Internet Appendix for “Reverse Survivorship Bias”

Juhani T. Linnainmaa*  

This Internet Appendix solves and calibrates a portfolio choice model with endogenous fund attrition to illustrate the economic significance of reverse survivorship bias.

I. Bayesian Portfolio Choice Model with Endogenous Fund Attrition

In this portfolio choice model a single mutual fund investor can invest in a risk-free asset, the market portfolio, and an actively managed mutual fund. The key features of the model are, first, the investor is uncertain about the fund’s alpha and, second, the fund’s survival depends on the investor’s continuing investment. The investor can always abandon the existing fund and, if the investor does so, the old fund disappears and the investor gets to draw a new fund with an unknown alpha. This switch “resets” the investor’s prior belief about the fund’s alpha.1

---

*Citation format: Linnainmaa, Juhani T., 2013, Internet Appendix for “Reverse Survivorship Bias,” Journal of Finance, DOI: 10.1111/jofi.12030. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the author of the article.
I use a different model in the main part of the paper for two reasons. First, the main model is more parsimonious because it directly specifies a (free) mapping from the posterior distributions to the exit probability. The portfolio choice model described here generates a qualitatively similar mapping, but it requires that one specify (or estimate from the data) many parameters that are not of direct interest, such as the properties of market returns, the risk-free rate, subjective discount rates, and so forth. Second, the model described here makes restrictive assumptions about the mutual fund industry and equilibrium: there is a single investor and the fund’s survival is contingent on the investor’s continuing investment; this is a partial equilibrium model with exogenously specified return processes; and the investor can invest in, and learn about, only one mutual fund at a time. By contrast, the free hazard function approach in the structural model sidesteps these concerns because as long as the functional form is sufficiently flexible, it can approximate the economic mechanism that drives fund disappearance.

A. Assumptions

I assume that an infinite-horizon investor maximizes log-utility over consumption,

\[
E \left[ \sum_{t=0}^{\infty} \beta^t \log c_t \right],
\]

where \( \beta \) is the investor’s discount rate. The wealth dynamics are given by

\[
W_{t+1} = (W_t - c_t)(1 + rf + \theta_{m,t}(\tilde{r}_{m,t+1} - rf) + \theta_{z,t}(\tilde{z}_{t+1} - rf)),
\]

where \( \tilde{r} \) and \( \tilde{z} \) are the return on market and the risk-free asset, respectively, and \( \theta \) are parameters.
where $r_f$ is the risk-free rate, $\theta_{m,t}$ is the proportion of wealth invested in the market portfolio, $\tilde{r}_{m,t}$ is the return on the market portfolio, $\theta_{z,t}$ is the proportion of wealth invested in the mutual fund, and $\tilde{z}_t$ is the return on the mutual fund. The investor cannot short the fund, so $\theta_{z,t} \geq 0$. The investor knows all the parameters except for the fund’s alpha. The date-$t$ return on the market portfolio is

$$\tilde{r}_{m,t} = \mu_m + \tilde{\varepsilon}_{m,t}, \quad (IA.3)$$

where $\mu_m$ is the expected market return and $\tilde{\varepsilon}_{m,t}$ is an i.i.d. draw from a left-truncated normal distribution with truncation at $x = -1$. The underlying untruncated distribution has a mean zero and variance $\sigma^2_m$. I assume that the fund’s beta is one so that its return is

$$\tilde{z}_t = \alpha + \tilde{r}_{m,t} + \tilde{\varepsilon}_{z,t}, \quad (IA.4)$$

where $\alpha$ is not known to the investor and $\tilde{\varepsilon}_{z,t}$ is also an i.i.d. draw from a left-truncated normal distribution with truncation at $x = -1$. The underlying untruncated distribution has mean zero and variance $\sigma^2_z$. These truncation assumptions give both the market portfolio and the mutual fund limited liability.

The investor updates his beliefs after each date as follows. First, after observing both fund and market returns, the investor backs out the signal $s_t \equiv \tilde{z}_t - \tilde{r}_{m,t}$ about the fund’s alpha. The investor then reverses the truncation by computing the signal realization $s'_t$ by mapping the truncated distribution to an untruncated distribution. The investor’s date-$t$ prior belief about the mean of this untruncated distribution is normal with mean $m_t$ and variance $v_t$. The investor knows that funds’
alphas are drawn from a normal distribution with mean \( \mu_\alpha \) and standard deviation \( \sigma_\alpha \), and so this is also the investor’s prior distribution about a new fund’s alpha. Letting \( m_0 \) and \( v_0 \) denote the mean and variance of the investor’s prior distribution, the belief dynamics are then given by\(^2\)

\[
\begin{align*}
    m_{t+1} &= (1 - w_t)m_t + w_t s_{t+1}, \\
    v_{t+1} &= \sigma_\alpha^2 w_t,
\end{align*}
\]

where \( w_t = \frac{v_t}{v_t + \sigma_\alpha^2} \). Each period, the investor can abandon the current fund (thus causing the fund to disappear) and, if the investor so chooses, draw a new fund with an unknown alpha. I assume that abandoning a fund and drawing a new fund costs the investor \( \kappa \% \) of wealth. This cost may represent components such as front-end loads, back-end loads, and the abnormal transaction costs incurred when assets are sold in the event of fund termination.

\[\textit{B. Solution}\]

The investor’s indirect utility is a function of three state variables: current wealth \( W_t \), the mean of the prior distribution \( m_t \), and the variance of the prior distribution \( v_t \). Because the investor maximizes log-utility, the wealth and belief terms are additive in the indirect utility function. I conjecture that

\[
V(W_t, m_t, v_t) = A + B \log W_t + g(m_t, v_t),
\]

(IA.7)
where $A$ and $B$ are constants and $g(\cdot)$ is a function of the investor’s date-$t$ beliefs. The investor’s optimization problem with this conjecture becomes

$$
V(W_t, m_t, v_t) = \max_{c_t, \theta_{m,t}, \theta_{z,t} \geq 0} \mathbb{E}\left\{ \log c_t + \beta V\left((W_t - c_t)(1 + \gamma_f + \theta_{m,t}(\tilde{r}_{m,t+1} - \gamma_f) + \theta_{z,t}(\tilde{z}_{t+1} - \gamma_f)), \tilde{m}_{t+1}, v_{t+1}\right) \right\}
$$

$$
= \beta A + \max_{c_t} \left\{ \log c_t + \beta B \log(W_t - c_t) \right\}
$$

$$
+ \max_{\theta_{m,t}, \theta_{z,t} \geq 0} \left\{ \beta B \mathbb{E}\left[ \log \left(1 + \gamma_f + \theta_{m,t}(\tilde{r}_{m,t+1} - \gamma_f)\right)\right] + \beta g(m_t, v_t), \right. \\
\beta B \mathbb{E}\left[ \log \left(1 + \gamma_f + \theta_{m,t}(\tilde{r}_{m,t+1} - \gamma_f) + \theta_{z,t}(\tilde{z}_{t+1} - \gamma_f)\right)\right] + \beta g(\tilde{m}_{t+1}, v_{t+1}), \\
\beta B \mathbb{E}\left[ \log \left(1 + \gamma_f + \theta_{m,t}(\tilde{r}_{m,t+1} - \gamma_f) + \theta_{z,t}(\tilde{z}_{t+1} - \gamma_f) - \kappa\right)\right] + \beta g(\tilde{m}_1, v_1) \right\}.
$$

(IA.8)

The last three lines of equation (IA.8) account for the fact that an investor can make a choice that disrupts the natural evolution of beliefs from $(m_t, v_t)$ to $(\tilde{m}_{t+1}, v_{t+1})$. The first possible choice is that the investor invests in neither the current nor a new mutual fund. The investor withdraws his money from the current fund, the fund disappears, and the evolution of beliefs stops. If $\kappa$ is low and if the investor has ever invested in a mutual fund, the option to draw a new fund must dominate this “permanent-quitting” choice. The second possibility is that the investor remains with the current fund. The investor updates his beliefs about $\alpha$ to $(\tilde{m}_{t+1}, v_{t+1})$ based on the fund’s return. The third possibility is that the investor abandons the current fund and draws a new one. The investor’s prior distribution about the new fund is the same as the original prior distribution, $(m_0, v_0)$, and so the investor’s beliefs “restart.” The investor updates to $(\tilde{m}_1, v_1)$ based on the new fund’s first return.
realization. The investor pays $\kappa$ to exercise this option.

The optimal consumption from the first-order condition for this problem is $c_t^* = \frac{1}{1+\beta} W_t$. Constants $A$ and $B$ can be solved by inserting the optimal consumption back into equation (IA.8) and matching the coefficients against the conjecture in equation (IA.7). The value function then becomes

$$V(W_t, m_t, v_t) = \frac{\beta \log \beta + (1 - \beta) \log (1 - \beta)}{(1 - \beta)^2} + \frac{1}{1 - \beta} \log W_t + g(m_t, v_t),$$

where $g(m_t, v_t)$ solves the following functional equation:

$$g(m_t, v_t) = \beta \max \left\{ \frac{1}{\beta(1-\beta)} \max_{\theta_{m,t}} \left\{ E \left[ \log \left( 1 + r_f + \theta_{m,t} (\tilde{r}_{m,t+1} - r_f) \right) \right] \right\}, \right. $$

$$\left. \frac{1}{1-\beta} \max_{\theta_{m,t}, \theta_{z,t}, \theta_{z,t} \geq 0} \left\{ E \left[ \log \left( 1 + r_f + \theta_{m,t} (\tilde{r}_{m,t+1} - r_f) + \theta_{z,t} (\tilde{z}_{t+1} - r_f) \right) + E \left[ g(m_{t+1}, v_{t+1}) \right] \right] \right\} + E \left[ g(m_1, v_1) \right] \right\}.$$

Equation (IA.9) verifies the conjecture about the form of the value function. The investor’s optimal investment decisions are determined by $g(m_t, v_t)$. The first argument of the outer-maximization problem is simplified here by noting that if an investor ever abandons a mutual fund and does not pick a new one, the investor’s beliefs must become “stuck.” The problem then becomes a standard log-utility investment problem with the solution shown here. Function $g(m_t, v_t)$ can also be solved, up to a static portfolio choice problem, in those belief states in which the investor has resolved all
uncertainty about a fund’s alpha. The value of \( g(m_t, v_t) \) at such a \( v_t = 0 \) boundary point is

\[
g(m_t, 0) = \frac{\beta}{(1 - \beta)^2} \max_{\theta_{m,t}, \theta_{z,t} \geq 0} \left\{ E \left[ \log \left( 1 + r_f + \theta_{m,t}(\tilde{r}_{m,t+1} - r_f) + \theta_{z,t}(\tilde{z}_{t+1} - r_f) \right) \right] \right\}. \tag{IA.11}
\]

Equation (IA.10) shows that an investor’s decision to remain with the current fund or switch to a new one depends crucially on the evolution of beliefs. The path of beliefs in turn depends on the level of uncertainty about the fund’s alpha. If the prior distribution about a fresh mutual fund’s alpha “dominates” the investor’s beliefs the current fund’s alpha (net of the strike price \( \kappa \)), the investor switches to a new fund. The log-utility assumption, which shuts down the intertemporal hedging demand component, cleanly isolates the value of this abandonment option.

I solve the problem in three steps. First, I create a mean-variance belief grid with 1,000 points for the posterior mean \( m_t \) and 1,200 points for the posterior variance \( v_t \).\(^3\) Second, I solve the static portfolio choice problems in equation (IA.10) for each grid point in the \( m_t \) dimension of the grid. Third, I generate initial guesses about the value of \( g(m_t, v_t) \) for each grid point and then start iterating over the grid, sweeping recursively from \( v_T \) toward \( v_0 \) at each iteration. I compute the value of \( g(m_t, v_t) \) from equation (IA.10) at each grid point given the current guesses of the value of this function at the next-period grid points. I iterate over the grid until the value of \( g(m_t, v_t) \) has converged at each grid point. The solution to the investor’s problem requires value function iterations because of the abandonment option. An investor who abandons the fund transitions back to \( g(m_0, v_0) \) and this value in turn depends on the investor’s optimal choices in the future, leading to a fixed-point problem.
C. Calibration

I set the model parameters to the following values. I set the mean and standard deviation of the true distribution of annual (net) alphas to $\mu_\alpha = -0.5\%$ and $\sigma_\alpha = 1.25\%$, the standard deviation of the idiosyncratic fund return component $\tilde{\varepsilon}_{z,t}$ to $8\%$ per year, the risk-free rate to $r_f = 5\%$, the expected return and standard deviation of the market portfolio to $\mu_m = 10\%$ and $\sigma_m = 30\%$, and the investor’s discount rate ($\beta$) to 0.9. I choose the cost of switching a fund, $\kappa$, so that the 10-year survival rate in the model is close to the empirical survival rate in the mutual fund data. A value of $\kappa = 0.064$ gives a survival rate of $69.3\%$, which is close to the $68\%$ survival rate reported in Table I in the paper. I note that survival rate increases monotonically in $\kappa$ because it only decreases the one-period expected return in equation (IA.10). I solve the model with the assumption that each period in the model represents one month.

[Table IA.I here]

Table IA.I Panel A reports average alphas for mutual funds that either survive or do not survive through year $T$. The underlying computations here are the same as those used in Table I, which is based on the CRSP sample. The alpha estimates in the model-based simulations exhibit the same upward-sloping pattern observed in the actual sample. The average alpha is negative for both surviving and dead funds in year one ($-0.5\%$ and $-35.5\%$, respectively), but both of these averages increase over time. For example, the average alpha estimate is $1.98\%$ per year for funds that survive through the tenth year but $-4.12\%$ for those that do not.

[Figure IA.1 here]
Figure IA.1 plots the critical posterior mean threshold that determines if and when an investor abandons a fund. The threshold is initially very low, approximately \(-10\%\) per year. This low threshold value suggests that the investor abandons a fund early only if the realized return is extremely low. The critical alpha threshold increases smoothly over time. After 10 years, the critical value is approximately \(-3\%\) per year, and after 25 years, the threshold is \(-1\%\) per year. The increase in the critical threshold leads to the increasing pattern in average alphas seen in Panel A.

Table IA.I Panel B measures the reverse survivorship bias within this portfolio choice model. The first column (“biased”) reports the distribution of observed alphas from the actual model. The second column (“unbiased”) changes the model so that every time a mutual fund disappears in the real model, a randomly chosen fund disappears in this alternative model. The idea here is the same as that of the randomized-exit simulations in Section 2.2 of the paper. The third column reports the true (noiseless) alpha distribution. I simulate at most 10 years of monthly return data from the model for 100,000 mutual funds to construct these distributions. Although the true mean of the alpha distribution is \(-0.5\%\) in the calibrated model, the observed distribution has a mean of \(-1.13\%\) because of the disappearance of poorly performing funds. Thus, in this rough calibration, the mean effect of reverse survivorship bias is 63 basis points per year.

The percentiles for the observed and true alpha distributions show the distortion in the shape of the alpha distribution. The worst 5% of funds in the distribution of estimated alphas have alphas of \(-11.62\%\) or lower. When funds disappear randomly, as they do in the second column, this percentile of the distribution is \(-9.09\%\). The biased and unbiased values for the 95th percentile of the distribution of estimated alphas are 7.76% and 8.08%, respectively. This calibration suggests that
reverse survivorship bias significantly distorts the shape of the observed alpha distribution within a model that uses a learning mechanism to connect fund survival to performance.
Notes

1Dangl, Zu, and Zechner (2008) study a model in which investors learn about mutual fund managers’ abilities and the management company can replace the manager. However, whereas they focus on the model’s implications on the relations between fund size, portfolio risk, and the fund manager’s tenure, I examine how the disappearance of poorly performing funds biases performance estimates.

The trick I use here to keep the posterior distribution closed under updating is more transparent in an alternative setup with log-normally distributed returns. Instead of working with a log-normal likelihood function, an investor could be assumed to be uncertain about the mean of the normal variate $\tilde{x}$ in $\tilde{r} \equiv e^{\tilde{x}} - 1$. The investor would then back out from each return observation the normal variate realization, $\log(1 + \tilde{r})$, and use this signal to update beliefs about the mean of $\tilde{x}$. I do not use this log-normal assumption because of its downside that the variance of $\tilde{r}$ also changes as the mean of $\tilde{x}$ changes. By contrast, the truncation of the normal distribution at $x = -1$ is an innocuous return-distribution twist because, for reasonable return volatilities, the amount of truncated mass is effectively zero. See Johnson, Kotz, and Balakrishnan (1994) for details on truncated normal distributions.

3I choose the grid for $v_t$ to match the deterministic evolution of the posterior variance in equation (IA.6). The last node thus corresponds to the variance of beliefs after having invested in the same fund for 100 years. (I calibrate the model so that one period corresponds to one month.) I assume that the variance of beliefs drops to zero after this date. I create the grid for $m_t$ so it covers 99.99% of the true population distribution of $\alpha$. 

11
REFERENCES


Figure IA.1. Critical posterior mean boundary for alpha (in percentage points per year) in a Bayesian portfolio choice model with endogenous fund attrition. This figure shows the critical level for the posterior mean for alpha in a Bayesian portfolio choice model below which a fund shuts down. In this model, an infinitely lived representative mutual fund investor with log-utility can invest in a risk-free asset, the market portfolio, and an actively managed mutual fund. The investor is uncertain about the fund’s alpha and updates beliefs using Bayes’ rule. The investor can abandon the existing fund by paying a cost of $\kappa$ and then draw a new fund to restart the problem. The mean of the true alpha distribution ($\mu_\alpha$) is $-0.5\%$ and its standard deviation ($\sigma_\alpha$) is $1.25\%$; the expected return of the market portfolio ($\mu_m$) is $10\%$ and its standard deviation ($\sigma_m$) is $30\%$; the risk-free rate ($r_f$) is $5\%$; and the investor’s discount rate ($\beta$) is $0.9$. (These are annualized parameter values. Each period in the model is one month long.) The standard deviation of the idiosyncratic return component is $8\%$. The cost of switching a fund is $\kappa = 0.064$ of wealth.
Table IA.I

Calibration Results for a Bayesian Portfolio Choice Model with Endogenous Fund Attrition

This table reports the calibration results for a Bayesian portfolio choice model. In this model an infinitely lived representative mutual fund investor with log-utility can invest in a risk-free asset, the market portfolio, and an actively managed mutual fund. The investor is uncertain about the fund’s alpha but updates beliefs using Bayes’ rule. The investor can abandon the existing fund by paying a cost of $\kappa$ and then draw a new fund to restart the problem. The mean of the true alpha distribution ($\mu_\alpha$) is −0.5% and its standard deviation ($\sigma_\alpha$) is 1.25%; the expected return of the market portfolio ($\mu_m$) is 10% and its standard deviation ($\sigma_m$) is 30%; the risk-free rate ($r_f$) is 5%; and the investor’s discount rate ($\beta$) is 0.9. (These are annualized parameter values. Each period in the model is one month long.) The standard deviation of the idiosyncratic return component is 8%. The cost of switching a fund is $\kappa = 0.064$ of wealth. Panel A reports average alphas conditional on fund survival. Panel B simulates from the model and reports the observed distribution in the correct model (“observed distribution”), the observed distribution in an alternative model in which funds disappear randomly (“randomized-exit distribution”), and the actual alpha distribution (“true alpha distribution”). This table is based on simulating 100,000 funds through the model with each fund generating at most 10 years of monthly returns.
Panel A: Mutual Fund Alpha Estimates Conditional on Survival

<table>
<thead>
<tr>
<th>Survive</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>−0.50</td>
<td>−0.33</td>
<td>0.05</td>
<td>0.46</td>
<td>0.84</td>
<td>1.14</td>
<td>1.41</td>
<td>1.63</td>
<td>1.82</td>
<td>1.98</td>
</tr>
<tr>
<td>No</td>
<td>−35.54</td>
<td>−21.70</td>
<td>−14.68</td>
<td>−10.88</td>
<td>−8.61</td>
<td>−7.10</td>
<td>−6.03</td>
<td>−5.24</td>
<td>−4.61</td>
<td>−4.12</td>
</tr>
</tbody>
</table>

Panel B: Observed versus True Alpha Distributions

<table>
<thead>
<tr>
<th></th>
<th>Observed Distribution</th>
<th>Randomized-Exit Distribution</th>
<th>True Alpha Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>−1.131</td>
<td>−0.502</td>
<td>−0.500</td>
</tr>
<tr>
<td>Percentiles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1%</td>
<td>−17.908</td>
<td>−12.694</td>
<td>−3.408</td>
</tr>
<tr>
<td>5%</td>
<td>−11.617</td>
<td>−9.087</td>
<td>−2.556</td>
</tr>
<tr>
<td>25%</td>
<td>−4.857</td>
<td>−3.993</td>
<td>−1.343</td>
</tr>
<tr>
<td>50%</td>
<td>−0.524</td>
<td>−0.505</td>
<td>−0.500</td>
</tr>
<tr>
<td>75%</td>
<td>2.862</td>
<td>2.986</td>
<td>0.343</td>
</tr>
<tr>
<td>95%</td>
<td>7.758</td>
<td>8.079</td>
<td>1.556</td>
</tr>
<tr>
<td>99%</td>
<td>11.174</td>
<td>11.784</td>
<td>2.408</td>
</tr>
</tbody>
</table>