

**The Role of Learning in
Dynamic Portfolio Decisions**

M.J. Brennan*

January 1997

*Irwin and Goldyne Hearsh Professor of Banking and Finance, University of California, Los Angeles, and Professor of Finance, London Business School. I am grateful to Eduardo Schwartz for comments on an earlier version of this paper.

Abstract

The Role of Learning in Dynamic Portfolio Decisions

This paper analyzes the effect of uncertainty about the mean return on the risky asset on the portfolio decisions of an investor who has a long investment horizon. Building on the earlier work of Detemple (1986), Dothan and Feldman (1986), and Gennotte (1986), it is shown that the possibility of future learning about the mean return on the risky asset induces the investor to take a larger or smaller position in the risky asset than she would if there were no learning, the direction of the effect depending on whether the investor is more or less risk tolerant than the logarithmic investor whose portfolio decisions are unaffected by the possibility of future learning. Numerical calculations show that uncertainty about the mean return on the market portfolio has a significant effect on the portfolio decision of an investor with a 20 year horizon if her assessment of the market risk premium is based solely on the Ibbotson and Sinquefield (1995) data.

The Role of Learning in Dynamic Portfolio Decisions

A central stylized fact about capital markets is that the equity risk premium is of the order of 8.5%. This figure, which is derived from the average annual stock market returns over the period beginning in 1926, is often treated as a parameter, rather than as an estimate. Thus Mehra and Prescott (1985) take the premium (which they estimate at around 6%) as a datum which gives rise to the “equity premium puzzle”¹. However, there are dissenting views: Brown *et al.* (1995) have argued recently that use of the realized mean return on the equity market as an estimate of the expected return is likely to involve a survival bias which, they suggest, could be as high as 400 basis points per year. Blanchard (1993), using a forward looking econometric approach, concludes that the premium is around 2-3%, and Scott (1992) places it even lower at 1-2%. Finally, of course, use of the realized *historical* average excess return on stocks as an estimate of the *current* risk premium presumes that the risk premium is constant over time. Without taking a position on the magnitude of the equity market risk premium, this paper explores some implications of uncertainty about the risk premium for optimal dynamic portfolio strategies.

Portfolio theory was originally developed to analyze the problem of an investor who faced a single period horizon, and a known investment opportunity set. Subsequent extensions have generalized the theory to allow both for a multi-period horizon, and for incomplete information about the investment opportunity set. Thus, the basic single-period theory was extended by Hakansson (1970), Merton (1971), Samuelson (1969), Breeden (1979) and others to allow for a multi-period horizon in which investment opportunities might either be constant, time-dependent, or even stochastic

¹ An interesting aspect of this puzzle is that few economists who recognize the puzzle hold the highly levered stock positions that they would hold if they behaved according to their own criteria for rationality.

- in the latter case, these early authors assumed that the investment opportunity set is observable and that the parameters of the stochastic process governing its evolution are known by the investor.

Brennan, Schwartz and Lagnado (1996) and Brennan and Schwartz (1996) have implemented the Merton-Breeden continuous time approach to dynamic portfolio planning by specifying a small number of state variables that describe the state of the economy and estimating their dynamics, and the relation between them and the returns on risky assets. An important limitation of these implementations is that the optimal investment program is determined while ignoring the fact that the stochastic process parameters on which it depends are only estimates.

It was recognized early that investors do not in fact know the parameters of the probability distribution from which returns are drawn, and that the errors caused by the need to estimate the parameters give rise to an additional ‘estimation risk’ that must be taken into account in portfolio analysis. Extensive work on the role of estimation risk in a single period context has been undertaken by Klein and Bawa (1976) and Bawa, Brown and Klein (1979). Kandel and Stambaugh (1996) consider the effect of estimation errors in the return prediction equation on the optimal portfolio strategy in a multi period setting; however, their investor is assumed to have a single period horizon which, as we shall see, neglects an important aspect of the dynamic problem that is induced by learning. In a fully dynamic setting, Williams (1977), Detemple (1986), Dothan and Feldman (1986) and Gennotte (1986) characterize the portfolio problem of an investor who is unable to observe the true state of the economy but who knows the stochastic process governing the variables that describe the state of the economy. An important property of the problem analyzed by these authors, which arises from the assumption of continuous trading, is that the problem may be separated into a filtering problem in which the investor estimates the current values of the state variables, and an investment

problem which is solved by treating the estimated values of the state variables as state variables themselves, and then proceeding as in the classical analysis of Breeden and Merton. Barberis(1995) has also considered the role of estimation risk in a multi-period setting, but under the restrictive assumption that the investor cannot revise his portfolio as he learns more about the stochastic process governing returns. Perhaps the most general setting that could be analyzed is one in which the investment opportunity set depends on a set of unobservable state variables whose evolution is governed by a stochastic process whose parameters must be estimated. Such a problem involves estimation risk arising from the unknown nature of both the parameters and the state variables, and it also involves learning, since the investor will learn more about the parameters of the stochastic process and the current values of the state variables as time passes and more data become available. This general problem presents a formidable computational challenge, and in this paper we analyze the simpler dynamic portfolio problem of an investor who knows that the investment opportunity set is constant but is uncertain about the mean return on the risky asset². We assume that the investor learns about the mean as he observes returns over time³. We can think of this as a special case of the analysis of Detemple (1986) or Gennotte (1986), in which the dynamics of the unobservable state variable, the mean return, are eliminated. In reality, there may be additional sources of information or signals about the mean return on the risky asset; however, it is difficult to find simple empirical constructs that correspond to such signals and therefore we start with the simpler problem in which the only source of

² We also considered the problem in which the true mean of the return process follows an Ornstein-Uhlenbeck process, but were unable to satisfactorily identify the parameters using a Kalman filter, suggesting that this is not a good model of the behavior of the risk premium.

³ Merton (1990, chapter 5) analyses a similar problem under restrictive assumptions, but does not present any numerical results.

information about the mean of the return process is the realized sample mean.

We assume that the volatility of the return process is known and constant. This is the simplest setting in which to analyze the role of learning in the dynamic portfolio allocation decision. There are advantages in starting from such a simple case. First, although as we have mentioned, the average return on common stocks since 1926 is often cited as the best estimate of the (current) mean return, there are good reasons to doubt that this parameter has remained constant for almost three quarters of a century which has witnessed the most dramatic economic, technological and social change of any comparable period in history. Therefore, as a practical matter, it is of interest to consider how the optimal investment allocation is affected by uncertainty over this important, but hard to estimate, parameter. Secondly, this simple case allows the development of simple and direct intuition, and easy calculation.

Like Genotte and Detemple we assume that the investor is able to trade continuously, and that the value of the risky asset follows a diffusion process. In this setting the analyses of Genotte and Detemple establish an important and surprising result: the variance of the instantaneous return on the risky asset that is used to determine the optimal portfolio is unaffected by the uncertainty about the mean of the process - uncertainty about this parameter has only a second order effect that disappears as the trading horizon shrinks to zero. Uncertainty about the mean of the return process does affect portfolio decisions, but only because the investor is able to *learn* about the mean return as time passes. This learning induces a hedging demand as the investor attempts to protect himself against learning bad news: that the mean is low. As one would expect from related results in a full information economy, hedging demands are equal to zero when the investor's utility function is logarithmic⁴: in this

⁴ See Feldman (1992).

case the investor's investment in the risky asset is the same as in the full information economy in which the mean of the risky asset return is equal to the investor's current assessment of the mean.

It might be thought that estimation risk would tend to reduce investment in the risky asset relative to the full information case. However, this turns out not to be the case. Depending on the investor's degree of risk aversion, the desire to hedge against changes in the assessed mean of the process which arise from observation of the instantaneous realized return on the asset may lead the investor to invest more or less in the risky asset than he would under complete information.

In the following section we present the investor's decision problem more formally, and in Section 3 present some solutions under representative parameter values.

II

The Investor's Decision Problem

Consider an investor who is concerned with maximising the expected utility of wealth at the end of a horizon, T, $U(W)$, where $U(W)$ is a twice-differentiable concave function. The investor is assumed to be able to invest in a riskless asset with instantaneous rate of return, r , and in a single risky security whose instantaneous rate of return is given by the stochastic differential equation:

$$\frac{dS}{S} = \mu dt + \sigma dz \quad (1)$$

where dz is the increment to a Brownian motion. Thus the stochastic process for the investor's wealth is given by:

$$\frac{dW}{W} = [r + \alpha(\mu - r)]dt + \alpha\sigma dz \quad (2)$$

where " is the fraction of wealth that is invested in the risky asset. The investor is assumed to know the diffusion parameter F , but to be unable to observe the drift parameter of the process, μ . Instead, at time zero the investor views the distribution of μ as normal with mean m_0 and variance v_0 . Denote the expectation and variance of μ conditional on observing the realized returns up to time t by m_t and v_t . Then it follows from the work of Lipster and Shirayev (1978)⁵ that changes in the conditional expectation are given by:

$$\frac{dm}{dt} = \frac{v}{\sigma^2} \left[\frac{dS}{S} - m_t dt \right] \quad (3)$$

while the conditional variance at time t , $v(t)$, is determined by v_0 and the differential equation:

$$\frac{dv}{dt} = \left[-\frac{v}{\sigma^2} \right] dt \quad (4)$$

so that $v(t)$ may be written as

$$v(t) = v_0 e^{-\frac{t}{\sigma^2}} \quad (5)$$

The investor's indirect utility function at time t depends on W , his current wealth, m , his current assessment of μ , and time, t : we write the indirect utility function as $J(W, m, t)$. The Bellman Principle implies that under the optimal policy $E[dJ] = 0$. Therefore, using Ito's Lemma, we have:

⁵ For a simplified exposition of the filtering problem see Gennotte (1986).

$$\max_{\alpha} \left\{ J_W[r + \alpha(m - r)] + \frac{1}{2} J_{WW} \alpha^2 \sigma^2 + \frac{1}{2} J_{mm} \sigma_m^2 + J_{Wm} \sigma_{Wm} + J_t \right\} = 0 \quad (6)$$

where $F_{Wm} = "Wv$, $F_m^2 = v^2/F^2$, and v is given by equation (5). In order to simplify the problem, we shall assume that the investor's utility function is of the iso-elastic class, so that the boundary condition for the indirect utility function may be written as:

$$J(W, m, T) = \frac{1}{\gamma} W^\gamma \quad (7)$$

where $\gamma < 1$ ⁶

Then it may be verified that $J(W, m, t)$ may be written as $(1/\gamma)W^\gamma u(m, t)$, where $u(m, t)$ is the solution to the control problem:

$$\max_{\alpha} \left\{ \frac{1}{2} \sigma^2 u_{mm} + \gamma \alpha v u_m + \left[\frac{1}{2} \gamma (\gamma - 1) \alpha^2 \sigma^2 + \gamma (r + \alpha(m - r)) u + u_t \right] \right\} = 0 \quad (8)$$

with boundary condition $u(m, T) = 1$. The optimal investment in the risky security is given by the first order condition from (8):

$$\alpha = \frac{\left[(m - r) + \frac{v u_m}{u} \right]}{(1 - \gamma) \sigma^2} \quad (9)$$

Note that if u_m is zero, expression (9) reduces to the familiar expression for the optimal investment in

⁶ $\gamma = 0$ corresponds to the logarithmic utility function.

the risky security when the stochastic process is known⁷. The term involving u_m arises from the incentive of the investor to hedge against unfavorable realizations of the unknown parameter μ . The induced hedging demand may be positive or negative. Non-satiation implies that the investor's expected utility is increasing in his current assessment of the mean return, so that $J_m > 0$. If $\zeta < 0$ this implies from the definition of $u(m,t)$ that $u_m < 0$; thus when $\zeta < 0$, uncertainty about the true value of the drift component of the stochastic process leads the investor to invest a smaller fraction of wealth in the risky security than he would if the stochastic process were known. Conversely, if $\zeta > 0$, uncertainty about the stochastic process leads the investor to *increase* his allocation to the risky security. The case $\zeta = 0$ corresponds to log utility, and it is well known in this case that the investor behaves myopically, so that the hedging demand is zero and the allocation to the risky security is the same as when the stochastic process is known.

III

Some Empirical Estimates

To gain some insight into the magnitude of the effect on portfolio allocations that is induced by learning, the optimal portfolio allocation was determined by solving a finite difference approximation to the control problem (7) for selected parameter values, when the investor has an iso-elastic utility function. The parameter values chosen for the prior distribution on the mean return, μ , were derived primarily from the Ibbotson and Sinquefield (1995) data on stock returns for the period 1926-94. These authors report that over this 69 year period the standard deviation of the annual return on the S&P500 index was 20.2%, and the mean annual return in excess of the risk free rate was 8.4%. We

⁷ See Merton (1971).

have taken the risk free rate as 5%, and in our central example we take the investor's current assessment of the mean market return, m_0 , as 13.5%. We consider two different values of v_0 , the variance of the investor's prior distribution of the mean market return: the first is $(0.0243)^2$, which is the variance of the sample mean for the 69 years of annual data; the second is $(0.0452)^2$, which is the variance of the sample mean of a series of 20 annual observations on a series with an annual standard deviation of 20.2%. The idea here is that the prior distributions should correspond roughly to using all the available Ibbotson and Sinquefield data, and using only 20 years' data while keeping the same series mean and standard deviation. We consider two values for F , the standard deviation of the return on the market: 20.2%, the Ibbotson and Sinquefield estimate, and 14.0% which corresponds more closely to recent experience. The optimal proportional portfolio allocation to risky securities, " π^* ", was calculated for different time horizons, T , and values of the risk aversion parameter, γ . " π^* " is the optimal proportional allocation allowing for the uncertainty about the current estimate of μ , and for the prospect of learning more about this parameter as subsequent returns are realized. We shall refer to it as the optimal dynamic allocation. We also calculated a second portfolio allocation,

" $\pi^* / (m - r) / (1 - \gamma) F^2$ " is the myopic or "full information" allocation that would be made if the investor knew that the value of μ was equal to m for sure, or ignored the possibility of learning about this parameter in the future.

Panel A of Table 1 relates to the prior distribution of the sample mean based on 69 annual return observations, while Panel B relates to a prior distribution based on a hypothetical 20 year period. Observe first that, consistent with the theory, for $\gamma < 0$ the optimal dynamic allocation that takes account of learning is always *less than* the myopic allocation, while the reverse is true for $\gamma > 0$. Secondly, note that the difference between the optimal dynamic allocation and the myopic allocation

declines monotonically as the investment horizon is reduced: this is because the influence of the true mean on expected utility also declines as the horizon is shortened, and therefore the incentive to hedge against realizations of this variable is correspondingly reduced. Thirdly, the difference between the two allocations is greater the smaller is F , the standard deviation of the market return; this is because a given change in the expected return on the market has a greater effect on the investor's expected utility the lower is the volatility of the market return⁸. Therefore the lower is F the greater the potential variation in the investment opportunity set against which the investor must hedge.

Considering $F = 14\%$ as representative of contemporary capital markets, we see that learning can have a very significant effect on the optimal portfolio allocation for investors with long horizons. For example, an investor with a risk aversion parameter, ζ , of 2, and a horizon of 20 years would optimally invest 136.1% of his wealth in risky securities if he takes the 8.5% risk premium as a known parameter. Taking account of the fact that the risk premium is a parameter estimated from 69 years of data reduces the optimal allocation to 99.5% of wealth; an investor whose prior distribution corresponded to only 20 years of data would further reduce the allocation to 77% of wealth, or a little over half the level that is optimal when learning is ignored.

Figure 1 plots the optimal allocation to risky assets over time of an investor with a risk aversion parameter of -3.5. The interest rate is again taken as 5%, and the investor is assumed to start the year 1976 with a prior distribution on the mean market return that is normally distributed with mean and variance $(0.135, 0.0243^2)$, corresponding to a prior derived from the data for the 69 year Ibbotson and Sinquefield sample period. The standard deviation of the market return is taken as 14%

⁸ This can be seen by considering the effect on the slope of the capital market line facing the investor of a given change in the market expected return for different levels of market volatility.

per year. The optimal portfolio allocations are plotted for three different policies. alp^* corresponds to the myopic policy which neglects the implications of future learning; alp is the policy which takes proper account of learning and reflects the fact that the investment horizon declines from 19 years at the start of the sample period to zero at the end; alp19 denotes the optimal proportion of wealth to be invested in risky assets allowing for learning, when the horizon is 19 years - thus this series does not reflect the shortening of the investment horizon as calendar time passes. Under all three policies the portfolio allocation changes over time as the parameters of the posterior distribution of the mean stock return (m_t , v_t) change in response to the new information about stock returns. The myopic policy, alp^* , is sensitive only to m_t . Both alp and alp19 change as v_t , the remaining uncertainty about the true mean declines over time, and finally alp changes as the remaining time to the horizon decreases; thus alp starts out equal to alp19 since the original horizon is 19 years, and then approaches the myopic policy alp^* as the time to the horizon approaches zero.

III

Conclusion

Most applied portfolio analysis takes the parameters governing the investor's opportunity set as known. Imperfect knowledge about the parameters gives rise to two quite distinct phenomena that we may conveniently label, 'estimation risk' and 'learning'. Estimation risk arises in a discrete time context, and uncertainty about the mean return will increase the risk and lead to an unambiguous reduction in the fraction of the portfolio allocated to the risky asset. In a continuous time context it is natural to assume that there is no uncertainty about the volatility of the process since this can be estimated precisely; then, as shown by Gennotte (1986) and others, the variance of the instantaneous

returns is unaffected by the estimation error. This means that with logarithmic utility, the allocation to the risky asset is precisely the same as if the mean parameter were known precisely. However, with other utility functions the prospect of learning more about the true value of the mean parameter, μ , as more returns are observed, induces an additional hedging demand for the risky asset. For iso-elastic utility, the hedging demand implies that the allocation to the risky asset is *more than* in the full information case if relative risk aversion is less than unity. However, if risk aversion is greater than that of the log utility function, then the prospect of learning reduces the allocation to the risky asset, just as in the case of estimation risk in a discrete single period setting.

The effect of potential learning on the optimal portfolio allocation increases in importance with the degree of uncertainty over the true value of the mean parameter and the investor's time horizon, and decreases with the volatility of the market return. For reasonable parameter values the effect is found to be significant for investors with horizons as short as 10 years.

The model we have considered is the simplest one in which the effects of dynamic learning could be assessed. The next stage is to investigate the effects of learning in contexts in which returns are partly predictable, as in Kandel and Stambaugh (1996). Finally, since learning can make investors behave in a more risk averse fashion than when this possibility is ignored, it seems natural to explore whether this phenomenon can account for the equity premium puzzle of Mehra and Prescott(1985), which has so far eluded satisfactory explanation.

$F = 0.202$									$F = 0.140$								
T =			"	"*) = " * -	"	"*) = " * -	"	"*) = " * -	"	"*) = " * -	"	"*) = " * -
$\zeta = -5$	20	0.266	0.327	0.161		0.470	0.680	0.210	$\zeta = -3$	10	0.293	0.327	0.034	0.550	0.680	0.130	
	5	0.309	0.327	0.018		0.608	0.680	0.072		20	0.406	0.490	0.084	0.725	1.020	0.295	
	10	0.443	0.490	0.047		0.840	1.020	0.180		5	0.466	0.490	0.024	0.921	1.020	0.099	
$\zeta = -2$	20	0.551	0.654	0.003		0.995	1.361	0.366	$\zeta = 0.2$	10	0.597	0.654	0.057	1.141	1.361	0.220	
	5	0.624	0.654	0.030		1.241	1.361	0.120		20	2.643	2.451	-0.192	6.045	5.102	-0.943	
	10	2.542	2.451	-0.091		5.520	5.102	-0.418		5	2.495	2.451	-0.044	5.300	5.102	-0.198	

Panel A: Period for estimating sample mean 69 years: variance of prior distribution of mean return, $v_0 = 0.0243^2$.

		$F = 0.202$			$F = 0.140$		
	$T =$	"	"*	$\lambda = "*/(1-\lambda)F^2$	"	"*	$\lambda = "*/(1-\lambda)F^2$
$\zeta = -5$	20	0.198	0.327	0.129	0.363	0.680	0.317
	10	0.237	0.327	0.090	0.409	0.680	0.271
	5	0.273	0.327	0.054	0.491	0.680	0.189
$\zeta = -3$	20	0.307	0.490	0.183	0.560	1.020	0.460
	10	0.364	0.490	0.126	0.634	1.020	0.386
	5	0.416	0.490	0.074	0.754	1.020	0.266
$\zeta = -2$	20	0.423	0.654	0.231	0.770	1.361	0.691
	10	0.499	0.654	0.155	0.875	1.361	0.486
	5	0.564	0.654	0.090	1.033	1.361	0.328
$\zeta = 0.2$	20	3.375	2.451	-0.924	16.336	5.102	-11.204
	10	2.808	2.451	-0.357	7.178	5.102	-2.076
	5	2.613	2.451	-0.152	5.878	5.102	-0.706

Panel B: Period for estimating sample mean 20 years: variance of prior distribution of mean return, $v_0 = 0.0452^2$.

Mean of prior distribution, $m_0 = 0.135$. F is the annualized standard deviation of the return on the risky security. T is the investor's horizon in years; ζ is the Pratt-Arrow coefficient of relative risk aversion. " $"$ " is the optimal proportional allocation to the risky security allowing for learning; " $"^*/(m - r)/(1 - \zeta)F^2$ ", is the optimal allocation ignoring learning.

The Effect of Learning on Optimal Portfolio Allocations

Table 1

References

- Barberis, N., 1995, How big are hedging demands? Evidence from long-horizon asset allocation. Unpublished paper.
- Bawa, V.S., S. J. Brown, and R. W. Klein, 1979, *Estimation Risk and Optimal Portfolio Choice*, Amsterdam: North Holland.
- Blanchard, O. J, 1993, Movements in the Equity Risk Premium, *Brookings Papers on Economic Activity* 1993(2), 75-118.
- Breeden, D.T., 1979, an Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities, *Journal of Financial Economics*, 7, 265-96.
- Breeden, D.T., 1984, Futures Markets and Commodity Options: Hedging and Optimality in Incomplete Markets, *Journal of Economic Theory*, 32, 275-300.
- Brennan, M.J. and E.S. Schwartz, 1996, The Use of Treasury Bill Futures in Strategic Asset Allocation Programs, *World-Wide Asset and Liability Modelling*, Ed. J.M. Mulvey and W.T. Ziemba, Cambridge University Press forthcoming.
- Brennan, M.J. , E.S. Schwartz, and R. Lagnado, 1996, Strategic Asset Allocation, forthcoming *Journal of Economic Dynamics and Control*.
- Brown, S.J., W.N. Goetzmann, and S.A. Ross, 1995, Survival, *Journal of Finance*, 853-873.
- Detemple, J.B., 1986, Asset pricing in an economy with incomplete information, *Journal of Finance*, 61, 383-392.
- Dothan, U., and D. Feldman, 1986, Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy, *Journal of Finance*, 61, 369-382
- Gennotte, G., 1986, Optimal Portfolio Choice under Incomplete Information, *Journal of Finance*, 61, 733-749.
- Hakansson, N.H., 1970, Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions, *Econometrica*, 38, 587-607.
- Ibbotson Associates, Inc. *1995 Yearbook*.
- Kandel, S., and R.F. Stambaugh, 1996, On the Predictability of Stock Returns: An Asset Allocation Perspective, *Journal of Finance*, 51, 385-424.

Klein, R.W., and V.S. Bawa, 1977, The Effect of Estimation Risk on Optimal Portfolio Choice, *Journal of Financial Economics*, 3, 215-232.

Lipster, R.S., and A.N. Shirayev, 1978, *Statistics of Random Processes*, Springer-Verlag, New York.

Mehra R., and E.C.Prescott, 1985, The Equity Premium: A Puzzle, *Journal of Monetary Economics*, 15, 145-161.

Merton, R.C., 1971, Optimal Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory*, 3, 373-413.

Samuelson, P.A., 1969, Lifetime Portfolio Selection by Dynamic Stochastic Programming, *Review of Economics and Statistics*, 51, 239-246.

Scott, M.F.G., 1992, The Cost of Equity Capital and the Risk Premium on Equities, *Applied Financial Economics*, 2, 21-32.

Williams, J.T., 1977, Capital Asset Prices with Heterogeneous Beliefs, *Journal of Financial Economics*, 5, 219-240.