Parameter Uncertainty and International Investment in a Multi-period Setting

by

Pierluigi Balduzzi*, Ludan Liu, and Zhe Xu*

This draft: March 11, 2009

Abstract

We examine the effect of parameter uncertainty on asset allocations in a dynamic setting, where the investor learns about the mean returns over time, employs cross-inference in the estimation of mean returns, and can invest in domestic and international equity indices. At the methodological level, we show that cross-inference affects the evolution of the mean return estimates and generates hedging demands for longer-history assets, which are absent in the context of learning without cross-inference. Moreover, in a context with multiple risky assets, the hedging demands induced by learning can be positive for some assets, unlike the single-risky-asset setting. A calibration exercise shows that ignoring cross-inference, learning, and estimation risk, following the naïve $1/N$ diversification policy, or not diversifying internationally, can lead to sizable utility costs.

JEL # G11, G15

* Carroll School of Management, Boston College. Corresponding author: Pierluigi Balduzzi, Finance Department, The Wallace E. Carroll School of Management, Boston College, Chestnut Hill, MA 02467. E-mail: balduzzp@bc.edu. Tel.: (617) 552-3976.
We thank Eric Jacquier, Darius Miller, Raman Uppal, Yangru Wu, Yihong Xia, seminar participants at Boston College, Rutgers University, Carnegie-Mellon University, Universitá degli studi di Brescia, and conference participants at the 2000 Financial Management Association Meetings, the 2001 Western Finance Association Meetings, the 2001 European Finance Association Meetings, and the 2004 ESF-CEPR Conference on Dynamic Portfolio Choice, Asset Pricing, and Mathematical Finance, for useful comments. A previous version of this paper was circulated under the title "Parameter Uncertainty and International Investment."

Parameter Uncertainty and International Investment in a Multi-period Setting

This draft: March 11, 2009

Abstract

We examine the effect of parameter uncertainty on asset allocations in a dynamic setting, where the investor learns about the mean returns over time, employs cross-inference in the estimation of mean returns, and can invest in domestic and international equity indices. At the methodological level, we show that cross-inference affects the evolution of the mean return estimates and generates hedging demands for longer-history assets, which are absent in the context of learning without cross-inference. Moreover, in a context with multiple risky assets, the hedging demands induced by learning can be positive for some assets, unlike the single-risky-asset setting. A calibration exercise shows that ignoring cross-inference, learning, and estimation risk, following the naïve $1/N$ diversification policy, or not diversifying internationally, can lead to sizable utility costs.

JEL # G11, G15

1 Introduction

Several studies have investigated the benefits of international diversification, and the role of emerging markets in particular, in the context of mean-variance spanning tests (see, for example, Harvey, 1995; Bekaert and Urias, 1996; Errunza, Hogan, and Hung, 1999; and De Roon, Nijman, and Werker, 2001). One shortcoming of these tests is that they do not translate immediately into optimal portfolio implications. In particular, given the uncertainty surrounding both the mean-variance weights (Britten-Jones, 1999) and the gains from international diversification, it is not clear what the normative implications for an investor should be. Moreover, while a few studies account for parameter uncertainty in computing the allocations of an internationally-diversified investor (see, for example, Jorion, 1985; and Stambaugh, 1997), none of these studies accounts for the multi-period nature of most portfolio problems.[1]

The behavior of an investor who accounts for parameter uncertainty in a dynamic multi-period setting has been studied by Brennan (1998), Barberis (2000), Xia (2001), and Brandt et al. (2005), for the case of investment in the U.S. equity index, but has not been explored for the case of multiple equity indices.[2] Moreover, when assets with histories of different length are available for investment, Stambaugh's (1997) combined-sample estimation, or cross-inference, uses data on longer-history assets to obtain more efficient estimation of the moments of the shorter-history assets.[3] Stambaugh applies his technique to a static mean-variance portfolio setting, but not to a multi-period dynamic setting.

To our knowledge, this is the first study to solve the problem of a multi-period dynamic investor who allocates his funds across various international equity indices, is aware of the uncertainty around parameter estimates of the return-generation process, anticipates revisions of those estimates, and employs cross-inference to efficiently use return histories of different length. This setup is used to investigate several issues, including the utility costs of not diversifying internationally.

We consider *three* possible combinations of domestic and foreign assets: The U.S.-WD case allows for investment in the U.S. risk-free asset, the U.S. market index, and an index of world-developed markets. The U.S.-EM case allows for investments in the U.S. risk-free asset, the U.S. market index, and an index of emerging markets. The U.S.-WD-EM case allows for investment in the U.S. risk-free asset, the U.S. market index, an index of world-developed markets, and an index of emerging markets.

We calculate optimal allocations at two different points in time: at the end of 1996 and at the end of

[1]See also Horst, De Roon, and Werker (2006) for an alternative derivation of the "shrinkage" approach; and see Li, Sarkar, and Wang (2003) for a Bayesian approach to the valuation of the gains from international diversification.

[2] While relatively few studies have investigated the dynamic portfolio problem of a Bayesian investor, several studies have investigated the myopic problem, starting with Bawa et al. (1979).

[3] Stambaugh (1997), like the present paper, takes a normative perspective. Brown and Barry (1984, 1985), on the other hand, investigate the effects of unequal sample lengths for financial market equilibrium.

2006. Since our sample of emerging-market returns only starts in 1989, the investor has little information prior to his portfolio decisions. This is precisely the challenge posed by short-history markets, and a similar situation is the one studied by Stambaugh (1997) in his analysis of 22 emerging markets at the end of 1995, with histories starting between 1989 and 1993. We also consider investment behavior under the optimal ("benchmark") policy, as well as a variety of sub-optimal policies. Specifically, we consider the effects of ignoring learning, estimation risk, and cross-inference, and we calculate the utility costs associated with the sub-optimal policies. We also calculate the utility cost of not diversifying internationally from the standpoint of optimal and sub-optimal investors.

A methodological contribution of this study is to show that for a dynamic investor, the presence of cross-inference affects both the levels of the parameter estimates *and* their evolution over time. Specifically, realizations of returns for the longer-history assets negatively affect mean estimates of the shorter-history assets. This generates an extra *positive* hedging demand for longer-history assets. These extra hedging demands, absent in the context of learning without cross-inference, are substantial when emerging markets are included in the portfolio. While hedging demands due to learning have been previously documented (see, for example, Brennan, 1998; and Barberis, 2000), this is the first paper to document the extra hedging demands induced by cross-inference. Moreover, while previous studies have documented *negative* hedging demands induced by learning in single-risky-asset settings, we show that in a multiple-risky-asset setting, the hedging demands induced by learning may be *positive* for some risky assets.

As to allocations and utility costs, at the end of 1996 a dynamic investor allocates a substantial fraction of his wealth to international markets, emerging markets in particular. Indeed, in the U.S.-WD-EM case, the allocation to the U.S. index is 39% of wealth, whereas foreign markets account for a total of 44%. Ignoring cross-inference substantially reduces the allocation to foreign developed markets, while ignoring learning leads to substantial overinvestment in emerging markets (up to 54%). Hence, learning generates a substantial negative hedging demand for emerging markets. The associated utility costs can be substantial as well: the ignorance of both cross-inference and learning implies a cost of 5% of portfolio wealth. Moreover, given the large international allocations, the cost of not diversifying internationally perceived by the benchmark investor is also substantial: 23%. The investor who ignores learning overestimates this cost, placing it at 26% of portfolio wealth.

At the end of 2006, the allocations of a dynamic investor are somewhat different, mainly reflecting the updating of the mean estimates for emerging markets. In the U.S.-WD-EM case, the U.S. index receives an allocation of 32%, while 18% is invested in developed markets, and 21% is invested in emerging markets. Utility costs are smaller for this longer sample: 2% when both cross-inference and learning are ignored. In line with the portfolio-allocation results, the costs of not diversifying internationally are also smaller: 14% for the benchmark investor.

Two articles studying the multi-period portfolio problem of an internationally diversified investor are useful benchmarks for our study: Ang and Bekaert (2002) and Das and Uppal (2004). Ang and Bekaert consider an investor facing a two-regime-switching process: correlations between markets and market volatilities can be high or low. They calibrate their model using data for the U.S., U.K., and German indices. They find that international diversification is still valuable, even when international co-movements strengthen in volatile markets. They also find that accounting for regime switching can be roughly as important as allowing for international diversification. Das and Uppal (2004) consider a related econometric setting, where international indices experience infrequent common jumps. Similarly to Ang and Bekaert, the presence of common jumps also has the effect of strengthening international co-movements at the time of high volatility. Their model is calibrated using two sets of assets: the U.S. index and five developed-market indices, and the U.S. index and five emerging-market indices. Consistent with the notion that jumps are infrequent, they find that the presence of correlated shocks has only a minor impact on allocations, relative to a situation where all indices follow diffusions.

Our paper complements these two studies by focusing on two other aspects of the econometrics of international stock markets: the different length of histories for different markets, i.e., the effect of cross-inference, and the uncertainty around parameter estimates of the return-generation processes, i.e., the effects of learning and estimation risk. We view these two features as being at least as worthwhile investigating as the strengthening of correlations that Ang and Bekaert (2002) and Das and Uppal (2004) focus on. Moreover, both of these studies make ad-hoc adjustments for the sampling variability of parameter estimates and portfolio allocations. These adjustments are not necessary in our setting, as the uncertainty about parameter estimates is accounted for by the investor.

The paper is organized as follows. In Section 2, we present the dynamic portfolio problem. In Section 3, we illustrate the data used to calibrate the portfolio problems. Section 4 discusses the parameter estimates obtained with and without cross-inference. Section 5 presents the results for the portfolio weights, while Section 6 presents the results for the utility costs. Section 7 concludes.

2 Parameter Uncertainty in a Multi-period Portfolio Setting

We consider the problem of a long-lived investor who, in the benchmark case, uses the combined sample of asset histories to estimate mean parameters and is aware of estimation risk and learning. We ignore estimation risk and learning for the second moments because second-moment estimates tend to be more precise than first-moment estimates, and the precision of second-moment estimates can generally be improved by increasing the sampling frequency. Moreover, accounting for learning about second moments would increase the dimension of the state space beyond numerical tractability.

For expositional convenience, the discussion in the paper deals with the case where three assets are available for investment: the risk-free asset, a domestic risky asset, and only one international risky asset.[4]

2.1 Preferences and Constraints

We are at time t_0, the beginning of the investment horizon, and the investor has constant-relative-risk-aversion preferences defined over time-T wealth, described by the utility function

$$U(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}, \tag{1}$$

where W_T is terminal wealth and γ is the relative-risk-aversion coefficient. In all the empirical exercises we set $\gamma = 5$. The investor's problem is to maximize his expected time-T utility

$$\max_{a_{1,t_0}, a_{2,t_0}} \; E_{t_0}\left(\frac{W_T^{1-\gamma}}{1-\gamma}\right) \tag{2}$$

subject to the constraint,

$$W_{t+1} = W_t \left[a_{1,t} \exp(r_{1,t+1}) + a_{2,t} \exp(r_{2,t+1}) + (1 - a_{1,t} - a_{2,t}) \exp(r_f)\right], \tag{3}$$

where $a_{1,t}$ and $a_{2,t}$ are the portfolio shares of the domestic and international asset, with log returns $r_{1,t+1}$ and $r_{2,t+1}$. r_f is the log return on the risk-free asset, which we assume to be constant.[5]

In our analysis, we impose non-negativity constraints on all portfolio shares, i.e., we impose short-sale constraints. This is consistent with De Roon, Nijman, and Werker (2001), and Li, Sarkar, and Wang (2003) who investigate the benefits of international diversification in the presence of short-sale constraints; moreover, Das and Uppal (2004) also consider the effects of imposing short-sale constraints. While not reported in the paper, we also produced results for the case where short-sales constraints are not imposed. These results are virtually identical to the case where the constraints are imposed and are available from the authors upon request.

The optimization problem described above can be solved by dynamic programming techniques. There are three state variables: the wealth level W_t and the means of the predictive densities of returns $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$; and two control variables: $a_{1,t}$ and $a_{2,t}$. As in Barberis (2000) and Brandt et al. (2005), we assume a non-informative prior, and therefore the means of the posterior density are simply the maximum-likelihood (ML) estimates.[6] Moreover, since we also assume lognormal returns, the ML estimates are also the means of the predictive density.

[4]The extension to the case of three risky assets (one domestic plus two international), which is conceptually straightforward, is illustrated in an appendix available from the authors upon request.

[5]Abstracting from time-variation in the risk-free rate is common in studies of dynamic asset allocation. For example, the same assumption is made by Das and Uppal (2004); and the case of a constant risk-free rate is also one of the two settings examined by Ang and Bekaert (2002).

[6]Brennan (1998) and Xia (2001), on the other hand, assume an informative prior, but then calibrate the parameters of the prior density based on historical returns.

Define the value function as

$$J(W_t, \hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) = \max_{a_{1,t}, a_{2,t}} \; E_t \left(\frac{W_T^{1-\gamma}}{1-\gamma} \right). \tag{4}$$

The Bellman equation is given by

$$J(W_t, \hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) = \max_{a_{1,t}, a_{2,t}} \; E_t \left[J(W_{t+1}, \hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t+1) \right]. \tag{5}$$

A simple induction shows that the value function can be written as

$$J(W_t, \hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) = \frac{W_t^{1-\gamma}}{1-\gamma} Q(\hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t), \tag{6}$$

where

$$\begin{aligned} & Q(\hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) = \\ & \min_{a_{1,t}, a_{2,t}} \; E_t \left\{ \left[a_{1,t} \exp(r_{1,t+1}) + a_{2,t} \exp(r_{2,t+1}) + (1 - a_{1,t} - a_{2,t}) \exp(r_f) \right]^{1-\gamma} Q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right\}. \end{aligned} \tag{7}$$

Since the risk aversion coefficient is generally believed to be greater than one, the maximization problem for the value function $J(.)$ turns into a minimization problem for $Q(.)$. The conditional expectation is taken over the predictive density for $r_{1,t+1}$ and $r_{2,t+1}$ at time t. Details of the numerical solution technique can be found in Appendix A.1.

2.2 Cross-inference

Historical time series of returns for foreign markets are typically shorter than for the U.S. market. Hence, we assume that the history for asset 1 (U.S.) starts at time 1, i.e., asset 1 has a history of t months. The history for asset 2 (foreign), on the other hand, starts at time $s_2 > 1$, i.e., asset 2 has a history of $t - s_2 + 1$ months.

As mentioned above, the investor treats the two returns as i.i.d. Moreover, the two log returns are assumed bivariate normal. The log-likelihood for the combined sample is given by

$$\text{constant} - \frac{s_2 - 1}{2} \ln(\sigma_1^2) - \frac{1}{2} \sum_{\tau=1}^{s_2 - 1} \frac{(r_{1,\tau} - \mu_{1,t})^2}{\sigma_1^2} - \frac{t - s_2 + 1}{2} \ln(|\boldsymbol{\Sigma}|) - \frac{1}{2} \sum_{\tau=s_2}^{t} (\mathbf{r}_\tau - \mu_t)' \boldsymbol{\Sigma}^{-1} (\mathbf{r}_\tau - \mu_t), \tag{8}$$

where $\mathbf{r}_\tau$ is the return vector, μ_t is the mean vector estimated at t, and $\boldsymbol{\Sigma}$ is the covariance matrix of the two returns.

As shown by Stambaugh (1997), combined-sample ML estimates can be obtained analytically. In particular, while the time-t ML estimate of the mean of asset 1 is simply the sample mean over the time-1 to time-t sample, $m_{1,t}(r_1)$, the ML estimate of the mean of asset 2 can be improved relative to its sample mean over the time-s_2 to time-t sample, $m_{2,t}(r_2)$. [7] We can write the "regression" model

$$r_{2,\tau} = \alpha_2 + \beta_{21} r_{1,\tau} + \epsilon_{2,\tau}, \ \tau = s_2, \ldots, t. \tag{9}$$

We assume that the investor does not update of variances and covariances of returns after t_0. Hence, the ML estimate of the slope coefficient is constant and is given by

$$\hat{\beta}_{21} \quad = \quad \frac{c_{2,t_0}(r_2, r_1)}{v_{2,t_0}(r_1)}, \tag{10}$$

where $c_{2,t_0}(.,.)$ and $v_{2,t_0}(.)$ denote sample covariance and variance for the shorter sample, running from s_2 to t_0. Given $\hat{\beta}_{21}$, we obtain the time-t estimate of α_2 as

$$\hat{\alpha}_{2,t} = m_{2,t}(r_2) - \hat{\beta}_{21} m_{2,t}(r_1). \tag{11}$$

Based on the ML estimates above, we obtain a combined-sample ML estimate of the mean of $r_{2,\tau}$, $\hat{\mu}_{2,t}$:

$$\hat{\mu}_{2,t} = \hat{\alpha}_{2,t} + \hat{\beta}_{21} m_{1,t}(r_1) = m_{2,t}(r_2) - \hat{\beta}_{21} \left[m_{2,t}(r_1) - m_{1,t}(r_1) \right]. \tag{12}$$

Hence, relative to the sample mean $m_{2,t}(r_2)$, the combined-sample ML estimate $\hat{\mu}_{2,t}$ is higher or lower depending on the difference $m_{2,t}(r_1) - m_{1,t}(r_1)$. For example, assume $\hat{\beta}_{21} > 0$. If the return on asset 1 is unusually high (low) for the time-s_2 to time-t sample, the estimate of the mean return for asset 2 is corrected downwards (upwards).

The ML estimate of the variance of asset 1 is simply the sample variance over the time-1 to time-t_0 sample, $\hat{\sigma}_1^2 = v_{1,t_0}(r_1)$. On the other hand, we can obtain combined-sample ML estimates of the variance of asset 2 and of the covariance between assets 1 and 2. Based on the regression model (9), and assuming no updating, we have:

$$\hat{\sigma}_2^2 \quad = \quad \hat{\beta}_{21}^2 v_{1,t_0}(r_1) + v_{2,t_0}(\hat{\epsilon}_{2,\tau}) = v_{2,t_0}(r_2) - \hat{\beta}_{21}^2 \left[v_{2,t_0}(r_1) - v_{1,t_0}(r_1) \right] \tag{13}$$

$$\hat{\sigma}_{12} \quad = \quad \hat{\beta}_{21} v_{1,t_0}(r_1) = c_{2,t_0}(r_2, r_1) - \hat{\beta}_{21} \left[v_{2,t_0}(r_1) - v_{1,t_0}(r_1) \right]. \tag{14}$$

As in the estimation of the mean, combined-sample ML estimates of the second moments depend on the difference $v_{2,t_0}(r_1) - v_{1,t_0}(r_1)$. Assume again $\hat{\beta}_{21} > 0$. If the variance of asset 1 for the shorter sample is unusually high (low), the estimate of the variance of asset 1, and the estimate of the covariance between asset 1 and asset 2 are corrected downwards (upwards).

[7] It is worth noting that the combined-sample ML estimates, conditional on the parameters of the covariance matrix, are the SUR estimates of the joint regressions of the two returns on a constant. Hence, when the series differ in length, SUR estimates differ from the OLS estimates, even when the equations have the same regressors; see Im (1994).

2.3 Law of Motion of State Variables and Predictive Density

As new information becomes available, the investor *learns* about the true parameters of the return density using the recursion:[8]

$$\hat{\mu}_{1,t+1} = \frac{\hat{\mu}_{1,t}t + r_{1,t+1}}{t+1} = \hat{\mu}_{1,t} + \frac{1}{t+1}(r_{1,t+1} - \hat{\mu}_{1,t}) \tag{15}$$

$$\hat{\mu}_{2,t+1} = \hat{\mu}_{2,t} + \frac{1}{t-s_2+2}(r_{2,t+1} - \hat{\mu}_{2,t}) - \hat{\beta}_{21}\left(\frac{1}{t-s_2+2} - \frac{1}{t+1}\right)(r_{1,t+1} - \hat{\mu}_{1,t}). \tag{16}$$

Two features of the recursion above are worth noting. First, conditional estimates of the means of the returns are positively correlated with the corresponding return realizations. This effect tends to generate positive correlations between asset returns and the value function, even though returns are treated by the investor as i.i.d (see Barberis, 2000). Second, in a setting where the two returns have histories with different lengths, realizations of the longer-history return affect the estimate of the mean for the shorter-history return. From equation (16) we have

$$\frac{\partial \hat{\mu}_{2,t+1}}{\partial r_{1,t+1}} = -\hat{\beta}_{21}\left(\frac{1}{t-s_2+2} - \frac{1}{t+1}\right). \tag{17}$$

If $\hat{\beta}_{21} > 0$, as it typically is across international stock markets, then high realizations of the U.S. return reduce expected foreign returns. This effect is stronger, the stronger the correlation between the two markets and the larger the difference in histories' lengths. This effect tends to generate a *negative* correlation between the domestic-asset return and the value function, and a positive component to the hedging demand for the longer-history asset. This effect is unique to this dynamic optimization context with cross-inference.

The covariance matrix $\mathrm{Cov}_t(\mathbf{r}_{t+1})$ of the predictive density is given by (see Appendix A.2)

$$\mathrm{Cov}_t(\mathbf{r}_{t+1}) = \begin{bmatrix} (\frac{1}{t}+1)\hat{\sigma}_1^2 & (\frac{1}{t}+1)\hat{\sigma}_{12} \\ (\frac{1}{t}+1)\hat{\sigma}_{12} & \left(\frac{\hat{\rho}_{12}^2}{t} + \frac{1-\hat{\rho}_{12}^2}{t-s_2+1} + 1\right)\hat{\sigma}_2^2 \end{bmatrix}, \tag{18}$$

where $\hat{\rho}_{12} = \hat{\sigma}_{12}/(\hat{\sigma}_1\hat{\sigma}_2)$. Here, it is worth noting that the predictive variance of $r_{2,t+1}$ exceeds $\hat{\sigma}_2^2$ by a factor that is a *weighted* average of $1/t$ and $1/(t-s_2+1)$. In other words, the predictive variance for asset 2, $[\hat{\rho}_{12}^2/t + (1-\hat{\rho}_{12}^2)/(t-s_2+1)+1]\hat{\sigma}_2^2$, is smaller than what it would be if we did not use the combined sample, $[1/(t-s_2+1)+1]\,\hat{\sigma}_2^2$. Thus, in the two-asset case, learning has a greater impact on the asset with a shorter history because both its own history and the history of the longer-history asset help resolve parameter uncertainty over time.

[8] The law of motion for $\hat{\mu}_{2,t+1}$ can be derived as follows: From equation (12), we have

$$\hat{\mu}_{2,t+1} = \hat{\beta}_{21}\hat{\mu}_{1,t+1} + \left[m_{2,t+1}(r_2) - \hat{\beta}_{21}m_{2,t+1}(r_1)\right].$$

Hence, we have

$$\hat{\mu}_{2,t+1} = \hat{\beta}_{21}\left[\hat{\mu}_{1,t} + \frac{1}{t+1}(r_{1,t+1} - \hat{\mu}_{1,t})\right] + \frac{\left[m_{2,t}(r_2) - \hat{\beta}_{21}m_{2,t}(r_1)\right](t-s_2+1) + (r_{2,t+1} - \hat{\beta}_{21}r_{1,t+1})}{t-s_2+2}.$$

Using the fact that $m_{2,t}(r_2) - \hat{\beta}_{21}m_{2,t}(r_1) = \hat{\mu}_{2,t} - \hat{\beta}_{21}\hat{\mu}_{1,t}$, and rearranging terms, one obtains equation (16).

2.4 Sub-optimal Policies

To evaluate the importance of cross-inference and learning, we consider three scenarios where the investor behaves sub-optimally.

2.4.1 *NC-L Policy*

In the first sub-optimal scenario, the investor does not make use of cross-inference, but is aware of future changes in parameter estimates. Parameter values are estimated separately using the longest sample available:

$$\begin{aligned}
\hat{\mu}_{1,t}^{NC} &= m_{1,t}(r_1) && (19)\\
(\hat{\sigma}_1^{NC})^2 &= v_{1,t}(r_1) && (20)\\
\hat{\mu}_{2,t}^{NC} &= m_{2,t}(r_2) && (21)\\
(\hat{\sigma}_2^{NC})^2 &= v_{2,t}(r_2) && (22)\\
\hat{\sigma}_{12}^{NC} &= c_{2,t}(r_2, r_1), && (23)
\end{aligned}$$

and the two mean estimates follow the recursion

$$\hat{\mu}_{1,t+1}^{NC} = \frac{\hat{\mu}_{1,t}^{NC} t + r_{1,t+1}}{t+1} = \hat{\mu}_{1,t}^{NC} + \frac{1}{t+1}(r_{1,t+1} - \hat{\mu}_{1,t}^{NC}) \tag{24}$$

$$\hat{\mu}_{2,t+1}^{NC} = \frac{\hat{\mu}_{2,t}^{NC}(t - s_2 + 1) + r_{2,t+1}}{t - s_2 + 2} = \hat{\mu}_{2,t}^{NC} + \frac{1}{t - s_2 + 2}(r_{2,t+1} - \hat{\mu}_{2,t}^{NC}). \tag{25}$$

Note that in this case realizations of $r_{1,t+1}$ do not affect $\hat{\mu}_{2,t}^{NC}$. Hence, the second component of the hedging demand for asset 1 is absent.

The predictive density of asset returns has covariance matrix

$$\text{Cov}_t^{NC}(\mathbf{r}_{t+1}) = \begin{bmatrix} (\frac{1}{t}+1)(\hat{\sigma}_1^{NC})^2 & \hat{\sigma}_{12}^{NC} \\ \hat{\sigma}_{12}^{NC} & (\frac{1}{t-s_2+1}+1)(\hat{\sigma}_2^{NC})^2 \end{bmatrix}. \tag{26}$$

Notice that since the two means are estimated separately, the investor ignores the correlation between the two estimates, and the predictive covariance is simply $\hat{\sigma}_{12}^{NC}$. We denote this sub-optimal policy the NC-L policy.

As is standard in the dynamic portfolio literature, we calculate utility costs by "feeding" the policies of an NC investor into the recursion for the value function of an investor who does account for cross-inference. One subtle aspect of this exercise, which is unique to our setting, is that the benchmark investor and the NC investor disagree on the realizations of the relevant state variables. Hence, in implementing the recursion, for each realization of $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$ for the benchmark investor we need to calculate the corresponding realizations for the NC investor, $\hat{\mu}_{1,t}^{NC}$ and $\hat{\mu}_{2,t}^{NC}$, since these, in turn, determine the choice of allocation.

Below, we derive the relation between $\hat{\mu}_{2,t+1}^{NC}$ and $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$. From equation (12), we have

$$\hat{\mu}_{2,t}^{NC} = \hat{\mu}_{2,t} + \hat{\beta}_{21}\left[m_{2,t}(r_1) - \hat{\mu}_{1,t}\right]. \tag{27}$$

Note that we can write

$$\hat{\mu}_{1,t} = \frac{1}{t}\sum_{\tau=1}^{s_2-1} r_{1,\tau} + \frac{t-s_2+1}{t} m_{2,t}(r_1), \tag{28}$$

and

$$m_{2,t}(r_1) = \frac{t}{t-s_2+1}\hat{\mu}_{1,t} - \frac{1}{t-s_2+1}\sum_{\tau=1}^{s_2-1} r_{1,\tau}. \tag{29}$$

Substituting (28) and (29) into (27), we obtain

$$\hat{\mu}_{2,t}^{NC} = \hat{\mu}_{2,t} + \hat{\beta}_{21}\frac{s_2-1}{t-s_2+1}\left(\hat{\mu}_{1,t} - \frac{1}{s_2-1}\sum_{\tau=1}^{s_2-1} r_{1,\tau}\right). \tag{30}$$

The expression above gives us, for given $\sum_{\tau=1}^{s_2-1} r_{1,\tau}$, the NC estimate of the mean return on asset 2 as a function of the benchmark estimates $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$. In the calibration exercises, we set $\sum_{\tau=1}^{s_2-1} r_{1,\tau}$ at its realized value over the sample. The intuition for $\hat{\mu}_{1,t}$ affecting $\hat{\mu}_{2,t}^{NC}$ is that as $\hat{\mu}_{1,t}$ exceeds $\sum_{\tau=1}^{s_2-1} r_{1,\tau}/(s_2-1)$, i.e., the mean return on asset 1 is unusually high during the second part of the sample, $\hat{\mu}_{2,t}^{NC}$ must exceed $\hat{\mu}_{2,t}$; i.e., by ignoring cross-inference the investor overestimates the mean return on asset 2.

2.4.2 *C-NL Policy*

In the second sub-optimal scenario, the investor makes use of cross-inference, but is unaware of future changes in parameter estimates. Hence, the predictive density of returns is the same as in the benchmark case, although it is taken as invariant over time. Since the investor's perceived investment opportunity set is constant over time, as shown by Samuelson (1969), the investor holds the *same* portfolio over his lifetime. We denote this sub-optimal policy the C-NL policy.

2.4.3 *NC-NL Policy*

In the third sub-optimal scenario, the investor ignores both cross-inference and learning. As in the first sub-optimal scenario, parameter estimates are obtained separately based on the longest available sample. As in the second sub-optimal scenario, the investment opportunity set remains constant over time. We denote this sub-optimal policy the NC-NL policy.

2.5 Utility Costs

To quantify the importance of cross-inference and learning, we calculate the utility costs associated with the three suboptimal policies. The utility cost for the NC-L policy measures the loss from ignoring cross-inference. The utility cost for the C-NL policy measures the loss from ignoring future learning. The utility cost for the NC-NL policy measures how much the investor loses ex ante by ignoring cross-inference and by ignoring future revisions in parameter estimates. Following Kandel and Stambaugh (1996), we calculate the utility costs based on the certainty-equivalent return (CER). The certainty-equivalent return (CER) measures the return on wealth that, if earned with certainty, would give the investor the same utility as the expected utility obtained for a given allocation policy. For given policy, a CER is obtained by solving the following equation:

$$\frac{[W_t(1+\mathrm{CER})]^{1-\gamma}}{1-\gamma} = J(W_t, \hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t), \tag{31}$$

where the program $J(.)$ is calculated under both optimal and suboptimal policies. The utility cost for a suboptimal policy is defined as[9]

$$\frac{\mathrm{CER}^* - \mathrm{CER}}{1+\mathrm{CER}^*}, \tag{32}$$

where CER^* is the certainty-equivalent return of the optimal policy. This ratio can also be interpreted as the fraction of wealth that an investor is willing to give up in order to use the optimal policy rather than the sub-optimal one.

3 Data

We implement the portfolio analysis using estimates from returns and dividend yields for three equity indices. The first index is the U.S. equity market. This is the value-weighted index for stocks traded on NYSE, AMEX and NASDAQ compiled by CRSP. We use monthly returns for the 1950-2006 period. The second index is the index of developed markets excluding the U.S. compiled by MSCI. We use monthly returns for the 1970-2006 period. The third index is the Emerging Markets index compiled by the International Finance Corporation (IFC); [10] We use monthly returns for the 1989-2006 period. All returns include dividends, are denominated in U.S. dollars, and are converted into real returns using U.S. CPI inflation. The risk-free rate is set equal to the average real rate of return on a one-month U.S. T-bill for the 1950-2006 period.

[9] See, for example, Balduzzi and Lynch (1999).

[10] The MSCI developed-market indices have been used by several other studies of international diversification, see, for example Harvey (1995) and Ang and Bekaert (2002). Similarly common is the use of the "investable" IFC indices for emerging markets; see, for example, Harvey (1995), Errunza, Hogan, and Hung (1999), De Roon, Nijman, and Werker (2001), and Li, Sarkar, and Wang (2003).

4 Estimates

The inputs for the dynamic portfolio optimization are the initial ML estimates of mean returns and the variance-covariance matrix of returns. Table 1 presents these estimates. The table is organized in two panels, corresponding to the samples ending in 1996 and ending in 2006. In each panel, we first present estimates for the three indices in the benchmark case (cross-inference) with three risky assets (U.S.-WD-EM). Second, we present estimates for the EM index in the benchmark case with two risky assets (U.S.-EM). Third, we present estimates for the WD and EM indices in the case where we ignore cross-inference (NC).[11]

4.1 Sample Ending in 1996

Consider first the effect of cross-inference on the WD estimates. As compared to the previous 1950-1969 period, during the 1970-1996 period (the period of overlap between histories of the U.S. and WD indices) the U.S. index experienced lower average returns (0.49% *vs* 0.83%) and higher volatility (4.56% *vs* 3.49%). Hence, given the positive correlation between returns on the U.S. and WD indices, cross-inference increases the mean estimate and reduces the volatility and covariance estimates for the WD index, making foreign developed markets more attractive.

The effects of cross-inference are similar for emerging markets in the U.S.-EM case. During the 1989-1996 period, when the histories of the U.S., WD, and EM indices overlap, the U.S. market displayed higher and less volatile returns (0.93% and 3.35%, respectively, compared to 0.58% and 4.28% for the previous 1950-1988 period). Hence, cross-inference reduces the mean and increases the volatility estimate for the EM index. In the U.S.-WD-EM case, the effects of cross-inference become more complex as it is the past performance (i.e., average and volatility) of *both* the U.S. and WD indices, together with their correlation, that matter: the cross-inference mean and variance estimates are described by *multi-variate* extensions of (12)-(14). As a result, the mean estimate for the EM index is higher than in the U.S.-EM case, while the volatility and covariance (with the U.S. market) estimates are slightly lower than in the U.S.-EM case.

Finally, it is worth noting how cross-inference leads to more precise estimates of both the WD and EM mean returns in all instances. For example, the standard error on the mean return for the EM index is 0.58 in the NC case, 0.55 in the U.S.-EM case with cross-inference, and 0.53 in the U.S.-WD-EM case with cross-inference.

[11]Note that the reported estimates for the U.S. index apply to all cases. Note also that the reported estimates for the WD index in the benchmark three-risky-assets case also apply to the benchmark two-risky-asset case: cross-inference only affects parameter estimates for the shorter-history assets.

4.2 Sample Ending in 2006

The same effects described above are at work for the sample ending in 2006. As in the shorter sample ending in 1996, cross-inference increases the mean estimate and reduces the volatility estimate for the WD index. For emerging markets, on the other hand, the main effect of cross-inference is to reduce the mean estimate in the U.S.-EM case, and to increase the mean estimate in the U.S.-WD-EM case.

5 Portfolio Weights

Table 2 presents the portfolio weights for a variety of policies and different combinations of the four assets available for investment (domestic index, two foreign indices, risk-free asset). Panel a presents results based on a sample up to December 1996. Panel b presents results based on a sample up to December 2006.

Each panel presents allocations corresponding to the benchmark case and to the three sub-optimal cases (NC-L, C-NL, and NC-NL). In addition to the optimal allocations, we report their 90% confidence intervals. These confidence intervals are computed by Monte Carlo, sampling from the posterior density of the mean returns evaluated at the beginning of the investment horizon, $t = t_0$. As a way to summarize the impact of the various aspects of sub-optimality on asset allocations, we calculate an index of deviation from the benchmark policy constructed as one-half of the sum of the absolute deviations of the sub-optimal allocations from the benchmark allocations. This index corresponds to the turnover in the portfolio if the investor were to change his allocations from the benchmark values to the sub-optimal values.[12]

5.1 Sample Ending in 1996

At the end of 1996, the benchmark investor is well-diversified internationally. This is true both in the two-risky-assets cases and in the three-risky-assets case. In the U.S.-WD case, the allocation to developed markets is 28%; in the U.S.-EM case, the allocation is 17% to emerging markets. When both developed and emerging markets are included in the portfolio, their shares amount to 18% and 26%, respectively, with 39% invested in the U.S. market.

As one would expect, based on the analysis of the previous section, ignoring cross-inference (NC-L policy) reduces the allocation to developed markets in both cases. On the other hand, ignoring cross-inference affects positively the EM allocation only in the U.S.-EM case.

Note that ignoring cross-inference affects allocations through *two* channels: First, different mean esti-

[12] One additional exercise that we performed, but did not report in the paper, is to follow the portfolio allocations over the investment horizon, while keeping the state variables (i.e., the mean estimates) at their initial values. The exercise allows us to verify that the allocations of policies that account for learning converge to the allocations of the corresponding policies that ignore learning, as time goes by, and as the hedging-demand component of the allocations tends to zero.

mates are used in the choice of asset allocations. Second, the link between returns in the longer-history markets and returns in the shorter-history markets, captured by the second term in equation (16), is missing (see the Appendix A.3 for a further discussion of the effects of cross-inference on the hedging demands). We separated the two effects by considering the optimization problem of a "benchmark-NC" investor, who updates the mean estimates following the *benchmark* recursion in equation (16), but then translates these estimates into the NC estimates (equation (30)) to calculate optimal portfolio allocations. Hence, the difference in allocations between this investor and the NC-L investor is driven only by the different law of motion of the mean estimates. In the U.S.-WD scenario, the U.S. index allocations of the benchmark-NC investor (not reported in the table) are quite similar to those of the NC-L investor: 58% *vs* 57%; hence, in this scenario, the extra hedging demands induced by cross-inference are small. In the other two scenarios, on the other hand, the differences are substantial. In the U.S.-EM case, the benchmark-NC investor allocates 45% to the U.S. index, as opposed to 28% for the NC-L investor. In the U.S.-WD-EM case, the benchmark-NC investor allocates 50% to the U.S. index and 5% to the WD index, as opposed to 41% and 1%, respectively, for the NC-L investor. Hence, when emerging markets are included, cross-inference generates substantial *positive* hedging demands for the longer-history assets.

Ignoring learning (C-NL policy) means that the investor ignores the positive correlation between return realizations and future estimates of mean returns. Hence, the hedging demands due to learning are absent and the investor adopts a myopic portfolio similar to that of a one-period mean-variance optimizer. In the U.S.-WD case, allocations to the two risky assets increase. In the U.S.-EM case, it is the allocation to emerging markets, which is substantially increased, whereas the U.S. allocation drops. In the U.S.-WD-EM case, while the EM allocation increases, both the U.S. and the WD allocations fall.

Hence, it is worth noting that the hedging demands induced by learning are not necessarily negative for all assets. Negative hedging demands due to learning arise necessarily in the one-risky-asset case, as demonstrated by Brennan (1998) and Barberis (2000). On the other hand, in the presence of multiple risky assets, the signs of the hedging-demand components of the allocations depend not only on the covariances between the mean estimates and the value function (which are positive and would generate negative hedging-demand components), but also on the inverse of the covariance matrix of returns. (See the Appendix A.4 for further discussion.)

When both cross-inference and learning are ignored (NC-NL policy), the two types of effect discussed above are combined. Note, though, that the interactions are complex and the NC-NL allocations do not necessarily lie between the NC-L and the C-NL allocations.

Portfolio reshufflings from the benchmark allocations to the sub-optimal allocations are more pronounced in the two cases where emerging markets are included. In the U.S.-EM case in particular, the turnover is as high as 46.9% for the NC-NL policy.

It is worth comparing these results to those of Ang and Bekaert (2002) and Das and Uppal (2004), who consider similar sample periods for the calibration of their models. Ang and Bekaert also consider a scenario with *four* assets available for investment: the risk-free asset with a constant risk-free rate, the U.S. equity index, the U.K. equity index, and the German equity index. The investor has an RRA coefficient of 5, a five-year horizon, and uses data from 1972 until 1997 to calibrate his portfolio model. When the investor accounts for regime switching (RS investor), he allocates 34% to the U.S. market and 22% to the U.K. and German indices combined in the high-correlation regime; and 88% to the U.S. market and 53% to the U.K. and German indices combined in the normal-correlation regime. Hence, our benchmark 1996 weights for the U.S.-WD case are between their RS weights for the high-correlation and for the normal-correlation regimes.

Das and Uppal (2004) consider two investment scenarios: a scenario where the U.S. market is combined with five developed markets and the risk-free asset, and a scenario where the U.S. market is combined with five emerging markets and the risk-free asset. In both scenarios, short-sale constraints are imposed. The investor has an RRA coefficient of 3 and uses data from 1980 until 1998 to calibrate his portfolio model. When the investor has access to developed markets, his allocations are roughly 55% to the U.S. market and 44% to developed markets. When the investor has access to emerging markets, his allocations are almost entirely skewed towards the U.S. market, with no investment in emerging markets. Hence, our results for the benchmark policy in the U.S.-W.D. case are comparable to theirs, albeit with lower allocations to developed markets. On the other hand, our benchmark allocations to emerging markets are higher, reflecting the fact that our sample ending in 1996 misses the 1997 and 1998 international financial crises.

Finally, some discussion of the statistical properties of the portfolio weights. Table 2 highlights the large variability of portfolio weights. For example, in the U.S.-WD case, the 5th and 95th percentiles of the distribution of the U.S. weight are 6% and 87%, respectively. Note, though, that this large variability of portfolio weights is not limited to the dynamic policies, but it is present also in the standard NC-NL case, where the investor is myopic and ignores cross-inference.

5.2 Sample Ending in 2006

At the end of 2006, benchmark allocations are quite similar to the 1996 allocations. In the benchmark cases, allocations to the U.S. market are of 44% (U.S.-WD), 49% (U.S.-EM), and 32% (U.S.-WD-EM), compared with 49%, 54%, and 39%, respectively. The allocations to developed markets are 28% (U.S-WD) and 18% (U.S.-WD-EM), the same as for the sample ending in 1996. Allocations to emerging markets are 14% (U.S-EM) and 21% (U.S.-WD-EM), to be compared with 17% and 26%, respectively.

As one may expect, these differences in portfolio weights are mainly driven by differences in the mean estimates, rather than by changes in correlation estimates. Indeed, we also computed optimal allocations using mean estimates at the end of 2006 while retaining the covariance estimates at the end of 1996. We detect only minor changes in both benchmark and sub-optimal allocations to the U.S. and WD indices. On the other hand, allocations to the EM index are consistently higher. This shows that the variance and covariance properties of emerging market returns at the end of 2006 contribute to making them less attractive, as compared to the earlier sample.

When it comes to sub-optimal policies, several of the effects are the same as for the other sample. For example, ignoring cross-inference (NC-L policy) reduces the WD allocation in both the U.S-WD and U.S.-WD-EM cases; and ignoring learning (C-NL policy) again increases allocations to the EM index in both the U.S-EM and U.S.-WD-EM cases.

Overall, deviations of sub-optimal policies from the benchmark are less pronounced than at the end of 1996, reflecting the fact that the longer sample reduces the effects of ignoring cross-inference and learning. Specifically, the largest turnover is now 22.8% for the NC-L policy in the U.S.-WD-EM cases.

6 Utility Costs

Table 3 presents two sets of utility costs associated with sub-optimal policies. First, we calculate the utility cost of employing a sub-optimal policy from the standpoint of the benchmark investor, when the investor *is* internationally diversified. In addition to the sub-optimal policies illustrated in Section 2.4, we also consider the $1/N$ policy, which allocates evenly across all available assets. Second, we calculate the cost of *not diversifying internationally*, from the standpoint of the benchmark investor, and from the standpoint of the sub-optimal investors. As with portfolio weights, we consider the three cases: U.S.-WD, U.S.-EM, and U.S.-WD-EM. Panel a presents results for the investment horizon starting in December 1996. Panel b presents results for the investment horizon starting in December 2006.

6.1 Sample Ending in 1996

It is worth noting that the ranking of utility costs of sub-optimal policies *does not* necessarily reflect the ranking of deviations of the sub-optimal portfolio policies. For example, in the U.S.-WD-EM case, the utility costs of the NC-L and C-NL policies are 2.08% and 1.47%, respectively, while the deviations from the benchmark policy are 20.4% and 27.4%, respectively.

The intuition for this apparently puzzling result is that the utility costs of sub-optimal policies can be essentially broken down into *two* components: the first component reflects only the current deviation of the sub-optimal portfolio policy from the benchmark policy; the second component reflects also the deviation of the next period's value function from the benchmark value function. The difference in ranking noted above is because the second component of the utility cost may be larger than the first one. (See the

Appendix A.5 for further discussion.)

Utility costs of sub-optimal policies can be substantial when both cross-inference and learning are ignored, amounting, for example, to 7.00% for the U.S.-EM case. It is the combination of ignoring both cross-inference and learning, together with the inclusion of emerging markets, that drives up the utility costs. Interestingly, the $1/N$ policy delivers the highest utility costs: as high as 10.31% for the U.S.-WD-EM case. This is to be contrasted with the out-of-sample results of the next section.

We can compare our results to those of Ang and Bekaert (2002) and Das and Uppal (2004). Ang and Bekaert estimate the utility cost of ignoring regime switching at about 6%.[13] As mentioned above, for the U.S.-EM scenario, our utility cost can be as high as 7.00% (NC-NL). Hence, accounting for cross-inference and learning can be almost as important as accounting for regime switching in correlations, in particular if emerging-market investment is included.

Similarly, Das and Uppal (2004) estimate the costs of ignoring systemic risk for an investor with various degrees of risk aversion, investment horizons, and asset combinations, under no-short-sale constraints. For an investor with a risk-aversion coefficient of 5 and a 5-year horizon, utility costs amount to 23 basis points for the case of international investment in developed markets and 1.5 basis points for the case of investment in emerging markets. Since their utility costs are proportional to the length of the investment horizon, the ten-year utility costs of ignoring systemic risk are mainly smaller than the costs that we estimate for the ignorance of cross-inference and/or learning.

We now turn to the costs of no international diversification. The cost can be substantial: 25.91% for the U.S.-WD-EM case, C-NL policy. Ignoring cross-inference and/or learning (NC-L, C-NL, and NC-NL) has mixed effects on the costs of no diversification, positive or negative depending on the assets available for investment.

As with the other sub-optimal policies, the costs of not diversifying internationally *are not* monotonically related to the allocations to international markets. The reason is that the costs of not diversifying internationally depend on both current *and* future restrictions on portfolio weights: the current deviation of restricted portfolio weights from the benchmark; and the deviation of the next period's restricted value function from the benchmark value function. Hence, it is possible that a given scenario displays lower costs of not diversifying internationally than a scenario with lower current international allocations.

Ang and Bekaert (2002) find that the utility costs of not diversifying internationally amount to about 6% of portfolio wealth. Their result is close to our 7.33% utility cost for the benchmark U.S.-WD case.

As in the case of portfolio weights, we computed the percentiles of the distribution of the utility costs,

[13]Note that Ang and Bekaert (2002) calculate utility costs as the percentage of wealth a sub-optimal investor should be *compensated with* to achieve the same expected utility as the benchmark investor; hence their utility costs can be converted to our measure of cost using the transformation: utility cost/(1 + utility cost).

drawing from the joint posterior density of the state variables evaluated at $t = t_0$. Utility costs, like portfolio weights, can display large variation. For example, the 5th and 95th percentiles of the distribution of the utility costs of not diversifying internationally, in the U.S.-EM benchmark case are 2.16% and 46.24%, respectively.

6.2 Sample Ending in 2006

Across the board, the costs of sub-optimal policies at the end of 2006 are smaller than for the shorter sample. For example, the cost of the NC-NL policy in the U.S.-EM case drops from 7.00% to 2.38%. This overall reduction in utility costs, going to a longer sample, is consistent with the fact that both the effects of cross-inference and learning tend to disappear as the sample lengthens. Interestingly, the utility costs of the $1/N$ policy are also lower than for the earlier sample.

7 Conclusions

This paper examines the effect of parameter uncertainty on asset allocations in a dynamic setting, where the investor learns about the mean returns over time, employs cross-inference in the estimation of mean returns, and can invest in domestic and international equity indices. At the methodological level, we show how the presence of cross-inference affects both the levels of the mean estimates and their evolution over time. This generates extra hedging demands for longer-history assets, which are absent in the context of learning without cross-inference. Moreover, we show that in the presence of multiple risky assets, the hedging demands induced by learning need not be negative for all risky assets. In a realistic calibration exercise, we show that ignoring cross-inference, learning, and estimation risk, or using the naïve $1/N$ policy, can lead to sizable utility costs. The absence of international diversification is also perceived to be costly.

Appendix

A.1 Solution Technique

The Bellman equation (7) is solved by backward iteration. Given the assumed joint normality of log returns, the density of log returns is approximated using a Gaussian quadrature rule with 20 quadrature points (see Tauchen and Hussey, 1991).[14] We use the product rule based on the Gauss-Hermite approximation to calculate the two-dimensional expectation, and we evaluate the expectation for different values of the state variables $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$.

For each pair $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$, we obtain analytically the values of $a_{1,t}$ and $a_{2,t}$ that minimize the expectation using a log-linear approximation of the portfolio return and assuming joint log-normality of portfolio returns and the value function. We approximate the log return on a portfolio as (see, for example, Campbell and Viceira, 2001, 2002)[15]

$$r_{p,t+1} = r_f + \mathbf{a}_t'(\mathbf{r}_{t+1} - r_f\mathbf{1}) + \frac{1}{2}\mathbf{a}_t'\mathrm{Var}_t(\mathbf{r}_{t+1}) - \frac{1}{2}\mathbf{a}_t'\mathrm{Cov}_t(\mathbf{r}_{t+1})\mathbf{a}_t, \tag{33}$$

where $\mathrm{Var}_t(\mathbf{r}_{t+1})$ is the *vector* of variances of the predictive density of returns.

We have

$$\begin{aligned} Q(\hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) &= \min_{\mathbf{a}_t} E_t \left\{ \exp\left[r_{p,t+1}(1-\gamma) \right] Q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right\} \\ &= \min_{\mathbf{a}_t} E_t \left\{ \exp\left[r_{p,t+1}(1-\gamma) + q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right] \right\}. \end{aligned} \tag{34}$$

Taking logs of both sides, we have

$$q(\hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) = \min_{\mathbf{a}_t} \ln E_t \left\{ \exp\left[r_{p,t+1}(1-\gamma) + q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right] \right\}. \tag{35}$$

We also have

$$\begin{aligned} q(\hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) &= \min_{\mathbf{a}_t} (1-\gamma) E_t(r_{p,t+1}) + E_t\left[q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right] \\ &\quad + \frac{1}{2}\mathrm{Var}_t\left[(1-\gamma) r_{p,t+1} + q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right] \\ &= \min_{\mathbf{a}_t} (1-\gamma) \left[r_f + \mathbf{a}_t'(E_t(\mathbf{r}_{t+1}) - r_f\mathbf{1}) + \frac{1}{2}\mathbf{a}_t'\mathrm{Var}_t(\mathbf{r}_{t+1}) - \frac{1}{2}\mathbf{a}_t'\mathrm{Cov}_t(\mathbf{r}_{t+1})\mathbf{a}_t \right] \\ &\quad + E_t\left[q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right] + \frac{1}{2}(1-\gamma)^2 \mathbf{a}_t'\mathrm{Cov}_t(\mathbf{r}_{t+1})\mathbf{a}_t \\ &\quad + \frac{1}{2}\mathrm{Var}_t\left[q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right] + (1-\gamma)\mathbf{a}_t'\mathrm{Cov}_t\left[\mathbf{r}_{t+1}, q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1) \right]. \end{aligned} \tag{36}$$

[14] The quadrature rule is applied to a bivariate standard normal with correlation coefficient equal to $\hat{\rho}_{12}$. We then use a linear transformation to obtain the mean vector and covariance matrix of the predictive density.

[15] As noted by Campbell and Viceira (2001), the approximation becomes exact in continuous time. In our implementation, we found the approximation to be very accurate, when compared to the maximizer obtained by grid search. The advantage of this approach is the speed of implementation.

Taking first-order conditions, we obtain

$$\mathbf{a}_t = \frac{1}{\gamma}\text{Cov}_t(\mathbf{r}_{t+1})^{-1}\left[E_t(\mathbf{r}_{t+1}) - r_f\mathbf{1} + \frac{1}{2}\text{Var}_t(\mathbf{r}_{t+1})\right] + \frac{1}{\gamma}\text{Cov}_t(\mathbf{r}_{t+1})^{-1}\text{Cov}_t\left[\mathbf{r}_{t+1}, q(\hat{\mu}_{1,t+1}, \hat{\mu}_{2,t+1}, t+1)\right]. \quad (37)$$

When the analytical solution above leads to negative allocations, we obtain the optimal constrained allocations imposing the standard Kuhn-Tucker conditions.

As for the state variables $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$, we consider 20 equally-spaced grid points over a range roughly centered around the realizations of the state variables at the end of 1994 and 2000.[16]

Also, note that the law of motion of the estimates $\hat{\mu}_{i,t}$ presents some interesting numerical challenges. First, the conditional density of the variables is time varying, hence the direct application of the Gaussian quadrature method to $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$ is not feasible. Instead, we approximate the standardized increments in the mean estimates, $(\hat{\mu}_{i,t} - \hat{\mu}_{i,t-1})/\text{Var}_{t-1}(\hat{\mu}_{i,t})$, which are normal with constant mean (zero) and variance (one). Second, since the mean estimates are not mean reverting, the approximation of the value function at the boundaries of the state space is problematic. We deal with this problem by choosing a range that goes well beyond the extreme realizations of the estimates, to ensure that the approximation of the value function inside that range is accurate.

The procedure above is repeated for each period from time $T-1$ to time t (note that the value function $Q(.)$ equals one at time T). For each period, we calculate the value of $Q(.)$ in between points of the grid by linear interpolation. When expectations are taken close and at the boundaries of the range for $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$, values of $Q(.)$ beyond the boundaries are approximated by the values *at* the boundaries. This approach proved to work better than the linear extrapolation of the value function beyond the boundary values.

A.2 Predictive Density

We derive the covariance matrix of the Bayesian predictive density of returns for the i.i.d. case. We start with the predictive variance of $r_{1,t+1}$:

$$\begin{aligned}
\text{Var}_t(r_{1,t+1}) &= \text{Var}_t(r_{1,t+1}|\mu_{1,t}) + \text{Var}_t(\mu_{1,t}|r_{1,1}, r_{1,2}, \ldots, r_{1,t}) \\
&= \hat{\sigma}_1^2 + \frac{1}{t}\hat{\sigma}_1^2 = \left(1 + \frac{1}{t}\right)\hat{\sigma}_1^2, \qquad (38)
\end{aligned}$$

where we replaced $\text{Var}_t(r_{1,t+1}|\mu_{1,t})$ with its ML estimate and we used the fact that the mean of the posterior density of $\mu_{1,t}$ is the ML estimate. The predictive variance of $r_{2,t+1}$ is given by

$$\text{Var}_t(r_{2,t+1}) = \text{Var}_t(r_{2,t+1}|\mu_{2,t}) + \text{Var}_t(\mu_{2,t}|r_{1,1}, r_{1,2}, \ldots, r_{1,t}, r_{2,s_2}, r_{2,s_2+1}, \ldots, r_{2,t})$$

[16] We explored the robustness of our results to changes in the width and the density of the grid. When we double the size of the grid, both allocations and utility costs are essentially unchanged. When we double the density of the grid, allocations are again unchanged, while there are some changes in utility costs, although the ranking of the costs is preserved.

$$= \quad \hat{\sigma}_2^2 + \frac{1}{t-s_2+1}(\hat{\sigma}_2^2 - \hat{\beta}_{21}^2\hat{\sigma}_1^2) + \beta_{21}^2\frac{1}{t}\hat{\sigma}_1^2 = \hat{\sigma}_2^2 + \hat{\sigma}_2^2\left(\frac{1-\rho_{12}^2}{t-s_2+1} + \frac{\rho_{12}^2}{t}\right), \tag{39}$$

where we replaced $\mathrm{Var}_t(r_{2,t+1}|\mu_{2,t})$ with its ML estimate, we used the fact that the mean of the posterior density of $\mu_{2,t}$ is the ML estimate, and we used the fact that $\mathrm{Cov}_t(\hat{\alpha}_{2,t}, \hat{\mu}_{1,t}) = 0$.[17] Finally, the predictive covariance between $r_{1,t+1}$ and $r_{2,t+1}$ is given by

$$\begin{aligned}
\mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1}) &= \mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1}|\mu_1, \mu_2) + \mathrm{Cov}_t(\mu_1, \mu_2|r_{1,1}, r_{1,2}, \ldots, r_{1,t}, r_{2,s_2}, r_{2,s_2+1}, \ldots, r_{2,t}) \\
&= \hat{\sigma}_{12} + \hat{\beta}_{21}\frac{1}{t}\hat{\sigma}_1^2 = \hat{\sigma}_{12} + \frac{1}{t}\hat{\sigma}_{12} = \left(1 + \frac{1}{t}\right)\hat{\sigma}_{12}. \qquad (40)
\end{aligned}$$

A.3 Cross-inference and Hedging Demands

We gain a better understanding of the extra hedging demands induced by cross-inference with the help of equation (37). Consider the hedging-demand component of the allocation to asset 1, the long-history asset. We have

$$\frac{1}{\gamma}\frac{1}{|\mathrm{Cov}_t(\mathbf{r}_{t+1})|}\left[\mathrm{Var}_t(r_{2,t+1})\mathrm{Cov}_t(r_{1,t+1}, q(., t+1)) - \mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1})\mathrm{Cov}_t(r_{2,t+1}, q(., t+1))\right], \tag{41}$$

where $\mathrm{Cov}_t(r_{i,t+1}, q(., t+1))$ is the predictive covariance. Using a linearization of $q(., t+1)$ around $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$, we can write the term in squared parenthesis as

$$\begin{aligned}
&\mathrm{Var}_t(r_{2,t+1})\left[\frac{\partial q(., t+1)}{\partial \hat{\mu}_{1,t+1}}\frac{1}{t+1}\mathrm{Var}_t(r_{1,t+1}) + \frac{\partial q(., t+1)}{\partial \hat{\mu}_{2,t+1}}\frac{1}{t-s_2+2}\mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1})\right] \\
&-\mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1})\left[\frac{\partial q(., t+1)}{\partial \hat{\mu}_{1,t+1}}\frac{1}{t+1}\mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1}) + \frac{\partial q(., t+1)}{\partial \hat{\mu}_{2,t+1}}\frac{1}{t-s_2+2}\mathrm{Var}_t(r_{2,t+1})\right] \\
&-\frac{\partial q(., t+1)}{\partial \hat{\mu}_{2,t+1}}\hat{\beta}_{21}\left(\frac{1}{t-s_2+2} - \frac{1}{t+1}\right)[\mathrm{Var}_t(r_{1,t+1})\mathrm{Var}_t(r_{2,t+1}) - \mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1})^2]. \qquad (42)
\end{aligned}$$

The last term in the above expression, which is positive as long as $\hat{\beta}_{21} > 0$, is due to the effect of cross-inference.[18]

A.4 Multiple Risky Assets. Learning, and Hedging Demands

Based on equation (37), the hedging-demand component of the allocation to asset 1 is

$$\frac{1}{\gamma}\frac{1}{|\mathrm{Cov}_t(\mathbf{r}_{t+1})|}\left[\mathrm{Var}_t(r_{2,t+1})\mathrm{Cov}_t(r_{1,t+1}, q(., t+1)) - \mathrm{Cov}_t(r_{1,t+1}, r_{2,t+1})\mathrm{Cov}_t(r_{2,t+1}, q(., t+1))\right].$$

Since realized returns and mean return estimates are positively correlated, $\mathrm{Cov}_t(r_{i,t+1}, q(., t+1)) < 0$. When the two rates of return are positively correlated and $\mathrm{Cov}_t(r_{2,t+1}, q(., t+1))$ is "large," the sign of the hedging-demand component of the allocation can be positive.

[17] We have $\mathrm{Cov}_t(\hat{\alpha}_{2,t}, \hat{\mu}_{1,t}) = \frac{1}{t}(\hat{\sigma}_{12} - \hat{\beta}_{21}\hat{\sigma}_1^2) = 0$, since $r_{2,\tau} = \alpha_2 + \hat{\beta}_{21}r_{1,\tau} + \epsilon_{2,\tau}$.

[18] $\partial q(., t+1)/\partial r_{2,t+1} < 0$, since $q(., t+1)$ is negatively related to the value function for $\gamma > 1$

A.5 Value Function and Utility Costs

Based on (36), we can write the $q(., t)$ function as

$$\begin{aligned} q(\hat{\mu}_{1,t}, \hat{\mu}_{2,t}, t) &= (1-\gamma)E_t(r_{p,t+1}) + \frac{1-\gamma}{2}\text{Var}_t(r_{p,t+1}) \\ &+ (1-\gamma)\text{Cov}_t\left[r_{p,t+1}, q(., t+1)\right] + E_t\left[q(., t+1)\right] + \frac{1}{2}\text{Var}_t\left[q(., t+1)\right], \end{aligned} \tag{43}$$

where $r_{p,t+1}$ is the return on the *optimal* portfolio. When we compare two value functions, as we do in the calculation of utility costs, we compare the different components of the value functions. The first two terms in the equation above reflect only current portfolio weights, whereas the last three terms in the equation above reflect also future realizations of the value function.

References

Ang, A., Bekaert, G., 2002. International asset allocation with regime shifts. *Review of Financial Studies* 15, 4, 1137-1187.

Balduzzi, P., Lynch, A., 1999. Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics* 52, 47-78.

Barberis, N., 2000. Investing for the long run when returns are predictable. *Journal of Finance* 55, 225-264.

Barry, C., Brown, S., 1984. Differential information and the small firm effect. *Journal of Financial Economics* 13, 283-294.

Barry, C., Brown, S., 1985. Differential information and security market equilibrium. *Journal of Financial and Quantitative Analysis* 20, 407-422.

Bawa, V., Brown, S., Klein, R., 1979. Estimation risk and optimal portfolio choice. Amsterdam: North Holland.

Bekaert, G., Urias, M., 1996. Diversification, integration, and emerging-market closed-end funds. *Journal of Finance*, 38, 925-984.

Brandt, M., Goyal, A., Stroud, J., 2005. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Review of Financial Studies* 18, 831-873.

Brennan, M., 1998. The role of learning in dynamic portfolio decisions. *European Finance Review* 1, 295-306.

Britten-Jones, M., 1999. The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance* 54, 655-671.

Campbell, J., Viceira, L., 2001. Appendix to: Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. Unpublished working paper, Harvard Business School.

Campbell, J., Viceira, L., 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. Oxford University Press, Oxford.

Das, S., Uppal, R., 2004. Systemic risk and international portfolio choice. *Journal of Finance* 59, 2809-2834.

De Roon, F., Nijman, T., Werker, B., 2001. Testing for mean-variance spanning with short sales constraints and transaction costs: the case of emerging markets. *Journal of Finance* 56, 721-742.

Errunza, V., Hogan, K., Hung, M., 1999. Can the gains from international diversification be achieved without trading abroad? *Journal of Finance* 54, 1999.

Harvey, C., 1995. Predictable risk and return in emerging markets. *Review of Financial Studies* 8, 773-816.

Horst, J., De Roon, F., Werker, B., 2006. Incorporating estimation risk in portfolio choice. *Advances in Corporate Finance and Asset Pricing*, Renneboog, L., editor, Emerald Group Publishing.

Im, E., 1994. Unequal numbers of observations and partial efficiency gain. *Economics Letters* 46, 291-294.

Jorion, P., 1985. International portfolio diversification with estimation risk. *Journal of Business* 58, 259-278.

Kandel, S., Stambaugh, R., 1996. On the predictability of stock returns: an asset allocation perspective. *Journal of Finance* 51, 385-424.

Kunsch, H., 1989. The Jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* 17, 1217-1241.

Li, K., Sarkar, A., Wang, Z., 2003. Diversification benefits of emerging markets subject to portfolio constraints. *Journal of Empirical Finance* 10, 57-80.

Liu, R. and Singh, K., 1992. Moving blocks jackknife and bootstrap capture weak dependence, in *Exploring the Limits of Bootstrap.* R. LePage and L. Billard, eds. New York: John Wiley & Sons.

Samuelson, P., 1969. Lifetime portfolio selection by dynamic stochastic programming. *The Review of Economics and Statistics* 51, 239-246.

Stambaugh, R., 1997. Analyzing investments whose histories differ in length. *Journal of Financial Economics* 45, 285-331.

Tauchen, G., Hussey, R., 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica* 59, 317-396.

Xia, Y., 2001. Learning about predictability: the effect of parameter uncertainty on optimal dynamic asset allocation. *Journal of Finance* 56, 205-246.

Table 1

Parameter Estimates

The table presents maximum-likelihood estimates of the parameters used for the dynamic portfolio optimization. Panel a presents estimates based on the sample ending in 1996. Panel b presents estimates based on the sample ending in 2006. In both cases, we report estimates obtained using the combined samples (Benchmark) and using separate samples (NC). We consider three stock indices: U.S. market index (U.S.), world developed markets excluding U.S. index (WD), and emerging market index (EM). $\hat{\mu}_{i,t}$ denotes the mean estimate, $\hat{\sigma}_i$ denotes the standard deviation estimate, and $\hat{\sigma}_{ij}$ denotes the estimate of the covariance between two market indices. In parenthesis, we report asymptotic standard errors of the estimates. Estimates of means and standard deviations are in percentage points per month. Estimates of covariances are in basis points.

Panel a: 1996

	Benchmark				NC	
	U.S.-WD-EM			U.S.-EM		
	U.S.	WD	EM	EM	WD	EM
$\hat{\mu}_{i,t}$	0.64	0.64	1.11	0.85	0.56	1.04
	(0.17)	(0.25)	(0.53)	(0.55)	(0.27)	(0.58)
$\hat{\sigma}_i$	4.14	4.80	5.83	5.91	4.94	5.73
	(0.13)	(0.18)	(0.44)	(0.46)	(0.19)	(0.42)
$\hat{\sigma}_{ij}$		9.89	10.29	10.95	12.02	7.16
		(0.80)	(0.87)	(0.88)	(0.97)	(1.01)
			11.91			11.29
			(0.96)			(1.28)

Panel b: 2006

	Benchmark				NC	
	U.S.-WD-EM			U.S.-EM		
	U.S.	WD	EM	EM	WD	EM
$\hat{\mu}_{i,t}$	0.61	0.61	0.91	0.74	0.53	0.82
	(0.16)	(0.21)	(0.38)	(0.39)	(0.23)	(0.44)
$\hat{\sigma}_i$	4.23	4.65	6.48	6.57	4.79	6.53
	(0.12)	(0.15)	(0.29)	(0.31)	(0.16)	(0.31)
$\hat{\sigma}_{ij}$		11.39	15.81	16.78	13.38	16.03
		(0.84)	(0.99)	(1.00)	(0.95)	(1.24)
			16.72			17.41
			(1.06)			(1.30)

Table 2

Portfolio Weights

The table reports optimal portfolio weights for allocations to four assets: U.S. market index (U.S.), world developed market excluding U.S. index (WD), emerging market index (EM), and the risk-free asset (RF). Below the weights, in parenthesis, we report the 5th and 95th percentiles of the distribution of the weights, drawing from the posterior density of the parameter estimates. Weights are calculated based on information from the sample ending in 1996 (Panel a) and the sample ending in 2006 (Panel b). We consider three combinations of the four assets: U.S., developed markets, and risk-free asset (U.S.-WD); U.S., emerging markets, and risk-free asset (U.S.-EM); U.S., developed markets, emerging markets, and risk-free asset (U.S.-WD-EM). We calculate optimal weights for four policies. The benchmark policy accounts for cross-inference and learning. The NC-L policy does not account for cross-inference. The C-NL policy does not account for learning. The NC-NL policy accounts for neither cross-inference nor learning. "Turn." denotes the overall turnover of switching from the benchmark to the sub-optimal policy. Note that in some cases the weights do not sum up to exactly one. This is because the weights, in percentage points, are rounded to the closest percentage point.

Panel a: 1996

	U.S.-WD				U.S.-EM				U.S.-WD-EM				
	U.S.	WD	RF	Turn.	U.S.	EM	RF	Turn.	U.S.	WD	EM	RF	Turn.
Benchmark	49	28	23		54	17	29		39	18	26	17	
	[6, 87]	[0, 70]			[8, 91]	[0, 59]			[0, 76]	[0, 61]	[0, 72]		
NC-L	57	13	31	15.0	28	35	37	26.1	41	1	23	35	20.4
	[14, 84]	[0, 66]			[0, 64]	[0, 89]			[0, 64]	[0, 45]	[0, 94]		
C-NL	55	36	9	14.4	51	39	10	21.1	30	16	54	0	27.4
	[4, 95]	[0, 82]			[0, 97]	[0, 100]			[0, 76]	[0, 65]	[0, 100]		
NC-NL	66	17	17	17.1	33	64	3	46.9	48	3	43	6	26.5
	[16, 94]	[0, 73]			[0, 76]	[0, 100]			[0, 74]	[0, 51]	[0, 100]		

Panel b: 2006

	U.S.-WD				U.S.-EM				U.S.-WD-EM				
	U.S.	WD	RF	Turn.	U.S.	EM	RF	Turn.	U.S.	WD	EM	RF	Turn.
Benchmark	44	28	28		49	14	37		32	18	21	29	
	[3, 81]	[0, 68]			[3, 86]	[0, 50]			[0, 72]	[0, 60]	[0, 56]		
NC-L	55	11	35	17.2	34	23	42	14.4	45	3	13	39	22.8
	[12, 79]	[0, 61]			[0, 67]	[0, 67]			[0, 64]	[0, 47]	[0, 67]		
C-NL	49	34	17	10.9	50	22	28	9.2	31	20	30	19	11.8
	[2, 89]	[0, 80]			[0, 94]	[0, 70]			[0, 76]	[0, 68]	[0, 78]		
NC-NL	62	13	25	17.8	40	33	27	18.8	51	5	18	25	19.1
	[14, 88]	[0, 71]			[0, 77]	[0, 93]			[0, 73]	[0, 55]	[0, 92]		

Table 3

Utility Costs

The table reports the utility cost of using a sub-optimal policy for a benchmark investor. The table also reports the perceived utility cost to a benchmark investor and a sub-optimal investor of not diversifying internationally. Below the utility costs, in parenthesis, we report the 5th and 95th percentiles of the distribution of the costs, drawing from the posterior density of the parameter estimates. Utility costs are calculated based on information from the sample ending in 1996 (Panel a) and the sample ending in 2006 (Panel b). We consider three combinations of the four assets: U.S., developed markets, and risk-free asset (U.S.-WD); U.S., emerging markets, and risk-free asset (U.S.-EM); U.S., developed markets, emerging markets, and risk-free asset (U.S.-WD-EM). The benchmark policy accounts for cross-inference and learning. The NC-L policy does not account for cross-inference. The C-NL policy does not account for learning. The NC-NL policy accounts for neither cross-inference nor learning. The $1/N$ policy allocates funds evenly across investment choices. Utility costs are in percentage points.

Panel a: 1996

Sub-optimal policy			
	U.S-WD	U.S.-EM	U.S.-WD-EM
NC-L	0.51	1.87	2.08
	[0.08, 1.79]	[1.03, 13.47]	[1.11, 10.73]
C-NL	0.28	1.10	1.47
	[0.04, 0.45]	[0.23, 3.46]	[0.24, 3.20]
NC-NL	0.71	7.00	4.97
	[0.21, 1.99]	[2.71, 37.20]	[2.54, 25.49]
1/N	4.84	8.79	10.31
	[1.75, 26.21]	[2.99, 40.06]	[2.10, 43.70]
No international diversification			
	U.S-WD	U.S.-EM	U.S.-WD-EM
Benchmark	7.33	12.09	22.90
	[0.51, 26.46]	[2.16, 46.24]	[5.48, 57.19]
NC-L	4.27	14.72	16.95
	[0.12, 21.48]	[0.78, 62.87]	[0.94, 64.49]
C-NL	6.58	11.95	25.91
	[0.00, 31.74]	[0.00, 65.34]	[0.66, 73.66]
NC-NL	3.63	20.51	21.86
	[0.00, 26.78]	[0.00, 82.90]	[0.00, 83.48]

Panel b: 2006

Sub-optimal policy			
	U.S-WD	U.S.-EM	U.S.-WD-EM
NC-L	0.47	1.11	1.44
	[0.08, 1.44]	[0.45, 7.87]	[0.64, 6.65]
C-NL	0.18	0.30	0.41
	[0.04, 0.28]	[0.08, 1.06]	[0.11, 1.12]
NC-NL	0.58	2.38	2.12
	[0.16, 1.55]	[0.90, 16.62]	[1.06, 11.32]
1/N	3.92	8.11	7.67
	[1.43, 23.46]	[2.33, 38.42]	[2.25, 36.54]
No international diversification			
	U.S-WD	U.S.-EM	U.S.-WD-EM
Benchmark	6.22	6.69	14.25
	[0.52, 22.79]	[0.76, 29.13]	[3.33, 39.94]
NC-L	3.51	6.82	9.09
	[0.13, 18.03]	[0.26, 39.16]	[0.61, 43.29]
C-NL	4.97	3.98	11.41
	[0.00, 25.18]	[0.00, 34.96]	[0.33, 45.95]
NC-NL	2.48	6.11	7.33
	[0.00, 20.52]	[0.00, 52.49]	[0.00, 55.74]