^f/k} - e^{-\sigma/\sqrt{k}}\right)} \tag{13}$$

Taylor expansions of the exponential terms lead to

$$\begin{aligned} w(p) &= \left(1 + \frac{r^f}{k} + O\left(\frac{1}{k^2}\right)\right) \times \\ &\quad \frac{p\left(\frac{\sigma}{\sqrt{k}} + \frac{1}{2}\frac{\sigma^2}{k} - \frac{r^f}{k} + O\left(\frac{1}{k^{\frac{3}{2}}}\right)\right) + (1-p)\left(-\frac{\sigma}{\sqrt{k}} + \frac{1}{2}\frac{\sigma^2}{k} - \frac{r^f}{k} + O\left(\frac{1}{k^{\frac{3}{2}}}\right)\right)}{\left(\frac{\sigma}{\sqrt{k}} + \frac{1}{2}\frac{\sigma^2}{k} - \frac{r^f}{k} + O\left(\frac{1}{k^{\frac{3}{2}}}\right)\right)\left(\frac{r^f}{k} + \frac{\sigma}{\sqrt{k}} - \frac{1}{2}\frac{\sigma^2}{k} + O\left(\frac{1}{k^{\frac{3}{2}}}\right)\right)} && (14) \\ &= \frac{(2p-1)\sigma/\sqrt{k} + \frac{1}{2}\sigma^2/k - r^f/k + O\left(k^{-\frac{3}{2}}\right)}{\sigma^2/k + O\left(k^{-\frac{3}{2}}\right)}. && (15) \end{aligned}$$

We are now ready to derive optimal portfolio weights. Begin with the Bayesian case:

Proposition S3. *The optimal Bayesian portfolio weight for horizon $T > 0$ is*

$$\lim_{k\to\infty} \omega^*_{t,k,T}(\bar{r}_t) = \frac{\bar{r}_t + \frac{1}{2}\sigma^2 - r^f}{\sigma^2} =: \omega_t^{bay}.$$

Proof. Consider the optimal portfolio weight for the Bayesian investor with horizon $T > 0$. There is no minimization step in (1); as a result, the objective function on the left hand side is the sum of (i) the objective for a myopic investor and (ii) continuation utility, which depends on the investment horizon T, but is independent of the optimal portfolio weight. The optimal weight is therefore independent of the horizon T. The optimal weight is now obtained by evaluating $w(p)$ at $p = \phi_t = \phi_k(\bar{r}_t)$:

$$\begin{aligned} \omega^*_{t,k,T}(\bar{r}_t) &= w(\phi_k(\bar{r}_t)) \\ &= \frac{(2\phi_k(\bar{r}_t) - 1)\sigma/\sqrt{k} + \frac{1}{2}\sigma^2/k - r^f/k + O\left(k^{-\frac{3}{2}}\right)}{\sigma^2/k + O\left(k^{-\frac{3}{2}}\right)} \\ &= \frac{\bar{r}_t/k + \frac{1}{2}\sigma^2/k - r^f/k + O\left(k^{-\frac{3}{2}}\right)}{\sigma^2/k + O\left(k^{-\frac{3}{2}}\right)}, \end{aligned}$$

where the third equality uses the definition of ϕ_k (recall (2)). Taking the limit as $k \to \infty$ delivers the Bayesian solution.■

Consider next the problem with a nondegenerate set of one-step-ahead conditionals, but assume $T = 1/k$. In (12) above, the belief set $\mathcal{P}_{t,k}(\bar{r}_t)$ was defined in terms of an interval $\left[\underline{p}_k(\bar{r}_t), \bar{p}_k(\bar{r}_t)\right]$ for the probability of the high state, with bounds bounds strictly increasing in $\bar{r}_t$.

Proposition S4. *The optimal portfolio weight of the myopic ambiguity averse investor ($T = 1/k$) is*

$$\begin{aligned} \lim_{k\to\infty} \omega^*_{t,k,1/k}(\bar{r}_t) &= \sigma^{-2}\max\left\{\bar{r}_t + \frac{1}{2}\sigma^2 - r^f - \left(\bar{\lambda} + t^{-\frac{1}{2}}\sigma b_\alpha\right), 0\right\} \\ &\quad + \sigma^{-2}\min\left\{\bar{r}_t + \frac{1}{2}\sigma^2 - r^f + \bar{\lambda} + t^{-\frac{1}{2}}\sigma b_\alpha, 0\right\} \\ &= \max\left\{\omega^{bay}_t - \sigma^{-2}\left(\bar{\lambda} + t^{-\frac{1}{2}}\sigma b_\alpha\right), 0\right\} \\ &\quad + \min\left\{\omega^{bay}_t + \sigma^{-2}\left(\bar{\lambda} + t^{-\frac{1}{2}}\sigma b_\alpha\right), 0\right\}. \end{aligned}$$

Proof. For given p, the optimal weight is $w(p)$. To solve the minimization step, substituting $w(p)$ back into the objective (1). We now need to find $p \in \left[\underline{p}_k(\bar{r}_t), \bar{p}_k(\bar{r}_t)\right]$ to minimize

$$\begin{aligned} g(p) &= p\log\left(e^{r^f/k} + w(p)\left(e^{\sigma/\sqrt{k}} - e^{r^f/k}\right)\right) \\ &\quad + (1-p)\log\left(e^{r^f/k} + w(p)\left(e^{-\sigma/\sqrt{k}} - e^{r^f/k}\right)\right) \\ &= r^f/k + p\log p + (1-p)\log(1-p) + \log\left(e^{\sigma/\sqrt{k}} - e^{-\sigma/\sqrt{k}}\right) \\ &\quad - p\log\left(e^{r^f/k} - e^{-\sigma/\sqrt{k}}\right) - (1-p)\log\left(e^{\sigma/\sqrt{k}} - e^{r^f/k}\right). \end{aligned}$$

The function g is strictly convex on $(0,1)$ and achieves a minimum at

$$\hat{p}_k = \frac{e^{r^f/k} - e^{-\sigma/\sqrt{k}}}{e^{\sigma/\sqrt{k}} - e^{-\sigma/\sqrt{k}}}. \tag{16}$$

The minimizer $\hat{p}_k$ is in the unit interval because we have assumed that $r^f < \sigma$. It is precisely the probability at which the one-step-ahead conditional equity premium $E\left[R\left(s_{t+1/k}\right)\right] - R^f$ is equal to zero. The conditional premium appears also as the bracketed term in the numerator of (13), so that $w\left(\hat{p}_k\right) = 0$. It follows that the solution to the minimization step is

$$p_k^*\left(\bar{r}_t\right) = \begin{cases} \bar{p}_k\left(\bar{r}_t\right) & \text{if} \quad \hat{p}_k > \bar{p}_k\left(\bar{r}_t\right) \\ \hat{p}_k & \text{if} \quad \hat{p}_k \in \left[\underline{p}_k\left(\bar{r}_t\right), \bar{p}_k\left(\bar{r}_t\right)\right] \\ \underline{p}_k\left(\bar{r}_t\right) & \text{if} \quad \hat{p}_k > \bar{p}_k\left(\bar{r}_t\right). \end{cases}$$

Substituting into (13) and using $w\left(\hat{p}_k\right) = 0$, we can express the optimal portfolio weight as a function of the sample mean:

$$\omega^*_{t,k,1/k}\left(\bar{r}_t\right) = \begin{cases} w\left(\bar{p}_k\left(\bar{r}_t\right)\right) & \text{if} \quad \hat{p}_k > \bar{p}_k\left(\bar{r}_t\right) \\ 0 & \text{if} \quad \hat{p}_k \in \left[\underline{p}_k\left(\bar{r}_t\right), \bar{p}_k\left(\bar{r}_t\right)\right] \\ w\left(\underline{p}_k\left(\bar{r}_t\right)\right) & \text{if} \quad \hat{p}_k < \underline{p}_k\left(\bar{r}_t\right). \end{cases}$$

We now compute $\lim_{k\to\infty} \omega^*_{t,k,1/k}\left(\bar{r}\right)$. Since the functions $\underline{p}_k$ and $\bar{p}_k$ are strictly increasing, the nonparticipation region of the state space can be represented by an interval of sample means $\left[\bar{r}_{lo}(k), \bar{r}_{up}(k)\right]$. If $\bar{r}_t > \bar{r}_{up}(k)$ the evidence about the equity premium is so positive that investment in stocks is positive even under the lowest probability for the high state. The upper bound $\bar{r}_{up}(k)$ is the unique solution to $\underline{p}_k\left(\bar{r}_t\right) = \hat{p}_k$ or, equivalently,

$$\left(2\underline{p}_k\left(\bar{r}_t\right) - 1\right)\sigma/\sqrt{k} = \left(2\hat{p}_k - 1\right)\sigma/\sqrt{k} \tag{17}$$

Use (12) and 11) above to rewrite the left hand side as:

$$\begin{aligned} \left(2\underline{p}_k\left(\bar{r}_t\right) - 1\right)\sigma/\sqrt{k} &= \frac{1}{k}\left(\bar{r}_t - \bar{\lambda} - \sqrt{\sigma^2\frac{-2\log\alpha}{t} + O\left(k^{-\frac{1}{2}}\right)}\right) \\ &= \frac{1}{k}\left(\bar{r}_t - \sigma b_\alpha/\sqrt{t} - \bar{\lambda}\right) + O\left(k^{-\frac{5}{4}}\right). \end{aligned} \tag{18}$$

Substitute (18) and (16) into (17) to obtain

$$\frac{1}{k}\left(\bar{r}_t - \sigma b_\alpha/\sqrt{t} - \bar{\lambda}\right) + O\left(k^{-\frac{5}{4}}\right) = \frac{2e^{r^f/k} - e^{\sigma/\sqrt{k}} - e^{-\sigma/\sqrt{k}}}{e^{\sigma/\sqrt{k}} - e^{-\sigma/\sqrt{k}}}\frac{\sigma}{\sqrt{k}}.$$

Taylor expansions of the exponential terms on the right side around 1 lead to

$$\frac{1}{k}\left(\bar{r}_t - \sigma b_\alpha/\sqrt{t} - \bar{\lambda}\right) + O\left(k^{-\frac{5}{4}}\right) = \frac{2r^f/k - \sigma^2/k}{2\sigma/\sqrt{k}}\frac{\sigma}{\sqrt{k}},$$

so that the upper bound of the nonparticipation region can be written as

$$\bar{r}_{up}(k) = \sigma b_\alpha/\sqrt{t} + \bar{\lambda} + r^f - \frac{1}{2}\sigma^2 + O\left(k^{-\frac{1}{4}}\right). \tag{19}$$

The lower bound $\bar{r}_{lo}(k)$ is the unique solution to $\bar{p}_k(\bar{r}_t) = \hat{p}_k$. By an argument similar to the one above, it can be written as

$$\bar{r}_{lo}(k) = -\sigma b_\alpha/\sqrt{t} - \bar{\lambda} + r^f - \frac{1}{2}\sigma^2 + O\left(k^{-\frac{1}{4}}\right). \tag{20}$$

Consider next the optimal weight in the case $\underline{p}_k(\bar{r}_t) > \hat{p}_k$. Evaluating $w(p)$ from (15) at $p = \underline{p}_k(\bar{r}_t)$ and replacing the first term in the numerator using (18) yields

$$w\left(\underline{p}_k(\bar{r}_t)\right) = \frac{\bar{r}_t - \sigma b_\alpha/\sqrt{t} - \bar{\lambda} + \frac{1}{2}\sigma^2 - r^f + O\left(k^{-\frac{1}{4}}\right)}{\sigma^2 + O\left(k^{-\frac{1}{2}}\right)} = \frac{\bar{r}_t - \bar{r}_{up}(k) + O\left(k^{-\frac{1}{4}}\right)}{\sigma^2 + O\left(k^{-\frac{1}{2}}\right)}. \tag{21}$$

The case $\hat{p}_k > \bar{p}_k(\bar{r}_t)$ is analogous, but with $\bar{r}_{up}(k)$ replaced by $\bar{r}_{lo}(k)$.

Combine the bounds (19)-(20) and the formulas for the optimal weight to deduce that the sequence of functions $\left\{\omega^*_{t,k,1/k}(\bar{r}_t)\right\}$ converges pointwise to

$$\lim_{k\to\infty}\omega^*_{t,k,1/k}(\bar{r}) = \begin{cases} \frac{\bar{r}_t + \sigma b_\alpha/\sqrt{t} + \bar{\lambda}\frac{1}{2}\sigma^2 - r}{\sigma^2} & \text{if} \quad \bar{r}_t < r^f - \frac{1}{2}\sigma^2 - \sigma b_\alpha/\sqrt{t} - \bar{\lambda} \\ 0 & \text{if} \quad |\bar{r}_t + \frac{1}{2}\sigma^2 - r^f| < \sigma b_\alpha/\sqrt{t} + \bar{\lambda} \\ \frac{\bar{r}_t - \sigma b_\alpha/\sqrt{t} - \bar{\lambda} + \frac{1}{2}\sigma^2 - r}{\sigma^2} & \text{if} \quad \bar{r}_t > r^f - \frac{1}{2}\sigma^2 + \sigma b_\alpha/\sqrt{t} + \bar{\lambda}. \end{cases}$$

■