

Dynamic Portfolio Choice with Parameter Uncertainty and the Economic Value of Analysts' Recommendations

Jakša Cvitanić

Division of Humanities and Social Sciences, California Institute of Technology

Ali Lazrak

Sauder School of Business, University of British Columbia

Lionel Martellini

Finance Department, EDHEC

Fernando Zapatero

Marshall School of Business, University of Southern California

We derive a closed-form solution for the optimal portfolio of a nonmyopic utility maximizer who has incomplete information about the alphas or abnormal returns of risky securities. We show that the hedging component induced by learning about the expected return can be a substantial part of the demand. Using our methodology, we perform an “ex ante” empirical exercise, which shows that the utility gains resulting from optimal allocation are substantial in general, especially for long horizons, and an “ex post” empirical exercise, which shows that analysts’ recommendations are not very useful. (*JEL* C61, G11, G24)

Intertemporal asset allocation is one of the cornerstones of financial economics, both for its theoretical and practical implications. After the very influential papers of Samuelson (1969) and Merton (1971), a number

A previous version of this article -without the empirical application- was entitled “Revisiting Treynor and Black (1973): an Intertemporal Model of Active Portfolio Management.” We express our gratitude to Noël Amenc, Avi Bick, Marc Cassano, Chris Jones, Nicole El Karoui, Marcin Kacperczyk, Bob Geske, Robert Grauer, Alan Kraus, Dion Roseman, Gordon Sick as well as seminar participants at the University of Arizona, Calgary, CEMFI, CIRANO, University of Colorado at Boulder, Courant Institute, HEC (Paris), Hebrew University, INSEAD, UC Irvine, LBS, Mc Gill, SFU, UBC, UCLA, University of Tel-Aviv, the Ninth Annual CAP Workshop on Derivative Securities and Risk Management at Columbia University, and the Pacific Northwest Finance Conference, for very useful comments and suggestions. We are also grateful to Mehdi Chaabouni and Dmitri Kantsyrev for outstanding research support. Jakša Cvitanić’s research was supported in part by NSF grants DMS-00-99549 and DMS 04-03575. Ali Lazrak’s research was supported by the Natural Sciences and Engineering Research Council of Canada under Grant 298575-04. Finally, we are grateful to an anonymous referee for many suggestions. All errors are, of course, the authors’ sole responsibility. Address correspondence to Fernando Zapatero, FBE, Marshall School of Business, USC, Los Angeles, CA 90089-1427, or email: fzapatero@marshall.usc.edu.

of different lines of research have been followed. One of the interesting directions, and the area to which this article belongs, is the literature on optimal allocation with incomplete information about security parameters, with Bayesian updating. The idea of using Bayes' rule to estimate security parameters goes back to Zellner and Chetty (1965), Klein and Bawa (1976), and Brown (1979) in a static setting. There are a large number of papers that extend this idea in a discrete-time setting. Recently, Kandel and Stambaugh (1996), Pástor and Stambaugh (1999, 2000), Barberis (2000), Pástor (2000), Baks, Metrick and Wachter (2001), Stambaugh (2003) and Jones and Shanken (2004) consider the portfolio problem with incomplete information, both on expected return and on variance of the securities, in discrete-time settings. These papers successfully address various important problems but the hedging component that is induced by learning does not figure in their optimal holdings. Finally, another recent, different approach followed by Harvey et al. (2004), studies optimal allocation with incomplete information and focuses on modelling the updated probability distributions, which in many cases will display relevant higher moments.

For the continuous-time setting used in this article, the literature starts with Detemple (1986), Dothan and Feldman (1986), and Gennotte (1986). These pioneering papers discussed the asset pricing implications of incomplete information on the economic output growth rate. Browne and Whitt (1996) solved the problem of optimal portfolio allocation in discrete-time and showed its convergence to a continuous-time solution. More recent work in this line of literature include Veronesi (2000), David and Veronesi (2002), and Pástor and Veronesi (2005).

In the context of the literature on continuous-time portfolio choice, to which this article is more closely aligned, investors have priors on the securities' expected returns [or abnormal returns with respect to some benchmark such as the capital asset pricing model CAPM], and they observe prices and update priors accordingly, in a Bayesian way. Financial econometricians seem to agree that it is feasible to obtain good estimates of variance parameters but much harder to estimate expected returns (Merton, 1980). Based on that, and on other reasons of a technical nature, most of the continuous-time literature focuses on the problem of incomplete information on securities' expected returns and considers the return variances as known constants. On the contrary, one important advantage of the continuous-time approach is that as long as the utility is not logarithmic, hedging demand due to parameter uncertainty emerges under optimal behavior. Lakner (1998) introduced the martingale approach to address the portfolio allocation problem with parameter uncertainty and gave some theoretical integral representations of the optimal portfolio policy. Brennan (1998) explored a similar problem to assess the magnitude of the hedging demand in a calibrated portfolio

choice problem. Brennan (1998) characterized the investor's solution through a partial differential equation that was solved numerically and quite accurately in the single risky asset case. Rogers (2001) found the closed-form solution for the case of a single risky asset. Stojanovic (2002) considered a more specific subclass of portfolio strategies and obtained similar results, as it turns out that the *true* optimal strategy belongs to this subclass. Brendle (2005) characterized in terms of partial differential equations the case in which the unobserved drift is a stochastic process.

Finally, the continuous-time methodology has been used by Brennan and Xia (2001) to assess the importance of the Fama and French anomalies for portfolio choice and by Xia (2001) to measure the effect on portfolio choice of learning about a predictive relationship in returns. Both papers consider investors with constant relative risk aversion (CRRA) preferences, possibly nonmyopic (resulting in hedging components in the optimal allocation), and solve for the optimal allocation through the use of numerical methods.

In this article we consider a setting which is similar to that of Brennan and Xia (2001) but less general.¹ We use martingale techniques that allow us to derive an analytic expression for the optimal portfolio of an investor with CRRA nonmyopic preferences. Our approach furthers previous work by allowing for an arbitrary number of assets and accommodating cross-correlation in the prior for expected return without losing the analytical nature of the solution. Furthermore, rather than expressing our portfolio weight in terms of the expected returns, we expressed them as a function of the standard alpha parameters that measure mispricing of a security in a CAPM (or multifactor) setting.² As usual in this literature, our investor is not meant to be a representative agent, and thus our approach is fully normative. The closed-form nature of our result allows the computation of optimal weights for portfolios with a large number of securities. In particular, we perform two exercises using analysts' recommendations as indicators of individual securities mispricing.

From our analytical formula, we can confirm some of the hedging demand signs and magnitudes documented in Brennan (1998) and Brennan and Xia (2001). For instance, we show that the hedging component is an important part of the optimal demand, especially for investors with high risk aversion or a long horizon. The hedging demand is the result of learning. More importantly, our analytical formula is helpful in

¹ Our prior on the expected return is Gaussian, whereas Brennan and Xia (2001) allowed in some cases a mixture of Gaussian distributions to model the lack of confidence in any particular asset pricing model. Our model corresponds to what they call the "pure prior distribution" in their paper.

² This represents an improvement over the Treynor and Black (1973) formula that is so widely used in the active asset management industry. Black and Litterman (1991, 1992) extended that model and introduced a methodology that allows investors to account for uncertainty in their priors on expected returns (expressed in terms of deviation from neutral equilibrium CAPM-based estimates), still in a static setting. Our result can be interpreted as a dynamic version of that approach, with Bayesian updating.

detecting some interesting economic mechanisms that explain the intuition of the optimal portfolio choice.

Owing to the perceived positive autocorrelation in return (good return realization is good news on expected returns), the long-term power utility investor will hold fewer shares of the index.³ We also show that in the presence of some positive or negative expected alphas, the investor will long/short the mispriced securities (as in Treynor and Black, 1973). Furthermore, we show that low-beta (resp. high-beta) securities are optimal substitutes for a portion of the risk-free securities (resp. the index).

Finally, we find that the economic mechanism by which the prior correlation affects portfolio allocation is in general complex, but within our context, we isolate two possible channels. On the one hand, regardless of its sign, the correlation among the alpha priors generates cross-learning and has the same effect as decreasing the uncertainty around both alphas. Consequently, it increases the incentives to invest in each underpriced security because the learning will be faster. On the other hand, for the sake of diversification of the estimation risk, a positive (negative) correlation between alphas generates an incentive to decrease (increase) the weight of the mispriced assets relative to the uncorrelated case. Thus, the implication of the correlation of alpha priors on the portfolio holdings is mixed: optimal behavior requires some counterintuitive holdings. For instance, it may happen that the investor goes long a fairly priced stock or even a negative expected alpha stock: this will happen if another stock has a negatively correlated alpha with higher expected alpha. It may also be optimal to short a stock with positive alpha in the presence of another stock that has a higher expected alpha and exhibits a positively correlated prior alpha.

From a practical perspective, under the assumption of (true) iid returns and Gaussian priors on expected returns, our results solve the portfolio dimensionality problem which represents a major obstacle to the numerical methods used for instance in Brennan and Xia (2001). This is something that could not be done in the previous work, and consequently, our results warrant further potential applications in the literature of portfolio choice with parameter uncertainty. For instance, within our specific context, our results would allow to carry a discussion of the Fama and French anomalies from a portfolio perspective with the same level of tractability of the discrete time literature (Pástor 2000) without giving up either the hedging demand in portfolio choice (Brennan and Xia 2001) or the commonality in the alpha priors (Jones and Shanken, 2004, and Stambaugh 2004).

We illustrate the benefits of our results in another context, where we test the usefulness of our analytical formula in a context of a large universe of stocks. We follow the idea of Kandel and Stambaugh

³ This result holds when the investors are more risk-averse than the log investor.

(1996): we estimate parameter values for individual securities and compute the “certainty equivalent” (or additional initial wealth, as in McCulloch and Rossi (1990)) that an investor with CRRA utility who follows a “naive,” suboptimal strategy would require to achieve the same level of utility that would result from the optimal strategy derived in this article. Investing in individual securities (above their participation in the market portfolio) is optimal when they have alphas that are different from zero. As estimates of securities’ alphas, we use the analysts’ recommendations collected in the IBES database. The type of “naive” strategies we consider involve taking long positions in securities with positive alphas and short positions in securities with negative alphas. We find that the utility gains (in terms of certainty equivalent) that result from optimal allocation are very significant, especially for long horizons.

We also perform a complementary, “ex post” exercise, in which we compute the average utility (as an approximation of expected utility) achieved by a CRRA investor, who uses the alphas implied by analysts’ recommendations, with rebalancing, and we compare it to the utility that the same CRRA investor would achieve if restricted to invest in the market portfolio and the risk-free security. Overall, we find that analysts’ recommendations seem marginally useful. Additionally, we find that analysts’ recommendations seem more useful when a large number of analysts are covering a particular firm.

The article is organized as follows. First we present the model, derive the formula and perform a comparative static analysis. In the following section, we present the results of our empirical exercises. We then briefly present some extensions. We close the article with some conclusions.

1. Optimal Investment in the Presence of Mispriced Assets

There are $n + 1$ risky assets (stocks) in the economy, the price of which is denoted by S_i , $i = 0, \dots, n$, and one risk-free asset (the bond or bank account), the price of which is denoted by B , which pays a constant interest rate r , such that its price dynamics are $dB/B = rdt$. The asset S_0 can be interpreted as the market portfolio or any traded benchmark. Its price satisfies

$$\frac{dS_0}{S_0} = \mu_0 dt + \sigma_0 dW_0. \quad (1)$$

where μ_0 is the constant expected return, and W_0 is a standard Brownian motion process. The parameter σ_0 is also constant. The other n assets are modelled as

$$\frac{dS_i}{S_i} = \mu_i dt + \sigma_i dW_0 + \sigma_{\varepsilon_i} dW_i, \quad i = 1, \dots, n, \quad (2)$$

where W_i are standard Brownian motion processes, independent of W_0 and also independent of each other. We denote by W the $(n + 1)$ – Brownian motion $W := (W_0, W_1, \dots, W_n)$. We assume that the parameters μ_i , σ_i , and σ_{ε_i} are also constant. W_0 is common to all risky securities, and the parameters $\sigma_i, i = 0, 1, \dots, n$ represent the *systematic* or *market* risk of the securities. This is the only source of risk for the market portfolio. On the contrary, individual securities also have an *idiosyncratic* source of risk, given by their respective σ_{ε_i} parameters. The individual securities S_i are part of the market portfolio. We assume that the company-specific risk component is diversified, and only the market component drives the price of the market portfolio. Although it can be argued that a financial market structure of that kind should be the result of some type of equilibrium, we do not consider that problem in this article and take the market structure as given.

We denote by σ the volatility matrix, that is, the matrix formed by stacking the row vectors of volatilities of the $n + 1$ stocks/portfolios. The rows of this matrix contain all zeros except the term corresponding to the market risk component and the term corresponding to the company-specific component. We assume that the volatility matrix is nonsingular. We also introduce the “risk premium” vector,

$$\theta := \sigma^{-1}(\mu - r \cdot \mathbf{1}), \quad (3)$$

where $\mathbf{1} = (1, 1, \dots, 1) \in R^{n+1}$ and $\mu = (\mu_0, \mu_1, \mu_2, \dots, \mu_n)$. We assume that r and σ are observable by agents in the economy. The investors also observe security prices, $(S_0, S_1, \dots, S_n)$, but observe neither the mean return vector $\mu = (\mu_0, \mu_1, \mu_2, \dots, \mu_n)$ nor the sources of noise W (otherwise they could immediately retrieve expected returns). This is motivated by the fact that expected returns are notoriously harder to estimate from finite samples than variances (Merton 1980, Jorion 1986). More formally, the investor’s information consists of the filtration

$$\mathcal{F}_t^S := \sigma(S_0(t), S_1(t), \dots, S_n(t)); \quad 0 \leq t \leq T$$

generated by the price process $(S_0(t), S_1(t), \dots, S_n(t); t \in [0, T])$.

Furthermore, investors have beliefs or “priors” about the vector of expected returns μ . We adopt a Bayesian approach and assume that the vector θ (equivalently μ , because r and σ are known) has a Gaussian prior distribution, independent of the Brownian motion W . We denote by $m = (m_0, m_1, \dots, m_n)$ the mean vector of the distribution of θ , and we denote by Δ its variance–covariance matrix. We point out that the only unobserved component of θ is the expected return μ . For that reason, we will use the variance–covariance matrix of the prior distribution of θ (instead of μ) to keep the notation as simple as possible.

When the investor described above has full information, Merton (1973) showed that, in equilibrium, the Intertemporal CAPM (ICAPM) holds,

$$\mu_i - r = \beta_i(\mu_0 - r) \quad (4)$$

where

$$\beta_i \equiv \frac{\text{cov}\left(\frac{dS_i}{S_i}, \frac{dS_0}{S_0}\right)}{\sigma_0^2} = \frac{\sigma_i}{\sigma_0} \quad (5)$$

Within the present context, we assume that security prices can deviate from the ICAPM Equation (4). That is, the expected return μ_i of security i admits the following decomposition,

$$\mu_i = \underbrace{r + \beta_i(\mu_0 - r)}_{\text{normal return}} + \underbrace{\alpha_i}_{\text{abnormal return}} \quad (6)$$

where β_i is as in (5), and α_i reflects a potential mispricing of stock i .

Finally, in our one factor setting with uncorrelated residuals, note that

$$\theta_i = \frac{\alpha_i}{\sigma_{\epsilon_i}} \quad (7)$$

for $i = 1, \dots, n$, and

$$\theta_0 = \frac{\mu_0 - r}{\sigma_0}. \quad (8)$$

Equations (7) and (8) show that, under the no residual correlation condition, the assumption of no correlation in the priors on the market premium (thetas) is equivalent to the assumption of no correlation in the priors on alphas and the expected market return (μ_0).

We now consider the problem of a risk-averse investor who has access to the financial markets described above. This investor is risk-averse and nonmyopic: preferences are characterized by a power utility over final wealth. Initial wealth is denoted by X_0 , and the investor optimally invests in the $n + 2$ securities described above so as to maximize utility of final wealth, given by the expression

$$u(X_T) = \frac{(X_T)^{1-a}}{1-a} \quad (9)$$

For $a = 1$ this would be logarithmic -myopic- utility. We focus on the more interesting case $a > 1$.

The market specification described in Equations (1)–(8) is convenient for our theoretical analysis and corresponds to the empirical problem considered in section 2. However, our result applies to a more general

setting which we describe next, in order to state our main theorem. We consider a model with one risk-free asset B , with $dB(t) = rB(t)dt$, and $n + 1$ risky assets S_i whose prices evolve according to the equations

$$dS_i(t) = S_i(t) \left[\mu_i dt + \sum_{j=0}^n \sigma_{ij} dW_j(t) \right], \quad S_i(0) > 0, i = 0, \dots, n, \quad (10)$$

where the volatility matrix $\sigma = \{\sigma_{ij}\}_{0 \leq i, j \leq n}$ is assumed to be nonsingular and known, whereas the mean return process $\mu = (\mu_0, \dots, \mu_n)$ is a random variable with normal prior distribution. We now introduce some notation, which is needed to state the main result. The portfolio strategy $\pi(\cdot)$ is a $n + 1$ -dimensional vector with elements $\pi_i(t)$, representing the proportion of wealth invested in asset S_i . The process π must be adapted to the price filtration. The $n + 1$ -dimensional process $W^*(t) := W(t) + \theta t$ denotes the risk-neutral Brownian motion. Denote by P the orthogonal matrix such that the variance-covariance matrix for the prior on θ can be expressed as

$$\Delta = P' D P \quad (11)$$

where D is a diagonal matrix whose i th element on the diagonal is denoted by d_i .⁴ We also define

$$\delta_i(t) = \frac{d_i}{1 + d_i t}, \quad (12)$$

$$A_i(t) = a - (1 - a)\delta_i(t)(T - t) \quad (13)$$

and, we denote by $A^{-1}(t)$ the diagonal matrix with diagonal terms $1/A_i(t)$. Finally, consistent with our notation above, we denote by $\bar{\mu}$ the row vector of conditional expected returns (updated priors), that is,

$$\bar{\mu}(t) = E[\mu | \mathcal{F}_t^S] \quad (14)$$

and similarly for $\bar{\theta}$, $\bar{\alpha}(t) = [\bar{\alpha}_1(t), \dots, \bar{\alpha}_n(t)]'$. We now present the formula for the optimal portfolio.⁵

⁴ We point out that matrix decomposition involves a simple immediate numerical procedure, similar to matrix inversion, and it is standard in computational software. Because of that, the formula is as explicit as that in Merton (1971), which in the multidimensional case involves the inversion of a matrix. As a result of this, the higher the dimension of the problem, the more advantageous the formula presented in this article versus the alternative numerical methods mentioned in the literature review.

⁵ The optimization problem we face is also studied in Stojanovic (2002). He considers only portfolio strategies with specific functional forms for the portfolio weights and uses a different calculus of variations approach. He obtains the same formula for the optimal portfolio, because the *true* optimal portfolio turns out to have the functional form he initially assumed. He also gives the optimal portfolio value at time zero only, as he does not consider learning through conditional means and variances.

Theorem 1. Assume that $a \geq 1$. The vector $\hat{\pi}(t)$ of optimal weights at time t for the portfolio optimization problem (9) in model (10) is given by

$$\hat{\pi}(t) = (\sigma')^{-1} P' A^{-1}(t) P \bar{\theta}(t) \quad (15)$$

where $\bar{\theta}(t)$ is the conditional mean of the risk premium given by

$$\bar{\theta}(t) = E[\theta | \mathcal{F}_t^S] = P' \bar{D}(t) \{ P W^*(t) + [\bar{D}(0)]^{-1} P m \}, \quad (16)$$

$m = \bar{\theta}(0)$, and the matrix $\bar{D}(t)$ is a diagonal matrix with diagonal elements $\delta_i(t)$. The conditional variance–covariance matrix for θ is equal to $P' \bar{D}(t) P$.

Proof. In Appendix A.1.

For the rest of the article, we focus on the model of (1)–(8). Investors fully incorporate that learning will occur in the future and make the appropriate adjustments in their risk aversion (through the term A_t) to hedge against estimation errors. For example, a *boundedly rational* investor, who treats her time t perception $\bar{\theta}(t)$ of the unknown θ as if it were the true parameter value, will adopt the *myopic* policy

$$\pi^m(t) = a^{-1} (\sigma')^{-1} \bar{\theta}(t).$$

Thus, we define the *hedging demand* as the residual component of the optimal portfolio net of the myopic demand, that is

$$\pi^h(t) := \hat{\pi}(t) - \pi^m(t) = (\sigma')^{-1} P' [A^{-1}(t) - a^{-1} I] P \bar{\theta}(t),$$

where I represents the identity matrix. Finally, taking the conditional expectation in (7) gives

$$\bar{\theta}_i(t) = \frac{\bar{\alpha}_i(t)}{\sigma_{\epsilon_i}}$$

for $i = 1, \dots, n$, and

$$\bar{\theta}_0(t) = \frac{\bar{\mu}_0(t) - r}{\sigma_0},$$

which allows us to express our portfolio holdings in terms of alpha, a quantity more familiar in the portfolio management industry. More specifically, if prior alphas of individual securities are uncorrelated with the prior expected return of the market portfolio, a direct analysis of the orthogonal matrix P shows that individual security holdings depend on the estimated return rates $\bar{\mu}_i$, $i = 0, \dots, n$, only through their estimated alphas, $\bar{\alpha}_i$, $i = 1, \dots, n$, and through $\bar{\mu}_0$. This observation is very useful from a practical point of view, because individual alphas are usually easier to estimate than μ_i 's.

In the rest of the section, we specialize this model for particular cases and derive the optimal selection problem of an investor who has nontrivial priors on alphas, possibly generated by better access to information, or better ability to process it. Without loss of generality, we only focus on the properties of the strategy at the initial date $t = 0$, and when there is no risk of confusion, we omit the time dependency of the variables of interest [for instance $\hat{\pi}_i(0)$, $\pi_i^m(0)$, and $\bar{\alpha}_i(0)$ will simply be denoted by $\hat{\pi}_i$, π_i^m , and $\bar{\alpha}_i$, respectively].

1.1 Uncorrelated priors

In this section, we assume that the prior alpha of each security is independent of the prior alphas of other securities (including the market portfolio). The expression for the optimal portfolio strategy takes on a simple form. We now introduce the result.

Proposition 1. *The optimal investment strategy of a risk-averse investor with incomplete information and uncorrelated priors is given by*

$$\hat{\pi}_i = \frac{\bar{\alpha}_i}{\sigma_{\varepsilon_i}^2 A_i}. \quad (17)$$

for individual risky securities. The optimal investment in the market portfolio is given by

$$\hat{\pi}_0 = \frac{\bar{\mu}_0 - r}{\sigma_0^2 A_0} - \sum_{i=1}^n \beta_i \hat{\pi}_i \quad (18)$$

Proof. In Appendix A.2.

Equation (17) refers to the extraordinary holdings in security i due to the abnormal return. Those would be in excess of the investment in i as part of the market portfolio. We will call these holdings *alpha-driven*. Clearly, when the observed abnormal return $\bar{\alpha}$ is zero, the investor does not deviate from the traditional ICAPM allocation (the *beta-driven* motive).

A myopic investor would hold

$$\pi_i^m = \frac{\bar{\alpha}_i}{\sigma_{\varepsilon_i}^2 a}, \quad (19)$$

and

$$\pi_0^m = \frac{\bar{\mu}_0 - r}{\sigma_0^2 a} - \sum_{i=1}^n \beta_i \pi_i^m. \quad (20)$$

As is well known, this myopic policy corresponds to an investor who perceives a constant investment opportunity set and invests in any

security with positive alpha. The weight of these underpriced securities in the optimal portfolio decreases with idiosyncratic risk and with risk aversion. The myopic weight for the market portfolio has a standard mean–variance component, as in the case of the other securities, which decreases with market volatility and risk aversion. Additionally, the market portfolio weight has a second component which penalizes it for positive beta underpriced securities (or negative beta overpriced securities). The intuition here is that when a security is strongly correlated with the market, a positive alpha for that security indicates that one should substitute some weight from the market to that asset. This adjustment is driven by a diversification motive and is therefore related to individual betas. In particular, zero-beta individual assets do not contribute to this effect.

Finally, as in Treynor and Black (1973), the total proportion of the portfolio invested in the risky securities by a myopic investor [i.e., the sum of (19) and (20)] is affected by the existence of risky securities with nontrivial alpha. Consider, for example, the market portfolio and a single security with positive alpha. If the beta of that security is lower than one, then total holdings in risky assets will be higher than they would be if the security had a zero alpha, and the opposite will hold if the beta of the security is greater than one. This effect also holds for the total proportion of the nonmyopic investor [i.e., the sum of (17) and (18)], but incomplete information will smooth this effect (through the presence of A_i).

1.1.1 Importance of the hedging demand. The nonmyopic investors of (17) and (18) incorporate future learning in their decision. In fact, as we shall illustrate, they perceive a positive autocorrelation in returns, although the true returns are i.i.d. Thus, from (17)–(20), the hedging demand in this context is given by

$$\pi_i^h = \frac{\bar{\alpha}_i}{\sigma_{\varepsilon_i}^2 a} \left[\frac{(1-a)d_i T}{a - (1-a)d_i T} \right], \quad (21)$$

and

$$\pi_0^h = \frac{\bar{\mu}_0 - r}{\sigma_0^2 a} \left[\frac{(1-a)d_0 T}{a - (1-a)d_0 T} \right] - \sum_{i=1}^n \beta_i \pi_i^h. \quad (22)$$

In the absence of mispricing ($\bar{\alpha}_i = 0$), the hedging demand for an individual security is zero, whereas the hedging demand for the market portfolio is

$$\pi_0^h = \frac{\bar{\mu}_0 - r}{\sigma_0^2 a} \left[\frac{(1-a)d_0 T}{a - (1-a)d_0 T} \right], \quad (23)$$

and is nonpositive. The intuition behind this result is that observing high (low) realized market returns leads to an upward (downward) revision of

expected future market returns. As a result, the investor perceives a positive serial autocorrelation in market returns. In fact, it can be proven that

$$\text{Var} \left[\log \left(\frac{S_0(t)}{S_0(0)} \right) \right] = \sigma_0^2 t + \sigma_0^2 d_0 t^2,$$

and thus the cumulative variance of market returns grows in a quadratic fashion with time, making the perceived market returns riskier in the long run than the true i.i.d. returns (the true cumulative return variance grows linearly with the horizon).⁶ Consequently, in the absence of mispricing, they invest less in the market portfolio than myopic investors.⁷

Furthermore, in the absence of mispricing, (23) shows that the magnitude (absolute value) of the market hedging demand increases monotonically with the horizon. This is because of the fact that the cumulative variance of the returns grows quadratically with the horizon, so the longer the horizon, the more magnified the risk of the investment opportunity. As a result, nonmyopic investors prefer to be conservative and postpone the investment decision. This is because the influence of the true mean on expected utility is higher for long horizons. Later on during the investing lifetime, as the horizon becomes small, the hedging demand converges to zero because estimation errors have a minor effect on expected utility.

Additionally, the smaller the volatility σ_0 of the market portfolio, the larger the hedging component of the demand for the market portfolio. This is because for small volatility in the market portfolio, a mistake in the estimation of its expected return implies a larger mistake in the estimation of the slope of the capital market line faced by the investor.

Now, when there is mispricing, the hedging demand expression in (21) shows that an interpretation similar to that for the market portfolio holds for any individual security. Regarding the market portfolio, (22) shows that, similarly to the myopic policy, the second component penalizes the hedging component of the demand for the market portfolio.

1.1.2 Impact of introducing a nontrivial alpha. An interesting question is the impact of the introduction of an investment opportunity (nontrivial alpha) on the optimal investment strategy of an investor. In particular, one would like to know whether the investor should optimally withdraw money from the market portfolio holdings or the risk-free holdings to

⁶ This feature of stock prices differs from Barberis (2000), where the predictability of returns induces mean reversion which in turn lowers the variance of cumulative returns over long horizons.

⁷ Note, however, that there are other aspects of the return distribution (higher moments) which are taken into account by a power utility maximizer. In particular, a power utility investor with a coefficient $0 < a < 1$ views the perceived market dynamic as more attractive and will adopt a positive hedging demand for the market.

allocate it to the perceived investment opportunity. The following proposition provides very simple insights into the question. For simplicity of exposure, we consider the case of a single investment opportunity, that is, a mispriced security with abnormal return α_1 and beta β_1 .

Corollary 1. *In the setting described above, denote the date t optimal holdings in the market portfolio and risk-free asset in the absence of the investment opportunity as $\pi_0^{\alpha=0}$ and $\pi_B^{\alpha=0}$. The changes in holdings due to the introduction of the mispriced security with abnormal return α_1 and beta β_1 are*

$$\begin{aligned}\Delta\pi_0 &:= \pi_0^{\alpha_1=0} - \pi_0^{\alpha_1 \neq 0} = \beta_1 \hat{\pi}_1 \\ \Delta\pi_B &:= \pi_B^{\alpha_1=0} - \pi_B^{\alpha_1 \neq 0} = (1 - \beta_1) \hat{\pi}_1\end{aligned}$$

with $\hat{\pi}_1$ as in Equation (17). Besides, when $\hat{\pi}_1$ is positive, (i.e., when $\bar{\alpha}_i$ is positive) we have that

$$\Delta\pi_B \geq \Delta\pi_0 \iff \beta_1 \leq \frac{1}{2}$$

Proof. Straightforward, from

$$\begin{aligned}\pi_0^{\alpha_1=0}(t) &= \frac{\bar{\mu}_0(t) - r}{\sigma_0^2 A_0(t)} \\ \pi_B^{\alpha_1=0}(t) &= 1 - \frac{\bar{\mu}_0(t) - r}{\sigma_0^2 A_0(t)}\end{aligned}$$

This result appears to have a natural interpretation for hedge fund investors: there seem to be two main reasons behind the success of hedge funds in institutional portfolios (see Schneeweis and Spurgin 1998 and Amenc, Martellini and Vaissié 2003, for a detailed study). On the one hand, hedge funds seem to provide diversification with respect to other existing investment possibilities (beta benefit). On the other hand, it is argued that hedge funds provide an abnormal risk-adjusted return (alpha benefit). A question that investors in hedge funds often ask is where they should take the money that they are planning to allocate to hedge funds from. If we consider the investment opportunity to be a hedge fund, we find, from the corollary above, that the introduction of a hedge fund with positive perceived alpha leads investors to optimally withdraw an amount from the money market account larger than that taken out of the market portfolio when the hedge fund has a beta lower than 1/2. Intuitively, this is because the hedge fund becomes less (more) comparable to the market portfolio as its beta decreases (increases). In other words, this result suggests that low-beta hedge funds (e.g., equity market neutral or convertible

arbitrage strategies) may actually serve as natural substitutes for a portion of an investor's risk-free asset holdings, while high-beta hedge funds can be regarded as substitutes for a portion of equity holdings.

1.2 Effect of the correlation of priors

In general, investors' priors on alphas for different stocks may well be correlated. Arguably, the investor uses similar algorithms to come up with priors on the different expected returns and when information about a particular stock triggers a revision of the prior on this stock, the investor will extrapolate the adjustment to priors that have been computed in a similar way. The relationship seems more obvious for priors on expected returns of stocks from the same country or industry, which might reflect a common factor structure. Correlation among priors is also assumed in Stambaugh (2003) and Jones and Shanken (2004). They argue that priors' correlation is positive. In our analysis, we study the case of negative correlation as well, because it helps understanding the hedging demand.

In the presence of correlation among priors, it is straightforward to apply (15) to compute the optimal weight in each security. However, it is more difficult to derive comparative static results as well as the economic intuition behind them. For that purpose, we concentrate in the next subsection on the case in which the prior alphas of mispriced securities are correlated with each other but uncorrelated with the prior on the expected return of the market portfolio. This specification seems appropriate for the case in which the mispriced securities are low-beta hedge funds, for example. In the following subsection, we replicate the example of Brennan and Xia (2001), with correlation among priors on expected returns of the individual securities and the market. This example serves to illustrate the importance of the hedging demand as well as the effect of correlation among priors.

1.2.1 Correlation between alphas. In this section, we focus on the case of two securities that are different from (but included in) the market portfolio. More precisely, there are two securities whose expected returns, and therefore respective α 's, are not observed by the investor. The investor has priors on those returns that are correlated with each other but uncorrelated with the prior of the expected market return. The assumption of two assets simplifies notation, but the ensuing results are robust to the case of more than two securities. However, the results depend on the assumption of no correlation between the priors on alphas and the prior on the expected market return.

As in the previous section, we use the priors on the market price of risk θ (rather than on expected returns). The variance-covariance matrix Δ of priors on vector θ is given by

$$\Delta = \begin{pmatrix} v_0 & 0 & 0 \\ 0 & v_1 & \gamma \\ 0 & \gamma & v_2 \end{pmatrix} \quad (24)$$

To apply Theorem 1, we first need to decompose Δ into a matrix D defined in terms of the nonnegative eigenvalues of Δ , which we denote $d_0(=v_0)$, d_1 , and d_2 , and in terms of an orthogonal matrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & \sqrt{1-p^2} \\ 0 & \sqrt{1-p^2} & -p \end{pmatrix}, \quad P'P = I,^8 \quad \text{such that } \Delta = P'DP \text{ where}$$

$$D = \begin{pmatrix} v_0 & 0 & 0 \\ 0 & d_1 & 0 \\ 0 & 0 & d_2 \end{pmatrix}. \quad \text{The relationship between the elements of } \Delta \text{ and } D \text{ is given by}$$

$$d_1 = \frac{v_1 + v_2}{2} + \frac{\sqrt{(v_1 - v_2)^2 + 4\gamma^2}}{2}$$

$$d_2 = \frac{v_1 + v_2}{2} - \frac{\sqrt{(v_1 - v_2)^2 + 4\gamma^2}}{2}$$

These parameters are necessary to compute A_i , whose definition is given by (12) and (13). Finally, we find parameter p from the following:

1. If $\gamma > 0$, then

$$p = \sqrt{\frac{1}{2} + \frac{v_1 - v_2}{2\sqrt{(v_1 - v_2)^2 + 4\gamma^2}}};$$

2. If $\gamma < 0$, then

$$p = -\sqrt{\frac{1}{2} + \frac{v_1 - v_2}{2\sqrt{(v_1 - v_2)^2 + 4\gamma^2}}};$$

Therefore, when the securities are uncorrelated and $v_1 > v_2$, we have $p = \pm 1$ (only p^2 matters), and when the securities are uncorrelated and $v_1 < v_2$, we have $p = 0$. In general, p is positive when the securities are positively correlated and negative when negatively correlated. If $v_1 > v_2$, p decreases in absolute value when the correlation becomes more positive or more negative. We are now ready to introduce the

⁸ Another possibility is to have nondiagonal terms in P with opposite signs.

optimal *alpha-driven* (in excess of their participation in the market portfolio) investment policy in these two securities.

Proposition 2. *In a setting with two securities whose priors are cross-correlated, but uncorrelated with the prior on the expected return of the market, the optimal holdings are given by*

$$\hat{\pi}_0 = \frac{\bar{\mu}_0 - r}{\sigma_0^2 A_0} - \beta_1 \hat{\pi}_1 - \beta_2 \hat{\pi}_2 \quad (25)$$

$$\hat{\pi}_1 = \frac{1}{\sigma_{\varepsilon_1}} \left[\left(\frac{p^2}{A_1} + \frac{(1-p^2)}{A_2} \right) \frac{\bar{\alpha}_1}{\sigma_{\varepsilon_1}} - p \sqrt{1-p^2} \left(\frac{1}{A_2} - \frac{1}{A_1} \right) \frac{\bar{\alpha}_2}{\sigma_{\varepsilon_2}} \right] \quad (26)$$

$$\hat{\pi}_2 = \frac{1}{\sigma_{\varepsilon_2}} \left[\left(\frac{(1-p^2)}{A_1} + \frac{p^2}{A_2} \right) \frac{\bar{\alpha}_2}{\sigma_{\varepsilon_2}} - p \sqrt{1-p^2} \left(\frac{1}{A_2} - \frac{1}{A_1} \right) \frac{\bar{\alpha}_1}{\sigma_{\varepsilon_1}} \right] \quad (27)$$

Proof. In Appendix A.3.

Optimal portfolios are explained by two terms, which depend on $\bar{\alpha}_i/\sigma_{\varepsilon_i}$, the *appraisal ratio* (see Treynor and Black 1973) perceived by the agent. Because $A_1 > A_2$, the sign of the term that includes the other security's appraisal ratio depends on the sign of p (this term is nonpositive when priors are positively correlated and nonnegative when priors are negatively correlated). It might be optimal to hold a security even when its alpha is negative, if the correlation of its prior alpha and the prior alpha of the other security is negative, and the alpha of the other security is positive and sufficiently large. We introduce the following corollary to analyze this result further.

Corollary 2. *In the setting described above, when the means of the priors on the alphas of both securities are positive ($\bar{\alpha}_i > 0$), and the priors are negatively correlated ($p < 0$), the following results hold:*

1. *Optimal investment in each security is higher than in the uncorrelated case.*
2. *Optimal investment in one stock increases with an increase in the perceived appraisal ratio of the other stock.*

Proof. In Appendix A.4.

The first part of the corollary is intuitive and shows that investors diversify their portfolios to take advantage of the correlation in the priors,

as is also the case with the correlation in their returns. The second part of the corollary suggests complementarity of the investment in the two securities.

The intuition behind these results is that prior correlation induces two effects.

The first effect is a “cross-learning” effect: learning about a given alpha is faster when learning occurs simultaneously for a correlated alpha, and, hence, it is equivalent to decreasing the uncertainty around each individual alpha. Consequently, cross-learning induces an incentive to overweight both securities relative to the uncorrelated case. For instance, the positive overweighting due to cross-learning for security 1 is defined as the difference between the first term in (26) and the corresponding term in the case of zero correlation and is equal to

$$\frac{1}{\sigma_{\varepsilon_1}} \left[\left(\frac{p^2}{A_1} + \frac{(1-p^2)}{A_2} \right) - \frac{1}{A_1^0} \right] \frac{\bar{\alpha}_1}{\sigma_{\varepsilon_1}},$$

where $A_1^0 = a - (1-a)Tv_1$. This term increases with the appraisal ratio of the security and decreases with idiosyncratic risk (see Appendix A.4).

The second effect represents hedging of the estimation risk. When the correlation between priors is negative, the investor holds more of one security the higher the appraisal ratio of the other security, in case of a mistake in the estimation of the appraisal ratio. For instance, the incremental weight of security 1 with respect to the uncorrelated case is given by

$$\frac{1}{\sigma_{\varepsilon_1}} p \sqrt{1-p^2} \left(\frac{1}{A_1} - \frac{1}{A_2} \right) \frac{\bar{\alpha}_2}{\sigma_{\varepsilon_2}}.$$

It turns out that when alpha priors are negatively correlated, both the cross-learning effect and the hedging effect give the same overweighting incentives. When alpha priors are positively correlated, the two effects imply opposite incentives, as the following corollary shows.

Corollary 3. *In the setting described above, when the mean of priors on the alphas of both securities are positive ($\bar{\alpha}_i > 0$), and the priors on the alphas are positively correlated ($p > 0$), the following results hold:*

1. *Optimal holdings in each security may be higher or lower than in the uncorrelated case.*
2. *Optimal holdings in one security decrease with an increase in the perceived appraisal ratio of the other stock.*

Proof. Straightforward to verify.

The first part of the corollary reflects the conflicting incentives of the cross-learning effect and the hedging effect. For example, it may be

optimal to short security 1 even if its expected alpha is positive when the expected alpha of security 2 is very high. Under such circumstances, optimal holdings in security 1 will obviously be lower than in the uncorrelated case. On the contrary, optimal investment in security 2 may be higher than in the uncorrelated case because the cross-learning effect dominates the hedging effect. In that sense, the relative expected alpha is more important than the absolute expected alpha.

The second part of the corollary shows that the hedging effect gives an incentive to replace investment in security 1 with investment in security 2.

1.2.2 An example. To gain some insight into the implications of the correlation of priors on optimal portfolio weights, we compute optimal portfolios in a setting that consists of the risk-free asset, the market portfolio and the Fama and French small minus big (SMB) and high minus low (book-to-market ratio) (HML) portfolios. We mimic the calibration from Brennan and Xia (2001): we take the statistics for the Fama and French portfolios, collected in Table 1, from their paper. Optimal portfolios for a twenty-year investment horizon are presented in Table 2. Following the approach in Brennan and Xia (2001), we use the returns variance/covariance matrix divided by the sample size as the prior variance/covariance matrix. The resulting correlation coefficients are reported in the last three columns of table 1. These are the correlations we use to derive the optimal portfolios of table 2. Because we have an explicit expression, we observe that the numerical method in Brennan and Xia (2001) overestimates the weight to the market (depending on the degree of risk aversion, from 3% to 8% additional proportion of wealth), and SMB (from 14% to 20% additional weight), whereas it underestimates the weight in HML (from 13% to 17%). We point out that the hedging magnitude for SMB is less than half the hedging demand if no prior correlation is taken into account. This suggests that the effect of prior correlation is substantial, especially

Table 1
Fama-French portfolio statistics^a

	Mean excess return (%)	Standard deviation (%)	Standard deviation of the mean (%)	Correlation market	Correlation SMB	Correlation HML
Market	5.21	15.7	2.94	1	0.32	-0.38
SMB	3.25	10	1.89	0.32	1	-0.08
HML	4.78	8.8	1.65	-0.38	-0.08	1

^aFama and French portfolio statistics for the period July 1963–December 1991, as reported in Brennan and Xia (2001). In Table 2, we compute the optimal allocation with parameter uncertainty across these portfolios. We use the mean excess return for the prior on the alpha of the corresponding portfolio. We use the standard deviation of the mean for the standard deviation of the prior on the corresponding alpha. Finally, we use the correlation among portfolios for the correlation among priors.

Table 2
Optimal weight in the market, SMB and HML portfolios^a

A: Myopic demand				
a	2	3	4	5
Market	1.7793	1.1862	0.8896	0.7117
SMB	1.0399	0.6933	0.5200	0.4160
HML	4.3871	2.9247	2.1935	1.7548
B: Hedging demand with prior correlation				
a	2	3	4	5
Market	-0.4622	-0.3781	-0.3068	-0.2559
SMB	-0.2701	-0.2210	-0.1793	-0.1496
HML	-1.1395	-0.9322	-0.7564	-0.6310
C: Optimal demand with prior correlation				
a	2	3	4	5
Market	1.3171	0.8081	0.5829	0.4558
SMB	0.7698	0.4723	0.3407	0.2664
HML	3.2476	1.9925	1.4371	1.1239
D: Hedging demand without prior correlation				
a	2	3	4	5
Market	-0.6827	-0.5371	-0.4286	-0.3542
SMB	-0.0938	-0.0940	-0.0821	-0.0712
HML	-1.4884	-1.1861	-0.9517	-0.7889
E: Optimal demand without prior correlation				
a	2	3	4	5
Market	1.0966	0.6491	0.4610	0.3575
SMB	0.9461	0.5993	0.4379	0.3448
HML	2.8987	1.7386	1.2419	0.9659

^aOptimal weight in the market, SMB and HML portfolios for a 20-year horizon investor with CRRA utility with risk-aversion parameter a . In determining the optimal allocation, we use the statistics reported in Table 1. For the optimal portfolios with prior correlation, the correlation coefficients are in the last three columns of Table 1. The numbers represent the proportion of the total portfolio invested in the corresponding portfolio. The balance to 1 of the total demand corresponds to investment in the risk-free security. We assume a constant interest rate equal to zero. Myopic demand is the optimal portfolio for an investor who acts as if the expected returns are known and equal to their mean. Myopic demand plus hedging demand equals the optimal demand. We consider two cases: correlated and uncorrelated priors.

for slow degrees of risk aversion. In our base case (Table 2), the magnitude of the hedging demand for the index could be as high as 35% of the myopic allocation. We also find that the correlation among the alpha priors is one important factor that affects the portfolio weight. In our base case, taking into account the prior correlation can magnify the magnitude of the hedging demand by as much as 188% (for the SMB) or reduce it by as much as 32% (for the index). More generally, this example illustrates the massive effect of cross-learning and hedging.

2. Applications Using Analysts' Recommendations

The key parameter in the optimal asset allocation formula derived in the previous section is the alpha of the stock. To test the usefulness of the formula, we need estimates of the alphas of available securities. An obvious candidate is analysts' buy/sell recommendations, whose objective is to point out stocks whose prices are out of equilibrium. In this section, we use analysts' recommendations and perform two different (and complementary) exercises, which we label "ex ante" and "ex post," respectively. For both types of exercise, we use analysts' recommendations as estimates of the alphas. Because analysts' recommendations are only expressed in qualitative terms (strong buy, buy, hold, sell, and strong sell), we need a mapping to transform the recommendations into numerical alphas, which we will explain later. Once we have the alphas, the formula provides us with an optimal asset allocation rule.

In the "ex ante" exercise, the question we ask is whether the formula, which computes the optimal asset allocation strategy for an investor with CRRA preferences, is actually useful for practical purposes. We follow the idea of Kandel and Stambaugh (1996), who studied the welfare implications of sample evidence of predictability in stock returns. More explicitly, the exercise we perform is the following: let us assume that preferences are of the CRRA type; then, we compute what the difference in utility would be between the optimal allocation strategy and some alternative "naive" strategy, for given parameter values, estimated from the data: the particular strategies we consider are versions of "go long the buy recommendations and short the sell recommendations." This exercise gives us a measure of the usefulness of the formula for actual financial values.

In our "ex post" exercise, we try to address the economic usefulness of analysts' recommendations more directly. Some papers on analysts' recommendations have shown that stocks favored by analysts outperform stocks disfavored by analysts based on ad-hoc portfolio construction rules (see for example, Barber et al. 2001, and Jegadeesh et al. 2004, among others). Although very informative on the quality of analysts' forecasts, these studies do not address the issue of their economic value. Because our "ex ante" exercise assumes that the alphas derived from analysts' recommendations are correct, we complement it with an "ex post" exercise. As in the static "ex ante" exercise, we take analysts' recommendations as estimates of alpha priors, but in this exercise, we update portfolios as analysts update their recommendations and we compute, "ex post," the actual utility resulting from following the dynamic strategy implied by the time series of analysts' recommendations.

The data, the mapping of the recommendations into alphas, and the parameter estimation, are common to both exercises. We describe them

next. We then explain our results for the “ex ante” and “ex post” exercises, respectively.

2.1 The Data set

We use IBES data on analysts’ recommendations, which are reported every month, to compute alphas. We explain the mapping of recommendations into alphas in the next subsection. We use CRSP monthly returns for individual stocks. We use CRSP value-weighted index monthly returns for the market portfolio. Finally, we use one-month Treasury bill returns, also from CRSP, as a proxy for the risk-free rate. The time frame for our study is constrained by the available time series from IBES, which covers the period November 1993 to December 2003.

The Summary History-Recommendation file from the IBES database contains a monthly snapshot of each company followed by sell-side analysts whose brokerage firm provides data to IBES. This database tracks, at mid-calendar month, the number of analysts following the stock, the average consensus rating level on a 1 to 5 scale (where 1 is a “strong buy,” and 5 is a “strong sell”) and its standard deviation for the stock, and the number of analysts upgrading and downgrading their rating in the month.

The total number of stocks in the database is 10,660. Of these, only 7895 have returns available in CRSP. We then filter out all stocks for which the number of covering analysts is too low. We exclude from the database all stocks for which the average coverage over the period is lower than 5.⁹ The average number of analysts for the remaining 2280 stocks is 9.62, and the standard deviation is 5.32. The average recommendation (usually called “the consensus”) is 2.03, corresponding approximately to a “buy,” (a “buy” recommendation is referenced as a 2), with a low dispersion (standard deviation equal to 0.38). The average dispersion around the consensus in our database is 0.63, with a standard deviation equal to 0.17.

2.2 Computing alphas and estimating other parameters

We use the time-series of analysts’ average recommendations, or consensus, as a proxy for the time dynamics of stock i alpha $[\bar{\alpha}_i(t)]$. One problem is that analysts do not provide the market with estimates for alphas but rather with qualitative recommendations. Consider the scaling factor ω (omega): We map analysts’ recommendations into alphas by applying the following scale: 1 = Strong Buy, corresponds to $+2\%/\omega$ annual alpha, 2 = Buy, corresponds to $+1\%/\omega$ annual alpha, 3 = Hold, corresponds

⁹ We want to have stocks covered by a large number of analysts so that we can obtain a good estimate of the standard deviation of the alpha prior, by using the dispersion across analysts. Using a threshold equal to 5 (respectively, 4, 3, 2, 1) leads to keeping 39% (respectively, 48%, 60%, 76%, 100%) of the stocks in the database. Five is not a very large number, but we get to use a good number of securities.

to +0% annual alpha, 4 = Sell, corresponds to $-1\%/ \omega$ annual alpha, 5 = Strong Sell, corresponds to $-2\%/ \omega$ annual alpha. The scaling factor ω allows us to perform a comparative static analysis to check the robustness of our mapping (see tables 3, 5, and 6). This method is referred to as “raw alpha” in the tables below.

Financial analysts’ recommendations are biased, which can be seen from the fact that they are “optimistic” as a whole: the average recommendation across stocks turns out to be 2.03, which approximately corresponds to a “buy” recommendation.¹⁰ Because the purpose of the “ex post” exercise is to assess the economic usefulness of the analysts’ recommendations, for that exercise we extend our analysis in two directions, to address this concern over analysts’ optimism.

We first subtract from each alpha the mean value of alpha across the 2280 stocks, so that we center alpha estimates

$$\alpha_i \rightarrow \alpha_i - \frac{\sum_{j=1}^n \alpha_j}{n},$$

where $n = 2280$. This method is referred to as “centered alpha” in Table 5. This method allows us to obtain an average alpha which is equal to zero by construction.

Secondly, we also consider changes in recommendations, rather than recommendations themselves, in the “ex post” exercise. In other words, rather than using the actual raw alphas, we use changes in raw alphas. By considering trading strategies that buy upgraded stocks and sell downgraded stocks, we follow the guidelines of some recent research (e.g., Jha, Lichtblau and Mozes, 2003, and Jegadeesh et al., 2004) that has actually shown that changes in recommendations are more informative than recommendations themselves. The rationale is that while the level of analysts’ recommendations on a given stock may be tied to the analysts’ own interests, and therefore may be less credible to investors, changes in analysts’ recommendations on a given stock are more likely to reflect their changing perception of the stock’s fundamentals. This method is referred to as “change in alpha” in Table 5.

Another set of parameters we need in our model is the variance–covariance matrix of the alphas’ priors. With respect to the variances, a possible estimate would be the dispersion of analysts’ recommendations, readily available from the IBES database. However, very often all

¹⁰ Several interpretations have been offered for this. For example, McNichols and O’Brien (1997) found that consensus recommendations are biased because optimistic analysts are more likely to provide recommendations than pessimistic analysts. Then, Dugar and Nathan (1995) and Lin and McNichols (1998) argued that analysts appear to favorably bias their recommendations for firms that have underwriting relationships with their brokerage firms, an interpretation that has been confirmed by recent concern over conflicts of interest in different lines of business within investment banks.

recommendations on a given stock are identical. For that reason, we use instead the average value of dispersion of analysts' recommendations over the whole sample period for each security. We estimate it at 0.066257%, which for $\omega = 15$ corresponds approximately to 1%, versus the 2% for a strong buy recommendation. Additionally, we assume that priors are uncorrelated.¹¹

We assume that analysts have superior information about company-specific risk and that is why investors would use alphas based on their recommendations. However, investors, like analysts, have incomplete information about market risk, and we therefore also have to deal with the estimates of $\bar{\mu}_0(t) = E[\mu_0|\mathcal{F}_t^S]$ and $\text{Var}[\mu_0|\mathcal{F}_t^S]$. At the initial date, we use the sample mean of the CRSP value-weighted index in the first three years' worth of data as an estimate for the uncertain expected return on the market. We then use the fact that this estimator is asymptotically normally distributed, with a mean equal to the true value and a standard deviation equal to $\frac{\sigma_0}{\sqrt{n}}$, where n is the sample size (36 monthly data points here), and σ_0 is the estimate for the market volatility. Finally, we use a Bayesian update of the market's expected return and the uncertainty around that value based on prices.

As we explained above, we use the time series of analysts' recommendations and assume that they are consistent with Bayesian updating based on the superior information on firms' specific risk.

Finally, for individual security parameter estimates (variances and covariances), we use the previous three years of data for each month.

2.3 “Ex ante” exercise

In this exercise, we wish to have an idea of the usefulness of the optimal investment strategy derived in this article by comparing the expected utility resulting from the optimal strategy with the expected utility from some simpler strategy, which we call “naive.” To determine the utility loss due to the sub-optimal strategy, we compute the certainty equivalent, as first suggested in McCulloch and Rossi (1990). The basic idea of the “naive” strategy we consider here involves taking a long position in securities with a “buy” recommendation and a short position in securities with a “sell” recommendation. For the optimal strategy, we take analysts' recommendations as indicators of alphas (the alphas we have defined as raw alphas). Additionally, we estimate the values of the other parameters in the model as explained before. We then compute the monetary compensation (“certainty equivalent”) which would make the expected utility of a risk-averse investor who follows the “naive” strategy equal to the expected utility that the same investor would

¹¹ Because the objective of the “ex post” exercise is to evaluate the usefulness of analysts' recommendations, we have performed several extensions where we have used different variances across securities and positive correlations, but the certainty equivalent only increases marginally, so we do not report these results.

achieve using the optimal strategy derived in this article. More explicitly, suppose that we have n_L securities with positive recommendations and n_S with negative recommendations. Additionally, we denote by x_L, x_S, x_M , and x_f , the constant proportions of the total wealth of the investor to be allocated in the securities with positive recommendations (long positions), negative recommendations (short positions), the market portfolio, and the risk-free security, respectively. Therefore, $x_L - x_S + x_M + x_f = 1$. We assume that the investor will invest a constant proportion, x_L/n_L , in each security with positive recommendation and will shortsell each security with a negative recommendation for an amount equal to a constant proportion x_S/n_S . Therefore, the expected utility of an investor who follows this strategy depends on x_M, x_f, x_L, x_S, n_L , and n_S . In the Appendix we derive the expected utility of an investor with CRRA and degree of risk aversion a , for given parameter values. Expected utility according to the “naive” strategy is given by Equation (A19). The expected utility resulting from the optimal asset allocation decisions from Equations (17) and (18), also for given parameter values, is given in Equation (A21).¹² For a given initial wealth $X(0)$, we find the amount ξ that equals the expected utility for $X(0) + \xi$ resulting from Equation (A19) to the expected utility that would result from (17) and (18) for $X(0)$. We normalize $X(0) = 1$ so that we report the percentage increase in wealth necessary for the “naive” strategy to match the expected utility that results from the optimal strategy.

We perform several comparisons (for different “naive” strategies). For all of them, to obtain robust results, we construct a time series of expected utilities according to each method, “naive” strategy and optimal strategy. Because we have monthly data, we repeat the exercise every month. So, after each set of monthly recommendations, for our parameter estimates, we have a measure of the certainty equivalent that equals the expected utility of the “naive” and optimal strategies. The numbers we report are the average of the two time series.

In Tables 3 and 4 we report, for different degrees of risk aversion, and different horizons, the increase in utility that results from switching from a “naive” strategy to the optimal strategy. We assume that the investor has perfect information about the market expected return (computed as explained before). Additionally, we assume that the proportion of wealth invested in the market portfolio in the “naive” strategy, x_M , is the optimal allocation in the Merton (1971) model.¹³ This makes the comparison between “naive” and optimal strategies more appropriate. Additionally,

¹² We point out that this “naive” strategy involves continuous rebalancing across securities, so as to keep proportions constant. A more natural type of “naive strategy” would be a “buy and hold” strategy. However, a “buy and hold” strategy would imply states in which wealth is negative, for which the CRRA utility is not well defined, and it is not feasible to compute the certainty equivalent.

¹³ This is also the allocation to the market portfolio in the optimal strategy resulting in the case of perfect information about the market portfolio parameters and no mispricing.

Table 3
Certainty equivalent in the “ex ante” exercise^a

A: positive <3, negative ≥3

a	Horizon (%) ($\omega = 15$)			Horizon (%) ($\omega = 25$)			Horizon (%) ($\omega = 40$)		
	1	3	5	1	3	5	1	3	5
2	23.93	94.81	215.73	8.63	28.64	53.06	3.97	12.55	22.07
5	11.26	37.96	71.49	5.80	18.47	32.72	4.06	12.74	22.24
8	9.85	32.58	60.02	6.57	21.10	37.72	5.52	17.62	31.29

B: positive < 2.5, negative ≥ 3.5

a	Horizon (%) ($\omega = 15$)			Horizon (%) ($\omega = 25$)			Horizon (%) ($\omega = 40$)		
	1	3	5	1	3	5	1	3	5
2	24.59	97.90	223.79	9.22	30.70	57.07	4.53	14.35	25.28
5	12.79	43.85	84.05	7.26	23.53	42.45	5.49	17.55	31.19
8	12.31	41.94	79.87	8.95	29.64	54.78	7.88	25.92	47.55

^aWe plot, for different degrees of risk aversion and different investment horizons, the utility gain of an investor with CRRA utility with risk aversion parameter “a” due to the use of the optimal investment strategy versus a naive strategy consisting of taking a long position in average “buy” recommendations and a short position in average “sell” recommendations, and fixed holdings in the markets and risk-free securities. The parameter ω determines the mapping of analysts’ recommendations into alphas, as described in section 2.2. For both panels A and B, the holdings in the market portfolio are those which would correspond to the Merton (1971) optimal allocation; additionally, the investor takes a long position of 80% (equally weighted across securities) in individual stocks with a positive recommendation and a short position of 20% (equally weighted) in securities with a negative recommendation; the balance is invested in the risk-free security. In panel A, we consider positive a consensus recommendation lower than three and negative a consensus recommendation equal or higher than three (for individual recommendations, one is strong buy and five strong sell). In panel B, we consider positive a consensus recommendation lower than 2.5 and negative a consensus recommendation higher than 3.5.

the certainty equivalent will be higher than the one we compute if the allocation to the market portfolio deviated from this optimal proportion.

In Table 3, we assume a “naive” strategy that invests a fixed proportion of 60% in individual stocks that have alphas different from zero (according to analysts’ recommendations), with a breakdown of a long position equivalent to 80% of the value of the portfolio and a short position equal to 20% of the value of the portfolio. The allocation to the market portfolio is as explained before. The allocation into the risk-free security (long or short) is the corresponding balance. We consider three possible mappings of the analysts’ recommendations into alphas, given by the values of the parameter ω , defined before. For panel A, the “naive” strategy involves taking a long position in all securities with a recommendation lower than 3 and a short position in all securities with a recommendation equal to or greater than 3 (we include value 3 here to compensate for the fact that most analysts recommendations are 1 and 2). In panel B, the “naive” strategy involves holding a long position on stocks with an average

Table 4
Certainty equivalent in the “ex ante” exercise^a

A: $x_L = 60\%$; $x_S = 20\%$			
	Horizon (%)		
a	1	3	5
2	8.04	26.60	49.15
5	4.21	13.22	23.09
8	3.95	12.34	21.44

B: $x_L = 40\%$; $x_S = 10\%$			
	Horizon (%)		
a	1	3	5
2	7.84	25.94	47.89
5	3.60	11.27	19.60
8	2.93	9.06	15.59

C: $x_L = 110\%$; $x_S = 20\%$			
	Horizon (%)		
a	1	3	5
2	10.03	33.58	62.74
5	9.51	31.47	58.02
8	12.74	90.30	84.46

^aWe plot, for different degrees of risk aversion and different investment horizons, the utility gain of an investor with CRRA utility with risk aversion parameter a due to the use of the optimal investment strategy versus a naive strategy consisting of taking a long position in average “buy” recommendations and a short position in average “sell” recommendations. A “buy” recommendation is a consensus recommendation lower than three, and a “sell” recommendation is a consensus recommendation equal or higher than three (for individual recommendations, one is strong buy and five strong sell). Both long and short positions in individual securities are equally weighted across the class (buy or sell) of securities. Additionally, the holdings in the market portfolio are those which would correspond to the Merton (1971) optimal allocation. The balance is invested in the risk-free security. We take a value of 25 for the parameter ω , which determines the mapping of analysts’ recommendations into alphas, as described in section 2. We consider different allocations. x_L represents the proportion of the total portfolio invested in a long position in securities with a buy recommendation and x_S the proportion of the total portfolio invested in a short position in securities with a sell recommendation.

recommendation equal to or less than 2.5 and a short position in stocks with an average recommendation equal to or greater than 3.5. Certainty equivalents are higher in panel B, due to the use of more securities, but just marginally, because their alphas are small (in absolute value). Overall, certainty equivalents are substantial in general and in some cases huge. Additionally, certainty equivalents increase with alphas (which is the result of a decrease in the parameter ω), because the optimal strategy maximizes the utility resulting from a better investment opportunity set. Certainty equivalents are the highest for the lowest risk aversion, $a = 2$.

Obviously, because the “naive” strategy is arbitrarily constrained, the opportunity cost is higher for the low-risk aversion investor, who would take greater advantage of the investment opportunities. As expected, the certainty equivalent is higher the higher the horizon, because the optimal strategy greatly benefits from learning.

In Table 4, we present the utility gain for different breakdowns of the “naive” strategy across stocks with positive or negative recommendations. Other assumptions are like in panel A of Table 3, for $\omega = 25$. The results are consistent with those in Table 3. However, we observe that when the long position on securities with positive alphas is very high (as, for example, in panel C of Table 4), the certainty equivalent is highest for the investor with the highest risk aversion, which indicates that the arbitrary allocation across individual securities greatly differs from what a high-risk averse investor would optimally choose.

2.4 “Ex Post” exercise

In the “ex ante” exercise, we tried to assess the usefulness of the formula derived in this article. In this exercise, we try to derive some conclusions about the usefulness of the analysts’ recommendations themselves. Our objective is to measure the effect of recommendations on the actual utility realization of an investor who uses the time series of recommendations for dynamic updates of the portfolio set up according to the formula derived in this article. The alternative we consider here is an optimal passive strategy that splits portfolio holdings between the market and the risk-free asset as described in Merton (1971).¹⁴ As in the “ex ante” exercise, we compute the certainty equivalent that added to initial wealth $X(0) = 1$ yields the same utility for the passive strategy we just described as the optimal strategy. Because the formula is expressed in expected utility terms, we compute the “ex post” expected utility, as we explain later.

The main theoretical problem we face to translate our model to this setting is the fact that in our model, the time series of alphas are the result of pure Bayesian updating using security prices. However, in practice, we observe the whole time series of analysts’ recommendations, and this time series does not have to match the time series that would result from taking the initial analyst recommendation, expressing it as the initial alpha, and updating it according to the information resulting from prices. To reconcile the alphas resulting from analysts’ recommendations with the time series that would result from pure Bayesian updating, we assume (investors assume) that the time series of analysts’ recommendations is the result of their Bayesian update of the specific risk of the securities. This assumption is supported by recent findings by Markov and Tamayo (2003) who reported evidence that the serial

¹⁴ When we say “Merton (1971), allocation to the market and the risk-free asset”, we actually mean the extension of Merton’s approach to a setting with incomplete information on the market’s expected return.

correlation pattern in analysts' quarterly earnings forecast errors is consistent with an environment in which analysts face parameter uncertainty and learn rationally about the parameters.¹⁵ We argue that analysts are more informed on the specific risk of the company, which we model as the Brownian motion process W_{σ_i} , than the investors. On the contrary, we assume that analysts, like investors, do not observe the market risk, the factor that we model as the Brownian motion process W_0 in Equation (2).

The remaining problem is the fact that we have data corresponding to a single realization of the random returns, while our model considers an expected utility setting. We address this problem by randomly forming portfolios and averaging across them. In particular, we randomly group securities into ten thousand portfolios of fifty securities each. For each portfolio, we compute the utility resulting from optimal allocation among the market, the risk-free security and the securities in the portfolio. We finally estimate the "ex post" utility as the average of the ten thousand realizations. With respect to the passive strategy, we use the single path we observe and compute the final utility to get the certainty equivalent.¹⁶

First, for the base case, we use the first three years of the sample to estimate parameter values and then the last six years to compute wealth realizations, with monthly updates. The horizon is therefore six years and the risk aversion, $a = 5$. We get ten thousand realizations (one for each portfolio), and we average across them. These results are presented in Table 5, for different values of ω , the parameter that characterizes our mapping of recommendations into alphas and for different computations of alphas, as explained above. Analysts' recommendations seem to generate small positive utility gains, with a maximum reached for some value of omega. The results are similar for all three methods. As expected, we verify that the certainty equivalent gain goes to zero as ω increases because the active portfolio converges to the passive portfolio.

We perform another exercise. For each of the portfolios we use in the previous exercise, we rank the securities according to the number of analysts covering each of them. We then split the portfolio in two, including the upper half of the list in one portfolio and the lower half in another. We then repeat the exercise with the portfolios that have securities with high coverage and then, independently, with the portfolios that have securities with low coverage. We use the parameters in our base case and do the exercise for all three methods described above to compute the alphas. We report the results in Table 6. When we use "centered alphas" and "raw alphas," recommendations for securities with high coverage are

¹⁵ Bayesian updating is also consistent with the fact that analysts tend to upgrade past winners and downgrade past losers, as reported in Boni and Womack (2004).

¹⁶ We also did some exercises using bootstrapping: the results are not very different, and constructing portfolios appears to make analysts' recommendations look better.

Table 5
Certainty equivalent in the “ex post” exercise^a

ω	Centered alphas (%)	Change in alphas (%)	Raw alphas (%)
15	0.334	0.237	0.288
25	0.233	0.144	0.376
35	0.175	0.104	0.311
45	0.140	0.081	0.258
55	0.116	0.066	0.219
65	0.099	0.056	0.190
75	0.087	0.049	0.167

^aCertainty equivalent utility gains resulting from using the optimal strategy derived in this paper with the alphas derived from analysts’ recommendations, versus following a passive strategy [Merton (1971) allocation to market and risk-free asset, with learning about the market portfolio expected return]. In the table, ω is the parameter that characterizes the mapping of the average analyst recommendation into alpha, as described in section 2. “Centered alphas” means that in order to transform analysts’ recommendations into alphas we normalize around the average alpha, as explained in section 3. “Change in alphas” means that instead of the actual alphas resulting from the mapping of analysts’ recommendations into numbers, we use the actual change in alphas. “Raw alphas” is the case in which we do not make any adjustment.

more useful than recommendations for securities with low coverage. The result is reversed when we use “change in alphas.” The difference in average utility is always statistically significant at the 1% level, except for the case in which we used “centered alphas” (see associated *p*-values in the last column of Table 5).

Our exercise indicates that, for securities with high coverage, using “raw alphas” is the best way to construct portfolios, whereas for securities with low coverage it is “change in alphas.” Our exercise also indicates that the certainty equivalent gain resulting from portfolios of securities with high coverage constructed with “raw alpha” is much higher than the certainty equivalent gain from portfolios of securities with low coverage constructed with “change in alphas.” Interestingly, although previous literature has documented that the predictive power of changes in alphas is stronger than that of alpha levels (for example, Jegadeesh et al., 2004), our exercise suggests that level alpha may be a better input when we consider securities with high coverage. It appears that the competition among analysts increases the quality of their recommendations. A possible explanation might be that in this case recommendations will not be as affected by possible conflicts of interest.¹⁷ This result is interesting because it would seem reasonable to expect higher value in the work of analysts for securities with less coverage.¹⁸

¹⁷ For a review on the literature on this topic see, for example, Bajari and Krainer (2004). They in fact found that for high-tech firms there is no bias on the recommendations of the firms in which the analyst has an investment bank interest.

¹⁸ Womack (1996) finds that analysts’ recommendations have greater value for smaller cap stocks.

Table 6
Certainty equivalent in the “ex post” exercise^a

A: Centered alphas			
ω	High coverage (%)	Low coverage (%)	p-value
15	0.369	0.148	< 0.01
25	0.224	0.155	< 0.01
35	0.161	0.127	0.0150
45	0.125	0.105	0.0696
55	0.103	0.089	0.1465
65	0.087	0.077	0.2200
75	0.075	0.067	0.2797

B: Change in alphas			
ω	High coverage (%)	Low coverage (%)	p-value
15	0.117	0.370	< 0.01
25	0.071	0.224	< 0.01
35	0.051	0.160	< 0.01
45	0.039	0.124	< 0.01
55	0.032	0.101	< 0.01
65	0.027	0.084	< 0.01
75	0.024	0.073	< 0.01

C: Raw alphas			
ω	High coverage (%)	Low coverage (%)	p-value
15	1.775	-1.283	< 0.01
25	1.072	-0.383	< 0.01
35	0.767	-0.188	< 0.01
45	0.598	-0.115	< 0.01
55	0.489	-0.079	< 0.01
65	0.414	-0.059	< 0.01
75	0.359	-0.046	< 0.01

^aCertainty equivalent utility gains resulting from using the optimal strategy derived in this article with the alphas derived from analysts’ recommendations, versus following a passive strategy [Merton (1971) allocation to market and risk-free asset, with learning about the market portfolio expected return]. In the Table, ω is the parameter that characterizes the mapping of the average analyst recommendation into alpha, as described in section 2. The alphas correspond to the “raw alphas” of Table 4 described in section 2. “High coverage” means that at every point in time we take into consideration the average recommendation on the top half of securities of each portfolio we use in the exercise (as described in section 2) in terms of number of analysts covering them. “Low coverage” is similar but for the bottom half of each portfolio. In the last column, we provide the p -value.

3. Extensions

In this section, we briefly examine several directions in which it is possible to extend our results related to the formula (15).

3.1 Multifactor models

The formulas derived in the theoretical part of the article can be extended to the case of more than two factors, which is frequently considered in the literature. More explicitly, we assume that there are n risky assets, with

prices S_i , $i = 1, \dots, n$, and one risk-free asset with interest rate r . The risky assets are driven by $n \geq 2$ independent Brownian motions W_i , so that the market is complete. The first k assets are considered to be the factors, and each of them is driven by a single and different Brownian motion. This entails no loss of generality since, in the complete market case, we can replicate any given portfolio (or factor). Each of the other assets is driven by these k factors and by one additional and different Brownian motion, representing idiosyncratic risk specific to the given asset. More precisely, we assume the following dynamics:

$$\frac{dS_i}{S_i} = \mu_i dt + \sigma_{ii} dW_i, \quad i = 1, \dots, k, \quad (28)$$

$$\frac{dS_{k+i}}{S_{k+i}} = \mu_{k+i} dt + \sum_{j=1}^k \sigma_{k+i,j} dW_j + \sigma_{k+i,k+i} dW_{k+i}, \quad i = 1, \dots, n-k. \quad (29)$$

We denote by $A(t)$ the diagonal matrix which has $1/A_i(t)$ as its diagonal elements, where $A_i(t)$ is as before. Also as before, we perform a matrix decomposition of the variance-covariance matrix of priors. Then, by Theorem 1,

$$\hat{\pi}(t) = (\sigma')^{-1} P' A^{-1}(t) P \bar{\theta}(t).$$

Assume that the priors on the drifts are uncorrelated, and, for simplicity, that $v_1 > v_2 > \dots > v_n$ so that P is the identity matrix. In this case,

$$\delta_i(t) = \frac{v_i}{1 + v_i t}$$

where v_i is the prior variance. For a triangular matrix σ , we can easily compute its inverse σ^{-1} . For illustration, here is σ^{-1} when there are two factors and two additional assets (the general case is an obvious generalization):

$$\sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_{11}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_{22}} & 0 & 0 \\ -\frac{\sigma_{31}}{\sigma_{11}\sigma_{33}} & -\frac{\sigma_{32}}{\sigma_{22}\sigma_{33}} & \frac{1}{\sigma_{33}} & 0 \\ -\frac{\sigma_{41}}{\sigma_{11}\sigma_{44}} & -\frac{\sigma_{42}}{\sigma_{22}\sigma_{44}} & 0 & \frac{1}{\sigma_{44}} \end{pmatrix}.$$

Also, we have $(\sigma')^{-1} = (\sigma^{-1})'$. From this, we can find the matrix $(\sigma')^{-1} A^{-1} \sigma^{-1}$ and, using formula (15), we can find the optimal portfolio proportions. In particular, let us define

$$\beta_{ij} = \frac{\sigma_{ij}}{\sigma_{jj}}, \quad \bar{\alpha}_i(t) = \bar{\mu}_i(t) - r - \sum_{j=1}^k \beta_{ij}(\bar{\mu}_j(t) - r), \quad i = k+1, \dots, n, \\ j = 1, \dots, k.$$

Then, the optimal proportions to be held in the nonfactor assets at time t are given by

$$\hat{\pi}_i(t) = \frac{1}{A_i(t)\sigma_{ii}^2} \bar{\alpha}_i(t).$$

3.2 Residual correlation

Following the empirical literature on Bayesian performance evaluation (See Baks, Metrick, and Wachter 2001 and Pástor and Stambaugh 2002), we have assumed that factor model residuals are noncorrelated. In this literature, these assumptions are made to facilitate the computation of the posteriors (or their simulations in the case where the specification of priors is not conjugate) and also to preclude the estimation problem of the large residual covariance matrix. Perhaps more importantly, the joint effect of the assumptions of no correlation among priors and no correlation among residuals decouples the learning of the alpha of one security from the learning of the alpha of any other security. This is a particularly desirable property from an empirical perspective, because it implies that the history of returns of a given security is sufficient to update its alpha, and therefore the survival bias may be neglected. Stambaugh (2003) and Jones and Shanken (2004) address this issue and incorporate some correlation among alphas and among residuals. We now extend our results to the case of correlation among residuals. Keeping the same dynamics (1) for the market portfolio, we model the n assets prices as

$$\frac{dS_i}{S_i} = \mu_i dt + \sigma_i dW_0 + \sigma_{\varepsilon_{i,1}} dW_1 + \dots + \sigma_{\varepsilon_{i,n}} dW_n, \quad i = 1, \dots, n,$$

where $\sigma_{\varepsilon_{i,j}}$ represents the residual risk covariance. We denote by $\sigma_\varepsilon = (\sigma_{\varepsilon_{i,j}})_{(i,j)}$ the $(n \times n)$ residual volatility matrix formed by stacking the row vectors of residual volatilities of the assets.

Due to residual correlations, the relationships (7) and (8) between alphas and thetas transform into

$$\theta = \begin{pmatrix} \sigma_0^{-1} & 0 \\ 0 & \sigma_\varepsilon^{-1} \end{pmatrix} \begin{pmatrix} \mu_0 - r \\ \alpha \end{pmatrix}, \quad (30)$$

where $\alpha = (\alpha_1, \dots, \alpha_n)'$ is the column vector of alphas. Thus, the assumption of no correlation among the priors of risk premia (thetas) is different from the assumption of no correlation among the priors of alphas and expected market return.

To isolate the effect of residual correlation, we state our results in a context of no correlation among the priors on alphas and expected market portfolio return. In the following proposition, we denote by η_0 the variance of the prior on μ_0 , by η_i the variance of the prior on α_i , for $i = 1, \dots, n$, and by D_η the $(n \times n)$ diagonal matrix $\text{Diag}(\eta_1, \dots, \eta_n)$.¹⁹

Proposition 3. *In the setting described above, with uncorrelated priors on alphas and the expected market return, the optimal investment strategy is given by*

$$\begin{pmatrix} \hat{\pi}_1 \\ \dots \\ \hat{\pi}_n \end{pmatrix} = (a\sigma_\epsilon\sigma_\epsilon' - (1-a)TD_\eta)^{-1}\bar{\alpha}, \quad (31)$$

$$\hat{\pi}_0 = \frac{\bar{\mu}_0 - r}{a\sigma_0^2 - (1-a)T\eta_0} - \sum_i \beta_i \hat{\pi}_i. \quad (32)$$

Proof. In Appendix A.5.

Equation (31) illustrates that the optimal weight in each mispriced security consists of an appropriate combination of the alphas. The structure of the weights is similar to the case analyzed in section 1.2.1. In particular, cross-learning is induced here by the residual correlation. Thus, even if the alpha priors are uncorrelated, the alpha posteriors will in general be correlated, due to residual correlation. The investor then incorporates the future correlation in the optimal allocation and in that sense residual correlations play a similar role to correlation among prior alphas. Equation (32) reflects the weight in the market portfolio and as in the correlated prior case, there is a penalty in the market weight due to mispricing.

3.3 Intertemporal consumption

While a portfolio problem with no intermediate consumption is a good model for the active management industry, consumption withdrawal may be of interest in other contexts. Our methodology may accommodate intertemporal consumption.

¹⁹ Note that the η s are related to the variance of the risk premium by

$$\begin{pmatrix} \eta_0 & 0 \\ 0 & D_\eta \end{pmatrix} = \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix} \Delta \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_\epsilon \end{pmatrix}'.$$

We assume that the utility function of the investor is

$$\int_0^T e^{-\kappa t} \frac{c_t^{1-a}}{1-a} dt,$$

where κ is a time discount rate. We show in Appendix A.6 that the optimal strategy for the model (1)–(8) at time $t = 0$ is

$$\pi^c = \int_0^T \phi(0, u, m, D) \pi(0, u) du, \quad (33)$$

where

$$\pi(0, u) = (\sigma')^{-1} P' A^{-1}(0, u) P m,$$

denoting by $A^{-1}(0, u)$ a diagonal matrix whose i -th element on the diagonal is given by

$$\frac{1}{a - (1-a)d_i u},$$

and where ϕ is a nonnegative weighting function defined in Appendix A.6 and satisfying

$$\int_t^T \phi(0, u, m, D) du = 1.$$

Note that Theorem 1 implies that $\pi(0, t)$ is the time zero optimal strategy of a fictitious terminal wealth investor who does not care about consumption before the horizon t and with identical risk aversion. Thus, expression (33) allows us to interpret the optimal strategy for the withdrawal consumption investor as a weighted average of the optimal strategies of a continuum of terminal wealth investors with a horizon that increases from 0 to T and with otherwise identical characteristics. Although it is outside the scope of this article, it is interesting to note that Wachter (2002) obtained the same relationship between the withdrawal consumption and terminal wealth investors' portfolio strategies when expected returns are mean reverting.

Finally, as the investing lifetime goes toward infinity, the optimal portfolio weight converges monotonically to a positive constant for the intermediate consumption case, unlike the terminal wealth case where the optimal weight converges to zero. This constant investment in the risky assets is required to support a subsistence consumption level.

4. Conclusions

In this article, we present an analytic expression for the optimal portfolio of a nonmyopic agent that has incomplete information about the abnormal expected returns (“alphas”) of the securities in the opportunity set. Our theoretical result can be interpreted as a dynamic version of the Treynor and Black (1973) formula. We show that the hedging component is an important part of the optimal portfolio, especially for long horizons. Our formula only requires simple computations and this allows us to do two empirical exercises, using analysts’ recommendations on a large sample of stocks. In an “ex ante” exercise, we show that using a naive strategy instead of the formula produces substantial utility losses, especially if the degree of risk aversion is low, and the horizon is long. In an “ex post” exercise, we find that analysts’ recommendations do not appear to be overly valuable. However, recommendations are more useful when a large number of analysts are covering a stock. Finally, we are able to extend our formula to the case of intertemporal consumption, multiple factors, and residual correlation.

Appendix A

A.1 Proof of Theorem 1. We denote by $X(t) = X^{x,\pi}(t)$ the wealth process of the investor using strategy π and starting with initial wealth x . From the definition of $W^*(t) := W(t) + \theta t$, we see that the risky asset prices satisfy the dynamics

$$dS_i(t) = S_i(t) \left[rdt + \sum_{j=0}^n \sigma_{ij} dW_j^*(t) \right]. \quad (\text{A1})$$

From this, we see that the process W^* generates the price filtration $\mathcal{F}^S$. We want to reduce the problem to the full information case. We replace the original Brownian motion with the so-called *innovation process* from the filtering theory and defined by

$$\bar{W}(t) := W(t) - \int_0^t [\bar{\theta}(s) - \theta] ds \equiv W^*(t) - \int_0^t \bar{\theta}(s) ds. \quad (\text{A2})$$

It is known that the $(n+1)$ -dimensional innovation process $\bar{W}(\cdot)$ is a $\mathcal{F}^S$ -Brownian motion. We are in the setting of the well-known Kalman-Bucy filter, and the filtering theory (Lipster and Shiryaev 2001, Thm 10.3) tells us that the conditional expectation $\bar{\theta}$ is given by Equation (16) and that $P'\bar{D}(t)P$ is the conditional variance-covariance matrix of the risk premium θ .

As is also well known, the wealth process in this model satisfies the dynamics

$$dX = rXdt + \pi' X[\sigma\theta dt + \sigma dW].$$

We see that this can also be written as

$$dX = rXdt + \pi' X[\sigma\bar{\theta}dt + \sigma d\bar{W}]. \quad (\text{A3})$$

Therefore, with this formulation we are in the context of full information, because $\bar{\theta}(t)$ is observable. Thus, we can use the methods of the theory developed for the full information case, simply by replacing θ with $\bar{\theta}$ and W with $\bar{W}$. To do this, we consider the “risk-neutral density” process

$$\bar{Z}(t) := \exp \left(- \int_0^t \bar{\theta}'(s) d\bar{W}(s) - \frac{1}{2} \int_0^t |\bar{\theta}(s)|^2 ds \right) \quad (\text{A4})$$

and the “state-price” process

$$\bar{\xi}(t) = e^{-rt} \bar{Z}(t). \quad (\text{A5})$$

It is easily shown by Itô's rule that $\bar{\xi} X^{x,\pi}$ is a P -martingale process with respect to the price filtration as long as $E \left[\int_0^T |\pi(s)|^2 ds \right] < \infty$, a.s. We may now recall and use the martingale/duality approach to utility maximization, as developed by Cox and Huang (1989), and Karatzas, Lehoczky, and Shreve (1987). First, note that we have

$$U'(x) = x^{-a}, \quad I(z) := (U')^{-1}(z) = z^{-\frac{1}{a}}.$$

Then, using the martingale/duality method, the optimal terminal wealth for our problem is given by

$$\hat{X}(T) = c e^{\frac{a-1}{a} r T} I[\bar{\xi}(T)] = \frac{x}{E[\bar{Z}^{\frac{a-1}{a}}(T)]} \bar{Z}^{-\frac{1}{a}}(T) e^{rT}.$$

Here, the constant factor $c := x/E[\bar{Z}^{\frac{a-1}{a}}(T)]$ is chosen so that the budget constraint $E[\hat{X}(T)\bar{\xi}(T)] = x$ is satisfied. Because $\bar{\xi} X$ is a P -martingale with respect to the price filtration, denoting E_t the expectation conditional on $\mathcal{F}_t^S$, this gives the following expression for the optimal wealth process:

$$\bar{\xi}(t) X^{x,\hat{\pi}}(t) = c \cdot E_t \left[\bar{\xi}^{-\frac{a-1}{a}}(T) \right].$$

Denoting by E^* the expectation under the probability P^* under which W^* is a martingale, the Bayes rule (Karatzas and Shreve 1991) can be used to show that

$$E_t^*[Y] = \frac{1}{\bar{Z}(t)} E_t[Y \bar{Z}(T)],$$

for a random variable Y measurable with respect to $\mathcal{F}_T^S$. From this, we get

$$X^{x,\hat{\pi}}(t) = c \cdot E_t^* \left[e^{-r(T-t)} \bar{\xi}^{-\frac{1}{a}}(T) \right] = c e^{rt} e^{-\frac{a-1}{a} r T} \bar{Z}^{-\frac{1}{a}}(t) \cdot Y_{-\frac{1}{a}}(t), \quad (\text{A6})$$

where

$$Y_{\alpha}(t) := E_t^* \left[\left(\frac{\bar{Z}(T)}{\bar{Z}(t)} \right)^{\alpha} \right]. \quad (\text{A7})$$

Thus, to find the optimal wealth we need to compute $Y_{\alpha}(t)$. This is done in the following crucial lemma.

Lemma 1. For all $\alpha \geq -1$, we have

$$Y_{\alpha}(t) = \prod_{i=0}^n g_{\alpha} \{T-t, [P\bar{\theta}(t)]_i, \delta_i(t)\}, \quad 0 \leq t \leq T, \quad (\text{A8})$$

where the function g_{α} is given, for $(\tau, x, y) \in [0, T] \times \mathbb{R} \times (0, \infty)$, by

$$g_{\alpha}(\tau, x, y) = \sqrt{\frac{(1+y\tau)^{\alpha+1}}{1+y\tau(1+\alpha)}} \exp \left\{ \frac{\alpha(1+\alpha)x^2\tau}{2[1+y\tau(1+\alpha)]} \right\}.$$

Proof of the lemma. Note that the process

$$\hat{W}(t) := PW^*(t)$$

is a $(P^*, \mathcal{F}^S)$ Brownian motion (in fact, $\hat{W}$ generates $\mathcal{F}^S$). We can express the process $\bar{Z}$ as

$$\bar{Z}(t) = \exp \left[- \int_0^t \hat{\theta}'(s) d\hat{W}(s) + \frac{1}{2} \int_0^t |\hat{\theta}(s)|^2 ds \right], \quad (\text{A9})$$

where the process $\hat{\theta}$ is defined by

$$\hat{\theta}(t) := P\bar{\theta}(t) = \bar{D}(t) [\dot{W}(t) + \bar{D}^{-1}(0)\hat{m}],$$

where $\hat{m} = Pm$. Recalling the expression for $\bar{D}(t)$, we see that

$$\hat{\theta}_i(t) = \delta_i(t) \dot{W}_i(t) + \frac{\delta_i(t)}{d_i} \hat{m}_i. \quad (\text{A10})$$

Thus, by independence, we have

$$Y_\alpha(t) = \prod_{i=0}^n E^* \left\{ \left[\frac{\bar{Z}_i(T)}{\bar{Z}_i(t)} \right]^\alpha \middle| \mathcal{F}_t^S \right\}, \quad (\text{A11})$$

where

$$\bar{Z}_i(t) = \exp \left[- \int_0^t \hat{\theta}_i(s) d\hat{W}_i(s) + \frac{1}{2} \int_0^t \hat{\theta}_i^2(s) ds \right].$$

To compute the terms on the right hand side of (A11), denote

$$F_i(s) := \frac{d_i}{2} \dot{W}_i^2(s) + \hat{m}_i \dot{W}_i(s),$$

for all $s \in [0, T]$. Then, by Itô's rule,

$$\int_0^t \hat{\theta}_i(s) d\hat{W}_i(s) = \frac{1}{d_i} \delta_i(t) F_i(t) - \frac{1}{d_i} \delta_i(0) F_i(0) + \int_0^t \left[F_i(s) \frac{\delta_i^2(s)}{d_i} - \frac{1}{2} \delta_i(s) \right] ds. \quad (\text{A12})$$

We substitute this last identity in the expression for $\bar{Z}_i$. We get lucky because the W_i terms cancel in the integrand of the ds integral. More precisely, we get

$$\bar{Z}_i(t) = \exp \left\{ - \frac{1}{d_i} \delta_i(t) F_i(t) + \int_0^t \left[\frac{1}{2} \delta_i(s) + \frac{\delta_i^2(s)}{2d_i^2} \hat{m}_i^2 \right] ds \right\}.$$

We are now in a position to compute $Y_{i,\alpha}(0) = E^* [\bar{Z}_i(T)]$. After some integration we obtain

$$Y_{i,\alpha}(0) = \sqrt{(1 + d_i T)^{\alpha+1}} \exp \left(\frac{\hat{m}_i^2 \alpha}{2d_i} \right) E^* \left(\exp \left\{ - \frac{\alpha \delta_i(T)}{2} [\dot{W}_i(T) + \hat{m}_i/d_i]^2 \right\} \right).$$

By integrating against the normal density, we obtain that

$$E \left(\exp \{ -\beta [W(T) + x]^2 \} \right) = \frac{1}{\sqrt{1 + 2\beta T}} \exp \left(- \frac{\beta x^2}{1 + 2\beta T} \right)$$

for each $\beta > -T/2$ and for each x . Using this, we see that $Y_{i,\alpha}(0) = g_\alpha(T, \hat{m}_i, d_i)$. By (A7), we finish the proof of the lemma for the case $t = 0$. A similar proof works for a general value of t , that is, we get $Y_{i,\alpha}(t) = g[T - t, \hat{\theta}_i(t), \delta_i(t)]$.

◇

Finally, we want to justify formula (15). On the one hand, we know that

$$dX = rXd t + X\pi' \sigma dW^*.$$

On the other hand, we can use the computations above to apply Itô's rule in (A6) and obtain another expression for dX . More precisely, we obtain

$$dX(t) = (...)dt + X(t) \sum_{i=0}^n \left[\frac{\alpha(1+\alpha)\delta_i(t)(T-t)}{1+(1+\alpha)\delta_i(t)(T-t)} \hat{\theta}_i(t) - \alpha \hat{\theta}_i(t) \right] d\hat{W}_i(t),$$

with $\alpha = -1/a$ or

$$dX(t) = (...)dt + X(t) \hat{\theta}'(t) A^{-1}(t) d\hat{W}(t).$$

Comparing the dW terms in the two expressions for dX , we easily check that $\hat{\pi}$ is indeed given by (15).

◇

A.2 Proof of Proposition 1. Note first that it is straightforward to compute the following inverse matrix:

$$\sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_0} & 0 & 0 & \dots & 0 \\ -\frac{\sigma_1}{\sigma_0 \sigma_{\varepsilon_1}} & \frac{1}{\sigma_{\varepsilon_1}} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{\sigma_n}{\sigma_0 \sigma_{\varepsilon_n}} & 0 & \dots & 0 & \frac{1}{\sigma_{\varepsilon_n}} \end{pmatrix} \quad (\text{A13})$$

Furthermore $(\sigma')^{-1} = (\sigma^{-1})'$. Using these results and formula (15) with $P = I$, we can find the optimal portfolio proportions at time t as follows, suppressing dependence on t :

$$\hat{\pi}_0 = \left(\frac{1}{\sigma_0^2 A_0} + \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_0^2 A_i \sigma_{\varepsilon_i}^2} \right) (\bar{\mu}_0 - r) - \sum_{i=1}^n \frac{\sigma_i}{\sigma_0 A_i \sigma_{\varepsilon_i}^2} (\bar{\mu}_i - r). \quad (\text{A14})$$

$$\hat{\pi}_i = -\frac{\sigma_i}{\sigma_0 A_i \sigma_{\varepsilon_i}^2} (\bar{\mu}_0 - r) + \frac{1}{A_i \sigma_{\varepsilon_i}^2} (\bar{\mu}_i - r). \quad (\text{A15})$$

We see that these optimal holdings can be expressed as

$$\hat{\pi}_0 = \frac{\bar{\mu}_0 - r}{\sigma_0^2 A_0} - \sum_{i=1}^n \beta_i \hat{\pi}_i \quad (\text{A16})$$

$$\hat{\pi}_i = \frac{\bar{\alpha}_i}{\sigma_{\varepsilon_i}^2 A_i}. \quad (\text{A17})$$

A.3 Proof of Proposition 2. The optimal proportions of wealth held in the three assets are given by

$$\hat{\pi} = \sigma'^{-1} P' \begin{pmatrix} A_0^{-1} & 0 & 0 \\ 0 & A_1^{-1} & 0 \\ 0 & 0 & A_2^{-1} \end{pmatrix} P \bar{\theta}. \quad (\text{A18})$$

where $\sigma'^{-1} = (\sigma^{-1})'$ is given by

$$\sigma'^{-1} = \begin{pmatrix} \frac{1}{\sigma_0} & -\frac{\sigma_1}{\sigma_0 \sigma_{\varepsilon_1}} & -\frac{\sigma_2}{\sigma_0 \sigma_{\varepsilon_2}} \\ 0 & \frac{1}{\sigma_{\varepsilon_1}} & 0 \\ 0 & 0 & \frac{1}{\sigma_{\varepsilon_2}} \end{pmatrix}$$

and

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & \sqrt{1-p^2} \\ 0 & \sqrt{1-p^2} & -p \end{pmatrix}.$$

A straightforward computation of $\sigma' - 1P'A^{-1}P\sigma^{-1}$ gives

$$\begin{aligned} \hat{\pi}_1 &= \left[-\beta_1 \frac{p^2 A_1^{-1} + (1-p^2) A_2^{-1}}{\sigma_{\varepsilon_1}^2} - \beta_2 \frac{p\sqrt{1-p^2}(A_1^{-1} - A_2^{-1})}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} \right] (\bar{\mu}_0 - r) \\ &\quad + \left[\frac{p^2 A_1^{-1} + (1-p^2) A_2^{-1}}{\sigma_{\varepsilon_1}^2} \right] (\bar{\mu}_1 - r) + \frac{p\sqrt{1-p^2}}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} (A_1^{-1} - A_2^{-1}) (\bar{\mu}_2 - r) \\ \hat{\pi}_2 &= \left[-\beta_1 \frac{p\sqrt{1-p^2}(A_1^{-1} - A_2^{-1})}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} - \beta_2 \frac{(1-p^2) A_1^{-1} + p^2 A_2^{-1}}{\sigma_{\varepsilon_2}^2} \right] (\bar{\mu}_0 - r) \\ &\quad + \frac{p\sqrt{1-p^2}}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} (A_1^{-1} - A_2^{-1}) (\bar{\mu}_1 - r) + \left(\frac{1-p^2}{\sigma_{\varepsilon_2}^2} A_1^{-1} + \frac{p^2}{\sigma_{\varepsilon_2}^2} A_2^{-1} \right) (\bar{\mu}_2 - r) \end{aligned}$$

or

$$\begin{aligned} \hat{\pi}_1 &= \frac{p^2 A_1^{-1} + (1-p^2) A_2^{-1}}{\sigma_{\varepsilon_1}^2} [\bar{\mu}_1 - r - \beta_1 (\bar{\mu}_0 - r)] + \frac{p\sqrt{1-p^2}}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} (A_1^{-1} - A_2^{-1}) [\bar{\mu}_2 - r - \beta_1 (\bar{\mu}_0 - r)] \\ \hat{\pi}_2 &= \frac{p\sqrt{1-p^2}}{\sigma_{\varepsilon_1} \sigma_{\varepsilon_2}} (A_1^{-1} - A_2^{-1}) [\bar{\mu}_1 - r - \beta_1 (\bar{\mu}_0 - r)] + \frac{(1-p^2) A_1^{-1} + p^2 A_2^{-1}}{\sigma_{\varepsilon_2}^2} [\bar{\mu}_2 - r - \beta_1 (\bar{\mu}_0 - r)] \end{aligned}$$

from which we obtain the announced result.

A.4 Proof of Corollary 2. Let us assume, with no loss of generality, that $v_1 \geq v_2$ and $t = 0$, and let us denote by A_i^0 the value of A_i for the case of zero correlation, that is

$$A_i^0 = a - (1-a)Tv_i.$$

Because the function $f(x) = 1/[a - (1-a)Tx]$ is convex, and because

$$v_1 = p^2 d_1 + (1 - p^2) d_2, v_2 = p^2 d_2 + (1 - p^2) d_1,$$

we see that

$$p^2 A_1^{-1} + (1 - p^2) A_2^{-1} \geq (A_1^0)^{-1}$$

$$(1 - p^2) A_1^{-1} + p^2 A_2^{-1} \geq (A_2^0)^{-1}.$$

Moreover, we also know that $A_1^{-1} \leq A_2^{-1}$ and that $p \leq 0$ if the correlation is negative. This implies, in case both $\bar{\alpha}_i$'s are positive and correlation is negative, that both $\hat{\pi}_1$ and $\hat{\pi}_2$ will be larger than in the case of zero correlation. The second statement of the corollary is obvious.

A.5 "Naive" strategy. Suppose that we have n_L securities with positive recommendations and n_S with negative recommendations. We are going to consider a constant proportion strategy which requires dynamic rebalancing. We are going to take a long position for a total wealth proportion $x_L = \pi_L n_L$ in the n_L securities, with equal weights across them (the wealth proportion invested in each security is π_L), and a short position for a total wealth proportion $x_S = \pi_S n_S$ in the n_S securities, with equal weights across them (the wealth proportion invested in each security is π_S). Weights in the long and short positions might be equal or not. Additionally, we might take a position x_M in the market portfolio, and a position $x_f = 1 - x_M - x_L + x_S$ in the risk-free security, so that $1 = x_f + x_M + x_L - x_S$.

First, let us compute the derived utility from any constant wealth proportion strategy defined by the weights $\varsigma = (\varsigma_0, \varsigma_1, \dots, \varsigma_n)$. From Itô's Lemma and the relation

$$\theta = \begin{pmatrix} \frac{\mu_0 - r}{\sigma_0} \\ \frac{\alpha_1}{\sigma_{\varepsilon_1}} \\ \cdot \\ \cdot \\ \frac{\alpha_n}{\sigma_{\varepsilon_n}} \end{pmatrix},$$

the terminal wealth derived from the constant proportion strategy ς is found to be

$$X_T = X \exp \left\{ \left[r + \sum_{i=0}^n \varsigma_i \frac{\sigma_i}{\sigma_0} (\mu_0 - r) + \sum_{i=1}^n \varsigma_i \alpha_i - \frac{1}{2} \left(\sum_{i=0}^n \varsigma_i \sigma_i \right)^2 - \frac{1}{2} \sum_{i=1}^n \varsigma_i^2 \sigma_{\varepsilon_i}^2 \right] T \right. \\ \left. + \left(\sum_{i=0}^n \varsigma_i \sigma_i \right) W_0(T) + \sum_{i=1}^n \varsigma_i \sigma_{\varepsilon_i} W_i(T) \right\},$$

where X is the initial wealth. Using the Laplace transform of Gaussian variables and the mutual independence of the Gaussian variables μ_0 , $(\alpha_i)_{i=1, \dots, n}$ and $[W_i(T)]_{i=0, 1, \dots, n}$, we get the derived utility

$$E[U(X_T)] = \frac{X^{1-a}}{1-a} e^{-(a-1)rT} \\ \times \exp \left\{ \left[\sum_{i=0}^n \varsigma_i \frac{\sigma_i}{\sigma_0} (\bar{\mu}_0 - r) + \sum_{i=1}^n \varsigma_i \bar{\alpha}_i - \frac{1}{2} \left(\sum_{i=0}^n \varsigma_i \sigma_i \right)^2 - \frac{1}{2} \sum_{i=1}^n \varsigma_i^2 \sigma_{\varepsilon_i}^2 \right] (1-a)T \right\} \\ \times \exp \left\{ \frac{1}{2} \left[\left(\sum_{i=0}^n \varsigma_i \frac{\sigma_i}{\sigma_0} \right)^2 \kappa_0 + \sum_{i=1}^n \varsigma_i^2 \kappa_i \right] (1-a)^2 T^2 \right\} \\ \times \exp \left\{ \frac{1}{2} \left[\left(\sum_{i=0}^n \varsigma_i \sigma_i \right)^2 + \sum_{i=1}^n \varsigma_i^2 \sigma_{\varepsilon_i}^2 \right] (1-a)^2 T \right\}, \quad (\text{A19})$$

where κ_0 (resp. κ_i) is the prior variance of μ_0 (resp. of α_i).

The naive utility is obtained as a specialization of (A19). Denoting I_+ (resp. I_-) is the set of securities with positive (resp. negative) recommendations, the naive utility is obtained by setting

$$\varsigma_0 = \pi_M, \quad \varsigma_i = \begin{cases} \pi_L & \text{if } i \in I_+, \\ -\pi_S & \text{if } i \in I_-. \end{cases} \quad (\text{A20})$$

in (A19).

In the Ex ante exercise (Tables 3 and 4), we gave the percentage increase of utility that results from switching from the naive utility (A19)–(A20) to the optimal utility given by

$$E[U(\hat{X}_T)] = \frac{X^{1-a}}{1-a} e^{-(a-1)rT} \left[\frac{Y_{\frac{r}{a}}(0)}{\frac{r}{a}} \right]^a \quad (\text{A21})$$

where we recall that $Y_{\frac{r}{a}}(0)$ is defined in (A8).

A.6 Proof or Proposition 3. From Theorem 1 and substituting (30), we get the optimal strategy

$$\hat{\pi} = (\sigma')^{-1} P' A^{-1} P \begin{pmatrix} \sigma_0^{-1} & 0 \\ 0 & \sigma_\varepsilon^{-1} \end{pmatrix} \begin{pmatrix} \mu_0 - r \\ \alpha \end{pmatrix}.$$

Inverting the matrix part of the above expression and substituting the expression for A gives

$$\hat{\pi} = \left(a \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_\varepsilon \end{pmatrix} \right) \begin{pmatrix} \sigma_0 & \sigma_0 \beta' \\ 0 & \sigma_\varepsilon \sigma_\varepsilon' \end{pmatrix} - (1-a) T \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_\varepsilon \end{pmatrix} \Delta \sigma',$$

where $\beta = (\beta_1, \dots, \beta_n)'$ is the column vector of betas. Now, we can use the relationship between Δ and D_η given in Footnote (A.7) and obtain, after some computations,

$$\hat{\pi} = \begin{pmatrix} (a\sigma_0 - (1-a)T\eta_0) & (a\sigma_0 - (1-a)T\eta_0)\beta' \\ 0 & a\sigma_\varepsilon \sigma_\varepsilon' - (1-a)TD_\eta \end{pmatrix}^{-1} \begin{pmatrix} \mu_0 - r \\ \alpha \end{pmatrix},$$

and a direct inversion of this matrix gives weights (31) and (32).

A.7 Intertemporal consumption. In the framework of section 3.3, under the observation filtration, the wealth dynamics are now given by

$$dX = rXdt - c_t dt + \pi' X (\sigma \bar{\theta} dt + \sigma d\bar{W}). \quad (\text{A22})$$

Following similar arguments to those above, we obtain the optimal consumption

$$\hat{c}_t = c_0 \exp^{\frac{r-a}{a}t} \bar{Z}^{-\frac{1}{a}}(t),$$

for an appropriate constant c_0 . Using Lemma 1, this expression in turn gives the optimal wealth process

$$\hat{X}_t = c_0 \exp^{\frac{r-a}{a}t} \bar{Z}^{-\frac{1}{a}}(t) \cdot G[t, P\bar{\theta}(t)], \quad (\text{A23})$$

where function G is defined for all $(t, X) \in [0, T] \times R$ by

$$G(t, X) = \int_t^T \exp\left(\frac{r-a}{a} \cdot (u-t)\right) \prod_{i=0}^n g_{\frac{1}{a}}[u-t, X_i, \delta_i(t)] du.$$

Now, applying Itô's lemma to (A23) one can obtain an expression for dX which gives, after identifying its diffusion term with the wealth dynamics (A22) diffusion term, the optimal strategy with intermediate consumption

$$\pi^c(t) = \int_t^T \phi[t, u, \bar{\theta}(t), D] \pi(t, u) du, \quad (\text{A24})$$

where

$$\pi(t, u) = (\sigma')^{-1} P' A^{-1}(t, u) P \bar{\theta}(t),$$

$$\phi(t, u, \bar{\theta}(t), D) = \frac{\exp\left(\frac{r-a}{a} \cdot (u-t)\right) \prod_{i=0}^n g_{\frac{1}{a}}\{u-t, [P\bar{\theta}(t)]_i, \delta_i(t)\}}{\int_t^T \exp\left(\frac{r-a}{a} \cdot (u-t)\right) \prod_{i=0}^n g_{\frac{1}{a}}\{u-t, [P\bar{\theta}(t)]_i, \delta_i(t)\} du},$$

and where $A^{-1}(t, u)$ is a diagonal matrix whose i -th element on the diagonal is given by $\frac{1}{a - (1-a)\delta_i(t)u}$.

References

- Amenc, N., L. Martellini, and M. Vaissie, 2003, "Benefits and Risks of Alternative Investment Strategies," *Journal of Asset Management*, 4(2), 96–118.
- Bajari, P., and J. Krainer, 2004, "An Empirical Model of Stock Analysts' Recommendations: Market Fundamental, Conflicts of Interest, and Peer Effects," working paper, University of Michigan.
- Baks, K., A. Metrick, and J. Wachter, 2001, "Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation," *Journal of Finance*, 56, 45–85.
- Barber, B., R. Lehavy, M. McNichols, and B. Trueman, 2001, "Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns," *Journal of Finance*, 56, 531–563.
- Barberis, N., 2000, "Investing for the Long Run when Returns are Predictable," *Journal of Finance*, 55, 225–264.
- Black, F., and R. Litterman, 1991, "Global Asset Allocation with Equities, Bonds and Currencies," working paper, Fixed-Income Research, Goldman Sachs.
- Black, F., and R. Litterman, 1992, "Global Portfolio Optimization," *Financial Analysts Journal*, 48(5), 28–43.
- Boni, L., and K. Womack, 2004, "Analysts, Industries, and Price Momentum," forthcoming in *Journal of Financial and Quantitative Analysis*.
- Brendle, S., 2005, "Portfolio Selection Under Partial Observation and Filtering," working paper, Stanford University.
- Brennan, M., 1998, "The Role of Learning in Dynamic Portfolio Decisions," *European Economic Review*, 1, 295–306.
- Brennan, M., and Y. Xia, 2001, "Assessing Asset Pricing Anomalies," *Review of Financial Studies*, 14, 905–945.

- Brown, S., 1979, "Optimal Portfolio Choice Under Uncertainty: A Bayesian Approach," in V. Bawa, S. Brown, and R. Klein (eds.), *Estimation Risk and Optimal Portfolio Choice*, North-Holland, Amsterdam.
- Browne, S., and W. Whitt, 1996, "Portfolio Choice and the Bayesian Kelly Criterion," *Advances in Applied Probability*, 28, 1145–1176.
- Cox, J. C., and C. -F. Huang, 1989, "Optimal Consumption and Portfolio Policies when Asset Prices Follow a Diffusion Process," *Journal of Economic Theory*, 49, 33–83.
- David, A., and P. Veronesi, 2002, "Option Prices with Uncertain Fundamentals: Theory and Evidence on the Dynamics of Implied Volatilities," working paper, University of Chicago.
- Detemple, J. B., 1986, "Asset Pricing in a Production Economy with Incomplete Information," *Journal of Finance*, 41, 383–391.
- Dothan, M. U., and Feldman, D., 1986, "Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy," *Journal of Finance*, 41, 369–382.
- Dugar, A., and S. Nathan, 1995, "The Effects of Investment Banking Relationships on Financial Analysts' Earning Forecasts and Investment Recommendations," *Contemporary Accounting Research*, 12(3), 131–160.
- Gennotte, G., 1986, "Optimal Portfolio Choice Under Incomplete Information," *Journal of Finance*, 41, 733–746.
- Harvey, C., Liechty, J., Liechty, M., and P. Müller, 2004, "Portfolio Selection with Higher Moments," working paper, Duke University.
- Jegadeesh, N., Kim, J., Krische, S., and C. Lee, 2004, "Analyzing the Analysts: When Do Recommendations Add Value?," *Journal of Finance*, 59, 1083–1124.
- Jha, V., D. Lichtblau, and H. Mozes, 2003, "The Usefulness of Analysts' Recommendations," *The Journal of Investing*, 12(2), 7–18.
- Jones, C., and J. Shanken, 2005, "Mutual Fund Performance with Learning Across Funds," *Journal of Financial Economics*, 3, 507–552.
- Jorion, P., 1986, "Bayes-Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis*, 21, 279–292.
- Kandel, S., and R. Stambaugh, 1996, "On the Predictability of Stock Returns: An Asset Allocation Perspective," *Journal of Finance*, 51, 385–424.
- Karatzas, I., J. P. Lehoczky, and S. E. Shreve, 1987, "Optimal Portfolio and Consumption Decisions for a 'Small Investor' on a Finite Horizon," *SIAM Journal of Control and Optimization*, 25, 1557–1586.
- Karatzas, I., and S. E. Shreve, 1991, *Brownian Motions and Stochastic Calculus*, Springer-Verlag, New York, NY.
- Klein, R. W., and V. S. Bawa, 1976, "The Effect of Estimation Risk on Optimal Portfolio Choice," *Journal of Financial Economics*, 3, 215–231.
- Lakner, P., 1998, "Optimal Trading Strategy for an Investor: The Case of Partial Information," *Stochastic Processes and their Applications*, 76, 77–97.
- Lin, H., and M. McNichols, 1998, "Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations," *Journal of Accounting and Economics*, 25, 101–128.
- Lipster, R. S., and Shirayev, A. N., 2001, *Statistics of Random Processes I*, Springer-Verlag, New York, NY.
- Markov, S., and A. Tamayo, 2006, "Predictability in Financial Analyst Forecast Errors: Learning or Irrationality?," forthcoming in *Journal of Accounting Research*.
- McCulloch, R., and P. Rossi, 1990, "Posterior, Predictive, and Utility-Based Approaches to Testing the Arbitrage Pricing Theory," *Journal of Financial Economics*, 28, 7–38.

- McNichols, M., and P. O'Brien, 1997, "Self-Selection and Analyst Coverage," *Journal of Accounting Research*, 35, 167–199.
- Merton, R. C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.
- Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–888.
- Merton, R., 1980, "On Estimating the Expected Return on the Market: an Exploratory Investigation," *Journal of Financial Economics*, 8, 323–362.
- Pástor, L., 2000, "Portfolio Selection and Asset Pricing Models," *Journal of Finance*, 55, 179–223.
- Pástor, L., and R. Stambaugh, 1999, "Costs of Capital and Model Mispricing," *Journal of Finance*, 54, 67–121.
- Pástor, L., and R. Stambaugh, 2000, "Comparing Asset Pricing Models: An Investment Perspective," *Journal of Financial Economics*, 56, 335–381.
- Pástor, L., and P. Veronesi, 2005, "Was There a NASDAQ Bubble in the Late 1990s?," forthcoming in *Journal of Financial Economics*.
- Rogers, L. C. G., 2001, "The Relaxed Investor and Parameter Uncertainty," *Finance and Stochastics*, 5, 131–154.
- Samuelson, P., 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51, 239–246.
- Schneeweis, T., and R. Spurgin, 1998, "Multifactor Models in Managed Futures, Hedge Fund and Mutual Fund Estimation," *Journal of Alternative Investments*, 1(2), 1–24.
- Stambaugh, R., 2003, "Inference About Survivors," working paper, University of Pennsylvania.
- Stojanovic, S., 2002, *Computational Financial Mathematics using Mathematica*, Birkhauser, Basel.
- Treynor, J., and F. Black, 1973, "How to Use Security Analysis to Improve Portfolio Selection," *Journal of Business*, 46, 66–88.
- Veronesi, P., 2000, "How Does Information Quality Affect Stock Returns?," *Journal of Finance*, 55, 807–837.
- Wachter, J., 2002, "Portfolio and Consumption Decisions Under Mean-Reverting Returns: An Exact Solution for Complete Markets," *Journal of Financial and Quantitative Analysis*, 37, 63–91.
- Womack, K., 1996, "Do Analysts' Recommendations Have Investment Value?," *Journal of Finance*, 51, 137–167.
- Xia, Y., 2001, "Learning About Predictability: The Effect of Parameter Uncertainty on Dynamic Asset Allocation," *Journal of Finance*, 56, 205–246.
- Zellner, A., and V. K. Chetty, 1965, "Prediction and Decision Problems in Regression Models from the Bayesian Point of View," *Journal of the American Statistical Association*, 60, 608–616.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.