MARKOWITZ MEETS KAHNEMAN:
PORTFOLIO SELECTION UNDER DIVIDED ATTENTION

Diego Nocetti¹

Abstract
This paper explores how the scarcity of cognitive resources affects portfolio decisions. I consider an economy where investors allocate mental effort to learn about the mean return of a number of assets, by retrieving information from a stock of memories. As a result, parameter uncertainty arises endogenously. I characterize the optimal division of attention and the optimal portfolios and I show that limited attention can provide interesting insights to the equity home bias.

JEL Classification: G11, G14, D81, D83
Keywords: Portfolio choice; Attention allocation; Memory deficits; Equity home bias puzzle.

¹ Fogelman College of Business and Economics, The University of Memphis, Memphis, TN38152, USA.
Tel: 901-759-0427. e-mail: dnocetti@memphis.edu
1. Introduction

This paper explores how the scarcity of cognitive resources affects portfolio decisions. I take as a starting point the ideas and experimental work of Kahneman (1973) on attention limitations and extend his framework to account for the interaction between attention and memory deficits in an economic setting. In particular, I consider an economy where individuals allocate their attention to estimate the mean return of a number of assets, retrieving information from memory. As a result, parameter uncertainty arises endogenously.

A large body of research has been devoted to study portfolio decisions under parameter uncertainty. A standard assumption in this literature is that investors learn about the true data generating process of asset returns using all available information. This assumption requires investors to have up to date databases of extremely large size. In this paper I argue that many investors do not use databases like econometricians, but make decisions based on the information currently available “in their minds”. Specifically, I conjecture that individuals rely on their memory to infer the parameter estimates. Beyond providing a more realistic flavor to the inference problem, the advantage of such treatment is twofold. First, since the sample size is possibly small due to scarce cognitive resources, parameter uncertainty remains significant even if the data available is large and there are no structural shifts. Second, the endogenous characteristic of the sample size allows disentangling the determinants of the deviations from Markowitz’ (1952) canonical portfolio selection analysis with perfect knowledge of the economy.
The next section reviews the standard portfolio selection problem with estimation risk, while section 3 describes the learning process of an investor with limited attention and presents the optimal division of attention and the optimal portfolios. In section 4 I show that limited attention and the endogeneity of estimation risk can provide interesting insights to the equity home bias puzzle (French and Poterba, 1991). Specifically, I demonstrate that if individuals are more productive in the retrieval of information that are from companies more familiar to them (e.g. domestic firms) they will face a lower predictive variance of returns and possibly higher expected returns for those firms. In addition, I show that the effect of any risk differential between domestic and foreign assets, for instance due to exchange rate risk, is augmented due to the endogeneity of estimation risk.

This explanation of the equity home bias is close to a growing literature that explores the effect of familiarity on investors’ portfolios. Recent studies (Huberman, 2001; Coval and Moskowitz, 1999) distinguish familiarity effects between pure behavioral biases and a “rational” or information-based bias (Brennen and Cao, 1997; Gehring, 1993) whereby more information in more familiar stocks leads investors to optimally allocate a larger fraction of their portfolio in those stocks. In the present model, however, such distinction is inappropriate because both effects are interlinked. I speculate that familiarity increases the ease with which individuals process and retrieve information from memory, a fact that is well established in the experimental and the neuropsychological literature. This behavioral bias, in turn, introduces an information asymmetry which makes it optimal to invest a larger fraction of the portfolio in more familiar stocks.
2. Bayesian approach to portfolio selection

I consider the simplest case where excess returns are i.i.d. and they follow a multivariate normal distribution with mean vector $\mu$ and known covariance matrix $\Sigma^2$. The investor does not known $\mu$ and has to estimate it using past data. As epitomized by Klein and Bawa (1976), the optimal portfolio with estimation risk is obtained by maximizing expected utility under the predictive distribution,

$$w = \arg\max_w \int U(w)p(R_{t+1}|\Theta_t) dR_{t+1}$$
$$= \arg\max_w \int \int U(w)p(R_{t+1}, \mu|\Theta_t) d\mu dR_{t+1}, \quad (1)$$

where $U(w)$ is the utility function, $p(R_{t+1}|\Theta_t)$ is the predictive density and $p(R_{t+1}, \mu|\Theta_t) = p(R_{t+1}|\mu, \Theta_t) p(\mu|\Theta_t)$ is the posterior density of $\mu$. Therefore, the Bayesian solution maximizes expected utility over the distribution of the parameters. As I shall demonstrate, the allocation of attention affects portfolio shares directly by changing the predictive density of excess returns.

Consider the standard case in the literature in which the investor has CARA preferences [i.e. $U = -\exp(-\gamma W)$] and uses $n$ observations of excess returns to estimate $\mu$. Under a diffuse prior of excess returns, the predictive p.d.f. is $R_{t+1}|\Sigma \sim N\left(\hat{\mu}_{t+1}, \Sigma \left(1 + \frac{1}{n}\right)\right)$, where $\hat{\mu}_{t+1}$ is the sample mean vector. In this framework, the Bayesian investor selects,

\footnote{As I argue in section 5, the main results of this paper also apply in non-i.i.d. settings.}
\[ w = \frac{1}{\gamma} \left[ \sum \left(1 + \frac{1}{n} \right) \right]^{-1} \hat{\mu}_{r+1}. \] (2)

It is therefore straightforward that, given a limited number of available observations, the Bayesian investor selects a portfolio with less risk than an investor who knows the mean of excess returns. However, when the sample size is identical for all assets, as is usually assumed, the composition of the efficient frontier portfolios does not change.

3. Portfolio selection with cognitive constraints

I imagine a representative investor that is endowed with a stock of memories of the entire history of excess returns. However, she relies on the retrieval of a subset of those memories to learn about \( \mu_i \). Memory retrieval, in turn, depends on the amount of attention (mental effort) that the investor exerts to learn about each asset \(^3\).

Following Kahneman (1973), mental effort has a limited capacity, \( k \), and is divisible (i.e. processing is parallel as opposed to serial) among activities which might differ in their demands. These activities correspond, in the present framework, to the inference of mean returns of each asset. Further, the effort exerted to a given activity determines a particular performance (larger retrieval of information from memory and, according to (1), higher expected utility). The transformation of mental effort exerted to activity \( i, e_i \), into a given performance is achieved with a cognition technology which takes the form

\[ \Phi_i e_i^\alpha = n_i, \] (3)

where \( \alpha \leq 1 \).

\(^3\) Shefrin (2005 p. 264-274) presents evidence that investors tend to rely in their memory to form return expectations.
I interpret $\Phi_i$ as a parameter that measures familiarity. Thus, equation (3) asserts that individuals are relatively more productive, in terms of mental effort exerted, in retrieving information about more familiar assets. This functional form captures a large amount of experimental evidence (Kahneman and Tversky, 1974; Mandler, 1980) showing that familiarity increases the ease with which different actions are performed. For example, individuals usually find more difficult to retrieve from memory a list of names randomly selected from the phone book than a list of famous people. Similarly, practice improves performance. Further evidence is provided by neuropsychological studies which show that brain activity decreases substantially once an activity becomes familiar. Schneider and Chein (2003) provide a selective review of experimental and neuropsychological evidence on the relation between information processing and familiarity.

Given the attention capacity constraint, $\sum_{i=1}^{m} e_i = k$, and the cognition technology, the investor selects the optimal portfolio shares and the optimal division of attention subject to the cognition/memory possibilities frontier

$$\sum_{i=1}^{m} \left( \frac{n_i}{\Phi_i} \right)^{\frac{1}{2}} = k .$$

I would like to emphasize that the allocation of attention should not be taken literally. Individuals do not make any calculations when they allocate their mental effort. However, I do not argue that individuals explicitly or consciously calculate how much mental effort each asset deserves. I assert rather that the allocation of mental effort is done as if individuals knew the optimality condition\(^4\). For instance, it seems reasonable

\(^4\) This sort of attention and memory management has been emphasized, for example, by Benabou and Tirole (2002)
that an investor that has ninety percent of her wealth in a security will devote a large part of her attention in acquiring information, from memory or otherwise, about this security. The next subsection formalizes this argument.

3.1. Optimal allocation of attention and portfolio shares

The endogeneity of the information set presents a new challenge to solve the portfolio problem. The reason is that the optimal attention level [i.e. the division of attention that maximizes expected utility in (1)] must obviously be established prior to determining the conditional expectation of excess returns, and the optimal portfolio choice. I therefore assume that the representative investor uses the following procedure [see Muendler (2003) for a similar characterization]. First, given an arbitrary prior estimate of expected returns for asset $i$, say $\hat{\mu}_{i,0} \forall i$ (recall the assumption of a diffuse prior), she decides the optimal level of attention and the ex-ante optimal portfolio shares jointly (the minimum variance portfolio); Second, given the optimal $n_i$ she finds the conditional estimate, $\hat{\mu}_{i,t+1}$, and the optimal portfolio shares$^5$.

For expositional simplicity I assume that the assets are uncorrelated. Then, the first order conditions are

$$\hat{\mu}_{i,0} - \gamma \sigma^2_i \left(1 + \frac{1}{n_i}\right) w_i = 0$$

for the (ex-ante) portfolio share of asset $i$, and, using the fact that $\text{Var}_i^* (R_{s+1})$ -the subjective covariance matrix-, has diagonal elements $\sigma^2_i \left(1 + \frac{1}{n_i}\right)$.

$^5$ An alternative to this assumption would be to follow the literature on econometric learning whereby the optimization and forecasting problems are separated. Such separation, however, has no theoretical or empirical basis. Forecast errors affect the allocation of wealth only insofar they affect expected utility.
Equation (6) states that the marginal benefit of retrieving additional memories, a decrease in the variance of the predictive density of excess returns, is equalized to its marginal cost. Solving equations (5) and (6) simultaneously and using the optimal level of attention to obtain expected returns leads to

**Proposition 1.** The optimal sample size and the optimal degree of attention to asset i is implicitly defined by

\[
\sigma_i^2 \Phi_i \left(1 + \frac{c}{n_i}\right)^2 \frac{1}{n_i} \sigma_i^2 \Phi_i - \sigma_i^2 \Phi_j \left(1 + \frac{c}{n_j}\right)^2 \frac{1}{n_j} \sigma_i^2 \Phi_j = 0 \quad \forall j
\]

and the (ex-post optimal) equity share in asset i is

\[
w_i = \frac{\hat{p}_{i,t+1}}{\sigma_i^2 \theta_i},
\]

where

\[
\theta_i = \gamma \left[1 + \frac{1}{n_i \{\sigma_i^2, \Phi_i, \sigma_j^2, \Phi_j, k\}}\right].
\]

Like the case with exogenous but finite samples, the investor with limited attention selects a portfolio with less risk (note, however that this holds even for infinite available observations). In the present setup, however, the adjustment factor is unobservable. Therefore, the equity shares are observationally equivalent to a higher degree of risk aversion, which is asset specific and endogenous. That is, an investor that takes into account estimation risk due to limited cognitive resources is indistinguishable from an
investor with asset specific risk aversion equal to $\gamma[1+(1/n_i)]$. I will therefore define $\theta_i$ as effective risk aversion.

4. Divided attention and the home bias puzzle

A vast literature has provided evidence for the lack of international (e.g. French and Poterba, 1991) and intra-national (e.g. Coval and Moskowitz, 1999) portfolio diversification. In the following subsections I consider how the divided attention framework might help explain the empirical regularity.

4.1. The home bias and familiarity

As argued before, it seems reasonable that individuals are relatively more productive processing and retrieving familiar information. This leads to

**Proposition 2.** If individuals are more productive in the retrieval of observations that are from companies more familiar to them, the holdings of those familiar equities will be larger.

Inspection of optimal attentiveness in (7) shows that the amount of information retrieved for asset $i$ increases with the familiarity of this asset and decreases with the familiarity of the other assets. That is, an increase in productivity produces a biased expansion of the cognition possibilities set and makes it optimal to increase the effort exerted to the now relatively more attention-intensive asset. Since more mental effort reduces the predictive variance of returns, investors allocate a larger share in more familiar assets.

Therefore, information asymmetries and the resulting lack of diversification arise endogenously in a model with attention constraints and memory deficits. Importantly,
although the lack of diversification is the result of a behavioral bias –the ease with which information is processed-, investors are perfectly rational in their decisions.

4.2. **Differential risk in domestic and foreign assets**

Foreign assets usually carry additional uncertainty in terms of exchange rate risk and political risk. If these risks cannot be hedged completely, domestic investors will tend to display a home bias. Empirical studies have found, however, that these additional risks can only explain a small bias towards domestic investments.

Assuming that the differential risk between domestic and foreign assets is reflected in the volatility of excess returns, how do the portfolios of investors with limited attention differ from those portfolios that incorporate all available information? Consider first the case of exogenous estimation risk [and for an investor with perfect knowledge of the economy and risk aversion $\theta$]. We have

$$
\frac{\partial w_{i}^{\text{Bayesian}}}{\partial \sigma_{i}^2} = -\frac{\hat{\mu}_{i,t+1}}{\theta \sigma_{i}^4}.
$$

(9)

By endogenizing effective risk aversion, the divided attention model disrupts the simple effect of volatility on the equity share. In particular, we obtain

$$
\frac{\partial w_{i}}{\partial \sigma_{i}^2} = \frac{\partial w_{i}^{\text{Bayesian}}}{\partial \sigma_{i}^2} + \frac{\hat{\mu}_{i,t+1}(1/n_{i})(\hat{n}_{i}/\hat{\sigma}_{i}^2)(\sigma_{i}^2/n_{i})}{\gamma \sigma_{i}^4 \left[1 + \frac{1}{n_{i}\left\{\sigma_{i}^2, \Phi_{i}, \sigma_{i}^2, \Phi_{i}, k, \mu_{0}\right\}}\right]^{c}}.
$$

(10)

From equation (7) one can show that attention to asset $i$ falls with the variance of excess returns, $(\hat{n}_{i}/\hat{\sigma}_{i}^2) < 0$, implying that the elasticity of attention with respect to risk,
\((\partial n_i/\partial \sigma_j^2)(\sigma_j^2/n_i)\), is negative. Therefore, the effect of risk on the equity share is augmented by the existence of limited attention. Intuitively, an increase in risk reduces the holdings on this asset, which in turn feedback to decrease attention and reduce even more the share.

This result underscores the importance of considering the portfolio selection and the cognition procedure jointly. In a framework where the statistical decision is separated from the economic decision, an increase in the variance of the variable under consideration would make the individual pay more attention to it due to the decrease in the efficiency of the estimate. Then, since portfolio shares are negatively related to the variance of excess returns we would obtain that those stocks with a lower share in the portfolio receive more attention. This is simply counterfactual. However, in the present framework, given an increase in the variance of the excess returns of asset \(i\), the investor reduces the holdings of this asset and invests a higher mental effort on other assets due to their relative increase in risk, i.e., adjusted by the portfolio weights.

The comparative static analysis gives us another interesting implication. Because the assets are uncorrelated, if estimation risk is exogenous the share of one asset is completely unrelated to the process of the other assets. With attention limitations, however, the variance of asset \(j\) is involved in the determination of the share of asset \(i\). In particular, we have \(\partial w_i/\partial \sigma_j^2 > 0\). As the variance of asset \(j\) increases the share on asset \(i\) also increases. Intuitively, an increase in the variance of one asset makes the investor reduce the share of this asset, pay less attention to it and more to the other asset. This
implies that, an increase in the risk of one asset changes the holdings in all other assets in the portfolio, even though they are uncorrelated.

In summary, the existence of cognitive constraints augments the effects of any differential risk that might exist between domestic and foreign assets.

4.3. The home bias and relative return optimism

French and Poterba (1991) suggest a simple explanation to the equity home bias; namely, that domestic investors may be relatively more optimistic about the domestic assets and therefore expect higher returns – the “relative optimism” hypothesis-. Shiller et al (1996) provide some evidence that supports the hypothesis. Do investors with limited attention display relative return optimism? In the framework that I have analyzed so far limited attention does not affect expected returns. This, however, is due to the assumption of diffuse priors. Consider, instead, that the prior belief about the expected return of asset $i$ is normal with mean $\mu_i$ and variance $\sigma_i^2/\phi_i$, where $\phi_i$ determines the prior precision.

With this prior, and again assuming that the assets are uncorrelated, a Bayesian investor’s belief about returns of asset $i$ is

$$ R_{i,t+1} | \Sigma \sim N\left(\frac{\phi_i}{n_i + \phi_i} \mu_{bi} + \frac{n_i}{n_i + \phi_i} \hat{\mu}_{i,t+1}, \sigma_i^2\left(1 + \frac{1}{n_i + \phi_i}\right)\right). $$

(11)

The presence of prior information does not change the signs of the comparative static analysis of the optimal portfolio allocations. However, expected returns are now a weighted average of the prior and the sample estimates. That is, investors shrink the optimal portfolio towards the minimum variance portfolio [Jorion (1985, 1986)].
The following proposition establishes the relation between familiarity and expected returns.

**Proposition 3.** *If individuals are more productive in the retrieval of observations that are from companies more familiar to them and the sample mean of excess returns is higher that the prior estimate, investors will appear to display relative return optimism.*

In addition to a lower posterior variance of returns due to familiarity (proposition 2) individuals now place a larger weight to the sample estimate of the more familiar asset. Therefore, domestic investors expect higher returns for domestic assets even if the actual mean of excess returns is identical to that of the foreign assets. Of course, if the mean of excess returns is lower that the prior estimate they will expect lower returns for the more familiar assets, reducing the home bias effects of the lower posterior variance.

Therefore, as before, one should be careful in interpreting the data on relative expected returns as providing evidence of limited rationality. The present framework establishes the possibility that differences in expected returns are perfectly rational given the investors’ limited cognitive resources. Of course, any positive sentiment towards more familiar assets will be exacerbated by the higher attention devoted to these assets.

**5. Evaluation and extensions**

This paper shows that introducing bounds to the attention/cognition process of an otherwise rational individual helps explain the equity home bias. However, the model has limitations that suggest a number of areas to pursue extensions.

First, in the model presented here memory retrieval is purely random. Casual introspection and a large literature in psychology reveal that this is far from reality.
Bringing memory biases can be done very simply. For example, it is well known that moods tend to cue memories that match in valence (e.g. positive mood, positive memories). This mood-congruency effect might have important implications for portfolio allocation and asset prices. For example, during periods in which investors are optimistic (pessimistic) they will tend to increase (decrease) their holdings on the risky assets, driving prices away from their fundamentals.

Second, I have not allowed for information sharing. One would expect that information sharing would drive the shares of the risky assets closer to the “standard Bayesian” case. Allowing for memory biases, however, might lead to completely different predictions. For example, in the context of mood-congruence effects, people might transmit a relatively larger amount of information that is consistent with their current mood. In such case, information sharing might exacerbate the effects of memory biases.

Third, although I have followed most of the literature on portfolio selection under estimation risk in considering the case of i.i.d. returns, the results presented here also hold in non-i.i.d. settings. For instance, Kandel and Stambaugh (1996) show that, even when returns are predictable (follow an AR(1) process), an investor that faces parameter uncertainty should use all available observations to form expectations. This is in contrast to an investor that knows the parameters of the generating process of returns who only uses the most recent observation. Importantly, they present explicit solutions for the predictive variance of returns and show that it decreases with the sample size. Therefore, the results presented in this paper apply directly to this setting. Future research should be
directed to studying the additional implications of limited cognitive resources under return predictability.

Finally, the assumption that excess returns are uncorrelated across assets is not intended to be realistic, but simplifies the exposition greatly. Stambaugh (1997) shows that, if the samples of return histories differ in length and the assets are correlated, investors can use the information from the assets with longer history to estimate more accurately the generating process of other assets. I therefore expect that correlation will strengthen the positive relation between attention and information retrieval of more familiar assets. Intuitively, investors will use the lower cost information to estimate the returns of less familiar assets. Whether this increases or decreases the home bias, however, is less clear. The decrease in information retrieval of less familiar assets will tend to increase their predictive variance, but now the observations from familiar assets can be used to reduce uncertainty in the other assets.
References


