Evaluation of Active Management of the Norwegian Government Pension Fund – Global

December 14, 2009

Andrew Ang
Ann F. Kaplan Professor of Business, Columbia Business School

William N. Goetzmann
Edwin J. Beinecke Professor of Finance and Management Studies, Yale School of Management

Stephen M. Schaefer
Professor of Finance, London Business School

[This page is intentionally left blank]

Preface

It is a great honor for the authors of this report to be chosen to evaluate the active management of the Norwegian Government Pension Fund – Global (the Fund). We are mindful that this is a service performed for the Norwegian people and that it addresses questions of importance to their economic future. The Fund is an extraordinary endowment created by the Norwegian people. The evaluation we undertake on their behalf seeks to provide input to help realize the full potential of this investment.

[This page is intentionally left blank]

Contents

Mandate

Preceding the report, we first review the mandate and the approach we took to fulfill it.

i) Evaluation of NBIM's historical track record:

- *Prepare a quantitatively based evaluation of the performance of NBIM's active management. The analysis should distinguish between the two asset classes (fixed income and equities) and, within the limits of available data, include a breakdown of performance by main active strategies that were used by NBIM in the evaluation period. The analysis should also cover the internally versus externally managed funds dimension. Methodologies for evaluating return, risk and risk adjusted return may be adapted to the particular area under scrutiny.*
- *Prepare a qualitatively based evaluation of NBIM's active management. This part of the report should include assessments of the strategic plans for active management, the risk budgeting process, internal and external reporting of active management risks and return, and whether or to what extent the specific strategies that have been used in NBIM's active management have been based on exploiting the characteristics of the Fund (cf. third bullet in iii)*

ii) Empirical studies of active management/tests of the efficient market hypothesis (EMH):

- *Describe briefly the efficient market hypothesis from a theoretical perspective.*
- *Present a survey of high-quality empirical studies of tests of the EMH in general with particular emphasis on tests of relevance to evaluations of active management performance, and discuss to what extent the relevance of the EMH varies across different markets and assets.*
- *Evaluate the relevance of empirical evidence for the evaluation of active management of the Fund.*

iii) Exploitation of the Fund's characteristics

- *Assess whether or to what extent a large long term investor like the Fund has comparative advantages and capabilities which justify utilizing significant resources on active management.*
- *Assess whether or to what extent such comparative advantages can be implemented successfully in all asset classes included in the Fund's investments,*

and discuss to what extent such comparative advantages and capabilities will affect the Fund's capability to enter into new asset classes.

- *Describe and evaluate strategies that utilize possibly comparative advantages. How can the risk/return-profile of each strategy be described in terms of a probability distribution of outcomes, and what is the verification horizon of each strategy? Address challenges in designing a proper (short term) incentive system for the active manager. In the case of strategies with long verification horizons, would it be better to include such strategies in the strategic benchmark (possibly by using a dynamic asset allocation approach)?*

Approach

In addressing the mandate for this report, we reordered the three main sections to begin with the review of the Efficient Market Hypothesis [EMH]. The rationale behind this change is that theory and empirical evidence provide a useful foundation for the evaluation of NBIM's (Norges Bank Investment Management) track record. Many of the techniques used for performance measurement result from tests of efficient markets. Thus, the report is organized as (i) a review of the theory and evidence on the EMH, (ii) an evaluation of NBIM's historical track record, and (iii) an assessment of the potential to use the Fund's comparative advantages to benefit the management process.

In each section we have sought to highlight the relevance of our findings to the topic of this report – the role of active management. Our intent is to offer as much practical insight as possible from the historical review, the quantitative analysis, and the strategic assessment.

Data and Other Inputs

In preparing the report we have reviewed the extensive published academic literature on the EMH (documented in the references). For our data analysis, we have relied upon historical monthly return numbers from NBIM on an aggregate fund basis as well as on performance broken down by fixed income and equity strategies and by external vs. internal performance. NBIM also provided information about fees, capital allocations to external funds, and benchmarks returns, all in NOK. Given the sensitive nature of this information it is not all detailed in our report. In addition we have used return information from global capital markets including fixed income and equity indices, volatility indices, liquidity indices, and other external data we deemed relevant to assess risk and return. These have been converted where necessary into NOK for appropriate comparison.

Process

We have had the advantage of several meetings with NBIM management to discuss the NBIM's investment approach, views on comparative advantage, processes for utilization of external managers, and other issues of importance in the assessment of its active

management. In addition, NBIM has been responsive to our regular requests for data and other technical questions. We thank them for their time, openness, and responsiveness during this process.

We had the opportunity to share some preliminary findings with the Ministry of Finance and NBIM in September and November 2009. We also thank the Expert Panel of the Ministry of Finance for comments. These meetings focused on the implications of the EMH, a preliminary report on NBIM performance, and discussion of the proposed recommendations for the management of the Fund. The feedback and further questions we received from these preliminary meetings were very useful in refining the analysis.

Overall Summary

The objective of this report is to evaluate the role of active management in the Norwegian Government Pension Fund – Global ("the Fund"). The project brief called for a review of the academic research on active vs. passive (index) management and the efficiency of markets, an historical analysis of NBIM's track record, and a prospective analysis of how to use the Fund's comparative advantages to benefit the management process.

In line with the brief, the report has three main sections. The first undertakes a review of the theory and empirical evidence on the Efficient Market Hypothesis [EMH]. If markets are completely efficient, active portfolio management has little potential to add to fund performance. Therefore, in assessing the potential role of active management for the Fund, it is important to reach a view on both the theoretical basis and empirical support for the EMH. The second contains a detailed empirical analysis of NBIM's active performance in managing the Fund, identifying both the average returns in excess of benchmarks and the risks to which the Fund has been exposed. The analysis of performance differentiates between internally and externally managed funds and between different asset classes and strategies. The final part of the report looks at the implications of the first two sections for the future management of the Fund. In particular, it identifies the comparative advantages of the Fund and asks, in the light of the review of the theory and evidence on the EMH and the Fund's performance, what types of portfolio strategies and performance benchmarks should be pursued.

The most recent expressions of the EMH in academic research recognize the existence of market frictions, information costs, agency, and capital structure constraints. In this setting there is no compelling theoretical or empirical evidence to recommend a pure strategy of passive indexing although finding active managers who consistently deliver excess risk-adjusted returns to investors is rare. In evaluating NBIM's track record, we find that active management has played a very small role in its performance to date. The incremental contribution of active management has been slightly positive overall, with notable negative returns during 2008 and the early part of 2009. Analysis of NBIM's active management style indicates that a significant component of performance is explained by exposure to systematic factors which fared very poorly during the financial crisis. We believe that exposure to such factors is actually appropriate for a long-term investor since the factors earn risk premiums over the long run, however, we recommend that, to the fullest extent possible, these factor exposures taken by active management be treated as part of the Fund benchmark. We propose a new framework to implement this change. In the same way the Fund has chosen to hold a significant

position in equities to capture the equity risk premium over the long run, we recommend that other factor risk premiums be accessed by the Fund to generate long-term positive performance. The proposed framework exploits the Fund's distinct long-horizon profile and allows the preferences of the asset owners to determine how much exposure to each risk factor should be taken.

Summary: Section I

In simple terms, the EMH asserts that, at all times, the price of a security reflects all available information about its fundamental value. A consequence of the theory is that, if true, it is impossible for an investment manager – and hence the clients of the manager – to consistently beat the market. The underlying principle driving the EMH is that, in a large, active marketplace for publicly traded securities, vigorous competition among scores of investors will drive speculative profits to zero. The implication of the EMH for investors is that, to the extent that speculative trading is costly, speculation must be a loser's game. Hence, an indexing strategy is bound eventually to beat a strategy that uses active management, where active management is characterized as trading that seeks to exploit mispriced assets relative to a risk-adjusted benchmark. In the world of the EMH, there are no mispriced assets because the invisible hand of the marketplace moves faster than any single agent.

In Section I of the report we review the extensive theoretical and empirical literature on the EMH. Our review of theory indicates that the EMH has been refined over the past several decades to reflect the realism of the marketplace, including costly information, transactions costs, financing, agency costs, and other real-world frictions. The most recent expressions of the EMH thus allow a role for arbitrageurs in the market who may profit from their comparative advantages. These advantages may include specialized knowledge, lower trading costs, low management fees or agency costs, and a financing structure that allows the arbitrageur to undertake trades with long verification periods. The actions of these arbitrageurs cause liquid securities markets to be generally fairly efficient with respect to information, despite some notable anomalies. As a result, the balance between indexation and active management is a choice variable for which the optimum depends on general beliefs about the existence and potential of manager skill, the pricing opportunities afforded within a given market, the time preferences and risk aversion of the investor, and the expertise and incentive contract of the specific manager.

Our review of the empirical tests of the EMH is divided into two parts: tests on prices and tests on investment managers and institutions. Tests of the theory using past price behavior in the stock and bond markets have occasionally produced evidence contrary to the conclusion of efficiency, suggesting that the EMH may not hold for all markets and all times. The logical foundation for these tests is a pricing model that represents the "fair" price of a security in terms of its exposure to a set of common risk factors. The simplest of these models is the Capital Asset Pricing Model [CAPM], and the most commonly used in recent times is a multi-factor model called the Arbitrage Pricing Theory [APT]. The APT holds that the investor will be compensated by higher returns for accepting the risk implied by exposure to these factors.

Both the CAPM and the APT stress the important role that risk factors play in determining the expected future return of investment in an asset. Tests of the EMH in this framework are implicitly joint tests of the pricing model and market efficiency, however. Much recent debate has focused on whether such violations should be interpreted as inefficiency, or simply the inability of researchers to correctly identify and specify the risk factors relevant to the market. Regardless, violations of the EMH based on pricing tests are of potential interest to the Fund because they may indicate sources of returns – whether these are factor-based or based on the pricing inefficiency.

If the benchmark is solely a market-weighted portfolio consisting of all traded securities, then active management (defined as deviations from these market weights) may be useful in accessing factor risk premiums which are not captured by market exposure. In the context of the APT this could also be interpreted as passive exposure to additional risk factors. Theory and empirical evidence suggests that investors are compensated for taking systematic risks – such as investing in "value" stocks vs. "growth" stocks and volatility risk – over the long term. In the presence of these multiple systematic risk factors, empirical tests overwhelmingly reject that the market portfolio is efficient and other static or time-varying combinations of assets result in higher reward-to-risk ratios.

The back-tests of trading strategies seeking pure alpha have suggested a wide array of potentially profitable investments. However, for a number of reasons these provide limited guidance to the Fund going forward. They represent simulated, not actual, returns and do not account for actual transactions costs, fees, and price impact. They also suffer from potential data-mining biases. Changing market conditions, including time-varying arbitrage activity, make it difficult to extrapolate future performance. Finally, many anomalies are not scalable and cannot be implemented in the large position sizes relevant to the Fund.

The second part of Section I reviews tests of the EMH using information on managers and institutions. These are more relevant to the Fund since they reflect the necessary role of an investment intermediary. Unfortunately, many of the studies of managerial skill focus on retail mutual funds, whose cost structure and value proposition differ markedly from the Fund's, and thus may not be a useful benchmark. Recent theory and empirical evidence suggests that some fund managers may have talent and out-perform market benchmarks before fees. However, the evidence does not support the conclusion that superior ability filters through to the ultimate investors in those funds. In the mutual fund industry, after-fee returns and alphas are, on average, zero or negative. While the average mutual fund typically underperforms a passive portfolio on an after-fee risk-adjusted basis, there is evidence that some managers are better than others.

Turning to the non-retail sector, there is some evidence of positive post-fee risk-adjusted returns in hedge funds where highly paid managers actively trade marketable securities. One caveat is that the quality and duration of these data, as well as the changing institutional marketplace for hedge fund services make it difficult to extrapolate such conclusions to future performance. By contrast, there is little convincing evidence of superior risk-adjusted returns to private equity and venture capital. Although some studies suggest skill persistence, the current data are not conclusive on this point. In the real estate sector there is simply not enough information to evaluate whether managers have added value on a risk-adjusted basis.

In other institutional investment sectors, such as large-scale endowments, pension funds and sovereign funds, there is even less evidence about the capability of active management to generate positive risk-adjusted returns. Some U.S. endowments performed exceedingly well prior to the recent crisis using alternative investments as the basis for their strategy. It is often noted that a long-horizon perspective allowed these endowments to focus on alternative asset classes. Most research suggests that pension fund managers are not able to identify top managers ex ante and the managers that serve the pension fund sector show little evidence of skill on a risk-adjusted basis. The few studies of sovereign fund trades in public securities provide evidence that, while stock prices respond positively when a sovereign fund invests, the long-term performance of these investments is not particularly good.

In summary, the EMH has been refined over the past several decades to reflect information, transactions, financing and agency costs. Tests of the theory on prices have produced violations suggestive of the potential for active management to add

value to a multi-asset portfolio, but finding consistent out-performing active managers is difficult.

Summary: Section II

Section II evaluates NBIM's historical track record of active management. It is based on quantitative analysis of data provided by NBIM and meetings with NBIM personnel to discuss their general strategy and philosophy of active management. We analyze quantitative data on performance for the whole Fund, the two asset class divisions fixed income and equities, and internal and external active management within each asset class division for the period from January 1998 to September 2009 at the monthly frequency.

In our view NBIM is set up to provide two services to its client, the people and future generations of Norway. First, it offers "passive" returns based on the benchmark from the Ministry of Finance. The components of this benchmark are determined by the desire to efficiently diversify the overall portfolio. The benchmark takes the form of a portfolio allocation by region and asset class and is intended to be a potentially tradable portfolio of securities, as opposed to a hypothetical benchmark. As such, it represents a realistically feasible alternative to active management. NBIM seeks to offer this benchmark return in an operationally secure and transparent manner at lowest possible cost to the client.

Second, NBIM offers active management that seeks to add positive, risk-adjusted return over the benchmark net of fees. NBIM pursues this goal through a combination of internal and external management, and a philosophy of outsourcing many aspects of its back-office operations. Thus, it has internal portfolio managers and traders as well as a number of external investment managers whom it engages via specific mandates tied to pre-determined benchmarks.

NBIM implements active management in two main ways. The first is the decision to deviate from long-term strategic loadings on factors (e.g. a temporary shift from the target allocation to equities) and second, the decision to hold securities in weights that differ from factor benchmark weights based on fundamental analysis. These roughly correspond to timing and selection, where the default, or the baseline case, is the Fund benchmark. This approach is employed for both internally and externally managed funds. Fundamental analysis and factor exposure operate quite differently: the former looks at investment opportunities from the "bottom-up" by finding attractive companies

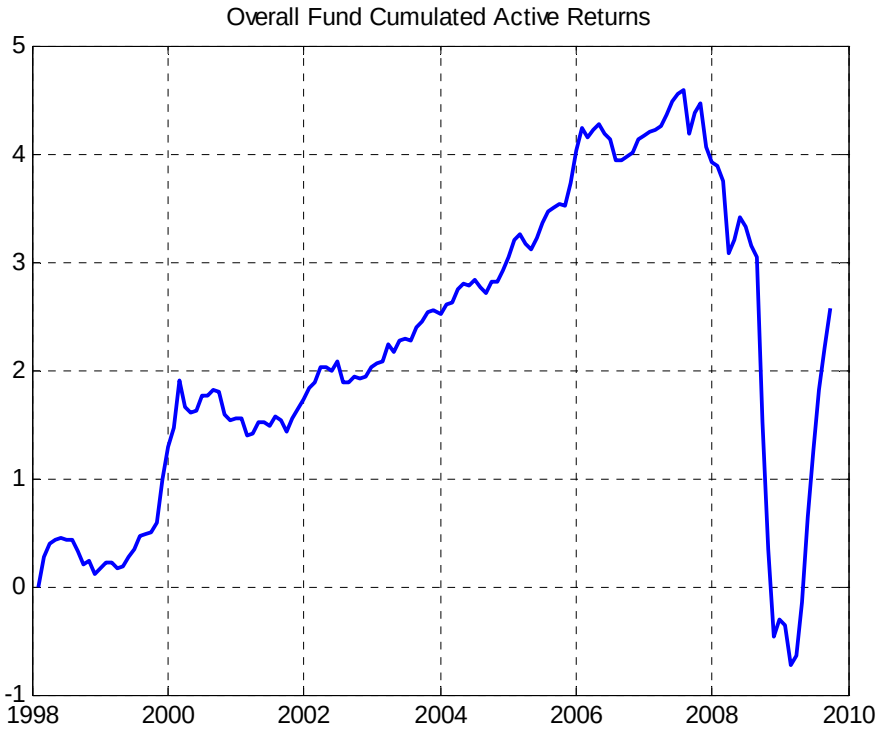
and securities and then aggregates to the portfolio level whereas the latter is a "top-down" investment technique that first chooses different factors and then implements them by trading securities. Given NBIM's stated approach to generating value, it is natural to ask whether NBIM has the capability of internally generating value as well as identifying external managers who add value. Our quantitative analysis addresses this question.

We define the "active return" of a fund, i.e., the part of the return that measures the contribution of active management, to be the difference between the return on the portfolio and the return on the corresponding benchmark. Thus, if the Fund were to hold the benchmark, the active return would be zero.

An alternative measure of the contribution of active management is the intercept, or "alpha", in a regression of returns on the benchmark. This measure and the average active return will be equal when the beta of the portfolio relative to the benchmark is one. In the case of the Fund the betas are in most cases very close to one but we nevertheless focus on the active return rather than alpha since a deviation in beta from one is also an active decision by the manager.

We address the question of whether the fund has been successful in its active management activities. Figure S1 plots cumulated active returns on the overall Fund, which are cumulated sums of the monthly active returns from the beginning of the sample (January 1998). Units on the y-axis are in percentages. The cumulated active returns generally trend upwards strongly and start declining precipitously at the beginning of 2008. Thus, there seems to be a break around the beginning of 2008. In our analysis we investigate the sample period up to December 2007 (the "pre-2008 sample") separately from the whole sample, where the last few months include the global financial crisis. Mindful of the biases which may result from conditioning on a particular "end-point", we also conduct analysis using information up to each date in the sample only.

Figure S1: Cumulated Active Returns on the Overall Fund



The figure plots cumulated active returns, which are cumulated sums of the monthly active returns from the beginning of the sample, on the overall Fund. Units on the y-axis are cumulated monthly percentage returns.

Table S1 below reports the average value of the active return for the total Fund and for the fixed income and equity sub-portfolios. Over the whole sample the active return on the total portfolio had a mean of 0.02% per month and the p-value, which tests whether this is equal to zero, is 0.57. Thus we cannot find any statistical evidence of significant active outperformance over the sample taken as a whole. In contrast, prior to 2008 the mean active return was 0.03% per month and this is highly significant. Note that even though this is statistically significant, the absolute value of the active return added to the fund, and thus the economic contribution, is relatively small.

Table S1: Average Returns (per month) to Active Management

	Total Fund		**Fixed Income**		**Equity**	
	Full Sample	**Pre-2008**	**Full Sample**	**Pre-2008**	**Full Sample**	**Pre-2008**
Mean	0.02%	0.03%	0.00%	0.01%	0.05%	0.06%
P-value	0.56	0.01	0.98	0.45	0.08	0.02

Within the overall portfolio, the results for fixed income and equity are quite different. For fixed income, the mean active return is zero for the full sample and 0.01% per month for the period prior to 2008. Neither is significantly different from zero and so there is no evidence that active fixed income management has added value. In contrast, the corresponding results for equity are 0.05% per month for the full sample and 0.06% per month for the pre-2008 period. While both are still relatively small, the latter is highly significant and the former marginally significant. Active management of the equity portion of the Fund has, in a small way and on average, added value since the inception of the Fund.

We next address the question of how much impact active management has had on the overall risk taken by the fund. Table S2 below reports the fraction of the overall variance of fund returns that is attributable to variability in the benchmark and to active management (the active return). As in Table S1, results are given for the total Fund and for the fixed income and equity sub-portfolios and for both the whole period covered by the data (January 1998 to September 2009) and the period up to December 2007.

Table S2: Variance Attribution of Fund Returns

	Total Fund		Fixed Income		Equity	
	Full Sample	Pre-2008	Full Sample	Pre-2008	Full Sample	Pre-2008
Benchmark Return	99.1%	99.7%	97.1%	99.8%	99.7%	99.7%
Active Return	0.9%	0.3%	2.9%	0.2%	0.3%	0.3%
Total Return	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

The results are striking and among the most important in this report: for the full sample, including the crisis, the active management activities of the Fund account for less than one percent of the overall variance. Excluding the crisis, this figure falls to a miniscule 0.3%. For the fixed income sub-portfolio, where many of the most extreme events in the crisis have been located, the figure including the crisis is still only 2.9%. The implication of this result is that, compared with the choice of benchmark, active management has an almost trivially small impact on the overall risk of the Fund. Table S2 shows that, to a first approximation, the Fund is more similar to an index fund than an actively managed fund.

In fact, we find that a significant part of even the very small component of the total Fund return represented by active return is linked to a number of well-recognized systematic factors. Thus the contribution of active management to the overall return that is genuinely idiosyncratic is extremely small indeed.

Our choice of factors is affected by three considerations. First, we select factors that are recognized as having significant pricing effects by the academic literature and industry practice. Second, the set of factors must be parsimonious and third, the factors must have benchmarks that as far as possible can be traded in large size relevant to the Fund and exist since January 1998 (start of the Fund returns). The ten factors that we use are shown in Table S3 below.

These factors capture the main systematic factors that have been identified in the literature as having significant explanatory power for the variation in the returns or their mean value, or both. The factors span the main dimensions of risk in financial markets, apart from the return on the market portfolio which is captured in the benchmark.

Table S3: Factors used in Analysis of Active Returns

TERM	Difference between long- and short-maturity U.S. Treasury bond returns.
CREDITAa	Difference between Aa and Treasury bond returns.
CREDITBaa	Difference between Baa and Aa bond returns.
CREDITHY	Difference between high yield and Baa bond returns
FXCARRY	Captures the carry trade of investing in currencies with high interest rates and shorting currencies with low interest rates.
LIQUIDITY	Reflects periods of high and low liquidity.
VALGRTH	Difference in returns between “value” stocks and “growth” stocks.
SMLG	Difference in returns between small and large stocks.
MOM	Captures the momentum effect of going long U.S. stocks with past high returns and short stocks with past low returns.
VOL	Captures differences between implied and realized volatility.

Table S4 shows the partial correlation between the factors listed in Table S3 and the active return in fixed income (left-hand panel) and equity (right-hand panel). For fixed income the equity related factors of VALGRTH, SMLG, and MOM are excluded. Similarly the fixed income related factors TERM etc. are excluded for the equity sub-portfolio.

Table S4: Exposure of Active Returns to Systematic factors

	Fixed Income		Equity	
	Partial Corr	P-value	Partial Corr	P-value
TERM	-0.21	0.01		
CREDITAa	0.35	0.00		
CREDITBaa	-0.33	0.00		
CREDITHY	-0.01	0.92		
FXCARRY	-0.04	0.64		
LIQUIDITY	0.35	0.00		
VALGRTH			-0.56	0.00
SMLG			0.41	0.00
MOM			0.02	0.80
VOL	0.37	0.00	0.28	0.00

The results reveal a highly important feature of the Fund's active returns. Both the fixed income and equity active returns are strongly related to many of the factors and this implies that a significant fraction of even the very small active return component of total returns could be replicated by a small number of portfolios constructed to mimic the behavior of the factors in Table S3. In fact, approximately 70% of all active returns on the overall Fund can be explained by exposures to systematic factors over the sample. It is appropriate that the Fund has exposure to these factors: these are associated with risk premiums that the Fund, as a patient investor, can seek to harvest over time, just in the way that, with an exposure to the market portfolio, it has sought to harvest the equity risk premium.

An important question for the risk management of the Fund is whether the very low active returns in 2008 could have been anticipated. The question is not whether they could have been accurately predicted – the answer to this is almost certainly "no" – but whether the possibility of such damaging returns could have been envisaged and, further, whether the probability of this event could be calculated.

One way to answer this question is to fit a multi-factor model to returns, i.e., to estimate the sensitivity of the active returns to the factors in Table S3. What we find, perhaps surprisingly, is that using only data that would have been available at the time just prior to the period of very bad active returns in 2008, the very poor results following the collapse of Lehman could have been predicted to a significant extent *conditional* on the realizations of the factors in Table S3. This is true both using estimated factor exposures as of December 2007 for the whole fund and estimating factor exposures as data became available through 2008. There is some evidence that NBIM may have

successfully altered factor exposures during the crisis and so avoided losses, particularly in the equities portfolio, that could have been more severe had the factor exposures remained constant in 2008 and 2009 at their December 2007 levels. Our inference, however, is based on an interpretation of the returns data over that period, rather any specific knowledge that these actions took place.

The key factors responsible for the strongly negative returns in the overall Fund, and especially in the fixed income portfolio, during the financial crisis are those reflecting liquidity and volatility. Factors such as liquidity and volatility have experienced infrequent but large negative returns in the past and recognizing this, together with the extent of the Fund's exposure to these factors, would have contributed significantly to assessing the Fund's overall risk exposure. Thus, we find the behavior of the Fund's active losses in 2008 and part of 2009 were broadly in line with what should have been expected given its exposure to liquidity, volatility, and other factors. Put another way, if the factor exposures pre-2008 and during the financial crisis had been communicated and the asset owner had some knowledge of the potential drawdowns of these factors, the Fund's losses in 2008 would not have been surprising and would have been within expected loss limits.

Section II also evaluates the Fund's risk budgeting process based on the behavior of the realized tracking error. The amount of active risk since 1998 has been 0.25% per month, which is substantially below the limit of 0.43% per month in the Fund's regulations. The tracking error limit was exceeded post-2007, but this was probably due to estimating forecasted variances with only historical data observed to date. A large part of the increase in tracking error post-2007 is due to increased volatility in the factor exposure of active risk.

The Fund's risk management is based on a process by which the asset owner contracts with its fund manager, NBIM, on a diversified portfolio representing an agreed benchmark, and then the fund manager is given boundaries with respect to the annual deviation of its return from the return of this portfolio, i.e. tracking error. The risk boundaries themselves are not targets. Thus, there is no expected amount of deviation from the benchmark and the risk-budgeting process. This in fact, implicitly rather than explicitly, defines a system of risk limits. This ambiguity, in our view, has two drawbacks. First it may lead to excessive risk aversion on the part of NBIM: without a target, it is difficult to take the manager to task for under-investing in active opportunities, but not the contrary. Second, it is not robust to changes in market conditions; periods of high volatility may cause the manager to trade when market risk increases. We recommend a tracking error target or bands rather than a strict risk limit. We also further

recommend that the tracking error concept only be used for market portfolios comprising liquid, tradable securities. We explore this further in Section III.

In evaluating the externally managed funds, we quantitatively investigate only the active management of the external managers and qualitatively comment on the selection of the external mandates by NBIM. We know little about the way in which the mandate benchmarks are chosen and so, while it is possible that some of these choices also reflect active decisions by NBIM, we are unable say whether these decisions add value. Our analysis of externally managed funds therefore focuses on their active returns, i.e., the return relative to the mandate benchmark. Overall external equity management has enjoyed a modest level of success but the active returns of external fixed income funds have been very poor. Although the fees paid to external funds are low, the large exposures of active external returns to systematic factors suggest that active external management has not reflected a large component of unique management ability.

In summary, active management has not detracted from the value of the Fund and accounts for only a very small amount of its overall risk exposure. Prior to 2008, active management contributed a small but highly significant portion to overall Fund returns. However, much of the behavior of the Fund's small active return can be explained in terms of systematic factors. Our recommendation is that these exposures are, in general, appropriate but that they should be brought into the benchmark and that the Fund's average exposure to these factors should be a "top-down" decision rather than emerging as a byproduct of "bottom-up" active management. These issues are discussed in Section III.

Summary: Section III

Section III assesses whether and to what extent the Fund has comparative advantages that justify investing significant resources in active management. This includes an evaluation of the opportunities, challenges and specific capabilities with respect to currently managed and new asset classes. In addition to this assessment, it proposes strategies that use the Fund's comparative advantages, and characterizes these strategies in terms of distributional characteristics, verification horizons, management feasibility compensation structures, and modes of evaluation.

We recognize several distinctive features of the Fund that contribute to its ability to fulfill its mission. First, the transparency and long-term mandate gives the Fund an advantage in ensuring responsible management and provides the on-going basis for

evaluating strategic decisions and measuring performance. Second, the large size of the Fund is an advantage that provides scale economies, the ability to influence corporate management, and potential access to the world's best investment managers and staff. Third, the Fund's long-term horizon and lack of immediate cash liabilities gives it a comparative advantage in accessing investment opportunities with long verification horizons. Fourth, NBIM as a Norwegian-based asset management entity organized within the central bank (Norges Bank) with a strong public service ethos represents a comparative advantage in that it helps mitigate the problem of agency which is the central challenge of delegated investment management. Finally, the use of segregated accounts for external managers allows NBIM detailed information for continuously monitoring its financial position and assessing its risk exposures.

We also note that each of these distinctive features has potential drawbacks. Public transparency and accountability necessitates active communication to align expectations about the Fund policy and explain performance results. This may lead to less risk-taking by management. The large scale of the Fund has drawbacks as well. Size makes it difficult to beat the market since every major decision is likely to have an adverse price impact. The need to diversify globally leads the fund into markets with higher transaction costs. The large scale necessitates the use of intermediaries and service firms, which represent further cost.

This assessment of the comparative advantages of the Fund provides the basis for a review of the current management approach and an assessment of potential strategies that utilize these advantages.

Our analysis suggests that NBIM takes a fundamental, bottom-up approach to active management focused on security selection. We are not strong advocates of this approach as a way to add value above and beyond factor-based benchmarks, as the empirical evidence (reviewed in Section I) in support of the contribution to returns via superior security selection is limited. However, assuming the costs of this style of active management are low, such activities do not detract from returns even though they may make a small overall contribution to Fund performance. Fundamental analysis may be beneficial for some other aims that the fund wishes to pursue such as socially responsible investing.

With respect to describing and evaluating strategies that utilize the Fund's comparative advantages, we suggest that the Fund focus more attention on factor risk. We recommend a more top-down, intentional approach to strategic and dynamic factor exposures. This approach relies on expected equilibrium compensation for taking non-diversifiable risk. It is similar to collecting an insurance premium for bearing risk

associated with those factors. The foundation of our proposal is the identification and construction of multiple factors that bear risk premiums and the transfer of these factors into the fund benchmark, rather than treating them as the byproduct of other active management strategies.

This approach is designed to utilize the Fund's comparative advantages. First, it relies upon the Fund's long-term outlook and lack of current liabilities. Factor premiums, like the equity risk premium of stocks, may exhibit consistent profitability only over long horizons, however they do not rely on identifying mispriced securities. The factor approach also utilizes the advantage of the governance structure of the Fund, in that there exists a transparent mechanism for setting the exposure to risk premiums – like the equity premium – and a common public understanding of the relationship between a policy decision and a market realization. It is natural and appropriate that the owners decide on the systematic exposure to well-established factors. This exposure needs to be communicated and included in the benchmark to the greatest extent possible. Finally, the factor approach utilizes the excellent capabilities of NBIM as developers and monitors of investable benchmarks. In practice, the top-down factor-based approach could be viewed as an extension of the Fund's asset allocation policy and NBIM's current operations. The Fund's asset allocation is determined by assessing the risk and return of investing in debt and equity. Consideration of additional factors is a natural extension. NBIM has clear competence in building indices and tracking them. Thus, NBIM is well suited to undertaking this task.

The need to implement the factor benchmarks in-house is driven by the lack of consistent, investable benchmarks for equity and fixed income factors around the world. NBIM should have major advantages in keeping implementation costs low. We argue that the factor benchmarks will provide clarity with respect to the source of return and the exposure to risk. In Section III we show how the risk-return profile of the factor-based allocation strategy can be described in terms of a probability distribution. Similar tools can be used to gauge the public's risk aversion with respect to different factors and after identifying the amount of desired exposure to each factor, the factor exposure is brought explicitly into the fund benchmark.

There are several major technical consequences of moving factor exposures into the benchmark of the Fund. First, the optimal benchmark is time varying and has non-market weights. These tilts away from market weights represent the loadings on factor risk premiums. Second, the benchmark is "passive" in the sense that it is constructed according to a set of rules and its composition is pre-determined through the definitions of the factors. But, it is "dynamic" in the sense that their composition is changing more

often than a standard S&P500 or FTSE portfolio. Third, and most importantly, the desired amount of factor exposure is set by the asset owner.

The top-down factor approach does not eliminate the potential use of active management. Rather, we define active management in terms of two decisions. First, the decision to deviate from long-term strategic loadings on factors (e.g. a temporary shift from the target allocation to equities) and second, the decision to hold securities in weights that differ from factor benchmark weights. These roughly correspond to timing and selection, where the default, or the baseline case, is determined by factor portfolios that comprise the Fund benchmark. Components of active management will be defined and measured in these two dimensions.

With respect to the current risk monitoring processes, we recommend setting a target for tracking error rather than a boundary, and we further recommend target flexibility that allows for changing market conditions. Tracking error as a risk budgeting and risk management tool should only be applied to a universe of liquid securities. For purposes of internal evaluation of the costs and benefits of active management, and capabilities with respect to timing and selection we recommend maintaining an internal record of decisions about both dimensions of active management. We also recommend that the factor benchmarks themselves be measured against a customized, market-weighted portfolio.

We recommend that different strategies be classified according to the horizon appropriate to the length of time it typically takes for profits from the strategies to be realized. For any strategy, theory may suggest long-run profitability while in the short run there may be extended periods of underperformance. We propose three "horizon buckets" that reflect relative liquidity and the different verification periods of various strategies. Reporting and performance for each class of investment should be segregated, although the management of the portfolio would use the factor model as a framework for integrating all three. Reporting of strategies and assets held in the long-horizon category should also emphasize metrics other than just returns, especially various cashflow-generating metrics. We point out that the lack of such a differentiation across horizons is a disincentive for NBIM to take active risks that generate positive long-horizon rewards.

To implement the process we recommend a series of steps to be taken via a transition process. This involves setting up a new division within NBIM responsible for creating, communicating, maintaining, and implementing factor portfolios and bringing factors into the benchmark for active management.

Summary: Conclusion

The brief for this report focuses on the question of whether the Fund should be actively managed and so the key finding of our report, one that drives the majority of our recommendations, may be a surprise. It is that, to a first approximation, the Fund is actually not an actively managed portfolio.

Active returns, the difference between the fund's actual return and its benchmark, constitute only a small fraction of both the mean and the volatility of the total return. Moreover, even this small component is itself substantially explained by the Fund's exposure to a number of systematic factors such as liquidity and volatility. Overall, therefore, the returns on the Fund are similar to those that could have been earned on a fund holding the benchmarks with additional, essentially passive exposure to these factors.

An important question raised by this finding is whether this pattern of risk exposures, dominated by the benchmarks and other systematic factors, is appropriate for the Fund. Our view, supported by our review of the academic literature and our assessment of the differentiating characteristics of the Fund, is that it is. The Fund may seek to continue or even enhance its exposure to purely idiosyncratic and information-based investment strategies. There is no evidence that this activity has hurt performance and may enhance some other aims that the Fund wishes to pursue such as socially responsible investing. However, if the Fund remains well diversified it is highly likely that a significant fraction of its active return will always be driven by one or more systematic factors.

A modern interpretation of the EMH does not deny the possibility that managers might possess skill that allows them to earn excess risk-adjusted returns. Indeed, even modest levels of skill should lead to at least some part of the portfolio being actively managed. Similarly, and more importantly in this context, we should not interpret passive management, the absence of active management, as meaning that the Fund should simply hold a static portfolio that replicates a combination of major stock and bond indices. In such an asset allocation approach the investor chooses the amount of stock versus bond exposure based on the investor's risk aversion. We advocate enhancing the foundation benchmark by moving to in-house, custom benchmarks that are market weighted.

Depending on its preferences and other characteristics, the Fund may wish to acquire fixed, and possibly dynamic, exposure to other factors. The chosen level of exposure is dependent on how the Fund's tolerance for a certain factor is different from that of

other market participants. For example, if it is felt that the Fund has greater tolerance to bear volatility risk – the risk that in volatile periods certain assets with high volatility risk perform poorly – then the Fund should have a positive exposure to a volatility factor. In holding such assets, the Fund is providing insurance to other investors: by holding more than the market fraction of assets with high volatility risk, the Fund allows other investors to hold a larger fraction of assets with less volatility risk and so the risk premium can be interpreted as an insurance premium that the Fund receives from other investors for reducing their exposure to assets which pay poorly when market volatility is high. Over the long run the Fund will be expected to earn a premium for supplying volatility insurance to other market participants, but during periods of high volatility the Fund will suffer relatively lower returns.

While the principle underlying this approach is the same as the principle motivating the 60%-40% equity-bond allocation, the difference is that there are other factors that are not strongly correlated to the equity premium that offer a reward for risk taking. These represent additional opportunities for return and should dominate the alternative of simply increasing equity exposure. That is, equity risk is only one dimension where the Fund earns a risk premium; the Fund should take advantage of other risk factors as well.

The key characteristics of the Fund that should influence the deviation of its investment policy from a market weight position are its absence of need for liquidity, its very long-term horizon and, at least while cash inflows continue from oil revenue, no explicit liabilities of the fund. These characteristics are important because it is along these dimensions that the Fund is most clearly differentiated from the average investor. The consequence is that the Fund should be able to earn risk premiums from taking exposure to assets that are, for example, somewhat less liquid. The long horizon also enables the Fund to ride out periods of short-term negative returns to factors that in the long run will have positive risk premiums, as predicted by theory and confirmed in data. These factors may include volatility risk, value-growth risk, credit risk, and liquidity risk.

Our main recommendation is that both the metrics used to monitor NBIM's performance and manner in which the Fund is managed become more closely aligned with the actuality of its risk characteristics. In reality the Fund's performance is largely determined by its exposure to a number of factors, especially volatility and liquidity risk, as described in Section II. We propose that the benchmarks be framed in terms of these factors – "factor benchmarks" – rather than portfolios defined in terms of geography and asset class as at present. In this framework the Fund's exposure to a particular factor is chosen by the investor, just as the current 60%-40% equity-bond asset

allocation is determined now. A factor benchmark approach would also, in principle, make it easier to include, evaluate, and monitor new asset classes, such as real estate.

One important lesson from the financial crisis is that some factors, e.g., liquidity and volatility, have distributions that are highly skewed with long left tails. In other words, returns on these factors may be relatively stable for much of the time but periodically experience very large negative returns. Our analysis shows that the Fund does have exposure to factors with this payoff pattern. Because of the characteristics of the Fund, these exposures are appropriate because the fund can expect to earn risk premiums in the long run. However, to understand both the risk premiums the Fund is likely to earn from these exposures and the risks it is running in doing so, it is necessary to take a very long term view. This is true for all estimates of risk premiums: even with 100 years or data the standard error on the equity risk premium is about 2%. In fact, over the sample period considered in this study, the realized equity risk premium has been close to zero.

Moving the management of the Fund to the factor benchmark framework will be challenging, but we believe well within NBIM's capabilities. In its management of the fund to date, NBIM appears to be a highly disciplined manager in that it is able to very precisely track a benchmark. Implementing factor benchmarks involves constructing the factors, selecting the factors to be included in the overall benchmark, and determining how much factor exposure should be taken. Given NBIM's world-class capabilities with respect to constructing and maintaining investable benchmarks, these challenges can be met. The low tracking error around the Fund's benchmark testifies to the operational efficiency of NBIM's management with respect to mimicking tradable factors.

In summary, our report finds that the risks that the Fund has been running, both before and during the crisis, are broadly appropriate given its characteristics and objectives. Our proposals do not therefore suggest a major change in these risk exposures. Active management has not detracted from Fund returns, and in fact has been a small but positive addition to performance. In light of the relative importance that factor exposures already play in the Fund's returns, we suggest that the Fund consider a framework that more explicitly recognizes the structure of its return generating process via investment in factor benchmark portfolios – and that both the way the Fund is monitored and the way it is organized on a day-to-day basis be adapted to this new framework.

Section I: Academic Evidence on Active Management

All modern investors are faced with the fundamental decision to use a passive management strategy, an active management strategy, or a combination of the two approaches. A passive management strategy is also known as indexing. Indexed assets are invested according to a pre-determined set of rules that seek to replicate the performance of an index of pooled securities whose positive historical performance and risk characteristics have been studied, and are known to match the goals of the investor. Passive indexation started in the late 1970's and grew very popular in the 1980's because of a theory prevalent in financial economics through a large part of the 20th century – the Efficient Market Hypothesis [EMH].

In simple terms, the Efficient Market Theory asserts that, at all times, the price of a security reflects all available information about its fundamental value. A consequence of the theory is that, if true, it is impossible for an investment manager – and hence the clients of the manager – to consistently beat the market. The underlying principle driving the EMH is elegant and intuitive. In a large, active marketplace for publicly traded securities, vigorous competition among scores of investors will drive speculative profits to zero. The implication of the EMH for investors is that, to the extent that speculative trading is costly, speculation must be a loser's game. Hence, an indexing strategy is bound to eventually beat a strategy that uses active management; where active management is characterized as trading that seeks to exploit mispriced assets. In the world of the EMH, there are no mispriced assets because the invisible hand of the marketplace moves faster than any single agent.

The driving question in financial economics since the inception of the EMH is whether it is correct. This question has led to some useful theoretical improvements on the Hypothesis that takes into account information theory and the frictional costs of intervention by arbitrageurs. It has also led to a wide range of empirical tests seeking to reject the Hypothesis. In this section we will examine the evidence provided by some of these tests in light of a key question: Is delegated active management a useful strategy for a large, diversified institutional investor such as the Fund? If so, is it equally relevant across markets and asset classes, or is there empirical evidence about the relative costs and benefits for its selective use?

The academic literature on the EMH is vast. A complete history of its theoretical development is intellectually interesting but not relevant for this analysis. On the other hand, there is some important and relevant recent scholarship on market efficiency in general and active investment management in particular. In this section we will begin with a brief discussion of current efficient market theory. Following this theoretical foundation we discuss the recent empirical evidence on efficiency as it pertains to a range of different markets – not simply the

large, liquid public securities markets but also the private capital markets. We take care to emphasize the relevant parts of the literature especially relevant to the Fund.

In Section IA we briefly outline the theory of the EMH. Section IB covers various aspects of empirical work testing efficiency at the individual security level. We address the question of performance in managed portfolios beginning with mutual funds in Section IC and other delegated vehicles in Section ID. Section IE covers fixed income and other asset classes. Section IF concludes.

IA: Theory

The intuition of the early Efficient Market Theory remains robust, however theory has developed to account for information costs, agency conflicts, financing constraints, and other real-world frictions.

IA 1. Early Theoretical Foundations

The early theoretical articulations of the EMH focused on arguments that future changes in security prices should be unpredictable. Cowles (1933) was the seminal paper that tested whether professional market forecasters could beat random stock selection. His follow-up paper, Cowles and Jones (1937) developed a theory of the random walk of stock prices. Among the first to develop the random walk theory rigorously was the iconoclastic mathematician and father of fractal geometry, Benoit Mandelbrot (1963) who showed that, even in a very general framework allowing for discontinuities and extreme events, changes in security prices should be unpredictable. Two years later, his protégé at the University of Chicago, Eugene Fama (1965) formalized and extended the argument using the law of iterated expectations, arguing that security prices should follow a random walk. The same year, Nobel laureate Paul Samuelson published a famous paper, "A Proof that Properly Anticipated Prices Fluctuate Randomly." In it, he refined the random walk model using the framework of futures prices, showing that spot market prices need not wander randomly, nor should the sequence of daily changes in prices even be uncorrelated with each other. Rather, the EMH implies only that "The market quotation …already contains in itself all that can be known about the future and in that sense has discounted future contingencies as much as humanly possible…" In short, futures prices should be unbiased, and that speculation should be a "fair game" with an expected reward of zero or, more generally, an amount that reflects a normal risk premium.

These early theories about market efficiency motivated a number of empirical studies of prices in various asset markets chiefly focused on whether security returns were indeed serially

uncorrelated – i.e. whether past price changes could predict future price changes. Although we will not go into these in any detail, the evidence resulting from these "random walk" tests was mixed. Empirical evidence of predictability frequently cropped up in market data, but it was generally dismissed as weak or unexploitable by a speculator due to transactions costs. To some extent, the theoretical logic of the EMH articulated by Cowles, Fama, Samuelson and Mandelbrot was so compelling – and ultimately so useful as a tool for the development of asset pricing models that it became the dominant intellectual paradigm for a generation of scholars.

Fama (1970) reviewed the evidence on the Efficient Market Theory using a taxonomy for levels of efficiency proposed by Roberts (1967). Weak form efficiency implies that past returns cannot predict future excess returns. Semi-strong form implies that public information cannot be used to predict future excess returns. Strong form implies that no information (even direct personal knowledge of a merger, for example) can be used to predict future excess returns. Fama concluded that the empirical evidence up to 1970 supported weak-form and semi-strong form market efficiency.

IA 2. Market Realism

More recent theory about the EMH has focused on making the theory more realistic. As the above quote from Samuelson points out, information is an essential feature of the theory. In effect the market price "impounds" all available value-relevant information about the future. This feature is common to all of the early theories. However, none of them explore either how the information is generated or the mechanism that causes the information to be reflected in prices. Nor do they provide a motive for information to be generated by the market. Why should a speculator do any research to evaluate the prospects for a company if trading on information is unprofitable? And, if no speculators actually collect information how can it be that prices nonetheless reflect all available information? Wouldn't this lead to a complete market failure and disequilibrium?

In the real economy, research is costly but potentially valuable if a speculator knows something no-one else knows. Indeed, empirical evidence on the gains to insider trading make it clear that illegally obtained private information can generate excess profits – which Fama (1970) would have classified as a strong-form efficiency violation. This has led to regulations preventing such activity in most markets. However, this argument extends to publicly available information since if publicly information is already impounded into prices, who would spend time and effort to collect and process this information allowing prices to be efficient? Grossman and Stiglitz (1976) address this paradox through a model of a market with costly information acquisition. In their model traders who invest in research are rewarded through speculative profits so that they at least recoup the cost of their investment. Their trading activity, in turn, pushes prices towards fair economic value. In effect, they become the first mover of the "invisible hand."

The Grossman-Stiglitz model portrays a "near efficient" economy in a constant state of controlled disequilibrium, but always moving towards equilibrium, driven by informed, active research and speculation. In the Grossman-Stigliz world, markets are by-and-large efficient but there are small pockets of inefficiency which are exploited by active managers with superior skills and resources.

This realistic picture of the investment market was mirrored by the contemporaneous development of the Arbitrage Pricing Theory [APT] by Ross (1976), who argued that the activity of arbitrageurs would naturally drive the expected return of assets towards a value consistent with an equilibrium trade-off between risk and return. The EMH was preserved, but it no longer narrowly hinged upon costless provision of information by the market, and it no longer ignored the role of arbitrageurs or speculators. Although the Grossman, Stiglitz, and Ross theories about asset prices portrayed a more realistic view of the asset markets, allowing for deviations from equilibrium prices and active arbitrage to correct these deviations, they all relied upon some basic assumptions about the arbitrageurs. In particular, the arbitrageurs in Ross's APT need to finance their purchases of undervalued stocks by borrowing cash. In order to exploit over-priced stocks, they need to borrow shares they do not have. What if these operations became difficult?

In 1997, Shleifer and Vishny explored the implications of these assumptions in a paper entitled "Limits of Arbitrage." Their paper was based on the old adage that the market can stay irrational longer than you can stay solvent. They constructed a model in which financing risk forced arbitrageurs to be cautious about exploiting mispricing. The implication of their model is that security prices might diverge from economic value for a long time if financing risk is high. The paper was particularly prescient: Long-Term Capital Management [LTCM], a very large, highly levered hedge fund collapsed in 1998. Among their major speculative positions was a bet on the convergence of U.S. vs. European and Japanese bond yields following the Asian currency crisis. This convergence eventually occurred, but in the short run the divergence between the bond yields increased and LTCM was forced to liquidate. The key implication of the Shleifer-Vishny paper for the EMH is that certain agents do not value assets according to rational asset pricing models and are instead driven by sentiment. This sentiment can significantly slow the diffusion of value-relevant information into security prices, and that both the capital structure and institutional framework for arbitrage matter. Such constraints do not need to arise from behavioral sources; financing constraints or leverage constraints in economies with rational agents can give rise to the same effects as shown by later researchers.

IA 3. Theory of Active Delegated Management

Thus far, the discussion of the theory about the EMH has focused on the potential for security prices to deviate from fundamental economic value, and the potential of an active manager to profit from this deviation. An equally important question from the perspective of an investor is whether a profitable delegated investment structure is possible. In other words, even if markets were not perfectly efficient, can an institutional investor take advantage of the inefficiency? This theoretical question is often referred to as the fundamental question of agency introduced by Ross (1973): a principal (the investor) retains an agent (the manager) and compensates them for generating a profit. Is there some combination of auditing and incentives that will result in the principal sharing significantly in the agent's gains, or will the price the agent charges for his/her service exactly equal the benefits generated? Put simply, suppose you hire a manager with a track record of generating positive risk-adjusted returns. Can you expect to beat the market after fees?

The most influential recent theory about this problem is Berk and Green's (2004) model of delegation. In their model, investors fail to earn positive risk-adjusted returns, even though they rationally invest with past successful managers. Their model allows some managers to be better than others and have talent on average, it rewards managers for information production, managers earn their fees, but the investment technology has diminishing returns to scale: fund flows push successful managers beyond optimal scale. Hence, in the Berk and Green model, prices may not be efficient, but the market for management services is. While there are gains for active management, these gains do not flow to principals (investors), they are captured entirely by agents (fund managers).

Another important recent theory about delegated investment management does not directly address the issue of price efficiency, but instead explains delegation as a response to changing market conditions. Mamaysky and Spiegel (2001) argue that the benefit of delegated management rests on the degree to which it is dynamic. Indexing provides only a very limited set of potential payoffs to investors. This range is grossly inadequate for most investor needs, which can only be met by dynamic adjustment of portfolio weights, and monitoring of the macro-economy. Mamaysky and Spiegel argue that managers are compensated for this active process. By the same token, investors who use only passive indexes give up the possibility of optimizing their investments with respect to their possibly complex goals.

These two recent theories are of course not mutually exclusive. It is useful to think of them as describing two different management capabilities: security selection and dynamic portfolio management. Since both are defined relative to a benchmark, this points to an important issue inherent in how active versus passive management is defined: There may exist skill in capturing returns beyond market-weighted passive indices. Moreover, the market-weighted benchmarks

themselves may poorly capture the desired risk-return trade-offs of investors. We address the question of appropriate benchmarks beyond standard passive market-weighted portfolios in Section III.

IA 4. The Swensen Approach

One additional conceptual framework for delegated investment management is worth including in this survey, despite it being a non-academic theory. David Swensen, the Chief Investment Officer for the Yale University Endowment published a highly influential book on institutional investing entitled "Pioneering Portfolio Management." This book has since become the bible for many U.S.-based endowment funds and has been credited with the broad-based trend towards alternative investing. Swensen posits major differences in efficiency across various asset classes. In highly liquid markets such as fixed income, he argues that the potential for making positive excess returns is limited due to competition and consequently in those markets there is little scope for fundamental research. By contrast, other markets such as venture capital and private equity have large potential payoffs to superior research and management skill. The gains in such markets are not competed away because of the Shleifer-Vishny problem – most managers have limited investment horizons. Swensen argues that perpetually-lived institutions such as college endowments can afford to play in these markets because their horizons are longer than those of their "competitors" for investment management services.

As empirical support for this theory, Swensen notes that the cross-sectional dispersion in manager performance for some markets is much higher than for others. Few fixed income managers differ from benchmarks by more than a few basis points, while hedge fund managers' track records vary widely. He thus counsels institutional investors with long horizons and sufficient resources to seek superior performance by careful selection of managers in the alternative space, and, if necessary for diversification, use indices for highly liquid asset classes. With the exception of 2008, the excellent track record of the Yale and other large University endowments over the past 15 years has provided some empirical support for his theory. Although the Swensen approach incorporates many of the subtleties of recent academic research, it leaves open a few questions. Among these are whether agency problems can be addressed through contracting and also what the role of dynamic asset management and allocation should be. Another issue is the limited tenure of endowment monitors. The horizon of the institution might be infinite but the horizon of its caretakers might not. Shorter term goals of university managers might induce risk aversion against short-term loss. Despite these caveats the Swensen perspective is a very useful foundation for considering the benefits of active management for the institutional investor.

Although our review of the theoretical development of the EMH is necessarily brief, the high points manifest an evolution from a relatively abstract model of rational expectations to a

framework incorporating financing, information, agency, and active management as crucial factors. While the original intuition of the EMH remains robust, i.e. that it is extremely difficult to earn excess returns in a competitive market, current academic theories no longer deny the existence of mispricing. They elaborate instead on the institutional framework for exploiting such mispricings, and conjecture a wider role for active management beyond beating the market.

In the next section we review the key empirical studies that test various implications of the efficient market hypothesis, concentrating on the results relevant to delegated investment management.

IB: Empirical Evidence Using Asset Prices

Testing the EMH requires an asset pricing model to represent fair economic value. Many apparent violations of the theory have been found using historical data, however these do not demonstrate that any economic agents profited as a result.

It is important to point out at the outset that much of the academic research on the EMH and active investment management has focused on a relatively narrow spectrum of assets: the U.S. equity market and equity mutual funds. This limitation is due almost entirely to data availability. Good price data for the U.S. stock market became available in the 1970's due to the creation, at the University of Chicago, of the CRSP database of equities.[1] Leading academic research institutions around the world now subscribe to the CRSP data and their scholars have extensively "mined" the data in search of violations of the EMH. Similar databases for mutual funds, hedge funds, private equity, venture capital, derivatives, and even conventional corporate debt are much less accessible and generally much more recent.

As a consequence of this focus on a few major asset markets, the question of the relevance of the myriad of empirical tests of the EMH over the past several decades is a natural one. Institutional investors typically diversify across many different asset classes and national markets, and rarely use retail equity mutual funds as managers. Thus, there is a mismatch between our desire for reliable empirical evidence about efficiency in many different markets

[1] Tests of the EMH in fixed income markets have taken a distinctly different form by testing deviations from the Expectations Hypothesis – that if markets are efficient, then, after adjustment for any risk premium, fluctuations in the forward rate for a given future period in calendar time should be unpredictable. A very large literature beginning from Fama and Bliss (1987) finds that forward rates are biased predictors of future sport rates and explains these deviations by the existence of risk premiums.

and the academic evidence to date. To address this uneven research landscape, following an overview, we divide this section by asset categories relevant to the Fund. Given the implicit contrast between indexation vs. active management, we also include a discussion of the empirical evidence about returns to investing in index funds.

IB 1. Price Studies vs. Manager Studies

Tests of the EMH can be divided into studies of prices and studies of investment managers. Studies of prices have generally focused on a search for trading rules that generate positive risk-adjusted investment returns when back-tested on historical data. The existence of such a rule would violate the basic tenet of the EMH, namely that current prices reflect all value-relevant information and are "fair" in the sense of an unbiased game. These studies are simulations of investment returns rather than actual performance – they show paper profits rather than actual profits.

Because of the widespread acceptance of the EMH in the latter part of the last century, any such rule reported in the academic literature was labeled an "anomaly" – a puzzle that challenged theory and could not be explained by economic theory.

IB 2. Methodological Issues

It is important to keep firmly in mind the statistical difficulty in distinguishing abnormal performance that is genuine from that which is spurious. With a long history of securities prices it is almost inevitable that some trading rules will appear to be profitable. If the researcher is given carte blanche to search over a wide range of trading rules then the statistical significance that can be attributed to even the best performing rules may be quite small.[2] Data mining issues in finance have been analyzed by Kosowski et al. (2006), among others. Given the numbers of analysts and professors actively studying asset prices, the data mining problem looms large.

Next to data mining the most prevalent methodological issue in studies of market efficiency and manager skill is the fact that all tests are in fact joint tests of a model of expected returns and of efficiency. In order to test for deviations from economic value, the researcher must specify economic value. For tests on stock returns, for example, this usually takes the form of an asset pricing model that relates expected returns to risk factors.

A subtle methodological issue is that most tests of efficiency are based on the Law of One Price, which posits that, in an efficient market, two economically equivalent assets will have the same

[2] These problems of data mining trace back to the "file drawer problem" of Rosenthal (1979) that the only articles journals publish are those that are "significant" while those that are "insignificant" are relegated to researchers' file drawers (or now hard drives).

value. This notion of economic equivalency in empirical studies ranges from the precise to the unmeasurable. For example, notions in the literature of the Law of One Price range from the two assets having precisely the same cash flows, to assets having only the same systematic risk exposure, to a more general notion of a present value relation that relies on a discounted stream of expected future cash flows being economically equivalent to a traded asset. While precise discrepancies in assets with cash flow equivalencies may be easily measured, deviations from the present value relation may not be, due to the impossibility of determining expectations at a moment in time.

IB 3. Early Price Evidence Against Efficiency: Anomalies

The earliest anomalies, documented in the 1970's and 1980's, showed that fundamental accounting and payout ratios as well as company characteristics could predict future stock returns. These effects were labeled anomalies because their patterns of returns were anomalous relative to the Capital Asset Pricing Model [CAPM], which predicted that only a stock's beta determined that stock's expected return. These studies measured stock returns relative to the market return as well as relative to other stocks and thus constituted evidence in the cross section of stock returns against the one-factor CAPM. Later studies uncovered stock return patterns that were anomalous relative to multi-factor models including size, value and momentum factors.

The earliest anomalies included the small firm effect (Banz, 1981) whereby small-cap companies performed better than large companies, the January Effect (Keim, 1983; Reinganum, 1983) which noted that the first few days in January provided very high returns, the earnings ratio effect (Basu, 1977) which showed that a firm's earnings yield predicted future excess stock return, the book to market effect (Stattman, 1980) where the ratio of a stock's book value to its market value predicted relative performance, and the short-term reversal effect (Rosenberg, Reid and Lanstein, 1985). Later research uncovered longer-term momentum and reversal effects: Jegadeesh and Titman (1993) showed that investing in the winning stocks over the past year and shorting the losing stocks (known as momentum investing) generated positive excess returns on the order of 8% per year. Debondt and Thaler (1985) documented an even longer-term reversal pattern at the three- to five-year horizon.

In the accounting literature, Bernard and Thomas (1989) discovered that investors tended to under-react to quarterly earnings announcements. A strategy of buying on positive earnings surprises and selling on negative earnings surprises yielded positive excess returns and constituted an important contradiction to the EMH. Hand (1990) found that investors reacted to information in annual reports that had previously been disclosed in quarterly reports, consistent with prices being set by unsophisticated agents. Another famous accounting

anomaly is the negative relation between accounting accruals (non-cash components of earnings) and future stock returns first documented by Sloan (1996).

IB 4. Multiple Factors

Because all efficiency tests are joint tests of a model – i.e. a benchmark, the specification of the benchmark is a vital step in the analysis. Virtually all of the early anomaly studies were based on observing return deviations from a beta-adjusted market model. Because most of these studies were on stocks, and the CAPM argued that the world market portfolio was the only relevant source of risk in the economy, the standard benchmark in early price studies became the U.S. market-weighted equity portfolio. Financial economists like Roll (1977) recognized early on that this was a poor approximation to the world wealth portfolio justified by theory, however the inability to measure the world wealth portfolio presented a practical barrier to implementing theoretically justified benchmarks.

A solution to this problem was provided by the recognition that the world wealth portfolio contained within it a relatively parsimonious set of risks, or factors. The APT developed by Ross generalized the single factor model to multiple systematic risk factors. His theory required that these factors have certain important characteristics. For example, they had to be widely regarded as systematic risks in the economy that, on average, people would wish to insure against. Many extensions of this framework were subsequently developed and some of the predictions of the model were foreshadowed in the intertemporal setting of Merton (1973).

A strong implication of the existence of multiple factors rather than just the single factor of the CAPM is that the market-weighted portfolio of traded assets is inefficient. That is, holding static or time-varying combinations of assets in non-market weights result in higher reward-to-risk ratios than the market portfolio. This was demonstrated by many studies like Kandel and Stambaugh (1987) who found that the market did not provide the highest possible mean return for its level of variance risk.[3] More recent studies like Hansen and Jagannathan (1997), which build on these so-called mean-variance spanning tests, also overwhelmingly reject the single market factor in favor of models with multiple factors. This evidence leaves clear scope for active management to identify benchmark factor portfolios that can price the range of security outcomes and describe the range of efficient portfolios.

In Chen, Roll and Ross (1986), Ross and his colleagues conjectured that these risks would include inflation shocks, unanticipated shifts in the yield curve, business risk represented by GDP shocks, and changes in default premiums (this last one was clearly an important factor in the recent crisis) in addition to market risk. Each of these risk factors, which appear in all

[3] Cf. Huberman and Kandel (1987) for the basic result and Kan and Zhou (2008) for recent review and econometrics.

investments, should thus command a premium from a risk-averse investor. Only a security with a completely riskless real return would require no premium. Securities with higher exposure to these factors would, on the other hand, be expected by investors to have higher returns going forward, although these returns would certainly vary because of exposure to economic shocks.

The APT did not specify what the factors were, however. Researchers have been trying to develop a parsimonious set of them for many years and the question is far from settled in the literature. Recent creative work has led to factor mimicking portfolios and benchmarks that include risk from higher moments and non-linearities. The key insight of the APT was that there were potentially many non-diversifiable sources of risk and that bearing these risks would give rise to earning risk premiums. Many of these risk premiums, in addition to market risk, are appropriate for a long-term investor such as the Fund to harvest. We discuss several of these factors in evaluating the active management of the fund in Section II and which factor risk premiums the Fund might collect in the future in Section III.

One interesting explanation for the excess returns generated by pervasive anomalies is that they actually capture economic risk factors. As such, their higher returns are compensation for higher non-diversifiable risk. Indeed it is now common to use portfolios derived from the size, value and momentum anomalies as systematic control factors in studies of equity asset returns, despite the fact that there is currently no broadly consistent economic theory linking these characteristics to risk. These have come to be called the "Fama-French" factors, after the researchers who, building on earlier factor models, popularized their use in a landmark paper in 1993. The logic for employing Fama-French factors as a benchmark is that these factors embody systematic risks, which are captured through size, value, momentum, and other factor returns, and a successful active manager must generate a return in excess of the exposures to these factors.

In using the Fama-French factors, academics make two important implicit assumptions. First, it is assumed that these factors can be replicated by a passive investment strategy. Although these factors are constructed from the prices of traded U.S. securities, they do not represent investable returns to a very large, global investor, a point emphasized by Cremers, Petajisto and Zitezewitz (2008). This is because the Fama-French factors include many small or micro-cap stocks and require high turnover.[4] The second implicit assumption is that the investor has an optimal pre-determined amount of exposure to each factor. In OLS regressions used to estimate Fama-French factor loadings, this is reflected in the assumption that the factor loadings do not vary over time, at least for the sample used to estimate the factor loadings. The investor is not paying the manager to access these factor risk premiums – the optimal factor loadings can already be implemented by the investor through passive portfolios – and

[4] We discuss this issue in Section II.

thus active managers should be rewarded only for adding average returns beyond the factor exposure of their strategies.

IB 5. Multiple Factors and the Scope for Management

The scope for active management can be expressed in terms of the factor models described above. In a CAPM framework, the benchmark portfolio is comprised of the capital-weighted market portfolio. That is, the security weights for each asset are equal to their proportionate percentage of all of the world's risky assets. The active manager has two ways to deviate from this benchmark. The first is to change the proportion of the benchmark itself – to shift some investment from the benchmark into a riskless asset, or alternately, to borrow and buy more of the benchmark. The second is to hold securities in different proportion than the capital weights. Both of these activities (sometimes called timing and selection) have the theoretical potential to add excess returns.

The same scope for active management applies to a multifactor portfolio. The benchmark portfolio in this case is a set of pre-specified portfolios of assets i.e. factor portfolios, each of which provides a premium to the investor. The manager's scope is now defined in terms of the variation in the relative proportions of factor proportions to hold and deviations from the individual security weights defined for each factor portfolio. Thus, as in the single-factor case, active management in the multi-factor context has both a dynamic factor exposure component as well as a security selection component.

The dynamic component may add to active management returns by opportunistically increasing the weights on a factor prior to an above-mean realization. The selection component may add value by identifying undervalued securities within a group that comprise a factor and over-weighting them prior to their increase to fair price. Together, these two dimensions comprise the scope of active management. The essential features of both involve the choice and definition of the benchmark itself. The benchmark presumably comprises low-cost, passive portfolios. In a multi-factor context, the relative proportions of factors may either reflect the composition of the market portfolio, or a composition that reflects the preferences of the investor with respect to the risks of each factor. Manager skill is therefore evaluated by subtracting the systematic factor returns.

If an active manager is benchmarked against only a single market factor when there are actually multiple factors driving returns, then active management may appear to add value (alpha) relative to the market factor by accessing these other sources of risk premiums, which are not included in the market factor benchmark. In fact, given the overwhelming evidence that the market portfolio cannot span the full range of security payoffs and there exist alternative portfolios with higher reward-to-risk ratios, it would be highly unusual for an active manager

not to harvest these alternative risk premiums if the manager's task were defined narrowly to outperform the market benchmark, not a multifactor benchmark. In this case we would expect to see high correlations between the returns of the active manager relative to the market benchmark and other systematic risk factors. However, once the additional factors are brought into the active manager's benchmark then an alpha with respect to the single-factor market portfolio may disappear.

In practice there is a final avenue for an active manager to add value in the presence of a benchmark incorporating multiple factors. Since theory and empirical work have not identified the complete set of factors, active management can also add value by identifying new factors not specified in a multifactor benchmark.

We build on these ideas to evaluate the fund's active performance in Section II and to advocate a broader set of factors than currently used by the Fund to benchmark active management in Section III.

IB 6. Recent Anomalies

Since the 1970's, researchers and practitioners have uncovered, and continue to find, anomalies that appear to survive the size, value and momentum factors. From January 2006 to August 2009, there were at least 20 articles in the Journal of Finance documenting new anomalies. These include Ang et al. (2006) who show that firms with higher idiosyncratic volatility have lower returns, Baker and Wurgler (2006) who document that investor sentiment is priced in the cross section of stock returns, and Kumar and Lee (2006) and Kaniel, Saar and Titman (2008) who find that the trading of small retail investors influences stock prices. Edmans, Garcia and Norli (2007) find that World Cup elimination losses result in next-day negative stock returns. Campbell, Hilscher and Szilgayi (2008) find that stocks with high distress risk have abnormally low average returns. Many of these anomalies were found not in datasets that are new to the profession but in the same datasets, except with longer samples, used by the early researchers. Naturally, this raises issues of data mining by hundreds of researchers combing the same datasets for apparent mispricing.

Some recent anomalies have an economic rationale based on the slow diffusion of information. For example, Cohen and Frazzini (2008) organize U.S. firms into supply networks, and find that shares of suppliers react later to economic shocks than do the primary manufacturing firms to which they serve as a supplier. A trading rule based on this information lag generated positive profits – contrary to the implications of the EMH. The study is particularly interesting because it makes clear the nature of the efficiency violation: there are limits to the capacity of the market to understand cross-company and cross-industry economic news. Behavioral explanations based on the under-reaction of agents to news shocks also motivate many recent anomalies.

This under-reaction may be caused by limited attention, agents incorrectly updating beliefs, or not taking into account all available information in making decisions. Such effects underlie the different market reactions to tangible and intangible information (Daniel and Titman, 2006), the reversal and subsequent out-performance persisting up to one year of stocks with high returns over the past week (Gutierrez and Kelly, 2008), and the recent anomaly of different price responses to earnings announcements on Fridays versus other days of the week (Dellavigna and Pollet, 2009), among many others.

Many trading strategies which have been and continue to be pursued by hedge funds have been back-tested by academia and found to be profitable; among them, pairs trading (Gatev, Goetzmann and Rouwenhorst, 2006), merger arbitrage (Mitchell and Pulvino, 2001), and convertible bond arbitrage (Agarwal et al., 2009). Similar back-tests of trading strategies have yielded positive risk-adjusted returns to trading in other assets, including fixed income (Duarte, Longstaff and Yu, 2005). Likewise many trading strategies by active managers are inspired by academic research.[5]

Although thought-provoking, studies on historical archival data are ultimately limited in their capacity to prove that markets are inefficient. In addition to the problem of data mining, it is also the case that many researchers generally do not consider realistic expenses such as transactions costs, management costs, and financing costs. Further, all analyses of past prices are, by definition, conditional upon market operations that do not include an investor's own presence. Estimates of actual price impact for certain popular equity strategies such as size, value and momentum suggest that this conditioning is important (Chen, Stanzl and Watanabe, 2002; Korajczyk and Sadka, 2004), but trading costs do differ widely across different types of investors. While transactions costs may be an impediment to arbitrageurs entering the market and removing a potential mispricing, they usually do not explain why prices are initially formed to give rise to an anomaly. Nevertheless, many anomalies tend to occur in less-liquid securities and these cannot be exploited in extremely large sizes.

The bottom line is that the study of "anomalies" and later price-based empirical tests of the EMH do not reflect the actions of real-time investors operating in actual markets with real price impact and expenses. While they are suggestive of market inefficiencies, they do not prove that any investor actually profited or even could have profited from such activity.

To the extent that NBIM seeks to exploit these documented anomalies through direct strategies or by employing managers who rely on them, it is important to consider that the investment

[5] Two pedagogical examples are provided in Harvard Business School cases 9-298-012 "numeric investors l.p." which considers investing in the well-known value and momentum effects and 9-209-047 "Martingale Asset Management L.P." developing a strategy based on low volatility stocks outperforming high volatility stocks.

community has a tendency to converge on the same set of quantitative strategies to generate alpha. For example, Lo (2007) studied the several U.S. equity market moves on August 7th, 2007, and found evidence that when one major hedge fund unwound large positions, it caused severe dislocations in the prices of small cap stocks and other securities used in convergence strategies. This was almost certainly due to the fact that many different active managers were using the same models – and hence holding the same securities – in their quest for alpha.

The implication for NBIM is that many of the documented strategies for generating positive risk-adjusted returns through exploitation of mispricing have already attracted considerable capital. While proprietary strategies based on unpublished anomalies are less likely to suffer from such coordination, it is worth noting that many of the leading quantitative managers and were trained in the same anomaly literature and methods. Some at major institutions are in fact authors of the studies cited above. Although they may be working separately, it is likely that many of these managers are operating within the same conceptual paradigm. Lo's (2007) findings suggest that this represents an additional source of systematic risk.

Among the set of short-term pricing anomalies, those for which NBIM has a comparative advantage will be those that reward rather than penalize scale – for example high fixed-cost research-based strategies in very liquid markets as opposed to strategies that focus on illiquid securities with high transactions costs. The Efficient Market Theory would suggest that, at least with respect to relative value mispricing, these opportunities will be difficult to find.

By the same token, the Shleifer-Vishny theory would suggest that a patient, well-capitalized investor may be able to find price discrepancies that take a long time to converge. There are fewer of these studies in the empirical literature because they are harder to benchmark and to track through time with statistical accuracy. For those that exist – such as value investing – investors with shorter horizons or capital constraints might avoid these opportunities and leave them to "value" investors. This would imply, however, an appropriate horizon for performance evaluation needs to be explicitly chosen to align incentives to find and invest in those opportunities. We discuss this further in Section III.

IB 7. Long-Horizon Forecasts as Evidence of Inefficiency

Most studies of pricing anomalies focus on departure from relative value; in effect whether one group of close economic substitutes deviates from another. An alternative question is whether whole asset classes like the stock market itself might be over or under-valued. In the late 1980's researchers began to document and debate the apparent predictability of the aggregate stock market at long horizons: low dividend yields and earnings-price ratios, for example, were found to predict lower market returns. The market itself appeared to have a mean-reverting component that seems to suggest time-variation in actual and perhaps expected returns. The

key papers demonstrating long-horizon predictability include Fama and French (1988a, 1988b), Poterba and Summers (1988), and Campbell and Shiller (1988). These papers show that past stock returns, dividend yields, and earnings price ratios predicted multi-year stock market returns.

Robert Shiller (1981) took a closely related approach to testing market efficiency at long horizons that was based on the question of whether the stock market might vary considerably from its fundamental value, as represented by the present value of expected future dividends. He argued that stock prices were too volatile compared to changes in future dividends. He tested this proposition of "excess volatility" by comparing the volatility of the stock market to the volatility of the discounted present value of future dividends and found standard deviation of the former greater than the latter – even though asset prices are expectations of future discounted values. This result essentially implied that investing at long horizons was less "risky" than at short horizons, and that using fundamental indicators of market value – like the dividend yield or the earnings price ratio might be a guide for timing the market.

These long horizon studies of market efficiency were not based on beta-adjusted equivalencies, but on a present value model. They thus relied upon a critique of the unobserved expectations of market participants, and the suggestion that the markets could be irrationally valued in aggregate. This was most succinctly expressed by Summers (1986). Many researchers, Marsh and Merton (1986), for example, questioned the ability of empirical tests to distinguish between the formation of rational expectations at a given time in the past, and irrationality defined ex post by academics.

Extensive research over the following decades has qualified the findings on long-horizon predictability of excess returns on econometric grounds. Richardson (1993) and others pointed out that the statistical tools used in long-horizon prediction studies led to problems in inference. Goetzmann and Jorion (1993, 1995) cast doubt on whether the dividend can be used to forecast the equity premium at longer horizons due to issues of statistical power. Recent studies demonstrating the weakness of the long-horizon predictability evidence are Ang and Bekaert (2007) and Welch and Goyal (2008). Welch and Goyal find no convincing evidence that fundamental ratios could have been used out-of-sample to beat the market.

It should also be noted that, apart from issues of statistical power, evidence of predictability in returns by itself does not necessarily violate efficiency: it could simply reflect changes over time in expected returns or the risk premiums associated with priced factors. In fact this is part of a broader problem, namely that no firm conclusions about efficiency are possible without the "correct" model of expected returns. Unless two return series are perfectly correlated, a difference in their average returns could be due to a risk factor which has not been properly

accounted for in the model of expected returns. This "joint hypothesis" problem affects analyses of efficiency in all asset classes and, in particular, in equities and fixed income.

Despite questions of statistical power, the insights from the efficiency studies at longer horizons have generated proposals for dynamic investment strategies for long-term investors. For example, Campbell and Viceira (2002) propose strategies for strategically varying exposure to interest rates and other macroeconomic variables over the business cycle. Recently, Wachter and Warusawitharana (2009) addressed the statistical uncertainty about long-horizon predictability of asset classes in a Bayesian framework and showed that even an investor who is relatively skeptical of predictability would vary their exposure to U.S. stocks and bonds using the dividend yield as a forecasting variable. However, despite these theoretical benefits it is not easy for an international fund as large as the Fund to shift at short horizons to capture these benefits. Moreover, Ang and Bekaert (2007) using longer data series, international markets and a robust, non-linear specification find that, "At long horizons, excess return predictability by the dividend yield is not statistically significant, not robust across countries, and not robust across different sample periods."

The academic debate over long-horizon predictability remains a lively one. The question at issue is whether deviations from the long-term policy portfolio based on macro-economic variables can add value, ex ante. In particular, should NBIM seek to add value through a policy of actively adjusting its exposure to the equity premium or to other priced factors through long-horizon forecasting models? There is no satisfying answer due to the fact that tests of long-horizon forecasting skill demand much longer time-series data than exist for active managers. Consequently, although it is useful to recognize the potential value to be added by correctly predicting long-term premiums, no convincing empirical evidence suggest that is a reliable basis for an active strategy.

There has also been considerable research into whether changes in risk premiums might be forecastable in the short term. In a widely cited article, Ferson and Harvey (1991), presented evidence that monthly stock and bond portfolio returns were predictable using variables such as the equity premium, inflation shocks, yield curve changes and changes in default spreads. In their framework, this predictability is not a source of excess, risk-adjusted return, but rather a reflection of the fact that factor risk premiums change through time. Ferson and Schadt (1996) and Ferson and Kang (2002) propose that the predictable component of the variation in factor premiums should be incorporated into the evaluation of manager skill. We cite some evidence on the application of this multi-factor conditional risk adjustment to mutual fund and pension fund management below.

IB 8. Crashes as Evidence

One additional popular challenge to the Efficient Market Theory is the belief that periodic crashes in asset prices are prima facie evidence against efficiency. The crash of 1987 engendered considerable discussion since the U.S. market fell by more than 20% over two days. If prices reflect fundamental values, and no relevant information about those values appeared to spark the crash, then how could the EMH be true? The answer is of course that the theory only requires that prices reflect the aggregate expectation of economic value. Prices do not have to be economically "right." They need only be consistent with rational economic reasoning at the time (cf. French, 1988).

The crash of 1987 is an interesting example. Immediately after the crash, the common wisdom was that prices had been too high due to market euphoria. The dramatic market rise in the months and years after the crash suggests that, if anything, the crash was a panic and that prices before the event were closer to the present value of the market than prices afterwards. All that market crashes tell us about the EMH is that expectations – whether optimistic or pessimistic – can deviate widely from ex-post realized economic value. The theory does not require humanity to be prescient, merely to use its knowledge in pursuit of speculative profits.

The financial and economic crisis during 2007 and 2008 is another case in point. The fact that prices in many asset classes simultaneously declined is not evidence against the Efficient Market Theory per se. In fact, the large negative commonality of asset price movements during this period is consistent with an underlying APT model where many asset classes are exposed to the same common factors. Systematic factors themselves, as we document in Section II, can be correlated and there is no reason why factor returns should be symmetric. Furthermore, it is well known that exposures to systematic factors vary over time and increase during bad times (cf. Ang and Chen, 2002) causing many asset values to decline together. Hence, the failure of diversification, in the Swensen sense, of a portfolio holding many alternative assets not providing adequate downside protection can be interpreted as not correctly measuring underlying factor exposure of those alternative asset classes.

In the following section we turn to studies of active management which utilize historical data on investment manager performance. While it also suffers some of the same limitations of price studies in that scale diseconomies are potentially significant – particularly for an entity such as the Fund, they are one step closer to realistic tests of the value of active management and its relation to the EMH.

IC: Mutual Fund Performance as Evidence

Studies of mutual fund performance show no evidence of aggregate skill after fees and persistent positive manager alpha is difficult to identify. However, recent models suggest some investment in active management is rational.

As noted previously, mutual funds are not commonly used by large institutional investors and thus are far from ideal for determining the potential value-added by active management for one of the largest institutional portfolios in the world. While the mutual fund literature is enormous most of it is not directly relevant to the Fund. However, there are several aspects of the mutual fund literature that are relevant to the fund, which we outline below. An important point to note is that since mutual funds are targeted primarily at the U.S. retail clientele, their relatively high costs represent a high hurdle for tests of skilled active management. The discovery of manager skill in the mutual fund industry would constitute compelling evidence to reject the EMH, however the converse is not true. Skilled managers might offer their services in venues with less regulatory burden and record-keeping expenses. We examine these other vehicles in Section ID.

IC 1. Active vs. Passive Performance Tests

Researchers have been studying the average performance of U.S. equity mutual funds for several decades. Michael Jensen (1968), for example, introduced the famous "alpha" formula as a means to adjust mutual fund returns for differential exposures to the market portfolio. He found no evidence of systematic manager skill. The philosophy behind these tests is that active managers have to outperform the benchmark and that the amount of benchmark risk (the factor betas) can be specified by the investor. In Jensen's original article the benchmark is the market portfolio and the approach has been extended to zero-cost portfolios capturing size, value, momentum and other CAPM anomalies (cf. Carhart, 1997). A related approach is introduced by William Sharpe (1992) who specifies low cost index funds as the factor benchmark. An important underlying assumption behind these tests is that the factor exposures can be replicated by the investor and most studies assume that these factor exposures are constant.

Rather than reviewing the series of studies since Jensen (1968), it is sufficient to cite some of the most recent and carefully performed studies. Fama and French (2008) use data from 1962 to 2006 to examine evidence for skilled active management. Using their Fama-French factors and a momentum factor as a benchmark, the authors find that equity mutual funds provided negative 85 basis points per year to the investor. In short, indexation would have been better than investing with an active equity manager, on average. Before fees, Fama and French find that active management added 29 basis points per year but do not find this amount statistically

significant. They also note that the cross-sectional variation in manager performance is inconsistent with the existence of a group of over-performers as there were fewer extreme positive performance managers than would be expected by chance.

Wermers' (2000) comprehensive study of mutual fund performance also uses stock-level characteristic controls as well as factor benchmarks. The former involves constructing control portfolios that hold stocks with the same book-to-market ratios, size, and other firm-level characteristics as the stocks held by the mutual funds. He finds that mutual funds outperform the S&P500 on a gross basis, but underperform on a net basis. Similarly, the average gross alpha is 0.79% per year, but the average alpha after risk controls is -1.16% per year. Ferson and Schadt (1996) argue that the evidence for significant underperformance by fund managers is mitigated after taking into account predictable time-variation in factor risk premiums and find, that the average risk-adjusted post-fee mutual fund returns over the period 1968-1990 are about zero. They suggest that raw mutual fund manager performance suffers from poor timing of factor exposures.

In his presidential address to the American Financial Association, Kenneth French (2008) goes further and computes the cost of active management to all investors – including individual investors, mutual funds, institutions, and hedge funds in the U.S. equity market. He concludes that the average investor would increase his average return by 67 basis points per year switching from active to passive management.

The results in Wermers (2000), Fama and French (2008), and others echo the common finding in the very large mutual fund performance literature that mutual funds beat, or equal the market, before costs but underperform their benchmarks after costs. One important difference between a mutual fund investor and the Fund is that the Fund should not be paying the same large fees. In fact, the Fund's large scale gives it considerable advantages in capturing a large part of the gross returns compared to the net returns reported in the academic literature. If NBIM's expenses are low and NBIM can attract talented people, it should be able to capture some part of the before-fee outperformance of active management reported in the mutual fund literature.

An interesting recent approach to the question of whether active management should be rejected as a strategy is to ask under what set of beliefs concerning market efficiency an investor should eschew all active management. Baks, Metrick and Wachter (2001) address this question using a Bayesian framework and their results suggest that virtually all investors – even very skeptical ones – would use at least some active management. In a related approach, Pastor and Stambaugh (2002) find that, "Investing in active mutual funds can be optimal even for investors who believe managers cannot outperform passive indexes." These Bayesian approaches make a strong case for active investing even in mutual funds. However, the

proportion of active management remains an open question. In NBIM's case, as we show below in Section II, the actual contribution of active management to the overall Fund return is extremely small, which is consistent with a Bayesian framework and very conservative prior beliefs about the existence of skill.

IC 2. Relative Performance Tests

Another approach to the question of active management is to see if winning managers repeat their superior performance. Tests of performance persistence have a long history beginning with William Sharpe (1966), who found some evidence of persistence. Since then, researchers continued to find some evidence of persistence, but funds that beat the market on a risk-adjusted basis consistently are hard to find. Often, performance persistence shows up in underperforming funds continuing to underperform rather than outperforming funds tending to continue their outperformance (cf. Brown and Goetzmann, 1995; Gruber, 1996; Carhart, 1997). Note that this literature does not support the conclusion that investing with winning managers will, on average, increase a mutual fund investor's probability of beating the market on a risk-adjusted basis. For NBIM the practical implication of these results is to validate the termination of poorly performing managers but does not necessarily support the hiring of top performers.

One finding that is important for the Fund is that flows into and out of mutual funds are strongly related to lagged performance measures (cf. Chevalier and Ellison, 1997; Sirri and Tufano, 1998). But once this money enters funds that have performed well in the past, these same funds tend not to outperform in the future. Such a phenomenon is explained by Berk and Green's theory where managers do have ability to add value but this ability decreases as funds increase in size. As investors learn from past returns they shift their money into funds with high past performance, but as these funds increase in size the manager's alpha disappears. While NBIM is a dedicated portfolio manager for the Fund, it does employ a large number of outside managers. Thus, NBIM does shift its money to and from external funds, which it must evaluate using data that include past return information. Given that many active strategies cannot be pursued in large size, NBIM's external allocations can be subject to this effect where as the amount of money allocated to external managers grows, the ability of external managers to add value in active management declines. We examine this in Section II.

There are many ways that active mutual fund managers can potentially deliver positive risk-adjusted returns. One that is relevant for the Fund is the technique of changing exposures to different systematic factors (cf. Jagannathan and Wang, 1996). In the standard factor control regressions the exposures to the factors are estimated to be constant. However, Mamaysky, Spiegel and Zhang (2008) find strong evidence that active management adds value by allowing dynamic exposure to systematic factors. Like the conditional performance metrics of Ferson

and Harvey (1991), however, the beneficial predictability in their analysis is mostly short term in nature. Timing skill, if it exists at all, is likely to be only measureable at short horizons. Thus, dynamic strategies are, by their very nature, difficult to measure with standard, static factor models. Indeed, recently Goetzmann et al. (2008) show that many dynamic strategies, or the use of derivatives by active managers, cause the standard linear model used to calculate alpha or risk-adjusted return to misrepresent performance. Nevertheless, changing factor exposures, relative to a market-weighted portfolio or a set of factors with fixed weights, is a potential source of alpha.

Finally, another area of the mutual fund literature directly relevant to the Fund is the finding, both theoretically and empirically, that incentives matter for performance. Chevalier and Ellison (1997) show that the strong response of investor flows to past performance tends to reward risk taking by mutual fund managers since mutual fund managers are compensated for assets under management. Elton, Gruber and Blake (2003) look explicitly at incentive fees in the mutual fund industry and find that funds that have compensation tied to performance exhibit stock selection ability. Khorana, Servaes and Wedge (2006) study funds in which the manager owns shares and thus has an economic alignment with customers. They find a slight but significant positive relationship. To the extent that NBIM has the scale to significantly increase the AUM of a manager, and once invested, has the implicit threat of decreasing the AUM, this clearly serves as a tool for incentivizing external managers. The literature further suggests that direct alignment through shareholding, and indirect alignment through incentive compensation has positive effects of realized performance.

ID: Other Managed Portfolios as Evidence

Institutional portfolio managers display little evidence of manager selection ability. There is past evidence of hedge fund alpha, however, changing market conditions would urge caution in presuming this to continue.

In this section we summarize the performance of other managed funds. Given the large number of constraints under which mutual funds operate and that superior talent may be attracted by the larger fees levied by other alternative asset vehicles, non-mutual fund portfolios constitute an interesting area to investigate active management. These managed portfolios are also far more relevant to the Fund.

ID 1. Institutional Funds

In this category we place institutional funds such as pension funds and endowments as well as large-scale private capital pools, although the latter differ in that they are taxable. Institutional

funds are distinct from mutual funds because they have much lower servicing and marketing expenses, and the decisions are made by professionals rather than retail investors. These lower costs should, in turn, represent less drag on performance.

There have been few studies on institutional funds and managers. Tonks (2005) uses UK pension fund data from 1983 through 1997 and finds positive risk-adjusted manager performance, accounting for survivorship and for momentum. Christopherson, Ferson and Glassman (1998) use a potentially survival-biased sample, however they find strong evidence of persistence in the performance of pension fund managers from 1979-1990, although the persistence is concentrated among poor performers. Goyal and Wahal (2008) show that in terminating managers, pension plan sponsors would have done just as well keeping the old managers as hiring new ones. Busse, Goyal and Wahal (2009) look at persistence in manager performance using more than 6,000 institutional accounts over 1991 to 2008. They find evidence of significant performance persistence, however it is attributable to the momentum factor – i.e. chasing returns with positive previous year performance.[6]

Stewart et al. (2009) use institutional manager data provided by Informa Investment Solutions from 1984 through 2006 and ask whether institutional pension plan sponsors add value through the manager selection process. In short, they ask if plan sponsor money is "smart." The answer is no. Consistent with Goyal and Wahal (2008) they found that products with the largest outflows outperformed those with the largest inflows.

In contrast, Bauer, Cremers and Frehen (2008) analyze the domestic equity portfolios of 955 pension plans compared to their specified benchmarks. They are able to examine gross and net returns and their focus is on understanding agency costs in the financial services industry. Their main finding is that pension funds out-perform style and size-matched mutual funds by 200 basis points per year. They attribute this differential to hidden agency costs in the mutual fund industry that are not captured by reported fees.

University endowments are of particular interest given the role that the top university endowment managers have played in changing the fundamental asset allocation approaches to institutional management. A systematic survey of college endowments is conducted regularly by NACUBO. The data from 1992 to 2005 were studied by Lerner, Schoar and Wang (2008) who looked for correlates to good performance. They found that the largest endowments and endowments for academically elite institutions out-performed the best. These institutions were also the group that relied most on alternative investments.

[6] Interestingly, in an earlier version of the paper with data extending to 2004, they find significant positive performance, accounting for factor exposures.

Brown and Garlappi (2009) and Brown, Garlappi and Tiu (2009) use the NACUBO data to study whether the asset allocation decisions of endowments were profitable, and whether they could add value through active management. The latter study specifically asks whether endowments took sufficient active decisions away from their benchmarks over the period 1989 to 2005. The approach essentially chooses a mix between active and benchmark strategies to maximize the risk return tradeoff ex post. They found that most endowments were too cautious in deviations from benchmark, and could have added value by being more active. They tested this proposition out of sample by using first half of the sample period to select an active allocation weight for each endowment then applied this to the second half of the sample period and found a significant increase in the Sharpe (reward-to-risk) ratio. Their conclusion is that the alpha-generating capabilities of endowment managers are typically underused.

One limitation of these recent studies of university endowments is that they all use data that end before the recent crash. This omission has advantages and disadvantages. The advantage is that a single, adverse event does not color analysis. A common human bias in data interpretation is to overweight the salience of recent events. To the extent that one believes that the basic economic system is in a long-term equilibrium state, one way to mitigate this bias is to omit recent data. Thus, a case can be made for placing considerable weight on these pre-crash institutional studies.

On the other hand, the disadvantage of these studies is that the crash of 2007-2008 revealed a lot about the long-run equilibrium state of the economy. In particular, it showed the dramatic effect of severe market distress on asset allocation strategies. In particular, the diversification expected by multi-asset managers based on low recent historical correlations did not protect against a truly major shock. By the same token, excess-return investment strategies that had performed regularly in the previous decade ceased to do so in this same interval, suggesting they were likewise vulnerable to market contractions.

The recent crash highlights a major limitation of the modern alternatives-focused endowment strategy. There is simply no long-term empirical data for most of the popular alternative investment classes. This makes estimates of risk and return imprecise at best, but more importantly, the lack of such data means that it is difficult to model performance during periods of macro-economic stress. The sudden shift towards investing in alternatives is thus truly remarkable, given the long-term perspective of endowments. Over the very long-term, endowments should expect major ruptures in markets similar to those that occurred in the 20th century. These include wars, recessionary shocks, demographic changes, market emergences and disappearances. None of the popular alternative asset classes – save real property – have been in existence in the current scale and institutional structure for more than forty years. The allocation away from well-understood asset classes like equities, fixed income, and real

properties should be regarded as a recent experiment, and well-documented successes such as the Ivy League endowments should be examined in the context of first-mover advantage, top manager special access, and a rare, quiescent period in financial history in which markets for normally illiquid assets flourished.

The success up to 2008 of major endowments like Harvard and Yale presents an important question for any large institutional investor. Should this approach be imitated? If so, what performance evaluation structure should be employed to address future shocks to market liquidity recently experienced by these endowments? The recent crash has led to reported losses in the 30% range and publicly disclosed liquidity problems by major institutions. For example, in 2008, the Harvard endowment announced an intention to sell its positions in venture capital. While the long-term performance of the large, leading endowments despite recent losses is still superior to that of smaller endowments when considered over the last decade, several questions should be considered. First, can the financial acumen of the leading endowment managers be imitated or acquired? Second, have recent market events forced a reconsideration of the diversification provided by alternative assets? Third, what is required to be a truly patient investor, and fourth would an alternatives-oriented portfolio have been robust to the financial and macro-economic shocks of the 20th century? These questions would need to be addressed before adopting an investment approach dominated by alternatives.

ID 2. Sovereign Wealth Funds

Other sovereign wealth funds [SWFs] represent a natural institutional comparison group for the Fund. However, the data on many of the largest SWFs, with one striking exception being Norway, is difficult to obtain and contains potential bias because the data, when available, is mostly voluntarily reported.

Bortolotti et al (2008) studied a voluntarily reported sample of 1,216 transactions by 35 SWFs over the period 1986 to 2008. A subset of these transactions were purchases of shares in public companies, which allowed the researchers to calculate an abnormal return. They find short-term positive performance and long-term underperformance when the purchases are benchmarked against an appropriate set of comparisons. Kotter and Lel (2008) and Dewenter, Han and Malatesta (2009) find similar positive announcement effects (1% to 2%) but essentially zero long-run risk-adjusted performance. The announcement effect is consistent with a positive market response to a SWF taking a major stake in a firm; the long-run flat to negative performance is consistent with the difficulty in identifying investment opportunities for out-performance over long horizons. Fernandes (2009), focusing on price not return, documents an apparently large premium (15%) due to sovereign fund ownership, also consistent with expected positive governance effects of a large external shareholder. On the other hand, Knill,

Lee and Mauck (2009) note that Sharpe ratios decrease due to sovereign fund ownership, which they interpret as potentially "destabilizing."

Although the academic literature on SWFs is very recent, the empirical evidence suggests that the potential gains by funds are related to issues surrounding expectations about governance and corporate control. To the extent that funds are able to acquire shares at market prices prior to announcement of a major stake, they may benefit from a modest announcement effect. The positive price adjustment appears to fully encompass the anticipated positive governance effects – in other words, the market free-rides on the SWFs' monitoring of company management.

ID 3. Price Impact and Indexation

One area of active research in institutional funds management is the price impact of fund flows. This line of study began in the 1980's with Shleifer's (1986) discovery that prices of stocks that moved onto the S&P index jumped on the day of inclusion – as index funds all bought on the same day. This led Shleifer to hypothesize that the demand for stocks sloped down and that as money chased particular investments, this would drive prices up and expected returns down. Extensions of these early studies pushed in several directions. Some looked for evidence that the shock was temporary – consistent with the Shleifer-Vishny limits-to-arbitrage model (cf. Wurgler and Zhuravskaya, 2002).

Others tested for causality between institutional flows and security prices. Gompers and Metrick (2001), for example, studied stocks held by institutions and found an association between valuations and institutional ownership. Sias, Starks and Titman (2006) showed that the price impact had both a permanent and a temporary component. Goetzmann and Massa (2003) used high frequency flows to S&P 500 funds to test the hypothesis of flows causing price changes, documenting an immediate impact on stock prices in the S&P 500 in the afternoon following positive order imbalances for index investors.

Taken together, these results suggest that indexation comes at a cost of its own due to institutions tracking these indices trading at the same time. The main cost is a one-time differential of 2% or more due to a sloped demand curve, another cost is the continuing response to flows. This evidence is particularly relevant for the Fund in setting its own benchmarks. The Fund would be better off not using standard market-weighted indices. We discuss this further in Section III.

ID 4. Hedge Funds

A natural arena to examine the question of market efficiency and the efficacy of active management is the hedge fund universe. Although there is no standard institutional definition of a hedge fund, they are generally actively-managed investment pools that seek to generate positive alpha for clients. They are sometimes referred to as absolute return funds because they do not benchmark themselves against specific long-only indices, but against an absolute standard such as LIBOR. The compensation structure in the hedge fund industry is consistent with this goal. The standard manager contract is a 2% fixed fee with a 20% incentive fee and a high water mark. The manager thus gets 2% of net assets per year and 20% of gains in excess of a fixed benchmark like 0%, or LIBOR. The high water mark provision means that previous year's losses must be recouped by the investor before the 20% incentive provision is in effect. Given these extraordinary incentives, one would expect the hedge fund industry to attract the most skilled managers. These very high fees also imply that if hedge fund performance adds value after fees and hedge fund strategies can be mimicked by a low cost manager, such strategies would lead to significant value.

One of the major challenges to studying hedge funds is that the industry is comparatively new and has changed dramatically in its short history. Performance data, particularly for early years, is of questionable quality and still voluntary making it subject to such problems as survival and selection bias. A deeper challenge to measuring hedge fund excess performance is that, unlike long-only managers, hedge funds use a wide range of securities, contracts, leverage and trading techniques, and their positions can shift quickly. For example, a global macro hedge fund might shift its exposure to a particular currency from highly positive to highly negative overnight and then back again within a few days. Statistical controls for this changing systematic factor exposure are imperfect at best. Because of their wide ranging strategies, hedge fund returns also may be skewed or have option-like characteristics that render standard linear techniques inadequate. This has motivated considerable research into how to capture these complexities with alternate statistical instruments.

With these qualifications, let us turn to the existing academic evidence. Empirical studies of hedge funds are comparatively recent. Fung and Hsieh (1997) studied a database of hedge funds and found that their trading strategies were extremely dynamic and not well-captured by standard index benchmarks. They introduced a number of additional systematic factors that are now commonly used as regressors in the performance measurement of hedge funds. Theirs was the first paper to show that hedge funds were a good investment. Following their pioneering study, Ackerman, McEnally and Ravenscraft (1999) found that hedge funds outperformed mutual funds over the period 1988 through 1995, but do not on average, provide positive risk-adjusted returns. In contrast, Brown, Goetzmann and Ibbotson (1999) found

evidence of positive risk-adjusted performance in a database of off-shore hedge funds over the same time period. Surprisingly, they failed to find evidence of performance persistence in funds. Later studies like Agarawal and Naik (2000) with longer time series data have since documented repeat-winners at short horizons among hedge fund managers, however, suggesting some skill differential. Jagannathan, Malakhov and Novikov's (2006) hedge fund study documents persistence in winners as opposed to losers – also strongly suggestive of skill.

More recently, Bailey, Li and Zhang (2004) found evidence of the average outperformance of hedge funds under the null of no arbitrage, even when non-linear factor payoffs are considered. Kosowski, Naik and Teoh (2007) examined the risk adjusted performance of hedge funds over the period 1990 to 2002 using fairly sophisticated measures. Their results concur that hedge funds over this extended period appear to deliver positive performance, and that performance is persistent at the annual horizon. Avramov, Barras and Kowsowski (2008) develop a predictive model for individual hedge fund managers similar in spirit to the conditional models discussed above used to predict mutual fund returns. They document a dynamic strategy of hedged fund allocation that generates positive alpha and attribute it to time-varying skill.

There is some contradictory evidence on hedge fund manager skill, however. Griffin and Xu (2007) examine hedge fund trades around the tech bubble of 2000 and cannot find evidence of differential or superior trading skill. Malkiel and Saha (2005) argue that hedge funds fail to deliver significant positive performance, and that the problems of survivorship bias and backfill bias loom large in any reliance on historical hedge fund data. Gibson and Wang (2009) add a liquidity risk variable to the Avarmov, Barras and Kowsowski framework and show that a lot of the alpha may be compensation for bearing liquidity risk.

On balance, however, some, but not all academic hedge fund studies have found evidence that hedge funds contribute positive excess returns. The only caveat is that, given the difficulties in performance measurement and the lesser quality and short history the hedge fund databases, the positive results about past performance need to be qualified with a concern that reported results might not truly reflect real-time investor experience.

Going forward, can these positive results be expected to continue? The changes in the industry such as the vast increase in scale, the competition for alpha within scale-constrained strategies, the re-institutionalization of proprietary trading, the orientation towards marketing, the introduction of indices and synthetics, and the shifting regulatory landscape, all suggest that the past cannot simply be extrapolated. So far, the cumulated evidence suggests that hedge fund manager skill exists and that the rewards to that skill can be passed on to fund investors, depending upon a judicious manager selection process. However, the mutual fund industry has succeeded in delivering negative active returns to investors while keeping any talent as fees. In

the long run, there is no reason why the hedge fund industry would not tend to the same Berk-Green equilibrium.

Furthermore, many hedge funds are small relative to the size of the Fund and many hedge fund strategies are simply not scalable. This is shown by Agarwal, Daniel and Naik (2004) who show high flows into hedge funds are followed by poor fund performance. Teoh (2009) is a comprehensive study of the size/performance relationship in the hedge fund universe. He finds strong evidence of a convex (diminishing) relationship between size and risk-adjusted returns, consistent with capacity constraints.

Because the hedge fund industry is undoubtedly an arena to which the best and brightest managers gravitate we compare NBIM active returns against hedge funds in Section II. While hedge funds are attractive as investment vehicles at this time, issues such as manager access, transparency, reporting, evaluation, and due diligence remain challenging, even for large-scale institutional managers. Any sophisticated hedge fund program should take into account such issues when deciding upon an allocation to specific managers and strategies.[7]

IE: Other Market Sectors

Fixed income, venture capital, private equity and real estate would seem to be potential areas for active managers to exploit skill. The evidence is mixed.

IE 1. Fixed Income

A recent study of the various fixed-income active strategies commonly taken by hedge funds was conducted by Duarte, Longstaff and Yu (2005). They looked at profits from trading on the interest rate swap spread, the yield curve (i.e. the Expectations Hypothesis), mortgage-backed security mispricing, arbitrage in volatility (i.e. differential pricing of risk in different securities markets), and capital structure arbitrage. In effect, they tested the efficiency of the fixed income market using a variety of strategies employed in practice. The concluded that several of these strategies were profitable ex post, and that those with greater complexity paid off more – consistent with the Grossman-Stiglitz theory. In many cases, however, the excess returns computed by Duarte, Longstaff and Yu were not statistically significant. Moreover, an open question is the exposure of these strategies to other systematic risk factors, especially liquidity and volatility risk, which we explicitly incorporate in Section II.

[7] Funds of hedge funds are set up to deal precisely with such challenges. We omit reference to this literature because it is less relevant to the Fund. If NBIM has sufficient expertise and resources it should invest directly with hedge funds rather than through funds of funds.

There are few studies of bond portfolio managers. An early study is Elton, Gruber and Blake (1993) who study bond mutual funds and find no evidence of manager skill. Ferson, Henry and Kisgen (2006) also study bond mutual funds using a different methodology and different data period and also report that performance on average is negative – although not statistically significant. In these studies of bond mutual funds, the issue of before-fee and after-fee performance is similar to equity mutual funds. Chen, Ferson and Peters (2009) find that bond mutual funds outperform several common bond index fund benchmarks on a before-cost basis but under-perform on an after-fee basis. They document a large average expense ratio of around 1.3%. Low-cost active management in fixed income markets could capture alpha close to those reported for gross returns. Repeat-winner tests on bond funds (cf. Huij and Derwall, 2007) find performance persistence among fixed income mutual funds, suggestive of differences in relative skill. Moneta (2009) recently studied holdings of bond mutual funds and documented some positive evidence of management skill, based upon timing ability.

As with active equity managers, one would expect to find skilled fixed income managers in the hedge fund universe because of the relative lack of strategy constraints and the attractive compensation. Following this logic, Fung and Hsieh (2002) studied the fixed income managers in the hedge fund universe up to 2001. The paper has a rich set of findings relevant to the analysis of active fixed income management. First, the authors show that the returns to various strategies within the fixed income category are related to not only standard long-only indexes but also to proxies for trend-following and convergence trading. They note that these factor exposures are different than those represented in the mutual fund universe. They also point out that the shocks to these factors in the 1990's and 2000's were relatively small. Back-testing the effects of Great Depression-scale shocks to credit spreads, the authors forecast the potential for large monthly declines, in the order of 8% for at least one style. Fung and Hsieh conclude that, "there exists cyclical exposure to risk factors inherent in most Fixed-Income Arbitrage funds that may be masked by the short existence of the funds themselves." How right they were! Although the paper is silent on the average risk-adjusted return to fixed income funds, the implication of their study is that even a positive and significant historical alpha for a fixed income return may actually the be the result of an exposure to rare risk events, and thus no more than fair compensation for risks – once they are all properly accounted for.

Okunev and White (2003) provide corroborating evidence by more intensive econometric work on "unsmoothed" monthly hedge fund returns. They uncovered a non-linear, option-like strategy inherent in the fixed income hedge fund index similar to a short position in a put on high-yield debt. This strategy would explain why many active fixed-income managers performed so poorly relative to indexes during the credit crisis. Thus, performance analysis of fixed income funds is highly dependent upon both the time-period of study and upon the choice or risk factors and their empirical proxies. Most interesting is the clear prediction by two

research teams in the early 2000's that fixed income hedge funds were asymmetrically exposed to serious credit shocks. The prediction of Fung and Hsieh is borne out in the experience of NBIM active management in 2008, especially in the fixed income portfolio, as we show in Section II.

Our discussion of fixed income has been necessarily brief because research on both bond market efficiency and delegated fixed-income portfolio management is much less extensive than for the equity market. This obvious large gap in the literature is unfortunate, both in general and specifically in the context of the Fund which is 40% invested in fixed income.

IE 2. Private Equity and Venture Capital

The EMH generally presumes a liquid capital market. Non-marketable alternatives such as real estate, private equity, venture capital – and even markets farther afield such as collectables – are thus considered to be less efficient. The Swensen approach to institutional investing stakes a considerable amount on these markets and thus it is worth considering the academic evidence on non-marketable alternative such as it is.

The most widely cited paper on private equity investment is Moskowitz and Vissing-Jorgensen (2002) which treats all non-marketable entrepreneurial investment in the U.S. as private equity. Their conclusion is that private equity is no more profitable than public equity. Phalippou and Gottschalg (2009) are even less sanguine. They claim that the reported returns in historical private equity databases are overstated and that private equity fund returns are 3% lower than those of the S&P 500 and 6% lower than a risk-adjusted benchmark. Conroy and Harris (2007), using data from 1989 through 2005 and adjusting for smoothing of returns, concur, pointing out that their portfolio characteristics are not particularly attractive. Franzoni, Nowack and Phalippou (2009) suggest that private equity funds have a major exposure to liquidity risk and this factor should be considered in the allocation decision.

The fees in private equity investments are very high. Most private equity funds charge a 2% management fee and 20% on the profits (called "carry"). The 2% fee is charged on committed, not invested, capital. That is, suppose the initial investment is $10 million on a fund where investors have committed $100 million of capital, which is drawn upon as needed by the manager. The management fee charged in the first year is on $100 million, not $10 million. Not surprisingly, even though the returns to investors are below the S&P500, or the risk-adjusted benchmark, Phalippou and Gottschlag find that the gross-of-fee alpha is positive at 4%. If NBIM contemplates private equity investments, these are best done in-house where most of the gross returns can be captured and the fees minimized.

Kaplan and Schoar (2005) also examine private equity funds and venture capital funds. They generally agree with the conclusion of average poor performance. Net of fees, the IRR of

private equity and venture capital investment differed little from that of the S&P500 over equivalent time periods. After appropriate risk-adjustment these investments are less attractive than an index of public equity. However, Kaplan and Schoar also note considerable cross-sectional variation, suggesting there may be differential manager skill. The fund offerings of general partners of previously successful private equity funds tend to also perform well. This lends some support to the Swensen approach, and suggests that market frictions such as access to top managers and economies of scale may come into play in achieving good returns in the non-marketable alternatives space.

Yale and Harvard may be able to invest in these top private equity funds along with certain other University endowments (cf. Lerner, Schoar and Wongsunwai, 2007). But, it is not clear that other investors can or that even past performers persist. Phalippou (2008) casts doubt on previous findings of persistence in performance for venture capital funds, arguing that the data support the decreasing returns to scale argument advanced by Berk and Green. On the other hand, Ljundqvist et al. (2009) argue that persistence in after-fee returns found by Kaplan and Schoar is driven by investors having some holdup power that prevents top funds from pricing their services to the margin.

Cochrane (2005) examines venture capital returns from 1987 to 2000, addressing a number of methodological problems. He finds that venture capital returned 15% in this period, roughly the same as the S&P 500, however risk adjustment for market exposure resulted in a negative alpha for the industry. These results are confirmed by Korteweg and Sorensen (2008) who also address econometric biases due to use of infrequently traded data and estimate low alphas for venture capital.

Metrick and Yasuda (2009) point out that the venture industry differs markedly from the private equity industry in that venture capitalist focus on small firms and thus the business is by definition labor-intensive and non-scalable. Buyout firms, on the other hand, can operate on billion-dollar companies. This suggests that the performance of venture capital as opposed to private equity should be viewed as strongly conditional upon scale, and that past percentage returns are not necessarily valid as a basis for projecting future performance. Venture capital returns depend critically on opportunities and capital in the industry. Thus, investing in a venture capital area that is currently “hot” for the Fund will probably lead to poor future performance.

IE 3. Real Estate

Basic theories of diversification would recommend holding a substantial portion on an investment portfolio in assets that constitute the majority of world wealth: real property. This asset class includes direct investment in housing, commercial buildings, agricultural land and

mineral and extraction rights. Constraints on real property investment are significant, however. They include issues of liquidity, agency, and nationalism. Securitized investment in property is, of course, possible via bonds and publicly traded property companies. Several attempts have been made to develop derivatives contracts based on real estate indices. Despite the widespread belief in demand for these derivative products, none have been successful enough to serve the needs of large-scale investors. Nor has the scale of publicly-traded real estate claims matched the relative wealth share of the asset class. As a result, it would be very difficult for the Fund to replicate the returns to real estate investment using passive investment products with one exception: real estate investment trusts [REITs], which will be discussed below.

The question addressed in this section is whether there are opportunities to exploit mispricing in real estate via active management. We thus review the academic evidence on the efficiency of the real estate market and consider how NBIM might or might not be able to profit from this.

1E 3a. Empirical Studies and Data Sources

Some institutional background: one set of real estate studies uses information from publicly traded firms (REITs) and other studies use pooled information from institutional investment via the National Council on Real Estate Fiduciaries [NCREIF]. The NCREIF database contains return information on U.S. properties held by insurance companies, pension funds, and endowments. These two sources of information are quite different, since REITs are priced by the market and the NCREIF data is appraisal-based. Finally, for residential real estate, the most widely observed market indexes, the Case-Shiller indexes, are based upon repeated-sales of houses. Only the REIT indices are replicable by an investor. There are similar indices for the UK and other countries, for example, Investment Property Databank provides an appraisal estimate for several countries, including the UK, broken down by property type. The same caveat applies. Tracking an index of properties you do not own is likely to be difficult.

Random walk tests on non-tradable real estate indices have generally shown high levels of predictability. Case and Shiller (1989), Goetzmann (1993) and Kou (1996) found that U.S. housing indices have strong autocorrelation. Goetzmann and Ibbotson (1990) calculated high autocorrelation for appraisal-based commercial property indices. Publicly traded real estate companies demonstrate less predictability, and may deviate in relative value from appraisal-based measures. Whether this deviation between two measures of commercial property values is due to stock market prices being wrong, or due to appraisals being wrong is currently an open research question. Pagliari, Scherer and Monopoli (2005) find that this deviation has narrowed through time, suggesting increased integration between the public and private real estate markets. The predictability of the non-market-based indices in any case does not suggest that the markets are inefficient in a no-arbitrage sense – frictions prevent speculators

from easily exploiting trends, a point made by Gatzlaff and Tirtiroglu (1995). A long-short trade on public vs. private real estate can be interpreted as a financing challenge in the Shleifer-Vishny framework.

Darrat and Glascock (1993) broadly ask whether real estate investments provide returns commensurate with their macroeconomic risks and find that they do. On the other hand, well before the burst in U.S. housing prices, some researchers were finding evidence that home prices deviated from "fundamentals" due to euphoric over-reaction. Capozza and Seguin (1999) report evidence from rent-to-price ratios suggesting that homeowners extrapolated temporary trends in income growth. Robert Shiller also famously predicted the recent housing bubble and collapse, suggesting it was due to market irrationality.

IE 3b. Skill

Do real estate managers have skill? With the broader accessibility of REIT pricing in the 1990's, real estate scholars generally followed the precedent of testing for trading rules and real estate manager skill. Cooper and Downs (1999) showed that a short-term contrarian strategy on REITs made excess returns after accounting for microstructure costs and frictions. Titman and Warga (1986) did an analysis of REITs and found little evidence of differential skill across managers, and noted that a single factor model was inadequate for analyzing real estate returns. Rodriguez (2007) looked at managers of REIT mutual funds from 1999 to 2004 and, in contrast to the results obtained by Gallo, Lockwood and Rutherford (2000) for an earlier time period, found no selection skill. More recently, Lin and Yung (2004) looked for evidence of performance persistence in REITs and found none, although they looked at returns on the publicly traded shares, as opposed to the performance of the underlying assets. Presumably the share prices would reflect investor expectations of manager skill as well as expectations of asset values.

Beyond REITs, there have been some studies of funds serving the institutional sector. Brueggerman, Chen and Thiodeau (1984) compared the performance of institutional real estate funds to a rational benchmark based on the CAPM, but the data series for analysis were short. Gallo, Lockwood and Rodriguez (2006) collected evidence from the co-mingled real estate fund sector serving institutional managers for the period 1985-2002 and tested for performance persistence. Among other things they found that tracking error was correlated to under-performance.

Ciocetti and Fisher (2002) and Goetzmann and Fisher (2005) examine the managed real estate assets represented in the NCREIF database – presumably indicative of a core investment style, as opposed to development or opportunistic styles. Their evidence suggests that over the period 1977 to 2004, the internal rate of return to a well-diversified portfolio of properties was

about equal to the performance of transactions-matched U.S. government bond investments over the same time period, suggesting that equilibrium expectations were not realized over the 28-year interval, additional factor exposures explained relative returns, or that this particular set of managed assets simply did not beat bonds.

Recent analysis of real estate investing by Bond, Karolyi and Sanders (2003) among others has explored the international structure of property returns. Bardhan, Edelstein and Tsang (2008) look at the relative efficiency of national real estate markets by using property stocks as a measure of the integration of global real estate markets. They find that real estate firms in less open economies have higher risk-adjusted returns; evidence in favor of local pricing and potential gains for the cross-border investor. These results might be relevant for NBIM because they represent a source of factor exposure to the asset class and also offer the potential for active management of marketable assets.

IF: Conclusion of Section I

The EMH has been refined over the past several decades to reflect information, transactions, financing and agency costs. Tests of the theory on prices have occasionally produced violations, suggestive of the potential for active management to add value to a multi-asset portfolio.

Tests of the theory on managers are of limited relevance to the Fund, since much of the efficiency literature has focused on retail mutual funds, which have a very different cost structure and value proposition. Nonetheless, it does appear that within mutual funds there is some evidence of differential skill. A Bayesian analysis of skill in the mutual fund sector argues that passive indexing is difficult to justify when there is a large population of managers.

Turning to other investment sectors, some studies of endowments have shown positive returns to active decisions over a period before the recent crisis. A limited number of studies report mixed results on other institutional manager performance. Studies using data prior to the mid-2000's find some performance persistence, but more recent evidence finds none. At best, the evidence for institutional manager skill is time-dependent.

There is evidence of positive returns in the hedge fund sector where highly paid managers actively trade marketable securities but the quality and duration of these data remain open to criticism and doubt. On the other hand, there is surprisingly little convincing evidence of superior risk-adjusted returns to private equity and venture capital. Despite claims that top managers can produce consistent high returns, the current data are not conclusive on this point. In the real estate sector there is simply not enough information to evaluate whether managers have added value on a risk-adjusted basis. Interestingly, the evidence does not

suggest a “pecking order” structure with respect to potential alpha ranging from highly liquid asset classes such as fixed income to equities to less liquid securities

To summarize the relevance of the efficient market theory and evidence with to the decision faced by the Fund:

1. The preponderance of empirical evidence from academic research suggests that liquid securities markets are generally efficient with respect to information. More specifically, prices are driven towards efficiency through active trading by market participants. Economic profits may accrue to managers with competitive advantages in the acquisition, analysis and trading on value-relevant information. As a result, the balance between indexation and active management is a choice variable for which the optimum depends on many factors. These include beliefs about the existence and potential of manager skill, the pricing opportunities afforded within a given market, the time preferences and risk aversion of the investor, and the expertise and incentive contract of the manager.

2. If the benchmark is solely a market-weighted portfolio consisting of all traded securities, then active management (defined as deviations from these market-weights) may be useful in accessing factor risk premiums that are not captured by market exposure. In the context of a multifactor model this could also be interpreted as passive exposure to additional risk factors. Theory and empirical evidence suggests that investors are compensated for taking systematic risks – such as value-growth and volatility risk – over the long-term. In the presence of these multiple systematic risk factors, empirical tests overwhelmingly reject that the market portfolio is efficient. However, often there are no standard industry benchmarks that could serve as passive benchmarks to systematic factors other than fixed income and equity risk factors.

3. Back-tests of trading strategies have suggested a wide array of potentially profitable investment strategies. However, for a number of reasons these provide limited guidance to the Fund going forward. They represent simulated, not actual, returns and usually do not take into account transactions costs, fees, and price impact. They suffer from data-mining biases. Changing market conditions, including arbitrage activity, make it difficult to extrapolate future performance. Finally, many CAPM anomalies are not scalable and cannot be implemented in the large position sizes relevant to the Fund.

4. Recent theory and empirical evidence suggests that some fund managers do have talent and out-perform market benchmarks before fees. However, little of that superior ability

filters through to the ultimate investors in those funds with after-fee returns and alphas being, on average, zero or negative.

5. Most studies of manager persistence suggest that negative performance persists, but that it is difficult to achieve positive risk-adjusted returns solely by investing with past winning managers. In the hedge fund arena there is some evidence in favor of positive performance persistence and hence past track records may help identify winning hedge fund managers. However, this evidence must be qualified by questions over data quality, the indications of diminishing returns to scale, and the negative impact of fund inflows.

6. For many alternative asset classes, the quality and duration of the empirical data on assets and managers is not sufficient to draw firm conclusions about the potential for profitable active management. Despite recent dramatic trends in institutional investing, little is known empirically about the performance of alternative assets over long business cycles. It seems that some alternative asset class managers did well in the relatively calm capital markets of the two decades before 2008, but many alternative asset classes performed very poorly in 2008.

[This page is intentionally left blank]

Section II: Active Management of the Fund

In this section we evaluate NBIM's historical track record of active management. Our focus is based on quantitative analysis of data provided by the manager and meetings with NBIM personnel to discuss general strategy and philosophy of active management.[8] We analyze quantitative data on performance for the whole Fund, the two asset class divisions fixed income and equities, and internal and external active management within each asset class division.

We summarize our results as:

1. NBIM's philosophy of active management is based on taking a modest amount of risk in many different securities and strategies on top of a base, passively managed portfolio. The major comparative advantage of the Fund is that it does not have capital and liquidity constraints facing many investors and it can ultimately profit from positions in strategies that may produce short-term losses but in the long run yield significant risk-adjusted profits.

2. Active returns represent an extremely small fraction of both the mean and the volatility of overall returns at the fund, asset class, and internal fund level. Thus, the contribution of active management to overall returns of the fund is minor compared to the benchmark decisions. This makes the Fund effectively an index fund, even with active NBIM management.

3. Active management has not detracted from the total return of the fund since inception. The average active return from January 1998 to September 2009 generated by NBIM has been statistically indistinguishable from zero. Prior to December 2007 active returns generated significant value (alpha). This was reversed during the financial crisis and global economic recession during 2008 and 2009.

4. Most of the decline in active management performance post-2007 was in the fixed income area and was caused by the high exposure of active returns to systematic factors, especially liquidity and volatility, which had tremendous declines over this period.

5. A large fraction of active returns at all levels, both internal and external, is related to exposure to systematic factors, especially volatility and liquidity. At the overall Fund

[8] We have not been provided with specifics with respect to internal and external investment strategies beyond monthly figures on returns, benchmark performance, and fees for external funds.

level approximately 70% of active returns are explained by exposure to systematic factors. Exposure to these factors is appropriate for a long-term investor but this indicates that the minority of NBIM's active management trades represent truly independent investment strategies.

6. The amount of active risk since 1998 has been 0.25% per month, which is substantially below the limit of 0.43% per month in the fund's regulations. The tracking error limit was exceeded post-2007, but this was probably due to estimating forecasted variances with only historical data observed to date. A large part of the increase in tracking error post-2007 is due to increased volatility in the factor exposure of active risk.

7. Given the factor exposure of active returns, if the full distribution of factor returns had been known, i.e. how large the negative factor movements could potentially be, the scale of the losses in 2008 might have been within anticipated loss limits. This is true both using estimated factor exposure as of December 2007 for the whole fund and estimating factor exposures as data became available through 2008 for fixed income and equity. Put another way, if the factor exposures pre-2008 and during the financial crisis had been communicated and the asset owner had some knowledge of the potential drawdowns of these factors, the Fund's losses in 2008 would not have been surprising. There is evidence that NBIM successfully altered some factor exposures and avoided losses, particularly in the equities portfolio, that could have been more severe had the factor exposures remained constant in 2008 and 2009 at their December 2007 levels.

8. Although correlated with factors such as volatility and liquidity, active returns at the overall Fund and fixed income and equity sublevels have low correlation to common hedge fund strategy returns. This is an attractive feature as it indicates that active management strategies are largely unique.

9. Active management by external fund managers have large exposures to systematic factors suggesting that active external management has not reflected a large amount of unique management ability. The performance of external fixed income funds has, on average, been poor, but is positive, on average, for external equity funds. The overall fees paid to external managers are low and are the same order of magnitude of the expense ratios for the entire Fund.

10. A qualitative assessment suggests that NBIM seeks to exploit the scale and market power of the fund to add positive excess return. The effect of its strategy on Fund performance is constrained by the current risk budgeting process.

Our analysis of active management takes the benchmarks of the Fund, internal fixed income and equity, and external funds as given. The Fund benchmark is set by the Ministry of Finance and ultimately endorsed by Parliament. We do not evaluate any changes in the benchmark itself. The process by which these benchmark changes were executed is also a timing decision that reflected transactions and operational costs which we do not analyze. As an example, the increase in equity portion from 40% to 60% is not included in the analysis. While we provide a qualitative assessment of how external funds were selected in the past, due to lack of data on the actual selection process there is no quantitative assessment of this procedure. We do, however, offer a limited qualitative assessment based on meetings with NBIM personnel. We defer discussions on how the current structure of the fund relates to its active management capabilities and proposals for extending this structure in Section III. In that section we also discuss future possible benchmarks for the fund.

In this section our analysis is driven by data and thus is necessarily backward looking. The data period covers two management regimes and an interval over which NBIM's management structure and capabilities have developed significantly. This means that conclusions based on the past active performance of the fund may not be relevant going forward. Our study also embodies an "end-point" problem. That is, we have been asked to analyze the historical active management of the fund after a period of negative performance. Thus the sample "conditions" on this end point. A study of active performance at another end point may lead to different conclusions. To mitigate this bias we also comment on active management prior to the period when large active losses were first beginning to be realized.

Another limitation of a retrospective empirical analysis is that the models are fit using calculations that include all data known at the end of the analysis period. For example, ex-post estimates of factor exposures on a given date in the sample period are not the same as those that an analyst would have found using information up to that date only. We have indentified cases in this report where this might be an important qualification, and when necessary, we have performed analyses using only historically available data as of a given date.

The remainder of this section is organized as follows. We begin by summarizing NBIM's investment philosophy in Section IIA. We define active and benchmark returns in Section IIB. We also describe the systematic factors. Section IIC analyzes active management at the overall fund level. Section IID details active returns for the fixed income and equity divisions. We investigate internal NBIM active strategies in Section IIE and external active strategies in Section IIF. Section IIG provides a summary of the qualitative assessment of NBIM's performance. Section IIH concludes.

IIA: NBIM's Approach to Active Management

NBIM's active management style seeks to take advantage of patience, scale and market power, judicious manager selection, and low cost, all of which exploit the comparative advantages of the fund.

In this section we assess the strategic plans for active management, the risk budgeting process, internal and external reporting of active management risks and return, and whether or to what extent the specific strategies that have been used in NBIM's active management have been based on exploiting the characteristics of the Fund. This assessment is based on conversations with NBIM management as well as published reports.

IIA. 1 Overview of NBIM's Value Proposition

In our view NBIM is set up to provide two things to its client, the people and future generations of Norway. First, it offers "passive" returns based on the Ministry's choice of benchmark. The components of this benchmark are determined by the desire to efficiently diversify the overall portfolio. The benchmark takes the form of a portfolio allocation by region and asset class. It is intended to be a potentially tradable portfolio of securities, as opposed to a hypothetical benchmark. As such, it represents a true alternative to active management. NBIM seeks to offer this benchmark return in an operationally secure and transparent manner at lowest possible cost to the client.

Second, NBIM offers active management that seeks to provide positive, risk-adjusted returns over the benchmark net of active fee. NBIM pursues this goal through a combination of internal and external management, and a philosophy of outsourcing many aspects of its back-office operations. Thus, it has internal portfolio managers and traders as well as a number of external investment managers whom they engage via specific mandates tied to pre-determined benchmarks. We understand that these external mandates are separate account contracts that allow NBIM direct custody and control on a daily basis.

Although we have been asked specifically to examine the potential for active management of the fund, as we demonstrate below, the management of the overall benchmark, and the component parts of the benchmark are important determinants of NBIM performance and should be considered integral to the management of the fund. The benchmark is a dynamic portfolio that has changed over time. Major changes include the decision to increase the allocation to equities from 40% to 60% via a gradual shift during a transition period, decisions to include small-cap and emerging market stocks, and a decision to include real estate in the portfolio. Other changes have included decisions to increase allocations to various other sub-classes of assets such as mortgage-backed securities. While the benchmark can be thought of

as a passive portfolio, it represents an explicit strategy or program of investment that requires active rebalancing, efficient trading, identifying liquid versus illiquid instruments, balancing the feasibility of cross-border ownership, and other essential characteristics. The maintenance of these benchmarks cannot be achieved at zero cost, but rather is an important, ongoing function of the Fund. In this section we take the benchmarks as given. We discuss in Section III proposals for expanding the benchmarks.

IIA 2. NBIM Strategic Plans for Active Management

NBIM management has articulated an approach to asset management based on a theory called the Fundamental Law of Active Management, which posits that skilled managers are able to generate positive risk adjusted returns through successful forecasts, and that, all else equal, the more independent forecasts they make, the greater the likelihood of achieving positive portfolio alpha. The Fundamental Law is not the basis for NBIM's value proposition. Rather it is a framework by which NBIM assesses investment opportunity. This theory is not inconsistent with the EMH as currently articulated in our exposition of Section I. Positive alpha may accrue to managers with superior (costly) research technology and managers with longer investment horizons. This relative advantage by managers is called the Information Coefficient [IC]. The fundamental law of active management encourages the manager to increase the number of independent investment opportunities – a feature called "breadth." This approach thus justifies broad diversification by a fund manager across managers and strategies.

NBIM implements active management in two main ways; these were communicated to us in meetings with NBIM management. The first bases investment decisions on fundamental analysis of individual companies relative to peer companies in the same sector or industry and the second manages exposure to traditional risk factors. These approaches are employed for both internally and externally managed funds. Fundamental analysis and factor exposure operate quite differently: the former looks at investment opportunities from the "bottom-up" by finding attractive companies and securities and then aggregates to the portfolio level whereas the latter is a "top-down" investment technique that first chooses different factors and then implements them by trading securities. In the context of the fundamental law, these are the main methods by which NBIM tries to find investment opportunities with positive IC. Given NBIM's stated philosophy, it is natural to ask whether the asset manager has the capability of internally generating positive IC as well as identifying external managers with positive IC. The quantitative analysis below addresses this question.

IIA 3. The Risk Budgeting Process

The Fund's risk budgeting is based on a process by which the asset owner effectively contracts with its fund manager, NBIM, on a diversified portfolio representing an agreed upon benchmark

and then the fund manager is given boundaries with respect to the annual deviation of its return from the return of this portfolio, i.e. tracking error. We observe below that the risk boundaries themselves are not targets. Thus, there is no expected amount of deviation and the risk-budgeting process is, in fact, implicitly a system of risk limits. This ambiguity, in our view, may lead naturally to excessive risk aversion on the part of NBIM: without a target, it is difficult to take the manager to task for under-investing in active opportunities, but not the contrary. Under the current risk-budgeting process the manager can only be penalized for exceeding a risk budget but not for underutilization of active opportunities. Our analysis of the internal risk budgeting process of NBIM focuses chiefly on the relative risk characteristics of internal vs. external management.

The risk budgeting process as currently practiced is also sensitive to changing factor volatility. The tracking error is relative to a market portfolio benchmark and thus active returns are measured relative to only that market portfolio. If active management has exposure to other risk factors, then the volatility of these other risk factors is not directly captured in the tracking error. The dramatic increase in factor volatility during the crisis is a circumstance in which a manager constrained by tracking error bounds would necessarily reduce factor exposure. This may not be advisable given that harvesting risk premiums over the long run is predicated on maintaining such exposures. At the very least, it is important to consider the mechanical effects of tracking error bounds on the expected premium when the rules force disinvestment or the reduction of exposure to other risk factors with desired long-run exposure in highly volatile periods.

We do not make an assessment of the risk budgeting process for NBIM active management strategies since we were not provided individual security positions or trades of the fund. We also cannot quantitatively assess the process by which capital is allocated to internal strategies or how external strategies are selected, but we do make some remarks on the basis of conversations with NBIM senior personnel. We were not given the identities and background information on external managers or internal management teams. Our review does not encompass NBIM's business operations, back office capabilities, trading costs, compensation, performance evaluation procedures, risk controls, or risk budgeting. We discuss some aspects of information that should be reported for risk-return analysis and better risk management in Section III. In that section we also propose some changes to NBIM's management structure to better do efficient risk allocation.

Thus, our comments on risk budgeting in this report center on realized tracking error performance. We do not believe this is the best way to monitor active risk relative to a benchmark and we discuss alternatives in Section III.

IIA 4. Exploiting the Characteristics of the Fund: Comparative Advantages

Section III of this report focuses extensively on the distinct characteristics and comparative advantages of the Fund and how NBIM takes them into account. Thus, in this section we only briefly enumerate them. The review of the theory in Section I makes it clear that a pre-condition for positive risk adjusted returns in an active, efficient market is some comparative advantage in research or trading technology. The management of NBIM has likewise communicated to us a value proposition for generating positive risk adjusted returns. This is based upon:

1. Patience

   In principle, the long-term orientation of the fund allows it to be a patient investor. Thus NBIM should be able to wait longer for investment opportunities to mature. Some active management strategies have large exposure to factors that have long verification horizons, e.g. strategies based on contrarian investment during times of distress, investment based upon convergence of fixed income securities due to equilibrium arguments, and even strategies based upon the equity risk premium. In addition, it is conceivable that this patience gives NBIM access to top external managers who prefer investors who will not immediately withdraw money during liquidity crises of the sort recently experienced.

2. Scale and Market Power

   NBIM manages a large fund capable of taking major stakes in corporations, within the upper limit on 10% ownership share in listed equities. This allows the firm to take advantage of opportunities to exploit activist management to generate returns and positive externalities. NBIM demands full transparency and knowledge of their investment process from external managers. Scale theoretically allows NBIM to attract and incentivize internal managers with excellent experience.

3. Judicious Manager Selection

   NBIM believes it can hire external managers with superior techniques and skills (i.e. high information ratios). Their approach is to avoid “outliers” in the distribution of external managers, and instead invest in solid performers within leading firms. Presumably this same approach is applied to internal management as well.

4. Low Cost

   NBIM believes it has a significant advantage in trading that provides low transactions costs. This advantage accrues to the passive benchmark measure as well as the active performance, however. Thus it is only partially relevant to evaluating active returns.

In the remainder of this section, we assess the empirical results of the NBIM value proposition based on the enumerated comparative advantages.

IIB: Returns and Factors

The sample period of evaluation is characterized by major negative drawdowns in 2007 and 2008 in many systematic factor returns.

The data for this analysis was provided by NBIM and is at the monthly frequency from January 1998 to September 2009. All returns are in NOK and, at the overall Fund level, the fixed income and equity asset class level, and for internal NBIM strategies, are all gross of fees. We were provided with net-of-fee returns on external funds and also data on fees paid to external managers. Fund returns include income from security lending and the benchmark returns do not exclude transactions costs from phasing in cashflows and changes to the benchmark.

IIB 1. Active Returns and Alphas

We define an active return of a fund to be the return relative to the fund benchmark. We denote the fund return as *Ret*, the benchmark return as *Bmk*, and define the active return, *ActRet* as

$$ActRet_t = Ret_t - Bmk_t. \quad (1)$$

Our analysis characterizes active returns. We also analyze residual returns, *ResRet*, which are defined as:

$$ResRet_t = Ret_t - \beta Bmk_t, \quad (2)$$

where the fund beta, β, is estimated by a regression of the fund return on a constant and its benchmark. We refer to the alpha as the mean of the residual return, which is equivalent to the constant in that regression. At the overall Fund, asset class, and internal strategy levels, the active return and the residual return are largely identical because the fund betas are very close to one. Thus, we focus our discussion on active returns, but for completeness we report

residual returns, as they yield the more traditional measure of alpha. Table 1 contains definitions of the fund return, benchmark return, active return, and residual return.

It is common to define the active return and alpha using a multivariate regression with various factors as regressors. Indeed, most of the academic literature on performance measurement defines an alpha as the constant in a regression containing usually the market factor (the CAPM benchmark), the other Fama-French factors in the case of equity fund performance assessment, or other systematic factors in the case of non-equity asset classes. We have chosen not to do this for several reasons. First, and most importantly, the fund manager is asked to outperform a benchmark and so generates returns relative to that benchmark. The fund manager was not benchmarked against other, possibly arbitrary, factors.[9] We examine whether the active returns are correlated with various systematic factors, and if the active returns could have been predicted by the exposure to these factors, but we do not use these factors to define the alpha or active return.

We report partial correlations of active returns with respect to the risk factors, which control for the effect of other risk factors when computing the marginal relationship between the active returns and one risk factor.[10] Coefficients in a multivariate regression of active returns on factors are less meaningful because the constant in that regression cannot be interpreted as a tradable return. Factors can be non-traded, and if the traded factors are long-short portfolios they can be scaled to any volatility. This free scaling variable affects the magnitude of the factor loading coefficients.

Third, the choice of systematic factors is also far from settled in the academic literature. For example, there is long-standing debate about whether the ubiquitous Fama-French model represents risk or mispricing. In addition, systematic factors are generally correlated to each other and the use of different factors as independent variables often leads to very different conclusions on whether active returns have positive or zero means, or alphas that are positive or equal to zero. This ambiguity supports our decision to examine exposure to different factors only after the active return is defined relative to the fund benchmark.

Finally, by defining the active return relative to the Fund's benchmark as opposed to a general factor model, we can examine the proportion of return variance attributable to different return

[9] In Section III we address the question of whether NBIM benchmarks should be made broader than the current bond-stock asset class benchmarks.

[10] Formally the partial correlations are defined as follows. If the set of factors is partitioned into $F = \begin{bmatrix} F_1 & F_2 \end{bmatrix}$ with corresponding covariance matrix $\Sigma_F = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, then the partial covariance of F_1 controlling for F_2 is given by $\Sigma_{F_1|F_2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

components that correspond to different levels of decision making. In particular, the Fund's benchmark is set by the Ministry of Finance, whereas the active return is generated by NBIM, the fund manager.

IIB 2. Variance Attribution

The variance attribution computes the percentage of variance due to the benchmark and active management using the identity

$$Ret_t = Bmk_t + (Ret_t - Bmk_t) = Bmk_t + ActRet_t.$$

Taking variances of both sides we obtain

$$\begin{aligned} Ret_t &= Bmk_t + ActRet_t \\ \frac{\text{var}(Ret_t)}{\text{var}(Ret_t)} &= \frac{\text{var}(Bmk_t) + 2\,\text{cov}(Bmk_t, ActRet_t)}{\text{var}(Ret_t)} + \frac{\text{var}(ActRet_t)}{\text{var}(Ret_t)} \\ 100\% &= Attr(Bmk) + Attr(ActRet), \end{aligned} \tag{3}$$

where *Attr(Bmk)* is the percentage attribution of the variability of the fund return to the benchmark and *Attr(ActRet)* is the percentage attribution due to active management. The covariance term, which is generally very small, is related to benchmark timing and we include it in the benchmark attribution leaving the attribution to active return as the pure part of the return generated from active management.[11]

We refer to the reward-to-risk ratio for active returns or residual returns as

$$Reward\text{-}to\text{-}Risk\text{-}Ratio = \frac{E[ARet_t]}{Std[ARet_t]}, \tag{4}$$

where *ARet* is either the active return, *ActRet*, or the residual return, *ResRet*. The reward-to-risk ratio for the residual return is the same as the Information Ratio. If the benchmark is a risk-free rate then the reward-to-risk-ratio coincides with the Sharpe Ratio. For comparison, the Sharpe ratio on U.S. equities is around 0.1 at the monthly frequency but only over very long time periods. In fact, at the monthly frequency the Sharpe ratio on the U.S. equity market portfolio over our sample period is only 0.03.

[11] These covariance terms are positive and very small.

IIA 3. Systematic Factors

We measure the exposure of active returns to several systematic factors. Our choice of factors is affected by three considerations:

1. We select factors recognized as having significant pricing effects by the literature and industry practice.

2. The set of factors must be parsimonious.

3. The factors must have benchmarks that as far as possible can be traded in large size relevant to the Fund and exist since January 1998 (start of the Fund returns).

We emphasize that the first consideration does not mean that the factors have to represent risk factors arising from a rational pricing model. In fact, the Fund should not care whether the origin of any factor premiums it collects is rational or due to irrational market behavior, but only that the factor premium should be persistent and not be expected to disappear in the near future. The second consideration means that we select factors that are systematic at a global level where possible. A strong argument can be made for using macro-based factors taking inspiration from the Chen, Roll and Ross's (1986) early macro factor model. However, there are no standard benchmark portfolios to trade macro risk. Instead, the effects of macro factors such as inflation and economic growth are strongly reflected in fixed income prices and term spreads as shown by Ang and Piazzesi (2003), Ang Bekaert and Wei (2008), and many others. We thus include these fixed income-based variables as factors. Finally, the element of tradability influences the choice of portfolios we use as factor returns. Table 1 describes the systematic factors and their data source.

We use the following factors at a monthly frequency:[12]

TERM: Difference between long- and short-maturity U.S. Treasury bond returns.

CREDITAa: Difference between Aa and Treasury bond returns.

CREDITBaa: Difference between Baa and Aa bond returns.

CREDITHY: Difference between high yield and Baa bond returns

FXCARRY: Captures the carry trade of investing in currencies with high interest rates and shorting currencies with low interest rates.

LIQUIDITY: Reflects periods of high and low liquidity.

[12] We thank Kenneth French, Sergi Gorovyy, Antti Ilmanen, and Min Wei for providing data. Further comments on some of these factors are made in Section III.

VALGRTH: Difference in returns between "value" stocks and "growth" stocks.

SMLG: Difference in returns between small and large stocks.

MOM: Captures the momentum effect of going long U.S. stocks with past high returns and short socks with past low returns.

VOL: Captures differences between implied and realized volatility.

Panel A of Figure 1 plots cumulated returns on the fixed income factors. These cumulated returns are simply cumulative sums of the underlying factor returns expressed in percentages per month. While they do not represent the actual holding period returns on trading strategies involving the factors, the graphs give a straight-forward picture of the overall pattern of returns without the distortions of compound interest and portfolio rebalancing.[13] The most notable pattern over the sample is the large decline in all fixed income factors in 2008 and their rebound in 2009. FXCARRY starts to decline in mid-2007 and the credit and term factors experience negative returns in 2008, corresponding to the global financial crisis.

Our liquidity factor is based on the difference between on-the-run and off-the-run U.S. Treasury 10-year bonds. Most recently issued U.S. Treasury bonds ("on-the-run" or benchmark issues) are more expensive and trade at higher yields than previously issued more seasoned bonds ("off-the-run" or old issues). This phenomenon has been long known by both academics and practitioners (cf. Amihud and Mendelson, 1991; Krishnamurthy, 2002) and has been documented in other countries (e.g. Boudoukh and Whitelaw, 1993). The on-the-run/off-the-run effect involves instruments with different trading volumes but identical and, effectively, zero credit risk having different prices. Thus, it reflects pure liquidity risk. Liquidity factors constructed from other assets such as Pastor and Stambaugh (2003) and Korajczyk and Sadka (2008) confound the effect of liquidity and credit and are correlated with other cross-sectional factors. The on-the-run/off-the-run series we use is provided by the Federal Reserve Board and is originally constructed by Gurkaynak, Sack and Wright (2007). We define the LIQUIDITY factor to be innovations in the on-the-run/off-the-run bond spread with a negative sign, so that periods of illiquidity are low realizations of the LIQUIDITY factor. We emphasize that while LIQUIDITY is the only factor that is not investable, we can still ask if active returns are correlated with this factor and if realizations in LIQUIDITY can forecast active returns.

[13] We choose to plot cumulated returns rather than plot accumulated wealth resulting from an initial investment in the long-short strategies because the active returns cannot be interpreted as investable long-short trading positions; one cannot obtain just the active return as usually the benchmark cannot be shorted.

Panel B of Figure 1 plots the on-the-run/off-the-run bond spread and shows two large increases in liquidity risk. The spread first increases during the latter part of 1998 corresponding to the Russian default crisis and the failure of LTCM. The second dramatic increase post-2007 corresponds to the current financial crisis. The liquidity spread reaches a peak in January 2009 and shrinks during the last few months of the sample corresponding to some stabilization of the world financial system. Schwarz (2009), shows that liquidity risk can explain more than two-thirds of the increase in interest rate spreads during the early months of the financial crisis using the spread between German federal government bonds and KfW agency bonds. Since KfW bonds are guaranteed against default by the German government, the German federal government/KfW bond spread also represents liquidity risk and is similar to the on-the-run/off-the-run spread on U.S. Treasury bonds. At a daily frequency over a common sample in 2007-2008, the correlation between the US Treasury on-the-run/off-the-run spread and the German federal government/KfW spread is over 90%. This is evidence of a strong common component in global liquidity risk.[14] Note that Figure 1 shows that the large peak of the on-the-run/off-the-run spread is very pronounced in 2008 and it is hard to measure exposure to volatility risk without including the post-2007 period.

One caveat with LIQUIDITY is that although it certainly captures some dimension of the illiquidity characteristics of less-liquid assets, LIQUIDITY will not capture all effects or rewards to holding illiquid assets over very long horizons, such as venture capital, real estate, and other non-listed vehicles. However, given that such investments were extremely small, or zero, in the Fund's portfolio over the sample, this probably makes LIQUIDITY a very good factor to use in this analysis, but perhaps not as good a factor to capture the returns to the more illiquid investments being considered by the fund in the future. We comment on future factor selection in Section III.

The equity factors VALGRTH and SMLG are from MSCI. Most academic studies use the value factor HML and the size factor SMB of Fama and French (1993), but these are not tradable even for moderately-sized investors as Cremers, Petajisto and Zitzewitz (2008) show convincingly. The HML and SMB factors do not account for free float and contain large numbers of illiquid micro-cap and small-cap stocks making them nearly impossible to replicate for large investors.

[14] Another illustration of the global commonality of liquidity is that the correlation between the on-the-run/off-the-run spread and the Bank of England's aggregate liquidity index is very high in absolute value. The Bank of England Financial Stability Report (BoEFSR) compiles an aggregate liquidity index from eight underlying measures which include bid-ask spread measures, volume indicators, and interest rate spreads. Data on this liquidity series from the latest BoEFSR dated 26 June 2009 has a correlation of -87% with the on-the-run/off-the-run spread over January 1998 to May 2009, which illustrates the global nature of liquidity premiums.

In contrast, the VALGRTH and SMLG factors from MSCI are investable, at least by a medium-sized institution, and are drawn from a global stock universe.[15] Because the Fama-French factors are extensively used in academic, but not industry, evaluation, we repeat some of our analysis for the Fund with HML and SMB instead of VALGRTH and SMLG in Appendix A. The models of overall fitted active returns are close to each other using VALGRTH and SMLG instead of the Fama-French factors.

The two last equity factors are MOM and VOL. Panel C of Figure 1 shows these factors experienced dramatic reversals in the post-2007 period with MOM starting to trend downwards in March 2009 and VOL experiencing large negative shocks starting in September 2008. The MOM factor is constructed by Kenneth French and is based only on U.S. stocks. Unfortunately, there is no recognized industry benchmark to capture the momentum effect but Kenneth French's series is at least well recognized academically. Momentum effects exist in all global asset classes and these effects are highly correlated (cf. Asness, Moskowitz and Pedersen, 2008) so a tradable portfolio capturing momentum effects in a wide variety of asset classes would be correlated with MOM.

The VOL factor essentially captures an investor's returns from selling volatility. Volatility has a negative price of risk. Investors long volatility exposure receive payoffs during periods of pronounced market turmoil, such as 2008, and are willing to pay insurance premiums to receive this protection. Over the long run, investors paying for volatility protection lose money on these insurance premiums, which reflects the positive return to investors supplying this insurance. The VOL returns we use are similar to the returns from other volatility strategies such as selling (delta-hedged) out-of-the money put options or covered calls. Panel C of Figure 1 shows the remarkable turnaround in VOL returns in 2008, which happened suddenly and wiped out ten years of past cumulated returns. Prior to 2008 the returns to VOL were persistent and very smooth. Consequently, trying to gauge exposure to VOL risk is difficult with only pre-2008 data.[16] As we will show below, VOL turns out to be a very important factor in explaining the active returns to the fund.

Table 2 reports correlations of the systematic factors. Given the remarkable events over the sample period, it is unclear how stable these correlations are going forward, especially the pronounced negative correlations between TERM and the credit factors and the generally

[15] Since the fund's equity benchmark is FTSE a better choice would be the value and growth FTSE indices. However, these are not available at the start of the sample.

[16] It is noteworthy that the net premium to VOL is flat over the ten year period, even though, in equilibrium, it must be positive because the average investor is averse to risk. This suggests that the period incorporating recent turmoil is not indicative of long-term equilibrium and should not be used to assess the expected returns to taking factor risk. This is likely to be true for other factors in the model as well.

negative correlations MOM has with the other factors. LIQUIDITY has very high correlations with FXCARRY and VOL. In fact the highest correlation in the table is the 60% correlation between LIQUIDITY and VOL. Periods of illiquidity are also periods of pronounced high volatility.

Finally, we convert all factors, except LIQUIDITY which is not tradable, to NOK as the Fund returns are expressed in NOK.

IIB 4. Summary

The period post-January 1998 had two intervals of global financial turmoil which straddled an intermediate period of low volatility. The late 1990's experienced the technology NASDAQ bubble and the Russian default crisis while the global financial crisis began in 2007 with the widening of credit spreads and price declines in selected markets, such as subprime mortgages and certain types of asset securitized debt. The systematic factor returns reflect these extraordinary events and many factors, in particular credit factors, LIQUIDITY, and VOL, had very large negative realizations in 2008 and the early part of 2009. This makes measuring the exposure of active returns to these factors more difficult in the earlier part of the sample even if active returns are highly correlated with these negative factor realizations. The extraordinary behavior of some of these factors, even of the equity premium, during the last 11 years may also mean that this data sample may not be representative of long-run expectations of factor risk premiums going forward.

IIC: Overall Fund Active Returns

Although representing a very small proportion of total returns, active management has not detracted from the total return of the fund and prior to December 2007 significantly added value by providing large exposures to risk factors.

This section analyzes the active performance of the whole Fund. Figure 2 plots cumulated active returns, which are cumulated sums of the monthly active returns from the beginning of the sample. The cumulated active returns generally trend upwards strongly and start declining precipitously at the beginning of 2008. Thus, there seems to be a break around the beginning of 2008. It is partly these active losses that have led to the commissioning of this report.

We analyze data prior to and including December 2007, which we refer to as the pre-2008 sample, as well as over the whole sample. We refer to the period from January 2008 to the end of the sample as the post-2007 sample. Appendix B reports some analysis on test for structural breaks and regime shifts to justify formally our choice of the pre-2008 and post-2007 sample

periods. Mindful of the end-point problem we also report the results of our analysis "rolling" through the sample which does not need to pre-specify break dates.

IIC 1. Overall Fund Variance Attribution

In Table 3 we report summary statistics of active returns on the fund. Panel A shows that active returns constitute an extremely small fraction of overall returns with the variance attribution being 0.9% of all fund return variance over the full sample and an even lower proportion at 0.3% over the pre-2008 sample. These numbers indicate the active return proportion is much smaller than typical funds reported in the literature starting with Brinson, Hood and Beebower (1986), who report that benchmark asset allocation decisions typically explain over 80-90% of return variability over time. Thus, active management decisions have contributed relatively little to overall fund returns and the benchmark decision is by far the most important decision. We revisit this important issue in Section III.

One reason the variance attribution to active returns is so small is the relatively small 1.5% per annum tracking error limit prescribed by the Fund's regulations. Furthermore, NBIM has overall evidently used relatively little of this tracking error budget. Assuming that returns are independent and identically distributed (IID), the annual 1.5% tracking error limit translates to a tracking error of $1.5\%/\sqrt{12} = 0.43\%$ per month. Over the whole sample the realized tracking error per month, which is the standard deviation of active returns, is 0.25% and NBIM has taken only approximately one third of the tracking error variance. The pre-2008 tracking error is even lower at 0.12% per month. Figure 2 shows the increase in the volatility of active returns at the end of the sample and this accounts for most of the 0.25% monthly tracking error over the 11-year sample.

The fact that the entirety of the risk budget was not used by NBIM, particularly in the early part of the period, represents an active investment decision. This may reflect an understanding by NBIM that the active risk budget represented an inviolable limit, rather than an average, in which case the small contribution of active management to the Fund's performance is a function of the client/manager contract. In fact, the post-2007 tracking error is 0.59% per month, which is above the 0.43% per month limit. If the tracking error risk budget was viewed as a strict upper bound, then NBIM's policy of not using the full risk budget in taking exposure to different factor risk was a very good decision. This could be interpreted as NBIM recognizing that the factor return profile is significantly negatively skewed and could become suddenly volatile. Certainly if the full tracking error had been taken the losses in 2008 would have been much more severe.

NBIM's annual reports state that part of the large negative active returns post-2007 may be due to mismeasurement, or unexpectedly large realizations relative to model forecasts, of return

covariances and value-at-risk statistics. If the global financial crisis had not occurred, the tracking error probably would have remained the same, or at least consistent with intended deviations. There is thus no evidence that NBIM were exceeding their risk budget on the eve of the crisis.[17] The large increase in tracking error post-2007 also points to the limitations of using tracking error to constrain or monitor the extent of active management. As systematic risk factors in all markets increased dramatically, real-time measurement of the scale and source of tracking error was a challenge to all managers. It also reveals the limitations of using tracking error with fixed threshold or target values as the basis for controlling risk. The dramatic increase in factor volatility during the crisis is a circumstance in which we would advise flexibility and the use of a tracking error target or range rather than a limit. The effect of a fixed threshold is that active positions will likely be scaled down as market volatility increases. We discuss other ways of setting active return benchmarks in Section III.

IIC 2. Overall Fund Active Returns

Panel B of Table 3 reports statistics of active and residual returns. Betas are almost one so there is little difference between active and residual returns. Over the whole sample the active return had a mean of 0.02% per month and the p-value to test whether this is equal to zero is 0.56. Thus we cannot find any statistical evidence of significant active outperformance over the sample taken as a whole.[18] In contrast, prior to 2008 the mean active return was 0.03% per month and this is highly significant. Note that even though this is statistically significant, the absolute value of the active return added to the fund is small and thus the economic contribution is small. In both cases, the amount of active management added falls short of NBIM's stated expectation of 25 basis points in annual net value-added over three-year rolling periods through active management.[19]

There are at least two remarkable facts about the active or residual returns reported in Table 3, Panel B. First, the autocorrelation of active returns is very high at 0.55 over the full sample. The pre-2008 autocorrelation is much lower at 0.16 and Figure 2 shows the sharp, mean-reverting drop to zero after January 2008. This caused the increase in autocorrelation when including the

[17] This assumes, as NBIM states in its annual reports, that they monitored their risk budget with a Value at Risk (VaR) framework using historical data as inputs. Furthermore, this assumes that all instruments in which the Fund was invested, were adequately captured by the risk management systems. VaR are subject to probability assessments and all VaR measures are expected to be exceeded once in 100 months. Thus there is no evidence that the risk budgets were exceeded on an ex-ante basis.

[18] We use a standard 95% confidence level and thus we state a statistic is significantly different from zero if the p-value is less than 0.05. All our p-values are two-sided and are computed using robust standard errors by GMM.

[19] Cf. Norges Bank Organisation and Accounts in the NBIM Annual Report 2006.

post-2007 data. Almost all of this was due to the fixed income returns (see below), so we defer a more extensive discussion of this to Section IID. The high autocorrelation implies that regular OLS standard errors are inappropriate; we take into account the high serial correlation and heteroskedasticity using robust Generalized Methods of Moments (GMM) methods in computing all p-values and standard errors. Second, the negative skewness of -2.41 over the whole sample was also caused only by the large negative active returns in the 2008 period. As Figure 1 shows, this also coincided with large negatively skewed returns on the factors and, in fact, a large part of the skewness of the active returns was due to systematic factor exposure, which we now demonstrate.

In Figure 3 we graph monthly active returns in red dots and overlay fitted active returns computed from a regression of active returns on all systematic factors:

$$ActRet_t = c + \beta_1 \cdot TERM_t + \beta_2 \cdot CREDITAa_t + \beta_3 \cdot CREDITBaa_t + \beta_4 \cdot CREDITHY_t + \beta_5 \cdot FXCARRY_t + \beta_6 \cdot LIQUIDITY_t + \beta_7 \cdot VALGRTH_t + \beta_8 \cdot SMLG_t + \beta_9 \cdot MOM_t + \beta_{10} \cdot VOL_t + \varepsilon_t.$$

The fitted values tell us the performance we would have expected, given only our knowledge of the portfolio factor exposures, i.e., setting the residual, ε_t, to have mean zero. The solid blue line computes fitted values using a full-sample regression and the dashed green line shows fitted active returns computed using the coefficients in a regression using data up to December 2007. All fitted returns are computed using the realized factors at each date. The R^2 for the full sample regression is a very high 68% and the pre-2008 regression R^2 is 48%, indicating that most of the variation in active returns is explained by variation in the factors. Thus, most of the active strategies pursued by NBIM represent exposure to systematic factor risk

Figure 3 shows that the large negative active returns in 2008 were largely replicated with the factor exposures. Remarkably, Figure 3 also shows that this is true in the pre-2008 regression. Put another way, if the factor exposure of the active returns were estimated using data up to December 2007 and if investors had knowledge of the full distribution of factor returns, they would have been prepared for a certain probability of large drawdowns in factor returns.

This is quite different from saying that the negative factor realizations could have been perfectly forecasted in December 2007. Rather, given knowledge of the tail distributions in these factors, an investor standing at the beginning of 2008 would have correctly assessed the probability of active losses given the factor exposure of the active management strategies. Such an investor would have been more prepared for the negative active losses that did occur

and, given the factor loadings of the active returns, would not have thought them "impossible".[20]

IIC 3. Overall Fund Factor Exposure

We report the exposure of active returns to the various factors in Panel C of Table 3. We compute partial correlations, which measures the degree of association between the active returns and factors controlling for the other factors. We report partial correlations rather than raw correlations because the factors are correlated. Thus, these partial correlations measure, for example, the relation between active returns and volatility recognizing that volatility is correlated with liquidity. In Panel C there are significant exposures of active returns to CREDITAa, CREDITBaa, LIQUIDITY, VALGRTH, SMLG, and VOL over the full sample. The exposures to LIQUIDITY, VALGRTH, SMLG and VOL are also significant in the pre-2008 sample. The large decrease in LIQUIDITY and dramatic fall in VOL (see Figure 1) caused the fitted returns to closely match the large negative losses in 2008 and the rebound in active returns in 2009 in Figure 3. Thus, the post-2007 active losses were an expected consequence of the observed factor realizations. In Appendix C we explore rolling partial correlations of certain factors in more detail.

It is interesting to contrast the large increase in tracking error relative to the Fund's benchmark in 2008 with the Fund's fitted factor exposure plotted in Figure 3. The tracking error relative to the Fund benchmark in the pre-2008 period is 0.121% per month compared to 0.586% per month in the post-2007 period. The increase in tracking error variance is a very large 23 times. In contrast, the residual standard deviation of active returns after taking out the systematic factor exposure is 0.103% per month in the pre-2008 period compared to 0.295% per month in the post-2007 period. The increase in residual variance across the pre-2008 vs. post-2007 period is eight times, which is an order of magnitude smaller than the increase in tracking error variance. This implies that a very large fraction of the increase in tracking error in 2008 was due to the Fund's exposure to systematic factors.

Naturally, factor loadings could be re-estimated by the investor receiving new information on active returns through time rather than just at December 2007. In Figure 4 we graph fitted active returns using an expanding window, which updates the factor regression coefficients using information up until the previous month. The fitted active returns are computed using those regression coefficients with the current month's factor realizations. Again, the full-sample and the expanding window fitted values are very similar. Interestingly, the fitted expanding window predicts much more severe losses in September 2008 than actually occurred. One area

[20] This argument is similar to the conclusions drawn by Fung and Hsieh (2002) in their analysis of fixed income funds reviewed in Section I.

where active management can add value is by changing factor loadings. The evidence reported in Figure 4 is consistent with the hypothesis that NBIM successfully changed factor exposures by changing either internal or external investment strategies, and so avoided the larger predicted losses in September 2008.

In our opinion the factor exposure of the Fund had not been properly communicated prior to 2008. Had the factor loadings, especially to LIQUIDITY and VOL, been publicly announced at the beginning of 2008 and the asset owner been made aware of the potential losses in the return distributions of those factors, then the events in 2008 would have been within anticipated loss limits. An appropriate comparison is the large losses experienced by equities in 2008. An investor with a given portfolio weight in equities who had knowledge of the downside potential in the equities return distribution from past negative equity events would not have found the losses from equity investments in 2008 surprising. Similarly, if the asset owner had been aware of the factor loadings of the Fund prior to 2008 and known of the potential drawdowns in those factor returns, the active losses of the fund would have properly anticipated.

IIC 4. Overall Fund Hedge Fund Correlations

A final question we address in this section is whether the active returns on the Fund were similar to the returns generated by other sophisticated active managers. In Table 3, Panel D we report correlations of active returns with HFRI hedge fund indices converted to NOK. These are divided into relative value (RV), event driven (ED), equity hedge (EH), and macro strategies.[21] Overall the active returns have significant correlations with EH strategies focusing on quantitative directional trades and an overall EH index. These correlations are around 20% and so are not very high. Below, we show that these all result from equity exposure. The correlation of the Fund's active returns with RV fixed income convertible arbitrage is borderline significant but has a slightly higher correlation of 0.25 compared to the EH correlations. Overall there is not a strong relationship between the overall active management style of the Fund and certain hedge fund traders. This indicates that most of NBIM's active investment management style is unique and this is an attractive feature.

IIC 5. Overall Fund Summary

In summary, there is no strong statistical evidence that active management added value over the full sample relative to the fund's benchmark. Nor, it is important to point out, is there evidence that active management over the long term, detracted from benchmark returns. Prior

[21] Detailed descriptions of these strategies can be found at http://www.hedgefundresearch.com. We report raw correlations, not partial correlations as we do not want to control for other hedge fund strategies in computing the correlation of the active return with one hedge fund strategy.

to 2008, NBIM generated significant value in its active management. We do find, however, active returns were very highly correlated with systematic factors. This is a positive finding, since these factors are appropriate long-term sources of risk premiums for the Fund. Although they were the cause of the severe decline in the crisis period, they were also the source or positive excess return in non-crisis periods. An issue discussed later in the report is the question of whether these exposures should be accounted for in the benchmarks themselves. An implication of the large factor exposures of the Fund is that only a minority of active management taken by NBIM represents truly independent bets relative to systematic risk factors.

The factor exposures also provide us with insight about risk assessment. Given the high exposures of active returns to certain risk factors, notably LIQUIDITY and VOL, had the full distribution of these factor returns been anticipated, the active losses post-2007 would have been within anticipated loss limits.[22] An appropriate analogy is that the asset owner was prepared for certain losses for a 60% equity weighting and losses of this magnitude occurred in 2008. Had the risk factor exposure pre-2008 been known and the asset owner been prepared for negative events in those factor returns, then the active losses in 2008 would not have been surprising. This conclusion is valid even when estimating factor loadings using pre-2008 data. The difference between ex-post and ex-ante factor exposure calculations suggests that NBIM succeeded in actively managing the portfolio factor exposure to avoid losses worse than anticipated by its factor loadings when the factor loadings are re-estimated through the sample using data observable at each date.

IID: Fixed Income and Equity Active Returns

There is no statistical evidence of active outperformance for fixed income but convincing evidence that active management in equity added value.

Panel A of Figure 5 graphs cumulated active returns of the fixed income asset class part of the fund. The downturn in fixed income active returns began in mid-2007, coinciding with accelerating declines in the sub-prime and certain other asset-backed securitized markets. This was about six months earlier than when the overall Fund started experiencing large losses (see Figure 2). The negative active returns persisted from mid-2007 to the end of 2008 and started rebounding in 2009. In contrast, active returns on the equity portion of the fund as shown in Panel B of Figure 5 do not share this marked downturn. There was a slight decline in active

[22] There is little evidence that any market had the foresight to calibrate VaR models using data from the 1930's. As noted above, Fung and Hseih (2002) were an exception.

equity returns in 2008 but it was much smaller than the negative performance of the fixed income active returns.

IID 1. Fixed Income and Equity Variance Attribution

In Table 4 we report various statistics on the active management of the fixed income and equity business units. Panel A shows that the variance decompositions of the total fixed income and equity returns are almost entirely due to benchmark decisions, with the active portion of these units accounting for 0.2% to 0.3% of the fund returns prior to 2008. Even with the dramatic downturn in fixed income active returns, the variance attributable to active returns is still only 2.9% in the full sample. Again this points to the fact that the overall biggest decision is the benchmark choice and the portion of the return due to active management is extremely small.

IID 2. Fixed Income Active Returns

Panel B of Table 4 lists statistics on fixed income active and residual returns. Neither is statistically distinguishable from zero in the full sample or prior to 2008, with p-values of mean active returns being 0.98 and 0.45, respectively.[23] The high autocorrelation of the overall Fund's active returns over the full sample reported in Table 3, which is 0.55, is due mostly to the high autocorrelation of fixed income active returns, which is an extremely high 0.72 in Panel B, Table 4. High autocorrelations in active returns are unusual because the autocorrelations of most asset returns are close to zero. Getmansky, Lo and Makarov (2004) argue that high autocorrelations are an outcome of exposure to illiquidity risk and non-marking to market of positions. Bollen and Pool (2008) argue that high autocorrelations are also possible indication of intentional return smoothing.[24] Note that the fixed income active return autocorrelations are much lower in the pre-2008 sample, at 0.11, compared to 0.72 over the full sample. This is consistent with the high fixed income autocorrelation arising from smoothing in the values assigned to illiquid fixed income instruments in the post-2007 period during the financial crisis. We discuss other aspects of the high fixed income active return autocorrelations in Appendix D.

One might expect that given the dramatic drawdown in fixed income active returns coinciding with the financial crisis, the skewness of fixed income active returns would be much more negative in the post-2007 sample. This is not the case, with full sample skewness and pre-2008 skewness being -0.80 and -2.51, respectively. There are three reasons for this. First, fixed

[23] Figure 5 shows that fixed income active returns start to decline in the latter half of 2007. Prior to June 2007 fixed income active returns have a mean of 1.48 basis points (0.0148%) per month with a robust p-value of 0.11 and so we cannot reject the hypothesis that mean active fixed income active returns are equal to zero.

[24] Since NBIM has full daily transparency and is able to monitor marks, the monthly autocorrelation of active returns may be a topic it may wish to investigate.

income returns start declining in mid-2007 so there was already some decline before January 2008. Second, and more importantly, fixed income active returns were also relatively volatile at the start of the sample. In 1998 Russia defaulted, which caused turmoil in fixed income markets and led to the failure of LTCM. These events show up in the relatively volatile returns in 1998-1999. However, it is important to note that these events simply affected the dynamics of fixed income prices – we examine active returns with the benchmark removed – so the active management volatility in 2008-2009 is presaged, but is much more severe, than the events in 1998-1999. Finally, the pronounced rebound in fixed income returns means there are large positive returns to balance out some of the large negative active losses post-2007.

IID 3. Equity Active Returns

In contrast to the fixed income performance the second part of Panel B, Table 4 shows that equity active returns were positive over both the full sample, at 0.05% per month, and over the pre-2008 sample, at 0.06% per month. The pre-2008 number is also statistically significant with a p-value of 0.02 and even the full sample significance level is close to 90%. Despite strong active performance, the numbers are very small in absolute terms and are overwhelmed by the overall variation of the benchmark. Generally, the amount of active management, in particular the successful equity management over the sample, would have to be orders of magnitude larger in order to make a meaningful difference to the total returns of the Fund. Also in contrast to fixed income active management, equity active returns exhibit relatively little autocorrelation or skewness even both the full and pre -2008 samples. In fact, the active management in equity clearly added value over the 2008-2009 period. This is remarkable given the turbulence in world equity markets at this time, and suggests that NBIM's equity management team at the time had demonstrable capability to adjust to shifting macroeconomic regimes.

IID 4. Fixed Income Factor Exposure

In Panel C, Table 4 we report partial correlations of active fixed income and equity returns with selected systematic factors. We select fixed income systematic factors for the fixed income analysis and equity factors for the equity data. We also include liquidity as a fixed income factor as liquidity has long been known to affect the prices of fixed income securities (see e.g. Chen, Lesmond and Wei, 2007) but do not include it for equity returns as the equity universe is restricted to listed equity markets which are always tradable, albeit with time-varying transactions costs that may also reflect both default risk, macro risk, and other cross-sectional equity factor risks. Although strictly speaking VOL is formed from equity (derivatives) prices, we include it as a fixed income factor. Volatility increases simultaneously across many asset markets. Collin-Dufresne, Golstein and Martin (2001), Pan and Singleton (2008), and other

authors show that volatility has been shown to be a driver of the pricing of many fixed income securities.

Panel C, Table 4 shows that fixed income active returns have very significant exposures to credit risk as well as TERM, LIQUIDITY, and VOL. The partial correlation with respect to LIQUIDITY and VOL are relatively large at 0.35 and 0.37, respectively. Exposure to these factors contributed the most to both the decline in active returns over 2008 and the subsequent rebound in 2009. The negative TERM partial correlation indicates that the Fund had an underweight position in long-term bonds. The positive CREDITAa exposure implies that most credit risk was taken in investment-grade bonds. The negative CREDITBaa partial correlation can be interpreted as the Fund having a relative underweight position in lesser investment-grade bonds compared to bonds with Aa ratings and better. The low and insignificant partial correlation with CREDITHY is consistent with non-investment grade bonds not being part of the Fund's investment universe.

IID 5. Equity Factor Exposure

In Panel C, Table 4, the equity active returns have highly significant partial correlations to all the equity systematic factors with all p-values well below 0.05 except MOM. The large absolute value partial correlation with respect to VALGRTH at -0.56 and SMLG at 0.41 indicates the active equity tilts are towards growth stocks and small stocks within the MSCI universe. There is also a highly significant partial correlation of 0.28 with VOL. Thus, both the fixed income and equity active returns have had large volatility exposure indicating that the fund has been – in effect – providing, or selling, volatility protection. In 2008 systematic exposure to this strategy, which performed very poorly in the second-half of 2008, accounts for a large part of the losses in active management, as we show next.

IID 6. Fixed Income Fitted Active Returns

Figure 6 graphs actual and fitted returns for the fixed income equity area. Panel A shows realized active returns in red dots, fitted active returns from a full sample regression in the blue solid line, and fitted returns from a regression using pre-2008 data in the dashed green line. There are two notable results from this plot. First, between 1999 and June 2007 realized active returns were very smooth. Over this period the standard deviation of active returns was 0.04% per month, which is extremely small compared to the standard deviation of 0.38% per month over the whole sample and the maximum permitted tracking error of 0.43% per month. Pre-2008, the standard deviation of fixed income active returns was 0.10%. The very low volatility of fixed income active returns in the middle of the sample causes the fitted returns using the regression over the whole sample to have greater variation than the realized returns over this period. This is consistent with the factor partial correlations in Table 4, as the large correlations with respect to LIQUIDITY and VOL imply that the active returns would be much more volatile

during the early and late periods of the sample where LIQUIDITY and VOL realizations were large and negative. There is less volatility of the fitted returns using the regression estimated with pre-2008 data.

Second, the R^2 of the whole sample regression is high at 0.54 compared to the R^2 of the pre-2008 regression at 0.18.[25] This is not surprising. There are only two periods to identify the effects of LIQUIDITY and VOL and the same two periods are important also for identifying exposure to TERM, CREDITAa and CREDITBaa: the fixed income turmoil in the late 1990's and the financial crisis post-2007. Factor payoffs are large, negative, and volatile during these periods. This makes exposures to these factors difficult to identify on an ex-ante basis but easy to pick up ex post from the whole sample. The fitted regression using pre-2008 data has a much lower R^2 for this reason and also displays little variation post-2007. Thus, without observing the post-2007 period and not knowing these tail distributions, the fixed income active returns after January 2008 could not have been foreseen using data only up to December 2007.

In Panel B of Figure 6 we estimate the fitted regression with an expanding window using data up to the previous month and compute the fitted active return using the factor realizations at the end of the current month. When the factor loadings are estimated through the sample, we can predict a large part of the negative active returns in September 2008. This fitted value is close to -2%, which is the lowest active return in the sample. The regression is less successful in fitting the neighboring months, August and October 2008. However, as already stated, conditional on knowing the distributions of the factors, the probability of a large part of the loss in 2008 could have been correctly assessed from the factor exposure of the active strategies computed through the sample. Thus, if the Fund's desired active strategies encompassed LIQUIDITY and VOL exposures, an investor who understood that the distribution of these factors included the possibility of large tail losses would have also understood the possibility of losses in the Fund on the scale of those actually realized in 2008. Put another way, if one could have known that the fund was loading on LIQUIDITY and VOL, then the active losses in 2008

[25] Note that the R^2 of the regression of active management returns at the overall Fund level on factors is 68%, which is the main variable of interest for overall performance. In our specifications, the systematic factors explain 54% of active fixed income returns and 29% of active equity returns (see below). These are highly significant but the proportion of variance explained at the asset class level is lower than for the overall Fund. This is due to several factors. First, none of the active return regressions by asset class uses the entire set of factors. In explaining the fixed income and equity active returns we used only bond factors and equity factors, respectively and naturally the proportion of variance explained will be lower for the smaller number of regressors. Second, the aggregate portfolio is not perfectly explained by the NAV-weighted asset class excess portfolios due to unobserved management overlays. Third, the estimation errors in the regressions using the asset class active returns are reduced through aggregation at the overall Fund level.

would have been recognized as being commensurate with the amount of the factor exposures, where the factor loadings were updated on a rolling basis.

In both the full sample regressions (Panel A) and the rolling regressions (Panel B) of Figure 6, the recent rebound in active returns in 2009 is greater than the fitted factor returns. This implies that the Fund's recent active performance has been better than predicted by the rebound in LIQUIDITY and VOL. This has also contributed to the overall Fund active returns (see Figures 3 and 4) being greater than predicted by the improvement in market liquidity and the stabilization of the global financial sector.

IID 7. Equity Fitted Active Returns

In Figure 7 we plot the corresponding graphs for equity active returns. The high partial correlations to the equity factors reported in Table 4 lead to reasonably high R^2s of the fitted regressions at 0.29 for the whole sample and 0.34 for the pre-2008 sample. An immediate consequence is that the active strategies pursued by NBIM's equity unit have been less based on factor exposures than the fixed income division, but the equity factor exposure is still highly significant and economically large.

The two panels of Figure 7 show there is little difference between the full-sample fitted values and the fitted values from the pre-2008 regression, except for a large discrepancy between the two fitted values in October 2008. During October 2008 VOL endured an extreme negative return (see Figure 1) and the equity active returns have very high exposure to this factor, as reported in Table 4. Interestingly, the fitted pre-2008 regression in Panel A, and the fitted expanding window regression in Panel B, both predict significantly worse drops in October 2008 compared to the full-sample regression. This is consistent with a hypothesis that NBIM successfully changing its active strategies in such a way as to have less VOL exposure prior to this period to avoid a calamitous outcome. Said differently, a passive strategy of holding the factor weights fixed at their pre-2008 levels and allowing for no discretion fared much worse during 2008 than the active strategy undertaken by NBIM during that time.

IID 8. Fixed Income and Equity Hedge Fund Correlations

In Panel D of Table 4 we report correlations of fixed income and equity active returns with several hedge fund strategies. Although the relative value strategies are more likely to be mimicked by active fixed income returns and the active equity strategies are more likely to be correlated with event driven and equity hedge strategies, we report all hedge fund strategies for both fixed income and equity returns as some hedge fund strategies straddle both asset classes. Over the full sample, fixed income active returns have no significant correlations with any hedge fund strategies. The only significant correlations for equity active returns are for EH quantitative directional and the EH index, most of which comes from the internal equity sector-

rotation strategy, which we examine below. There are some p-values just above 0.05 for RV fixed income arbitrage and the ED index, however. Overall the individual fixed income and equity active returns do not overlap with most hedge fund strategies.

IID 9. Fixed Income and Equity Summary

Based only on testing whether mean active returns are equal to zero, there is no statistical evidence of active outperformance for fixed income, but there is convincing evidence that active management in equity added value. The active returns had almost no economic impact on overall fixed income or equity performance as active returns constituted a tiny fraction of total returns. This is consistent with active management playing a minor role at the overall Fund level. Ex post there were high exposures of both fixed income and equity active returns to systematic factors, especially LIQUIDITY and VOL for fixed income and all equity factors for the equity business line. The proportion of returns explained by the factor exposures was much larger in fixed income than in equity. These large factor exposures imply that the large, active losses in 2008 for fixed income could partially have been anticipated, if the full distribution of factor shocks had been known and if the exposures had been updated as data became available through the sample. Part of active management involves changing the weights on different active strategies that result in different weights to systematic factors. There is evidence that equity active management avoided a much worse outcome during 2008 than following a strategy of holding the factor weights fixed at their pre-2008 levels.

IIE: Internal (NBIM) Active Returns

Internal (NBIM) active returns have added value in both fixed income and equities.

The internal funds management at NBIM in both the fixed income and equity business units involve an enhanced index fund and alpha satellite vehicles. We could not evaluate any internal alpha strategies on a disaggregated basis as no individual strategy returns data were provided and consequently this report contains few details about them. We were also not given a breakdown of capital allocations between the index mandates and the other active strategies. We also make no comments on the extent of how internal vs. external active strategies are decided as we were not provided with data or other information on the allocation process. Our inferences about the capabilities of NBIM to manage alpha-generating strategies are based largely on the historical aggregated performance provided and general conversations about investment philosophy with senior NBIM management.

IIE 1. Internal Fixed Income Returns

We were given internal fixed income returns that combine both an enhanced index mandate and alpha-generating strategies.

Not surprisingly, the active performance of the internal fixed income funds largely mirrors that of the overall fixed income unit as most of the funds are managed internally (see Section IIF for comments on externally managed funds). For example, the plots of internal fixed income active returns in Figure 8 look nearly identical to the corresponding pictures for overall fixed income active returns in Figure 6. Given this similarity we only comment in this section on aspects of internal fixed income active performance that are qualitatively different from the overall fixed income active returns. However, for completeness, we include in Table 5 and Figure 8 the same analysis as in Table 4 and Figure 6 in Section IID.

Table 5 reports statistics on internal active returns. In Panel B, an important difference with overall fixed income performance in Table 4 is that internal active returns are much less skewed. Internal fixed income active returns have a positive skewness of 0.22 compared to a negative skewness of -0.80 for the whole fixed income unit. Thus, the external mandates contributed most to the large overall fixed income active return negative skewness, which was manifested in large negative returns post-2007. Note that the low mean active returns cannot be interpreted as unsuccessful internal active management as the proportion of internal active strategies relative to the passive index strategies is unknown. Presumably most of the internal fixed income returns represent enhanced index performance. If the proportion of index versus active is constant across the pre-2008 and post-2007 periods, then the small difference in the mean active returns pre- and post-2007 in Panel B, Table 5 indicates that the internal active fixed income strategies have been quite successful. This also implies that the large fixed income losses have been mostly concentrated in externally managed funds; we examine this further in Section IIF.

Panel C, Table 5 reports another difference with the overall fixed income active returns. In Panel C, Table 4, the partial correlation of the fixed income returns with TERM is -0.21 and highly significant. For internal fixed income returns the partial correlation with TERM in Panel C, Table 5 is close to zero and statistically insignificant. We confirm below that the short TERM exposure for the overall fund is caused by very significant negative partial correlations for the external funds.

In summary, internal fixed income active returns are very similar to overall fixed income returns because most fixed income funds are internally managed.

IIE 2. Internal Equity Returns Variance Distribution

We comment on internally generated equity returns which combine an enhanced indexing strategy, AS, and other long-short active strategies.[26] We report statistics on these active return strategies in Table 6. Panel A reports the variance attribution of the AS strategy.[27] Not surprisingly since AS is an enhanced index fund, the benchmark accounts for almost all return variation with variance attributions of the benchmark being rounded to 100.0% over the full sample and 100.0% pre-2008.

IIE 3. Internal Equity Active Returns

In Panel B, Table 6 we report the performance of the active returns of internal equity and the AS index fund. Internal equity active returns are positive in both the full sample and pre-2008 at 0.05% and 0.09% per month, respectively. The pre-2008 performance is highly statistically significant. The pronounced negative skewness of -3.35 for internal equity returns over the full sample is due entirely to the active equity losses post-2007. A similar picture holds for the AS active returns, which is consistent with most of the internal performance constituting index performance. The good track record of the internal equity strategies suggests that, if scaled up, they could significantly enhance the Fund's total performance.

IIE 3. Internal Equity Factor Exposure

We report partial correlations of internal equity and AS active returns to the equity systematic factors in Panel C of Table 6. Internal equity returns exhibit significant partial correlations to all systematic equity factors. NBIM's active equity returns are overweight growth stocks, indicated by the -0.33 partial correlation with VALGRTH, overweight small stocks, as shown by the 0.26 partial correlation with SMLG, underweight past winners with a -0.25 loading on MOM, and the positive VOL exposure of 0.26 indicates internal equity returns, like the overall Fund, are long volatility exposure. The enhanced index AS returns have the same signs as the overall equity active returns, except the SMLG and MOM partial correlations are indistinguishable from zero. The partial correlation with respect to VOL is particularly high at 0.53.

Selling volatility protection is a strategy that economically over the long run should provide a positive return but, as Figure 1 shows, can experience pronounced short-term losses. The large

[26] NOK amounts for a few other internal equity strategies were provided, but no notional capital or risk budget allocations were made available. We did not analyze these returns because we cannot translate them to returns without further information, particularly without information on their mandate, risk profile, or background.

[27] We were given only relative internal equity returns in excess of benchmark and so cannot compute a variance decomposition for AS. Similarly, we can only compute residual returns for AS as we were only given a benchmark for AS.

VOL exposures indicate that the selling of volatility insurance was not restricted to the fixed income portion of the Fund. As we now show, the amount of VOL exposure has been more successfully managed in the equity portion of the fund than in fixed income.

IIE 4. Fitted Internal Equity Active Returns

In Figure 9 we plot actual returns of the internal equity and AS active returns along with fitted values from a regression with the equity systematic factors. We graph the fitted values from a regression using the whole sample in the solid line and estimate a rolling regression through the sample in the dashed line. For the latter we compute the fitted active return using the factor realizations at the end of the current month with the coefficients estimated at the end of the previous month.

We plot fitted values from full sample and rolling regressions for internal equity active returns in Panel A. The graph shows there was a large negative return of -1.48% in September 2008. During this month, the volatility factor experienced a large decline, after many years of steady gains of -19%. The rolling regression estimated up to September 2008 implies large positive loadings to VOL. During the next month, October 2008, the fitted return from the rolling regression is a very large -3.91%, so large that we do not plot it on the scale of Panel A, Figure 9. The VOL factor performs extremely badly in October 2008 with a return of -56%, which causes the very large negative fitted return in this month. Changing this factor exposure very quickly to avoid this loss can be interpreted as evidence of superior factor timing ability.

Nevertheless, it is possible that the rolling regression estimates factor loadings with large error and that NBIM had not changed any factor exposure during the post-2007 period. This is unlikely given our conversations with NBIM and the fact that in Panel A, Figure 9, the actual return in October 2008 was better than the ex-post fitted return estimated over the whole sample (shown in the solid blue line), which has the benefit of look-back bias.[28]

IIE 5. Internal Equity Active Hedge Fund Correlations

We report correlations of the internal equity active returns with hedge fund strategies in Panel D of Table 6. Like the overall Fund returns (see Table 3) and the equity asset class returns (see Table 4), the internal equity returns are positively correlated with RV fixed income convertible arbitrage, the EH quantitative directional strategy, and the overall EH index. Thus, the internal equity strategies are mostly responsible for these hedge fund correlations. In contrast, the AS strategies are not significantly correlated with any hedge fund returns

[28] See also Appendix C.

IIE 6. Summary of Internal Active Returns

Internal fixed income returns are very similar to overall fixed income returns and have not detracted from performance over the sample period. The internal fixed income returns have large factor exposure, especially to credit, LIQUIDITY, and VOL factors.

The internal equity management has been more successful than internal fixed income management. Unfortunately due to the small active risk budget, this good performance in internal equity strategies did not make a large impact on the equity unit as a whole and certainly not on the overall Fund. There is also evidence that NBIM managed the factor exposure of the internal equity strategies to avoid a much larger loss during October 2008 than what was realized. This is evidence of skillful active management.

IIF: External Funds

External managers have, on average, large exposures to systematic factors and this indicates they possess, on average, only a small amount of unique management ability.

IIF 1. The Number of External Funds

Over the sample period there are a total of 49 external mandates in fixed income and 166 external mandates in equities. Figure 10 plots the number of external mandates for fixed income in Panel A and for equities in Panel B in the solid blue line on the LHS scale. The fraction of the fixed income and equity portfolios is displayed in the dashed red line on with the RHS scale and averages just below 10% for fixed income and 30% for equities. The number of external fixed income accounts reached a peak of nearly 40 just before 2008. Partially as a response to the poor fixed income active performance post-2007 NBIM terminated many of these mandates. In contrast, the number of equity external mandates continued to grow post-2007 and had only a modest reduction in 2009.

External managers presumably allow NBIM to access expertise not available within its organization. NBIM selects these managers and awards them a mandate benchmark. The external manager's task is to outperform that benchmark with low tracking error. Thus, there are two levels of decision making involved: (1) the selection of the mandate and the external manager and (2) the portfolio decisions by the external manager and the consequent value added relative to the mandate benchmark. These two normally independent decisions may at times by confounded, particularly in circumstances in which the assets are unique, or in cases in which NBIM must seek advice from the external manager about which assets to use in the benchmark.

A relevant question is why there are so many external managers for NBIM. While external managers give NBIM access to strategies and talent that is not available in-house, a large number of external managers increases operational risk and the additional delegation brings with it increased complexity in risk management as well as another layer of incentive problems. Indeed, from the literature covering flows into managed funds in Section I, the Berk-Green model predicts that funds receiving large infusions of cash will experience a decline in active returns. Given the scale of the Fund, this theory may be relevant to the decision to engage active external managers.

Our conversations with NBIM have given us a better understanding of the process by which external managers are chosen. Their aim is to identify an individual rather than an organization and NBIM therefore makes contact with established asset management companies to try to identify "up-and-coming" managers within the organization. In terms of the Fundamental Law (see Section IIA), the manager's skill is represented by the information coefficient, IC. NBIM says that they are not necessarily looking for someone in the top quartile of IC skill but, perhaps, someone in the second quartile. NBIM explain that they do this because they are not trying to find a manager to add 500 basis points per year but around 60 basis points. Thus, our understanding for the rationale for multiple external managers is the desire of NBIM to seek breadth of independent positive information-ratio strategies. By extending beyond their internal capability, they hope to access returns relatively uncorrelated with their own.

They also cite the dynamic process of learning about new investment technologies as another potential reason for using external managers. For example, when mortgage-backed securities [MBS] were added to the benchmark, it made sense to add one or more managers with specific expertise in MBS. Thus, expansion of the set of assets in the benchmark is also likely to expand the manager pool. We make further comments on the use of external funds in Section III.

It is important to note the following observations in our analysis of external funds. First, NBIM counts some additional internal overlay, sources of funding, and other portfolios in its official external business reporting units. We purposely exclude these as we want to measure active returns only by external active managers. Second, beginning in December 2007, fixed income external funds underwent a large restructuring with numerous terminations bringing externally managed funds in-house. While there is an argument for looking only at fixed income external funds prior to December 2007 we include the post-2007 period. Fund performance after this period still represents investment decisions that were made, at some prior point, by external managers and these are still separate from the main internal portfolio. Thus, they are still not internal portfolios in the sense of Section IIC. After the restructuring the funds still have benchmarks and while they may not be active in the same way as they were prior to December 2007, they still have benchmarks that are different from the main fixed income benchmark.

Finally, the decision to terminate a fund, just like the decision to start an external mandate, is also an active decision and how these funds are treated after their termination is a continuation of that active decision.

We also note that although some fixed income markets like subprime mortgage and other asset-backed securitized markets started to decline in 2007, the worst episodes of the financial crisis occurred in 2008. Many external fixed income funds performed extremely badly over this time with one-month active returns reaching -15% for some funds, several multiples greater than the standard deviation of the active returns of these funds prior to 2007. These returns were probably unexpected and should be included in the evaluation of external fixed income funds.

In the next section we analyze the active returns of external managers, i.e., their performance relative to their mandate benchmarks. As described earlier, the choice of a mandate benchmark by NBIM is also an important decision but we have insufficient information about the nature of the mandates awarded by NBIM, or the way in which these are decided to carry out any formal analysis on the mandate selection process.

From conversations with NBIM we understand that the mandates typically represent sub-portfolios of the main equity or fixed income benchmarks, for example, a particular region or sector in the equity benchmark or a particular asset sub-class in fixed income. The mandates could be chosen in one of two ways. First, they may simply represent a desegregation of the overall benchmark. In this case, there is nothing "active" about the choice of particular benchmarks; it is simply an organizing mechanism for delegating the management of the total portfolio. Alternatively, the internal and external mandates may aggregate to a portfolio that does not correspond to the overall benchmark and, in this case, the mandate decisions themselves represent active choices by NBIM. Our understanding is that the choice of mandates is, in at least some cases, more than a simple desegregation but we know little more than this. In particular, we have no way of identifying a "benchmark for the mandate benchmark" and so cannot comment on whether or not the active component of the mandate choice adds value.

IIF 2. External Active Returns

In this section we examine the external active return, which is the external fund return relative to the mandate benchmark. We treat each external fund with equal, not NAV, weights in our analysis because each separate fund is making its own decisions. Of course, the contribution of the external funds to the overall Fund return involves a decision about the capital committed to that external fund. NBIM may have added value by the decision to weight some external managers more than others. The amount and timing of the capital are decisions made by NBIM

and are not analyzed here as we focus only on the active management of the external funds relative to the mandates given to the funds.

We start by reporting statistics of the active returns across funds in Table 7. Panel A reports the average and median active return across the fixed income external funds are negative at -0.45% and -0.06% per month, respectively. External active fixed income management has added, on average, negative value. The average skewness is also negative at -0.89 which is caused by large drawdowns in 2008 as we report below. The average active management for equity external funds has been more successful than for fixed income, providing a mean return of 0.09% per month and a median return of 0.05% per month.

Figure 11 succinctly summarizes the mean active returns with robust 2 standard deviation bounds for fixed income (Panel A) and equity (Panel B) external funds. The funds are ordered in increasing mean active returns. The majority of fixed income external funds underperformed their benchmarks with less than a third of external funds generating positive active returns. The performance of equity external funds is much better, with approximately two-thirds of funds beating their benchmarks on average. Note that in most cases the standard errors are wide and even for the better-performing equity funds it is hard to reject the hypothesis that a typical external fund adds no value beyond the benchmark.

Naturally, there is a large degree of heterogeneity in the management styles of all the funds listed in Figure 11. However, there also may be commonalities in behavior across the external managers. In order to capture these potential common components in active returns we specify the following model:

$$\begin{aligned} ActRet_{it}^{*} &= \lambda_i F_t + \sigma_i u_{it} \\ F_t &= \phi F_{t-1} + \sigma_F v_t, \end{aligned} \tag{5}$$

where *ActRet** is the de-meaned active return for external fund *i*, and *F* is a latent common factor. The model in equation (5) captures a common component, *F*, in the active returns of externally managed funds. The coefficient λ is fund *i*'s loading on that common factor and would be zero if the way the active returns are generated is independent of other funds. The estimation of this model, which allows for the different sample lengths of each fund, is discussed in Appendix E.

Figure 12 shows the estimated active common factor. The common factor for external active fixed income returns in Panel A displays many of the overall properties of the return pattern for the fixed income asset class (see Figure 6): it is more volatile at the beginning and end of the sample and very stable between 2002-2007. We overlay the fitted returns obtained by regressing the common factor onto the fixed income systematic factors. The R^2 from this

regression is high at 0.46 and like the overall fixed income returns in Figure 6, the identification is driven largely from the last two volatile years of the sample. In Panel B, the systematic equity factors track the common active factor for the external equity funds uniformly well over the whole sample and the regression has an R^2 of 0.56. For both the fixed income and equity external funds, the large decline toward the end of 2008 in the fitted factor returns is due to the large negative return for VOL.

We explore the systematic factor exposure in more detail in Panel B of Table 7. External fixed income active returns have significant exposure to TERM and VOL. These partial correlations are particularly large in absolute value at -0.38 and 0.36, respectively. External active fixed income managers undertook strategies that were underweight long-term Treasuries and strategies that sold volatility insurance. The external active equity factor has significant partial correlations with all equity systematic factors. Perhaps most notably, both fixed income and equity active managers have been short volatility, which combined with the internal fixed income and equity strategies, have given the Fund a large overall VOL exposure (see Table 3).

In Panel C of Table 7 we report correlations of the external active return common factor with hedge fund indices. With two exceptions, the active returns are not correlated with other hedge fund players. The first exception is that active fixed income external managers have significant correlations with fixed income relative value strategies, particularly those pursued in convertible arbitrage. The second exception is that equity external active managers tend to correlate highly with equity hedge quantitative directional strategies. These low hedge fund correlations are somewhat surprising but can be interpreted as good news; the external mangers are generating their alpha by methods not widely mimicked by other informed investors.

In summary, the active returns of external managers have large exposures to systematic factors indicating that a large part of external active management is not providing returns different from just loading on factor risk. The external fixed income funds have, on average, had relatively poor active returns but the record of active management in the external equity funds has, on average, generated value.

IIF 3. Net vs. Gross Returns

Section I notes that fees are an important component in the performance of actively managed portfolios: often there is evidence of skill before fees, but the after-fee returns to investors is often negative relative to a benchmark. Active fees are generally also high. Is this experience shared by NBIM?

Panel D, Table 7 suggests that the answer to this is, “No”. NBIM provided data on fees for a select sample of funds and over a shorter sample period, beginning in January 2005 and ending

in December 2008. This data reported fees in NOK terms every quarter for management and performance. We converted these to (management) expense ratios and performance fee ratios by assuming the fees were spread out equally over the quarter and used the NAV at the end of each month as a denominator. Like our previous analysis on external funds, we only consider fees paid on funds that are managed externally and do not count internal overlay and other portfolios.

We first provide some statistics on the cross section of expense ratios and performance fees across funds in Panel D, Table 7. All numbers are annualized. The mean expense ratio across external funds is 0.87% for fixed income and 0.76% for equity, but there is wide variation in these fees. The median expense ratio is much lower at 0.04% for fixed income and 0.14% for equity. The performance fee ratios across funds are also skewed to the right with the median performance fee ratio being 0.01% for both fixed income and equity external funds and the mean performance fee ratios are 0.17% and 0.29%. Thus overall the fees are low and are the same order of magnitude of the expense ratios for the entire Fund.[29]

An alternative way to measure fees is to look at the aggregate fees paid by NBIM to its external managers rather than looking at the typical external fund. The average expense ratio for NBIM's total external managers is 0.06% for the fixed income and 0.17% for the equity divisions. The performance fee ratios are similar at 0.07% and 0.17%, for fixed income and equity, respectively. The external fees reported this way are also low. However, since there is large exposure of the active external returns to common factors (see Figure 12), an argument can be made that it may be possible to obtain these systematic factor tilts at even lower cost, especially since the overall active management tracks the systematic factors so closely.

The last two lines of Panel D, Table 7 reports the correlation between gross and net average active returns and gross and net alphas. These correlations are high and there is little quantitative change in the analysis if gross returns are used instead of net. Thus, we do not repeat any analysis using gross returns, especially since they are on a much shorter subsample.

IIF 4. Summary of External Active Returns

The fixed income and equity external fund managers retained by NBIM on average contributed little excess return above their benchmarks. Fixed income external managers had, on average, a negative excess return, while equity external managers had, on average, a positive excess return. The cost of external management has been relatively low, although not as cheap as internal NBIM active management. The external funds on average have significant exposures

[29] NBIM Annual Report 2008 reports an annual 0.09% management costs on average since 1998 (including performance based fees).

to systematic factors that could presumably be replicated through a passive investment strategy.

IIG: Qualitative Assessments

NBIM seeks to exploit the scale and market power of the fund to add positive excess return. The effect of its strategy on Fund performance is constrained by the current risk budgeting process.

In this section we qualitatively assess NBIM's active management in light of the empirical results of the previous sections and based on our meetings with NBIM senior management. Elements of this evaluation include assessments of the strategic plans for active management, the risk budgeting process, internal and external reporting of active management risks and return, and whether or to what extent the specific strategies that have been used in NBIM's active management have been based on exploiting the characteristics of the Fund.

1. Assessment in Terms of NBIM's Value Proposition

   With respect to NBIM's stated strategy, our ex post empirical analysis suggests NBIM was able to generally generate positive excess returns in equity but less successfully in fixed income. With respect to breadth we found that both fixed income and equity residual returns were correlated to realized risk factors, suggesting that the investment decisions generating active returns were not independent of common factors. We revisit this evidence in more detail in Section III.

2. Assessment of Risk Budgeting

   We find that the overall risk budgeting framework for active management as regulated in the management guidelines set by the Ministry of Finance may have the feature of discouraging active management and that the risk budget as currently structured by tracking error limits does not take significant advantage of the potential for active management. This is especially true that the factor exposure of active management is very large and the volatility of these factors is not captured in the Fund's benchmark. The implication of our findings is that the risk budgeting process may result in the Fund underutilizing its capacity for active management.

3. Assessment of Internal and External Reporting of Active Management Risks and Return

   The use of a tracking error boundary rather than a target may lead to excessive risk aversion on the part of NBIM. Under the current risk-budgeting process the manager can only be penalized for exceeding a risk budget but not for underutilization of active

opportunities. The risk budgeting process is also sensitive to changing factor volatility. This can mechanically force disinvestment at inopportune times.

Due to limited access to information about external and internal managers, risk budgeting procedures and controls we cannot quantitatively assess the process by which capital is allocated to internal strategies or how external strategies are selected. Our review did not encompass NBIM's business operations, back office capabilities, trading costs, compensation, performance evaluation procedures, risk controls, or risk budgeting

4. Assessment of the Exploitation of the Comparative Advantages of the Fund:

NBIM is aware of their advantage of a long horizon perspective and uses it. We also infer some effects of this in their willingness to make factor allocations to obtain exposure to risk premiums. The current review, to the extent that it is a response to poor performance during the financial crisis, risks forcing management to question whether the client is, indeed as patient as presumed. In Section III we propose further means to make fuller use of this valuable characteristic of the Fund.

We find that although NBIM has been pursuing active strategies with large exposure to factors which exploit the comparative advantages of the Fund, these factor exposures and the potential downside distributions of these returns were not properly communicated to the asset owner. Had the factor exposures been known prior to 2008 along with the distribution of potential drawdowns of the factor returns, then the asset owner would have been able to calculate correctly the probability of the active losses incurred post-2007. If the probability of these losses was deemed acceptable then the losses themselves would have been within anticipated loss limits.

NBIM market power may have contributed to its track record, but we do not specifically attribute any measured result to this feature of the fund. The mixed performance of external managers suggests that the effect is possibly uneven by asset class, at least historically. It would be unwise to infer too much from recent events with respect to fixed income manager norms, however. Internal performance, particular in equity strategies, supports the assertion that NBIM has the ability to hire strong internal management staff.

Our analysis of NBIM costs suggest that this relative advantage was realized with respect to internal management. We have not been asked to review NBIM's overall fees in a comparative framework.

5. Aggregate and Disaggregate Performance Assessment

   Our key finding is that, due to risk budgeting policy, active returns had a very small influence on overall performance. Over the history of the Fund, active management has neither hurt nor helped total return, although there were extended periods when it significantly enhanced performance. There is also evidence suggesting that NBIM successfully altered some factor exposures and avoided losses that could have been more severe had the factor exposures remained constant in 2008 and 2009 at their December 2007 levels.

   Another key finding is that that excess performance was due in part to exposure to systematic factors. We regard this as an appropriate basis for seeking return, as it is consistent with both the long-horizon mandate of the Fund and with capital market theory, reviewed in Section I, which posits compensation for accepting systematic risk.

   We consider the possibility that NBIM has not been active enough – or put another way, the Fund may be paying for active management capabilities that it does not fully utilize. The internal active capabilities seem to have been superior to external capabilities. With respect to external management, many of the external managers have not added value, with the active returns of external fixed income funds being especially poor.

IIH: Conclusion of Section II

By far the most important influence on the performance of the Fund is the choice of benchmark. This accounts for over 99% of the total variance of the Fund's returns so the contribution of active returns to the overall Fund performance has been small. However, a significant faction of even the small component of total Fund returns represented by active returns is explained by exposure to a limited number of common factors. The overall behavior of the Fund is therefore very similar to an index fund with a small overlay of exposures to systematic factors such credit, value-growth, liquidity, volatility, etc.

Although small on average, some of the factors to which the active returns of the fund were exposed experienced large negative realizations during the recent crisis. Understanding the factor exposure of the Fund is essential in understanding the Fund's active return losses in the crisis. Many of the factor exposures present in active returns – even those that were responsible negative active returns in the crisis – are, we believe, appropriate for the Fund but the amount of exposure should be deliberately chosen by the asset owner, rather than determined by the fund manager. We now address these issues in Section III.

[This page is intentionally left blank]

Section III: Comparative Advantages and Capabilities

The central question in the report is whether the Fund should be actively managed. To that end, the goal of this section is to assess whether and to what extent the Fund has comparative advantages that justify investing significant resources on active management. This includes an evaluation of the opportunities, challenges and specific capabilities with respect to currently managed and new asset classes. In addition to this assessment, we have been asked to propose strategies that capitalize on the Fund's comparative advantages, and to characterize these strategies in terms of distributional characteristics, verification horizons, management feasibility, compensation structures, and modes of evaluation.

These are ambitious goals and they define a perspective on the strategic framework for the management of the Fund going forward, based upon its particular needs and strengths. Our intent in this section is first to respond to the mandate by evaluating the Fund's comparative strategic advantages and disadvantages. Second, we sketch out some broad options for the future which derive immediately from the strategic assessment. Given that these strategies would be pursued through NBIM, we draw upon our assessment of the NBIM track record and general approach to investment discussed in the previous section as the foundation for our proposals. We re-iterate the caveats noted at the beginning of Section II. As noted in the previous section, our analysis is performed with limited information about the internal strategies used by NBIM. We have not been provided with sufficient information about internal management performance. We have also not been given information about internal risk-budgeting processes, procedures and measures. Thus we are only able to offer a limited opinion about NBIM's internal capabilities to generate returns via specific strategies. We are aware that the Ministry of Finance, as part of the evaluation of active management in the Government Pension Fund – Global, has asked NBIM to prepare a business plan for active management going forward.

Our summary of key findings is as follows:

1. The Fund has several distinct features that give it a potential comparative advantage with respect to the pursuit of active returns. These include its governance structure, scale, and long-term investment horizon.

2. Our survey in Section I suggests that only limited opportunities lie in the traditional quest for undervalued securities around the world via direct or delegated management, particularly for a fund of such large scale. NBIM takes a fundamental, bottom-up approach to active management focused on selection. Section II showed that fundamental analysis actually has had only a small effect on overall Fund performance. However, assuming the costs of this style of active management is low, such activities

do not detract from returns and may be beneficial for the Fund's other aims, such as socially responsible investing, even though they may have small overall contributions to Fund value.

3. The Fund should focus more attention on factor risk. NBIM's current active management strategy, intentionally or not, is primarily based upon strategic (fixed) and dynamic exposures to factors. We recommend a more top-down, intentional approach to strategic and dynamic factor exposures. This approach relies on expected equilibrium compensation for taking non-diversifiable risk. It is similar to collecting an insurance premium for bearing risk associated with those factors.

4. The factor approach is based upon development of tradable factor benchmarks. Some of these are straightforward extensions of those already in use by NBIM, e.g. equity benchmarks. Others, such as a volatility factor are not yet in use for decision-making purposes but the Fund already has an exposure to them. Factor benchmarks allow better communication of factor risks taken by the Fund and enable the preferences of the ultimate investor, the Norwegian people, to determine how much exposure to each factor will be taken, analogous to how the trade-off between equity versus bond risk is now determined. The asset owners' choices of factors and the amount of exposure need to be included in the benchmark to the greatest extent possible.

5. In practice, the top-down factor-based approach is an extension of NBIM's current operations. NBIM has clear capabilities in building indices and tracking them. This makes the manager well suited to undertaking this task. We present some intermediate steps towards development of the factor-based approach and discuss some salient factors.

6. The top-down factor approach does not eliminate the potential use of active management. We define active management in terms of two decisions. First, the decision to deviate from long-term strategic loadings on factors (e.g. a temporary shift from the target allocation to equities) and second, the decision to hold securities in weights that differ from factor benchmark weights. These roughly correspond to timing and selection, where the default, or the baseline case, is determined by factor portfolios. Components of active management will be defined and measured in these two dimensions.

7. For purposes of internal evaluation of the costs and benefits of active management, and capabilities with respect to timing and selection, we recommend maintaining an internal record of decisions about both dimensions of active management. We also recommend that the factor loadings chosen by the asset owner be themselves measured against a customized, market-weighted portfolio.

8. If NBIM's management service were changed to a factor based approach based on harvesting risk premiums, this would change some aspects of performance evaluation from a metric based on beating benchmarks to ones delivering factor benchmarks cheaply. This would require some investigation into effective compensation structures.

9. We make one suggestion with respect to measurement and compensation. A horizon-based differentiation of trading strategies can help moderate any tendencies towards "short-termism." We recommend "horizon buckets" that will be reported and assessed at different frequencies. This allows strategies which tend to pay off over many years, possibly decades, and may be subject to large short-term drawdowns in returns a longer time to fruition and mitigates the temptation to cut back or add to allocations to these strategies in response to short-term losses or gains.

10. If the Fund seeks to further diversify into alternative assets, we believe it should evaluate these in a factor-based framework consistent with traditional assets. It should also consider using its scale to buy the best expertise or operations available.

The rest of this section is organized as follows. Section IIIA reviews the strengths and weaknesses of the Fund. Section IIIB discusses active management currently practiced by NBIM. We propose a new factor benchmark approach for the Fund in Section IIIC. Section IIID proposes setting up different horizon categories for reporting and evaluating performance and Finally, Section IIIE concludes.

IIIA: Strengths and Weaknesses of the Fund

The Fund has several distinct features that give it a potential comparative advantage with respect to the pursuit of active returns. These include its governance structure, scale, long-term investment horizon, and its relationship to its management company.

IIIA 1. Governance Structure

The Fund has a transparent, long-term mandate with broad public support that gives it legitimacy. By stressing that the Fund benefits the ultimate owners, the Norwegian people, this mandate frames how the Ministry, NBIM, and other entities associated with the Fund act. NBIM as investment manager also has a stated mandate to maximize returns within the parameters provided by the Ministry of Finance. These explicit guidelines provide a basis for evaluating strategic decisions and measuring performance. In short, the clarity of its foundation, the separation of duties of the asset owner and the fund manager, the carefully

proscribed guidelines for conduct, and the ultimate accountability to the public all represent comparative advantages for the Fund.

The current governance structure whereby the responsibility to fulfill the fund goals is delegated to the Ministry of Finance and the implementation is performed by NBIM in a rigorous, transparent environment is truly a model for other sovereign funds. A fund of the scale of the Norwegian Fund in another governance context would be a great temptation, politically and bureaucratically, to direct funds away from the future benefit of all citizens to enrich the present consumption of, at worst, only a select few people. Similar entities, including other sovereign wealth funds, large pension funds, and endowments, are faced with how best to structure such a fund in order for it to fulfill its public mandate, as opposed to becoming the basis for influence and unequal benefit. In our view, the Fund's governance structure sets an example for the rest of the world in this respect.

Among other things, the public, transparent mandate, and governance structure reduce the potential for wasteful misdirection of assets and promote rational investment decision-making. In very practical terms, asset allocation decisions, manager selection decisions, and even individual investments can be driven solely by the optimization process. While the preferences and objectives of the Norwegian public should be reflected in how these decisions are made, the actual decisions that result from these processes are not subject to political influence. Although this issue might be overlooked in a report focused on active management, it is worth recognizing this comparative advantage from the outset.

The transparency also represents a potential limitation of the current governance structure, however. To the extent that communications associated with NBIM, the Ministry of Finance, Parliament, or other public entities associated with the Fund are potentially in the public domain, this may limit informal communications with respect to strategy. A major institutional investor would normally expect to have considerable formal and informal communication with a manager. These communications include discussions about the specific trading strategies employed, changes in operations and personnel, estimates of the factor risk exposures, time-series performance data for external review and analysis, discussions about macro-economic outlook, and the strategy the manager employs in response to changes in investment opportunities. This information seems restricted principally to the manager, NBIM. A major endowment, for example, would expect to have an ongoing understanding of changes in its systematic risk exposures, and immediate explanations for how decisions by its managers generated realized returns. At present, it appears to us that NBIM receives this information but it is not regularly communicated to, or discussed with, the Ministry of Finance. The requirement for transparency in terms of reporting and benchmarking also hinders the Fund taking advantage of some of its comparative advantages as it seeks to invest in asset classes

and strategies that do not have well-defined, liquid, and passive benchmarks. We suggest some ways to modify the reporting framework and governance structure below so that the Fund can better take advantage of appropriate investments.

Another key feature of the governance structure that is a major benefit is that the Fund's framework is designed to give the beneficiaries a sense of confidence. Because performance is reported regularly, the Norwegian people are able to trade off the benefit of current consumption with the expectations about long-term performance based on a set of reliable, well-grounded principles of investment management. These include diversification, minimization of fees and costs, and prudent active decision making. One drawback of this bond of trust, however, is the potential for loss of confidence if the Fund underperforms. The governance structure relies on the proper establishment of public expectations with respect to performance delivered by management. In short, NBIM should be careful not to set high expectations for active management performance if indeed it is unlikely to meet those goals as the public may not be well prepared for that eventuality. The flipside of the governance structure reporting regularly to the ultimate asset owners is that the asset owners should determine what types of active management should be pursued by the manager of the Fund. Our proposal below has the distinguishing feature of allowing the preferences of the asset owner to determine which, and how much, factor risk premiums should be collected in the active management of the Fund.

IIIA 2. Scale

The most striking feature of the Fund is its size. There are few funds of the same scale in the world. Theory suggests that size has potential advantages: these include economies of scale, potential to influence corporate management, access to the best external investment managers in the world, and potential partnerships with other large-scale entities to generate value. The scale of the NBIM's budget has the potential to attract the best internal staff in the world and the development of proprietary and distinct management technology.

IIIA 2a. Economies of Scale

The Fund's large size allows it to benefit from economies of scale in the organizational costs of fund management, especially in passive management. It also means that the fund is not prevented, as many smaller funds are, from trying to build capacity that is world-class in every aspect of asset management, including active management. Size also allows the fund to negotiate low fees with external managers and this is a potential benefit in both active and passive management. The fact that it has no need for marketing to attract customers may also result in NBIM having lower costs than other large managers.

The Fund's size and relative budget allow its manager NBIM to potentially attract the best management talent in the world. This potential is moderated by social norms that impose constraints on compensation, but hypothetically a fund of this size should be able to attract the very best internal staffers. However, we have not been asked to review the personnel in the NBIM organization. Absent other issues that would make the compensation structure non-competitive, NBIM should expect to attract and retain world class investment staff.

This expectation extends to external managers as well. Through NBIM, the Fund undoubtedly has access to the leading investment managers in every market in the world and should be able to hire the best in class. As we have documented in Section I, however, best in class does not guarantee positive risk-adjusted returns. As always in these situations, the problems come from asymmetric information, adverse selection, and incomplete monitoring.

A particular opportunity presents itself in emerging markets. The Fund's size means that, in some parts of the world, allocating a relatively small fraction of the Fund's portfolio may represent a very substantial opportunity to a local manager and this may allow the Fund to get access to the best in class in a market with relatively more opportunity to exploit off-equilibrium asset pricing.

IIIA 2b. Diseconomies of Scale

The Fund invests globally in many different markets. This requires the manager to maintain offices around the world and a complex back office operation adept at cross-border investing. It must monitor risk for a global portfolio on a daily basis. This is a challenging management task that requires considerable professional resources and expertise. Put simply, as the Fund extends its global reach the commensurate costs increase. These costs will also increase with the expansion into other, less liquid asset classes.

The major disadvantage to scale is that the forces of diversification conflict with the capability to identify and exploit under-valued assets. The empirical evidence on investment management skill reviewed in Section I suggests that superior performance is often concentrated in strategies that are not easily scalable. It is very difficult to create large, diversified portfolios with positive alphas. Historically, large, diversified positive alpha portfolios have the feature that they are exposed to common factors such as size, value, momentum, or other factors featuring in the voluminous academic work in the Fama and French tradition. Or, alternatively, these portfolios are composed of securities with common fundamental characteristics. While diversification is the enemy of pure alpha, it is vital to the minimization of risk and to the efficient pricing of risk. Completely diversified investors are at a distinct advantage in competition for profitable risky opportunities. At the margin, they require compensation only for systematic risk exposures.

The second major diseconomy of scale is due to price impact. Large block sales and purchases of corporate securities move prices. Part of this is due to the concern by counterparties that they are relatively less informed. Some institutional investors have sought to mitigate this adverse selection discount through the development of crossing networks in which trades are executed among major players at prices set in other markets and the development of trading strategies such as algorithmic trading. However, it is unlikely that these trading venues will ever completely eliminate the relative disadvantage that large-scale investors face due to micro-structure effects. Certain large markets, like corporate markets and real estate, can only be accessed through broker-dealers. NBIM does, however, have both the incentive and the capital to take the lead in the development of the technology for alternative trading venues with lower trading costs and this may be a development it could consider in the future.

A third potential disadvantage of the Fund's large size is the fact that, precisely because it is so large, it represents a potential source of very substantial fees for financial service firms of all types. The Fund has successfully defended itself from these pressures so far but, as the Fund is expected to remain large and grow larger for the foreseeable future, the potential for these "agency" problems to arise will not recede. Indeed, Section I reviews the Grossman-Stiglitz, Berk-Green, and other equilibrium models where although significant alpha exists in the investment community, none of that alpha filters to investors and is instead captured by management. A large, aggressive community of financial intermediaries, which includes investment banks, broker-dealers, and other financial service providers, have incentives for steering investors to expensive products that provide little benefit. The fact that NBIM does not have to attract investment funds in the market, an advantage in many respects, means that there is no automatic correction mechanism for poor performance that results from overpaying for services.

One line of defense is to constrain the policy of the Fund so severely that it excludes many of the services that financial service firms might otherwise wish to provide. An extreme solution would be to carry out all investment management internally (although this gives rise to an agency problem of a different type). Another would be to adopt a strategy of pure indexation which restricts investment to markets for which there are providers of index funds. Committing the Fund to a strategy of pure indexation is a draconian solution and would mean that Fund would miss out on the gains from many investment areas where no index funds exist. Moreover, many strategies with long verification horizons are inherently active because of the lack of a well-formulated, passive benchmark.

It appears to us that all stakeholders in the Fund are keenly aware of this problem and the culture of careful cost management seems deeply ingrained. Nonetheless, the problem will not go away over time and, indeed, is likely to become more troublesome as the fund moves into

less liquid and less populated markets. Guarding against this problem is a key function of the senior management of NBIM and the stakeholders that monitor the fund. The external cost assessments and high transparency of the Fund greatly help in this goal.

IIIA 3. Long-Term Horizon

The Fund is likely to be cash positive for many years and does not have to plan for unexpected redemptions. Thus it has a low requirement for liquidity; indeed, one potential strategy for the fund is to serve as a provider of liquidity in the global capital markets. Even when net contributions to the fund from commodity resources dwindle, the fixed spending rule combined with the Fund's perpetual horizon implies that only the real (inflation-adjusted) return will be monetized.

The Fund, as a consequence of being a long-term investor, according to NBIM's annual report, "will therefore, generally speaking, have a higher risk bearing capacity than other funds that it would be reasonable to compare it to." This implication stems directly from the "Swensen Model" of university endowments reviewed in Section I, which effectively funds the perpetuity in real terms. The Fund has an even greater horizon advantage than university endowments. Its liability stream will not begin until many years in the future. Thus, it has an even greater risk-bearing capacity than many other large-scale endowments of pension funds that have at least some current cash needs.

One implication of the regular flow of funds into the portfolio is that the Fund becomes an a-cyclical investor, in contrast to many funds that are pro-cyclical. This advantage, clearly identified in conversations with NBIM, aids the Fund since it can expect to earn positive returns from continuing to invest when asset prices are low as a result of short-term liquidity shocks or temporarily high risk premiums. The a-cyclicality also allows them to access top managers who are sensitive to liquidity-induced withdrawals. In contrast, a pro-cyclical investor "buys high" by tending to increase holdings of risky assets precisely when their prices are relatively high.

In terms of finance theory the horizon of the marginal investor is much shorter than the long horizon of the fund. Many rational theories of time-varying risk aversion, such as the theoretical canonical consumption models like Campbell and Cochrane (1999), predict that risk premiums change over the business cycle and are highest during economic downturns. Although we have not predicated our recommendations on the implementation of a strategy that relies on a dynamic risk premium model, the framework we propose certainly allows such an approach. For example, its long investment horizon would potentially enable the Fund to switch to riskier assets with higher expected returns during periods of high risk premiums, thus absorbing risk from marginal investors who have high risk aversion. The Fund could expect to earn higher returns for engaging in this transfer of risk. In behavioral models, periods of high risk premiums

can arise from elevated investor sentiment. Although the research on behaviorally-based strategic trading is still in its infancy, it is conceivable the Fund could profit from periods of short-term changes in the risk premium by counter-cyclically selling assets, or decreasing exposure to factors with low risk premiums during these times. The long horizon permits the Fund to take advantage of temporal variations in the price of factor risk until normal risk-return relations are established. In either the rational or behavioral framework, the long horizon enables the Fund to invest a-cyclically in order to collect long-run risk premiums.

The current rebalancing rules, which have been derived partly with public consultation, permit an automatic optimal allocation of risk in response to time-varying risk premiums. The asset owner, through the Ministry of Finance, sets very long-horizon targets (60%-40% equity-bonds) but the rebalancing rules imply dynamic responses to price movements to maintain these targets. The rebalancing rule is counter-cyclical and loads on equities when equities have considerably declined in price relative to bonds, when equity risk premiums are likely to be high, and reduces the weight on equities after large gains, when equity risk premiums are likely to be low.[30] The rebalancing rule and framework is a tremendous advantage for the Fund that prevents time-inconsistent and arbitrary actions of changing asset allocations in response to short-term noise. Below, we propose an extension of this framework to multiple factors.

While in principle the Fund can benefit by being a very patient investor, the public nature of the Fund and the particular arrangements that exist for monitoring its performance means that it has a high need for transparency. This, in turn, may reduce its tolerance for the type of episodic risk that accompanies investment in less liquid assets such as structured credit, private equity, venture capital, etc.

Going further, the need for those with fiduciary responsibility for the Fund to undertake an extensive review of its management and performance in the wake of a major systemic dislocation in capital markets, suggests to us that the actual capacity of the Fund to bear significant liquidity risk may be limited at present. To the extent that the Fund is in fact a natural provider of liquidity and could expect to earn a risk premium from doing so, the current structure of the Fund is not optimal to take advantage of liquidity and other factor risk premiums. It is possible that this problem might be alleviated by changes in the way in which the Fund's strategy is communicated to its stakeholders and we return to this issue below.

[30] Even if returns are independent and have constant risk premiums, the rebalancing rule is optimal in a utility-based framework (cf Samuelson, 1969; Merton, 1971). In this case, an investor's risk aversion determines the optimal holdings of risky assets, which are static. Rebalancing back to this fixed holding ensures that utility remains optimal.

IIIA 4. Management Company

The Fund clearly benefits from having a Norwegian-based asset manager. NBIM has developed as an organization within the framework of a central bank. This has cultural advantages; most significantly a commitment to government service and an ethical workplace. Second it has developed an institutional capability for learning that has seen its technical capabilities extend from a specific segment of fixed income to a wide range of internal and external investment activities. In operating over a period of rapid evolution of investment strategies and mandates, NBIM's staff have been learning by doing. This institutional experience has considerable value in active management specifically and more broadly. In part as a result of its size and prestige, NBIM has access to world-class advisors as well as strong links to the research community. At the same time, most of NBIM's senior staff have a commitment to the success of the fund that appears to extend beyond their direct compensation as salaries at some comparable institutions are much higher than at NBIM.

One limit of the structure of a closely-held, domestic manager is that change and adaptation to new market conditions is not as easy. In addition, the ethos of learning by doing represents a type of gradualism that may limit the ability of the Fund to change its capabilities quickly. For example, adding new asset classes like real estate cannot be done either immediately or as quickly as other investment organizations. We observed above that the scale of the fund means that it should be able to attract top management talent. If the public nature of NBIM's role limits its capabilities in terms of human resources, that is a potential concern. Without more information, and comparative information about the qualifications of NBIM's personnel and compensation in the investment industry, we cannot directly explore this issue and merely flag it as a potential question.

IIIA 5. External Funds

One of the distinct advantages of NBIM's organization structure and approach is that, because it insists on segregated accounts, it is able to see the decisions of its external active managers in complete detail. The data available from observing the entire set of trading decisions by the managers is very valuable in estimating the risk factors to which the fund is exposed on a daily basis. Even when the track record of the manager is insufficient to reveal the nature of the systematic factor exposures, the underlying portfolio provides that information. Although this detailed information is also very useful for calibrating the NBIM approach to active management, even absent a focus on seeking positive alpha returns, the security-level information available on a daily basis should be an important input to management. By the same token, an organizational structure established to monitor and analyze this level of data is also a significant comparative advantage of the fund.

A potential conflict of interest arises in the use of segregated accounts between the Fund and the external manager because the Fund's internal management capability means that, armed with detailed information about the transactions, it has the capacity to compete with the external manager and/or reverse-engineer any consistent, quantitative strategy pursued by the manager. Some external managers may not be prepared to accept a mandate from the Fund under these conditions and this reduces the range of options for the Fund. Given the size of the accounts NBIM will place with external managers, and the large (absolute levels) of the fees which accompany the large invested amounts, this is probably a small concern. According to NBIM management the firm maintains a strict "Chinese wall" between their own internal managers and staff who interact with external managers; that external managers accept this suggests that the problem is not serious in practice.

This Chinese wall is a good present solution to the problem. However, over time, there needs to be flow of legitimate information about risks and factor exposures across the Chinese wall. The strategies chosen in external funds must complement and enhance the active and index mandates pursued internally. Section II uncovers large exposures to systematic factors in the external fund managers which sometimes are opposite in sign to the exposures of internal strategies. It is likely that active internal managers are in the best position to evaluate and set external mandates so that they do not overlap and enhance the internal return strategies.

The Fund is relatively distinct in having focused to a large extent on public capital markets and management towards benchmarks. In order to serve this mission, NBIM has had to develop a technology for tracking internal and external portfolios against a "passive" benchmark. As a result, we believe that NBIM has a major potential advantage with respect to maintaining daily positions and daily benchmarks and the capability of keeping a very low level of tracking error. In other words, the small amount of active management we document in our analysis quite likely makes the Fund one of the most disciplined benchmark managers in the world.

IIIA 6. Summary

The comparative advantages of the Fund – its governance structure, scale, and long-term investment horizon, confer significant advantages. However, all these advantages can be pursued in either a passive framework, or an active framework, or some combination of the two. The passive framework consists of market-weighted benchmarks, a set allocation of those benchmarks, and automatic rebalancing rules. The pure passive approach does not rule out external funds; external managers are chosen to manage certain segments of the benchmarks at a lower cost than could be achieved internally. The active framework accommodates discretion to deviate from those benchmarks. The current NBIM investment style blends passive and active management by combining enhanced index funds with internal and external

active strategies. We now discuss the advantages and disadvantages of the current active management framework of the Fund.

IIIB: NBIM Active Management

Active management capabilities as currently defined add only marginally to the performance of the fund.

IIIB 1. Investment Philosophy

NBIM's current active investment philosophy focuses on the assessment of fundamental value, i.e., identifying assets that are mispriced. In the past, and possibly currently, its approach has included tactical asset allocation, factor-based position taking, and relative value as well as fundamental value strategies.[31] As Section II shows, these and other active strategies account for an extremely small fraction of the fund's overall performance and until recently NBIM has used very little of its permitted risk limits.

As discussed in Section II, NBIM employs an analytical framework commonly referred to as the "fundamental law of active management" which is predicated on the existence of manager skill, defined as the correlation between manager forecasts and outcomes. It is worth stating this more formally in this section in order to show how a focus on active management at the individual security level may not translate to a large degree of active management, in excess of systematic factors, at the aggregate level.

Grinold and Kahn's (2000) fundamental law of active management is a quantitative, statistical model. As such, it says little about the potential source of positive risk-adjusted returns, merely how to systematically evaluate them. It takes the form:

$$IR = IC \times \sqrt{BR}. \tag{6}$$

The fundamental law states that a portfolio manager should seek to maximize the information ratio [IR] which is the ratio of average residual return to residual risk.[32] The IC is the correlation of the manager's forecast with actual outcomes [how good the forecasts are] and BR is the breadth of the strategy [how many bets are taken]. High IRs can be generated by finding

[31] See "Highest Possible Excess Return at Lowest Possible Risk", Norges Bank, Updated 26 May, 2004.

[32] Section II shows that the residual return is almost identical to the active return because the beta with respect to the benchmark is almost one at the overall Fund, asset class levels, and internal levels for the Fund.

opportunities where the manager has skill, i.e. ICs greater than zero indicate the manager has forecasting ability, and by finding many such opportunities, i.e. BR is ideally large. Breadth can be thought of in terms of the number of securities, managers, or the frequency of bets over time. The classic example of narrow breadth is stock market timing because, even when a timer has skill, a decision to rebalance occurs rarely and consists of making investments in few assets (e.g. bonds versus equities). In contrast, a cross-sectional equity strategy that yields very low ICs, often as low as 2-5% can be highly profitable because the number of stocks used in the strategy may be in the thousands. The cross-sectional strategies of value, size, momentum, and other classic CAPM anomalies covered in the finance literature in Section I fit into this category. This approach thus justifies broad diversification across managers and strategies.

There are two limitations of the fundamental law that are relevant to NBIM. One limitation is that the IC may be declining as the breath increases. The first manager you hire may have a high IC but the tenth manager you hire may not. The 100th manager might have zero IC but still charge a fee. Similarly, as assets under management increase, the ability to generate high ICs diminishes. This is especially true for many of the traditional CAPM anomalies which are strongest in the small and middle capitalization universes. Our ex post empirical analysis in Section II suggests that over most time periods since the beginning of the sample, NBIM was able to generate positive IC in the equity market.

The second limitation is that it is difficult to measure breadth. Manager decisions, both internal and external, may be correlated. For example, U.S. equity managers seeking undervalued equities might all independently identify the same opportunities – stocks with low prices relative to other variables, such as a the book value, for example. These correlated decisions reduce the breadth. Put another way, having overweight positions on two thousand stocks with high book-to-market ratios, low price-to-earnings ratios, or other valuation metrics tilting towards value stocks is not two thousand distinct bets making up BR – it is one bet on the VALGRTH factor.

The evidence in Section II suggests that NBIM's ability to find large breadth is less convincing than its ability to generate high ICs. We found that both fixed income and equity active returns were correlated to realized risk factors, suggesting that the investment decisions generating active returns were not independent of common factors. Approximately 70% of the active return variation of the Fund in the sample was explained by exposure to systematic factors. Thus, employing many different managers and strategies is not providing the desired benefits predicted by the fundamental law. In fact, it suggests that equivalent results might be achievable with fewer managers and strategies. This might raise the average IC and reduce monitoring costs, however, we are certainly not privy to the on-going investment opportunity set of the manager and thus this is at best ex-post speculation.

In pursuing active management NBIM is effectively taking on large factor exposures and NBIM has demonstrated some skill in managing the time variation of these factor exposures. The focus on fundamental valuation strategies at the individual asset level seems to aggregate to provide largely factor risk exposure at the overall portfolio level. However, we cannot confirm this hypothesis without position-level information. The aggregate-level factor exposure is certainly appropriate for the Fund's long-term horizon and we propose a way to formalize and control this factor exposure below.

IIIB 2. Further Benefits of Active Management

The fact that in large diversified portfolios such as those held by the Fund the choice of benchmarks and factor risk exposures is likely to dominate does not mean that there is no role for asset-level fundamental analysis, or other active strategies pursued at the firm level, only that this role is fairly limited. Many of the ultimate aims of the fund – to ensure capital is being effectively deployed by managers, to increase the efficiency, liquidity, and transparency of markets, and to further other aims consistent with the fund's objectives such as a socially responsible investment, or advocating an environmentally-friendly agenda, require actively tracking the corporate actions and investments of thousands of companies. While these benefits are outside the scope of our report, we note that fundamental-level active management at the individual company level should help such activities.

Active management at the company level also complements the comparative large scale advantage of the Fund. A large shareholder in a publicly traded firm has the capability of influencing and changing management. NBIM uses this capability to pursue good corporate governance and socially responsible corporate behavior. We are unable to assess the extent to which NBIM has the capability to exercise the salient role of an activist investor, however research by Becht et al. (2009) suggests that this may be a potential source of value for large institutional investors. Certainly for small, activist hedge funds such activities yield significant risk-adjusted profits as shown by Brav et al. (2008). This is in contrast to an older literature summarized by Romano (2001) that activism conducted by large institutional shareholders has had little impact on firm performance.

Other sources of rents combine the scale factor with reputation. For example the Yale endowment has benefitted greatly from its (at least pre-crisis) aura of success by being offered highly select private equity participation in exchange for its implicit manager endorsement. We have not assessed the extent to which the Fund has taken advantage of this positive reputational effect, but the potential exists to do so. It will depend, however, on the external perception of the quality of the management team. It is more likely that such opportunities are brought to the attention of the Fund through company-level active management. Furthermore,

identifying and monitoring external managers is best done by skilled active internal management teams.

The potential scale benefit due an influential shareholder extends to investment in young firms as well. The Fund may be able to capture rents (or risk premiums) by making large commitments that significantly affect the likely success of the deal. For example, by taking a large pre-IPO stake, it may significantly improve the likelihood of success of the IPO and may therefore be offered its stake on favorable terms. Similarly, taking an initial seed interest in an up and coming external manager may enable investment on more favorable terms than investments in more established funds. These and other activities are enhanced by fundamental active management.

While active management certainly confers further benefits, the very large and diversified portfolio of the Fund means that these activities will still have modest contributions on the overall performance of the Fund. Furthermore, aggregating individual active activities is likely to create, and certainly ex-post has resulted in, large overall exposures to a few systematic factors. Thus, while we believe bottom-up fundamental analysis may have the potential to add to returns, and certainly has not detracted from returns over the last 10 years, we advocate a new framework where the main drivers of returns – factor risk and factor exposure – is made explicit and decisions on the type and extent of factor exposure are made from top-down decisions. We now describe this framework.

IIIC: Factor Risk Opportunities

Opportunities for the Fund lie in the potential to seek return through both fixed and time-varying exposure to factor risk commensurate with the Fund's distinct profile.

The opportunity we identify in this section stems from the strength NBIM has with respect to development and use of benchmarks. As we have noted above, if we define active management in terms of the behavior of active returns (the difference between actual and benchmark returns) then its impact on fund performance has been small. The impact of the choice of benchmark, however, is critical.

Both active and passive management involve setting a benchmark. Although active in name, the Fund has effectively been a passively managed portfolio even with the large active losses in 2008 and the early months of 2009. On the one hand, if active management is defined as taking positions in other than market-weighted bond or stock portfolios then the fund certainly should be active; the overwhelming evidence in the academic literature reviewed in Section I is

that the market portfolio is not efficient.[33] If managing a fund relative to a portfolio with time-varying dynamic exposures to various priced risk factors, specifying how the set of factors and exposures to those factors change over time, and being able to capture the full set of priced risk premiums is defined as passive management, then the fund could be entirely passive. Since specifying these prerequisites completely for an entirely passive fund is not possible, there remains some scope for active management, but relative to a benchmark that captures as much factor risk as possible.

The bottom line is that, as currently structured, regardless of whether NBIM has the capabilities to add value through active management defined solely as deviations from benchmarks, the risk guidelines in place render that activity marginal to the overall performance of the Fund. But, a broader definition of active management that recognizes the development and use of benchmark portfolios as an active investment process leads to a different conclusion.

Our recommendation is to incorporate the average factor exposure of active management into the benchmark portfolio and define the scope of active management to be two dimensional. The first dimension is the temporary deviation from the policy benchmark factor exposure and the second dimension is the traditional choice to deviate from the composition of any of the "passive" factors in the benchmark by choosing to weight some securities more than others. While both of these activities have been recognized as traditional potential sources of excess returns, the review in Section I suggests that there is only limited evidence that any given manager can consistently use them to produce positive risk-adjusted returns. However, moving the average systematic risk exposure of the current active component of the portfolio into the benchmark has the effect of properly accounting for factor exposures, and makes them a choice variable of the investor, rather than a byproduct of bottom-up active management. By the same token, this re-structuring will have the effect that selection and timing skill is not confused with simply an on-going policy of holding more factor risk than the benchmark.

Given this new definition of the scope of active management, the decision about the extent of active management should reflect its costs and potential benefits. The statistical history of the active performance of NBIM as documented in section II suggests that the benefits of active management, defined in excess of benchmarks *and* factor exposures have been small. However there is a legitimate question of whether the factor exposures in that active component could have been accessed strictly through a passive approach. That is, it is possible that some factor premiums may only be "captured" by the fund through external or internal managers whose *stated* goal is positive alpha. This issue of implementation is important to consider in evaluating the practical challenges of implementing a "top-down" factor-based

[33] Note the Fund does not hold the market portfolio: it has a fixed equity-bond mix and within the equity and bond asset classes has fixed regional allocations.

management philosophy. However, we suggest that it is possible to create passive, low-cost, factor risk portfolios that can be managed independently of generating alpha. Certainly creating pure factor risk portfolios greatly helps in communicating the risks associated with factor exposures and should be done to separate true "alpha" from "beta".

The top-down approach is thus not at odds with the current bottom-up management of NBIM and does not require a radical change in practice. Although it importantly provides the Fund's owners with the opportunity to choose the level of factor exposure, in terms of the investment process it is more in the nature of a change in cost-accounting method. This change seeks to achieve two things. First, it properly attributes added value in the production process. Second, it seeks to reduce costs in the production function. The first is achieved by moving the current factor exposure of active management into the benchmark and the second is achieved by seeking to capture factor exposures, wherever possible, by passive, transparent indices comprising liquid, publicly traded securities. In short, if a legitimate risk factor with an expected risk premium, for example, the bond horizon premium, can be captured by holding long-term bonds as opposed to paying a fee for investing with a fixed income hedge manager, the former is more attractive than the latter.

IIIC 1. Factors as a Source of Return

The financial crisis of 2008 has had a major effect on common beliefs about the source of returns to alternative investments. Before the crisis, the common assumption was that alternative investment managers focused on inefficient markets where opportunities existed to gain positive risk adjusted returns. The crisis revealed that many of these opportunities were actually exposed to a factor that many loosely recognize as liquidity risk.[34] This has caused a re-assessment of the model of relying heavily on alternative investments. In this section we suggest that, while priors about returns on these types of investments should be adjusted, they should not necessarily be avoided. Rather, they should be viewed as sources of risk premiums. These risk premiums are potentially important because they are not captured by statistical factors estimated from higher frequency returns on more standard fixed portfolios of stocks and bonds.

The crisis thus fits nicely into standard multifactor models of equilibrium compensation. Going from model to application requires parameterization of models: the specification of factor exposures, factor premiums, and common market understanding of both. In practice, the longer the horizon of the risk factor realization process, the less accurate is the assessment of

[34] We proxied this in Section II with LIQUIDITY, which was the differences in yields between liquid versus illiquid U.S. Treasury securities. This factor, together with VOL, played the largest role in explaining the declines of the fund during 2008 and its rebound in 2009.

the associated premium – or perhaps the less homogeneous is the market assessment of the premium, and hence different agents in the economy are more apt to charge different prices for insuring that risk – even when their risk aversions for that factor are equal.

The premise of the Fund's portfolio is that investing in financial assets over the long term will grow the assets in real terms. The foundation for these expectations is over five hundred years of historical experience in the Western capital market tradition and compelling financial theory developed to explain the long-term growth of investment capital. The heart of this theory is simple. Investors are compensated for risk – at least for risk that they cannot otherwise avoid through diversification. In this section we develop a framework for using this approach as a guideline for formulating the future investment policy of the Fund. This perspective does not depend on identifying mispriced assets but rather in estimating and efficiently allocating across a parsimonious set of priced risk factors.

IIIC 2. Capstone Intellectual Foundation

The neo-classical framework of the CAPM asserts that all the world's assets are held collectively in the same proportion by all the world's investors. In this simplest setting, investors use borrowing and lending to adjust their exposure to the market portfolio according to their risk tolerance. More risk-averse investors are happy to hold small amounts of the market portfolio; the market risk premium accrues to more risk-tolerant investors who harvest higher expected returns because they are paid to bear the market risk given up by more risk-averse investors. Another way of looking at the CAPM world is that negative returns on the market define periods of "bad times". A risk-tolerant investor is prepared to bear greater losses during bad times and in compensation she earns higher returns, on average, for taking on this risk. The current Fund benchmark sits in this world: the passive benchmarks have market-capitalization weights and the 60%-40% equity-bond proportions are determined by the preferences of the Norwegian people expressed through Parliament.

IIIC 3. Multiple Factors

The theoretical literature in Section I points to more than one risk factor than just the market portfolio being reflected in asset prices. The multifactor approach of Ross, Merton, Fama, French, and many others, advocates additional risk factors. As Section I summarizes, this is, in fact, the current paradigm of financial theory and empirical work. Indeed, we followed the tradition of using mimicking factors to represent some of these risk premiums in Section II. In a multifactor world, investors care about more than just low market returns, that is, the definition of "bad times" is extended to encompass other events disliked by the average investor. Economists call these periods of "high marginal utility." For example, suppose periods of low liquidity are also bad times, which are correlated with, but not exactly the same as,

periods of low market returns. This extra source of liquidity risk cannot be diversified away by investors trading with each other and liquidity risk is reflected in asset prices along with market risk.

A risk-tolerant investor prepared to bear greater losses during times of high illiquidity can earn a premium. Such an investor supplies liquidity insurance to investors unwilling to bear illiquidity risk. During normal times the liquidity-providing investor collects insurance premiums; during periods of illiquidity this investor faces losses, but overall the insurance provided in normal times outweighs the payouts during bad times and thus the investor earns a positive premium, on average, for providing liquidity. Collecting risk premiums in this multifactor world is similar to selling insurance because the risk premiums are compensation for making payouts, or enduring losses, during bad times. A similar explanation of portfolio allocation in the presence of multiple factors is given by Cochrane (1999).

Active management undertaken by NBIM partially fits into this multifactor framework. In Section II we found that active management has successfully captured exposure to many of these risk factors, especially liquidity and volatility risk, although, once again, these have represented only a small fraction of the overall return on the fund. Relative to the market benchmark, these additional factor risk premiums should be collected by a long-term investor. In fact, a zero exposure to these factor risk premiums would not be optimal for a long-term investor such as the Fund. However, like the 60%-40% equity-bond optimal asset allocation decision, the choice of which risk factors to which the fund should be exposed and the amount of exposure to each risk factor is a question for the ultimate owners of the fund.

Since ultimately financial assets are vehicles to earn risk premiums, the question of which assets to hold becomes a question of optimal factor allocation, as opposed to standard asset allocation. Our new proposed framework for the Fund involves specifying various sources of risk factors and their optimal holdings.

IIIC 3a. Benefits

Viewing a benchmark in terms of underlying factors rather than asset classes brings a large number of benefits, which include:

1. Better understanding of risk-return trade-offs.

   Given that assets are essentially conduits for factor risk premiums, certain asset classes, like hedge funds, private equity, or venture capital are not very meaningful descriptions. Understanding the factor risk behind these heterogeneous asset categories is a more fruitful way to estimating the risk-return trade-offs of the fund. Factors have different payoff structures over different horizons and allowing each factor to be harvested in an

investment style with an appropriate horizon and evaluation method mitigates tendencies to change these strategies in the face of short-term losses when their premiums are of a long-term nature.

2. Weighing the addition of new factors or asset classes to the portfolio.

   Financial innovation continually brings new asset structures to investors. Recognizing the underlying factor risk exposures allows an investor to gauge if the new asset class is truly adding a new factor risk premium, or just repackaging an existing factor exposure. The introduction of options and other volatility derivatives is an example of the former, whereas the (untranched) CDO market was perhaps an example of the latter (long credit risk, short fixed income factors, and short liquidity risk). As another example, the review of private equity in Section I implies this asset class might be classified as long small-cap equities and short bond horizon risk and risk-free rates.

3. Allowing the investor to determine which factors should have large or small exposure.

   In an asset allocation procedure the investor decides on an optimal amount of market risk (e.g. a 60%-40% equity-bond split) and in a factor allocation procedure the investor can separately decide on the amount of volatility risk, value-growth risk, liquidity risk, etc. held in the portfolio. At the moment NBIM active management effectively engineers significant factor exposure but the amounts chosen are decided by NBIM and not made at the level of the investor. Importantly, by specifying factor benchmarks, the choice of what types of, and how much, factor exposure can be tailored to exploit the comparative advantages of the fund.

4. Better benchmarking for active management.

   If a factor could be obtained alternatively at low cost, then bringing that factor into the benchmark raises the bar for active portfolio evaluation. This raises standards for both internal and external active management.

5. More robust portfolios.

   Many factors have skewed risk-return trade-offs which have large negative returns after stable periods of small but steady gains. These are probably appropriate investments for a long-run investor. Norway's process of specifying the benchmark would mean that various factors would only be included after public debate and a general consensus is reached. Thus, periods like the low returns in 2008-2009 may be properly anticipated if the factor risk profiles are correctly assessed. An important conclusion from Section II is that, had the risk profile of NBIM's active strategies been assessed from a factor

exposure standpoint, the possibility that losses such as those in 2008 could occur would have been communicated and understood by the asset owners.

The construction of proposed benchmarks will lead to a good public discussion about which factors are appropriate. This will make the asset owners more aware of the risks, particularly if they are represented as risk factors with premiums to compensate risk taken to bear these factor risks instead of positive-alpha opportunities without properly accounting for such risk. The downside is that the public would need to be further educated in order to understand better the factor risk-return trade-offs.

It is important to note that this view of the world is not new and, as Section I emphasizes, has been the prevailing academic paradigm since its development in the 1970's. However, the institutional adoption of the implications of multiple factors and multifactor models for asset allocation in industry has been much slower.

IIIC 3b. Drawbacks

Despite its theoretical attractiveness, there are a number of impediments to this idealized theoretical goal, which include:

1. The lack of investable or widely recognized factor benchmarks.

   One major difficulty is that factor benchmarks would involve more trading than market-weighted passive portfolios and leverage through long-short positions, although, in the case of the Fund, some or all of the leverage might be accommodated internally through adjustments to holdings of riskless debt. Many of these factor benchmarks need to be built from scratch. Risk management capabilities may need to be enhanced to deal with greater turnover and dynamic leverage.

2. No long time series to estimate factor risk premiums and factor risk-return trade-offs.

   The downturns in 2007 and 2008 of many factor returns, including the tremendous decline of the equity market portfolio in wiping out ten years of gains, is a good illustration that factor risk premiums are hard to estimate and easily mis-measured. Long-term, historical global market data could be used to assess the distribution and characteristics of low-frequency factors, but it has not yet been put to this purpose.

   The Fund's desired levels of exposure to each factor are decisions that are analogous to deciding the exposure to the equity market portfolio. It will be informed by estimates of the risk and return characteristics of each factor about which more is known for some factors than others. For example, there is extensive time series evidence of a risk

premium for the value-growth and momentum factors and less history for estimating the risk premium associated with volatility.

3. No consensus on what constitutes the complete set of factors.

   Theory is silent on what the set of factors is, but statistical studies suggest a limited number, usually fewer than 10, are enough to capture most variation in expected returns. However, even if the set of factors identified is incomplete – and at some level this is inevitable – including the Fund's chosen factor exposure in the benchmark is important because it removes from the active return a component that is, in fact, unrelated to active management and should be obtainable much more cheaply than true active strategies. Thus, this approach is justified even if the set of factors is "incomplete".

   Choosing the most appropriate set of factors and assessing their risk and return characteristics is more difficult. However, it should be noted that, even in the case of the one factor that has been subject to more analysis than any other – the equity market factor – it is not possible to assess the equity risk premium with great precision. It is therefore unrealistic to expect high precision in the case of factors, e.g., volatility, for which available return histories are much shorter.

4. Lack of a widely accepted robust procedure to do factor allocation.

   The same critique exists for ordinary asset allocation. Just as asset positions need to be rebalanced, factor positions must also be rebalanced. The rebalancing procedure for asset classes, currently equities and bonds, must be extended to incorporate factor risk positions.

5. Need for a new governance structure to implement the process.

   We propose one possible vehicle below.

Despite these disadvantages, we emphasize that factor exposure is already been taken by the Fund and represents substantially all the risk in passive returns and the majority of the risk in active returns. Currently, however, the asset owner has no way of deciding on the type or extent of exposure. The creation of factor benchmarks will help immensely in communicating the types of active management already being pursued by NBIM. Furthermore, NBIM's past annual reports and past reports to the Storting on the management of the Government Pension Fund already employ some single-factor and multifactor models in decomposing active returns. Thus, the factor benchmarks are a natural extension of the factor decompositions that have already been carried out by NBIM and the Ministry of Finance.

Recommending a complete structure to facilitate such a move from purely market-weighted passive benchmarks to factor benchmarks is a large enough task that it is outside the scope of this report. Nevertheless, below we give some short term recommendations of how the Fund could move towards this goal within the next one to three years.

Factor benchmarks build from the foundation of the market portfolio. In fact, if no factor tilts are desired then it is optimal to hold the market. Thus, the factor benchmarks must build from a benchmark portfolio of tradable securities weighted by their market capitalization.

IIIC 4. New Market-Weighted Benchmarks

Recommendation: Move to in-house, custom benchmarks

Currently the Fund uses passive benchmarks: FTSE for equities and Barcap (formerly Lehman) for fixed income. Section I reviews a number of distortions and costs for passive followers of these portfolios, with costs arising from many investors forced to track the same index changes. NBIM reports in its 2008 annual report transactions costs of 11 basis points (annualized), some of which are due to being forced to adhere to index weights. The Fund is a natural provider of liquidity to investors who wish to rebalance to the weights of published indexes and, as a result, should collect these premiums. Although NBIM made it clear that it avoids the well-known index recomposition effects, it does not avoid the persistent pricing differential between index and non-index securities.

Another important motivation for creating its own market benchmark is that several widely used existing fixed income indices are, in fact, not market weighted. These include the current Barcap index used by NBIM but also Merrill Lynch, Citi, Credit Suisse, and other widely used fixed income indices. Most fixed income indices are weighted by the nominal rather the market value of debt and therefore overweight more distressed companies. Many securities in these indexes are also thinly traded and are better off held in structures that are designed to accommodate illiquid securities better suited for long-term horizon strategies. A difficulty in creating a market-based fixed income index is that, unlike equities markets, fixed income securities do not trade in centralized exchanges where every trade must be reported. Only recently with the advent of TRACE reporting in the U.S. have corporate bond prices been accessible to all market participants. There is no corresponding vehicle for corporate bond price dissemination in Europe. In the recent financial crisis, several fixed income markets ceased to function (many asset-backed securitized markets as well as those for tranched and bundled fixed instruments) and prices for many constituents of fixed income indices were unreliable. Placing these types of securities in longer horizon categories, rather than short-term liquid buckets, would better fit their liquidity profiles. We comment further on this below.

The new market-weighted capstone portfolios should

1. Comprise only liquid securities.

   This does not limit securities to traded exchanges, which are mainly equities – U.S. Treasury and other sovereign issues trade in very liquid over-the-counter markets. However, certain corporate and asset-backed fixed income securities present in the Barcap fixed income indices will be excluded. Perhaps the long-term horizon of the fund may allow securities with prices or trades occurring at intervals longer than the daily frequency to be included.

2. Minimize turnover and have minimum size restrictions.

   We also recommend imposing minimum liquidity restrictions such as turnover or volume. These should be more relaxed than liquidity restrictions for the FTSE index because of the Fund's long-term orientation and few immediate needs for liquidity.

3. Adjust for free float.

   This is the proportion of the issue that is available for trade and not held by investors who never trade. This is rarely done in existing fixed income indices.

To a certain extent the Fund already creates its own custom indices: indices are modified by excluding companies on the basis of ethical considerations and the country weights of securities held are partially determined by the import mix of Norway. The market-weighted foundations also allow the decisions of the ultimate benchmark – which may incorporate specific country weights, sector weights, holdings of non-liquid assets, and other deviations from market-weighted, liquid securities – to themselves be benchmarked. That is, these portfolios should be the zero point of measure for all deviations taken by the asset owner, which then becomes the benchmark for active management by which the fund manager is measured.

A further advantage of creating true market-weighted benchmarks is that they would not include illiquid securities currently present in the present fixed income indices. The concept of tracking error as a risk measurement tool assumes that the underlying assets can be easily traded to manage risk. This is not the case for illiquid securities present in equity and especially fixed income benchmarks. Removing these illiquid instruments makes tracking error a more effective risk budgeting and risk management instrument. Below, we discuss an alternative structure that allows the illiquid instruments to be benchmarked in a manner more appropriate to their long-term payoffs.

Starting from these market-weighted benchmarks allows the Fund to specify active tilts that would be built into the benchmark. These would permit deviations from market value that reflect factor risk premiums built into the benchmark. For example, in harvesting a value premium companies with low prices relative to book value (or low prices relative to other accounting variables such as earnings, sales, or cash flows), would be given higher weights relative to companies with high prices relative to book value. However, the extent of these deviations and which factors should be taken cannot be answered without the existence of traded portfolios designed to collect these factor risk premiums. We now turn to this issue.

IIIC 5. Factor Benchmarks

Recommendation: Develop factor benchmarks

There currently exist no widely-used, recognized indexes that reflect factor risk premiums. Although Section II used several tradable portfolios to attempt to capture some systematic factors, they cannot be used as benchmarks to determine the factor style orientation of the Fund or as a basis for benchmarking NBIM's active management as they are not tradable in the large size and are not constructed to minimize trading costs. For example, the MOM factor used in Section II capturing momentum risk is certainly not tradable for the Fund's size.

These factor portfolios will, by construction, involve dynamic leverage. Simply buying a portfolio of value stocks gives an investor limited access to a value premium since the main exposure of this portfolio is equity market risk. To isolate the value premium component requires looking at the difference between the returns on value versus growth companies. This is why all the size and value mimicking portfolios used in the Fama-French style benchmark models to evaluate equity active managers are long-short portfolios. These portfolios also involve turnover: a value stock does not always remain a value stock; after a large enough price increase that stock will exit from the value category. Finally, these portfolios may involve derivatives to gain exposure to certain factor risk premiums. In the case of the volatility factor portfolio, this portfolio will need to trade primarily in derivatives like options and volatility swaps as the prices of these derivative instruments are very sensitive to movements in volatility. In practice a credit mimicking portfolio may also use derivatives, such as credit default swaps, to facilitate access to credit risk. Because of dynamic leverage, trading pure factor benchmark portfolios imposes high demands on risk management.

IIIC 6. Simple Example

In this section we present a simple example of how these factor benchmarks would work. Appendix F details how the Fama-French model commonly used in academic research fits into this framework.

Consider three stocks, which we label “Growth”, “Neutral” and “Value” with market weights 20%, 50%, and 30%, respectively. We construct a VALGRTH factor which is long the value stock and short the growth stock. Suppose the desired portfolio is 100% equities with a loading of 5% on the VALGRTH portfolio. The optimal benchmark is the market portfolio plus the VALGRTH portfolio, which is

	Market Portfolio	VALGRTH Factor	Benchmark Portfolio
Value	0.20	0.05	0.25
Neutral	0.50		0.50
Growth	0.30	-0.05	0.25
	1.00	0.00	1.00

The baseline benchmark is the market portfolio – this is the present benchmark currently used by the Fund. The VALGRTH factor picks up the differences in returns between the value stock and the growth stock and is a zero-cost mimicking portfolio, i.e., a portfolio whose weights add to zero. The weight of exposure to VALGRTH, which is 0.05, is chosen by the investor.

The desired portfolio is a combination of the market weights now adjusted for the VALGRTH exposure. Note that the desired portfolio is not a market-weighted portfolio; it is tilted towards the value stock and tilted away from the growth stock to reflect the positive loading of the portfolio on VALGRTH. In practice, the desired portfolio may have negative weights if the loading on VALGRTH is high enough. Thus, risk management to restrict the short positions and to track the dynamic leverage will be required.

Now suppose that the value stock has done well and its gain in price now makes it “Neutral”. “New Neutral” (formerly “Value”) moves from a market capitalization weight of 0.20 to 0.40. The old “Neutral” stock has declined in price from a 50% weight to 30% now making it “New Value”. The “Growth” stock remains a growth stock with the same weight of 30%. The VALGRTH factor is dynamic and reflects the change of the new value stock. The composition of VALGRTH changes and is rebalanced to drop the old value stock (which is now neutral) and hold the new value stock (which was previously neutral). Now we have:

	Market Portfolio	VALGRTH Factor	Benchmark Portfolio
New Neutral (formerly Value)	0.40		0.40
New Value (formerly Neutral)	0.30	0.05	0.35
Growth	0.30	-0.05	0.25
	1.00	0.00	1.00

Note the desired portfolio reflects the change of the identities of the value and growth stocks. In a pure market portfolio benchmark, there is no rebalancing to take into account of the dynamic value-growth effect. The new benchmark does. Certainly this brings higher transactions costs; there would be no trading in the purely passive market-weighted benchmark. With the desired portfolio there is a rebalancing because certain stocks have changed their underlying value-growth status. Minimizing costs and turnover will be essential in the construction of the factor benchmarks.

Despite being highly stylized, this simple example illustrates three main points of introducing the factor benchmarks:

1. The desired amount of exposure to the VALGRTH factor is set by the investor (here a weighting of 5%).

2. The benchmark portfolio is no longer a market-weighted portfolio and now reflects the VALGRTH exposure desired by the investor.

3. The benchmark portfolio is "passive" in the sense that it is based on a set of systematic rules, but is "dynamic" in the sense that its composition changes over time.

An important point illustrated by this example is that the ultimate benchmark reflects the combination of the market portfolio capstone plus the desired amount of VALGRTH risk taken by the investor. This makes the benchmark non-market weighted, but it is very different from non-market benchmarks constructed, for example, by Arnott, Hsu and Moore (2005) and others. Arnott and co-authors construct a portfolio weighting by various metrics of value. These "fundamental indexation" portfolios capture some part of the VALGRTH premium, but the exposure is not controlled by the investor and it is not separately specified from the market-weighted benchmark. Other important differences are that the factor portfolios we advocate can be easily extended to other factors (see below) and would also be globally constructed within and across all asset classes.[35]

The simple example also demonstrates several important differences between the factor benchmark portfolio and the current factor risk models used in industry provided by BARRA, Northfield, Axioma, and other factor model providers. These factor models tackle factors from a risk perspective whereas the factor portfolios we advocate are designed from a return-

[35] These same comments also apply to "factor" indices recently created by MSCI BARRA, and other index providers (see, e.g. http://www.mscibarra.com/products/indices/thematic_and_strategy/factor/performance.html). In addition to not being able to specify the Fund's desired amount of factor exposure and not being constructed using all asset classes, the Fund should be able to do much better by using efficient portfolio construction techniques.

enhancing perspective. Factor models from commercial vendors are typically used to create optimized mean-variance portfolios and are used to model return covariances and take the benchmark as given; our factor approach builds the desired factor exposure into the benchmark. The industry factor models are not tradable; the factor benchmarks we propose are explicitly tradable in large size with low transactions costs and low turnover. The industry factors are often not based on economic theory and do not necessarily permit risk-return analysis; our factor benchmarks are designed so that the investor can optimally choose which factors should be taken into the benchmark and how much exposure is optimal, which we now explain.

IIIC 7. Choice of Factors

A central question is which factors should be included in the Fund's portfolio. We recommend that the chosen factors should

1. Be justified by academic research.
2. Have exhibited significant premiums in the past which are expected to be maintained in the future.
3. Have return history available for "bad periods".
4. Be implementable in liquid, traded instruments.

We recommend requirement (1) so that the factors have an intellectual foundation. Only the factors with the strongest support in academic research should be included in the Fund's benchmark. This requirement does not preclude factors that have demonstrated ability to provide excess returns over long periods, but lack a theoretical explanation which is accepted by the majority of scholars. Our position is that long-term empirically documented anomalies like momentum and the value premium should be treated as risk factors. Research has demonstrated a rational and behavioral logic to the pricing of these factors, but the precise mechanism of how these premiums are generated is still unresolved. The usual debates in academic research mean that it will be rare that all factors will have a universally accepted economic story and that there will be consensus on the long-run risk premium of an individual factor (this is true even of the equity premium). If academic research has indicated that the factors are associated with risk premiums over the long term, then they should be considered for incorporation into an investment management policy. As continuing academic research identifies new potential factors, or qualifies earlier consensus on existing factors, these findings should inform investment policy accordingly.

Requirement (2) means that time series of reasonable length are available to permit risk-return analysis and (3) ensures that bad surprises are mitigated, but not eliminated. However, even (3) does not mean that a complete distribution of factor returns will be observed: volatility also performed very badly during the 1987 stock market crash, but performed even worse during the 2008 crisis. Nevertheless, having some data points to measure extreme drawdowns is desirable for assessing risk-return trade-offs and risk management. The four requirements also mean that because the factors are motivated from equilibrium arguments and are systematic, they will persist no matter how many large investors like the Fund are investing in them.

The three requirements exclude bringing a "new" factor into the benchmark. This is by design. A factor should be included in the benchmark only if it is widely recognized as being a risk factor. The many strategies that are excluded by (1)-(3) are best pursued by active management in NBIM's day-to-day activities. In Section I we mentioned the tendency of certain strategies to be first recognized as anomalies – this is "alpha" which active risk may capture – but as more players enter these strategies, underlying commonality in asset price movements may result – if the expected return patterns persist this is now "beta" and captured by a factor risk benchmark.

We emphasize that the set of factors is not static. For example, only with the introduction of traded options in the late 1960's and the publication of Black and Scholes' seminal paper in 1973 could investors gain exposure to volatility risk. In the past century, foreign exchange risk only became tradable after the removal of the gold standard and the breakdown of Bretton Woods in the mid-1970's. Momentum risk only became widely known after the publication of Jegadeesh and Titman's work in 1993. The high yield bond market only rose to prominence in the 1980's. Factors may also disappear. While the size effect was very prominent after the publication of Banz (1981), since the 1990's the size effect, as measured by Fama and French's size factor, has had weak returns in the U.S. (but still is strong globally) and is sometimes eliminated from current industry factor models. As the capability of hedging risks in the economy evolves, risk premiums can be expected to change as well. For example, the weakening of the size effect in the U.S. is consistent with efficient markets where the actions of informed investors cause a CAPM anomaly to be arbitraged away.

Our factor approach also goes beyond accessing risk premiums through holding assets in long-only positions. The current conceptual models for decomposing asset returns into priced factors are generally based on a "building block" approach (cf. Ibbotson and Chen, 2003) and models reflect the fact that the equity premium is a composition of several underlying factors, including short and long term interest rates, credit risk, and inflation. The "building block"

model thus develops expected returns for asset classes based on a hierarchy of asset risk premiums:

Expected Return on Equity = Inflation + Real rate + Term premium + Credit premium
+ Incremental equity premium.

The first three of these are effectively included in a riskless long-term bond yield, the credit premium reflects corporate credit risk, and the final component is an incremental premium for holding equity.[36] All of these factors are conceptually captured in the market-weighted fixed income and equity portfolios. The risk factors we advocate go further in permitting the investor to tilt towards assets capturing different types of factor risks and some of which, like the VOL factor used in Section II, are primarily accessed through derivatives. Our risk factors are specified on top of the market-weighted capstones as long-short deviations just as the example of the previous section.

We recommend that the Fund consider building some of the factors we used for analysis in Section II and consider the following candidate factors to include in the benchmark:

1. Term structure risk
2. Credit risk
3. FX Carry risk, which could also be folded into a value-growth factor
4. Value-Growth risk
5. Small-Large risk
6. Momentum risk
7. Volatility risk

The academic literature on each of these factors is extensive. It is not the goal of this report to provide a complete survey of the theory and empirical evidence regarding their compensation for risk. Nevertheless we provide a brief review of the empirical basis for the candidate factors in Appendix G, which demonstrates that long data is available for many of these candidate factors and gives some suggestions about how these factors could be implemented by the Fund.

[36] Ibbotson and others have sought to estimate the historical premiums for each of these for U.S. data. Dimson, Marsh and Staunton (2002) seek to provide similar evidence in the international cross section of stocks.

Constructing these factors will be challenging but we believe well within NBIM's expertise and competence. It will require additional data collection and analysis, as well as addressing important micro-structure issues. For example, practical development of factors will involve limiting turnover, positions, and rebalancing given the size of the Fund's portfolio and the desire to control costs. In many markets short positions are not permitted or can be prohibitively expensive, and this must be taken into account in factor construction. They must be limited to liquid assets that can be traded at low cost. They will also involve harnessing the Fund's advantages of being able to invest in global markets in many different asset classes.

For example, the VALGRTH factor we used in Section II has only value-growth differences in market-capitalization weights in global equity markets. Value-growth effects exist in commodities (measured by convenience yield or demand relative to inventories) and other asset classes. The FXCARRY factor is a form of value-growth – it goes long currencies with high yields and short currencies with low yields and could be folded into a VALGRTH factor. The MOM factor in Section II does not hold any international equities. An ideal momentum portfolio would also capture momentum in fixed income, commodities, currencies, and other asset markets and not just equities. As a final example, the VOL factor measured volatility risk using only one security – variance swaps on U.S. equities – but a full volatility factor would trade volatility risk across many different asset classes and possibly pursue volatility risk in cross-sectional strategies within asset classes.

The most difficult factor benchmark to develop is liquidity and we deliberately do not recommend it as a candidate factor, at least initially. In Section II we used a liquidity factor, LIQUIDITY, which reflected price differences between securities with the same credit risk (U.S. Treasuries) and equivalent in almost all aspects except for trading volume. This captured liquidity risk without any credit risk but captures liquidity effects only in very liquid markets. The long horizon of the fund makes it a natural collector of liquidity premiums in very long horizon, very illiquid markets. By definition these assets do not meet the requirement that the factor premium can be implemented in liquid instruments. It will be extraordinarily hard, perhaps impossible for the next few years, to build a satisfactory liquidity factor that reflects the return of these types of assets, which include real property, infrastructure investments, and venture capital investments. Unfortunately, while the need for such a factor is great, it is not captured by current benchmarks.

IIIC 8. Optimal Factor Allocation

Placing factors into the benchmark consists of three steps:

1. Risk-return analysis on each factor.
2. Determining how much factor exposure to bear.
3. Combining the factor exposures with the capstone market portfolio to yield an overall benchmark portfolio.

Note that this procedure can only be undertaken once tradable factor risk benchmarks have been created. Then, a debate can take place on the question of how much value-growth risk, volatility risk, or exposure to other factor benchmarks should be taken and whether they are appropriate drivers of long-term value. This expands the current debate on how much equity versus bond risk the Fund should take to include factor benchmarks. Once the factor exposures have been specified, it is straightforward to build an aggregate benchmark that adjusts the market-weighted foundations by the positions in each of the factor benchmarks, as illustrated by the simple example above.

The overall benchmark portfolio is "passive" in the sense that it is constructed according to a set of rules and its composition is pre-determined. But, it is "dynamic" in the sense that their composition is changing more often than a standard S&P500 or FTSE portfolio. It is important to note that these factors are correlated and that their correlations change through time, e.g., the simultaneously negative returns of many factors during 2007 and 2008. The benchmark asset constituent weights will change over time as called for by the changing asset positions in the dynamic factors. There is also no reason why the benchmark positions in underlying assets need to be positive if the amount of desired factor risk is large enough.

With the factor benchmarks, the question of whether to add different asset classes to the Fund's portfolio is now better phrased as the question of how much exposure to different factors the Fund should bear. Furthermore the evaluation of any new asset class to be added should be viewed as whether that new asset class brings about a new source of factor risk which is not spanned by existing factors and significantly adds to the risk-return profile of the Fund, rather than adding new asset classes per se. Arguably some recent alternative vehicles in the hedge fund and private equity areas simply repackage certain systematic factors in much more expensive forms.

The question of which factor risks the Fund should hold is better phrased as which losses the Fund is better prepared to bear than other investors. All these factors represent equilibrium compensation for risks and negative realizations for these factors hurt all investors with

exposure to these factors. The difference is that other investors dislike these negative outcomes even more than the Fund; the Fund is in a better position, because of its long horizon and other comparative advantages, to withstand higher short-term losses. It is in supplying this "insurance" that the Fund earns high returns over the long run. In deciding on the factor risk, the asset owner is specifying "beta" and this beta risk is built into the benchmark. The asset manager is now supplying "alpha" relative to these factors. The optimal factor decision is deciding how large the betas should be.

It is of course not possible to present a full risk-return analysis and recommendations for how much factor exposure the Fund should choose without having constructed the benchmark factors. Nevertheless, we can illustrate some of the tools which could be presented to help the Fund choose its exposures to the factors.

In Figure 13 we graph the market returns and the factor returns of the overall Fund over the sample period. The market returns are defined as the current Fund benchmark, which has changed its equity-bond proportion over the sample. We propose above to change the market return to the true market-weighted benchmarks. The factor return is defined as the fitted portion of the active return to a full-sample regression using the systematic factors of Section II. The period-by-period fitted returns are graphed in Figure 3. These are the returns that could be generated assuming that NBIM did not change its factor weights and if the factors were tradable. Note that since both assumptions are violated this is simply an illustration of a procedure that could be used to gauge the desired amount of factor exposure.

If we assume the combination of the market factor and the factor exposure actually taken as a candidate choice, then Panel A of Figure 13 shows the cumulated total portfolio return, absent any additional active return, that would have resulted over the sample in the solid blue line. For comparison, the cumulated market return is graphed in the dashed red line. The two lines almost lie on top of each other because the amount of risk in excess of the market has been very small (see Section II). We can vary the amount of market exposure and factor exposure to produce similar plots. We plot one such variation in the dashed-dot magenta line, which holds the market exposure constant and increases the factor loadings by a multiple of five. Repeating this process with different scenarios, combined with extensive simulations in addition to realized returns, can provide different distributions. These can be used to construct risk-return trade-offs, distribution functions, and prepare investors for potential downside risks in choosing different combinations of factor exposures.

One such example distribution is drawn in Panel B of Figure 13. We plot the probability density function of bearing market exposure and five times the factor loadings taken by NBIM in the sample. This combination of factors produces returns that are slightly more left-skewed (negative skewness) and more fat-tailed (excess kurtosis) than a normal distribution with the

same mean and variance. Various statistics from this distribution, especially statistics summarizing the left-hand tail drawdowns, can be used to calibrate an investor's desired risk aversion to the different factor exposures.

We emphasize that this procedure is no different from the process undertaken by Parliament and the Ministry of Finance in selecting the current 60%-40% equity-bond asset allocation and preparing investors for the possibility of downside risk with that current allocation. The extension is that various amounts of factor risk are now incorporated and the asset owner now makes decisions on the appropriate amount of factor exposure. We do not believe that standard, single-period optimization based on mean-variance utility will be especially useful in understanding first-order relationships among factor portfolios. A better approach, similar to that currently employed by the Ministry of Finance, is to prepare extensive simulations, particularly emphasizing potential downside losses, to calibrate risk tolerances to each factor of the asset owner.

IIIC 8. Active Management and Factor Benchmarks

There are several roles to be played by active management within the proposed multifactor benchmark structure. First, the factors are active strategies. Thus, better benchmarks will be built by managers with intimate knowledge of active strategies. In turn, building predictable, systematic variation into the benchmarks improves the active strategies. In short, active management such as that currently practiced by NBIM is a highly desirable, but not essential, prerequisite for building the factor benchmarks.

Second, active management can change the factor weights. Like the static 60%-40% equity-bond allocation currently set by Parliament, the factor benchmark allocations will be relevant at very long horizons. There exists scope to add value by changing the factor weights in the short term. NBIM currently engages in this practice, at least in reduced form, which we identified in rolling factor regressions in Section II.

Although it is possible to develop an overall benchmark that itself has time-varying weights to each of the factors, we recommend against doing this and recommend that the asset owner, through Parliament and the Ministry of Finance, set static factor exposures relevant in the very long run. Current models of dynamic asset allocation, such as those summarized by Campbell and Viceira's textbook, are not very practical for the Fund because they yield extremely leveraged allocations that vary tremendously over time. Even the more recent developments in the literature on dynamic asset allocation, which take into account learning and uncertainty in the predictive dynamics (cf. Chen, Jun and Miao, 2009), still advocate very large changes in asset mixes that are infeasible for a very large investor and cannot be easily extended to a large number of asset classes, factors, or predictive variables. Adding value by factor timing should

be done at the fund manager level as direct market knowledge and trading experience is necessary to react quickly to changing market conditions.

Third, active management exploits mispricing after accounting for factor risk exposure. By taking into account the factor exposures this raises the standards for active management. Over time it also makes the factor construction more efficient and lowers cost as insights from active management skills in trading technologies, analysis, and collecting and processing information inform the method by which the benchmarks should be constructed.

Active management may also identify ways to exploit other factor risk premiums, for example, those factors not included in the benchmark because they are not scalable, capacity is limited, or capturing these risk premiums require rare skills and specialized resources. Some of the risk premiums in the latter category include event-driven anomalies and high-frequency trading.

Finally, the set of factors probably varies over time. An important area for active management would lie in identifying new risk premiums currently not meeting the benchmark criteria.

We recommend that the scale of active management be much larger than in the present context, where active management is defined relative to the market-weighted benchmark. However, in the factor framework the systematic portion of the current active returns are now captured in the benchmark. Once the factor benchmarks are built and the amount of factor exposure is determined, there will be a residual tracking error. This is the dynamic counterpart of the existing tracking error which is at present tied to a purely passive, market-weighted benchmark. For simplicity we recommend that the same tracking error concept be used and that a small tracking error target or expected band, rather than a strict limit, be imposed. The imposition of an absolute limit, as is presently the case, can unduly constrain the manager. If NBIM's total discretion (i.e. risk budget) is fixed at the same amount as over the past 10 years and some risk is transferred to the benchmark as factor exposure, then the remaining discretion must be lower. In our sample period for the study of active management in Section II, the Fund's tracking error relative to the Fund's market benchmark, which did not explicitly incorporate factor risk, was 0.25% per month. With a factor benchmark, a reasonable tracking error relative to the factor benchmark would be lower than the 0.25% per month because now a large part of the systematic factor exposure in active returns is brought into the Fund's benchmark.

We summarize the scope of active management to add value beyond a factor-based benchmark in Table 8. Note that the indexation of the market portfolio and risk factors that command premiums have already been built into the benchmark. In the current framework, the (constant) exposure to risk factors is now counted as part of active management whereas in the

proposed factor benchmark case this would be moved from active management to the owner's decision.

IIIC 9. Implementation via a Transition Process

As noted from the outset, our proposal can be thought of as a re-definition of the current actual investment process of NBIM, rather than a radical departure from practice. The key step in our proposal relates to the development of factor portfolios. The research and development for this should be undertaken immediately, since full implementation requires a considered assessment of the robustness and utility of the factors. The initial step is to form a separate division of NBIM charged with "Factor Benchmark Development" to build the new market-weighted foundation portfolios and the factor benchmarks. This department should be rewarded for creating portfolios that can be traded in size at lowest possible cost. Note that in many institutions these factors would constitute active management; they are active strategies relative to the market portfolio and expertise in active management is highly recommended for personnel hired into this division. Thus, employees should be rewarded with the same incentives as regular active managers but with the additional charge that the factors must be built with simple, "systematic" and "non-discretionary" rules, at lowest possible turnover and cost that can be easily included in the overall Fund benchmark.

We recommend that the first factors to be included should be term structure risk and credit risk in the fixed income portfolio and size, value-growth, and momentum in the equities portfolio. These factors do not need the procurement of additional data beyond the characteristics of the security, current prices, past returns, and valuation ratios that should already be regularly collected and constructed as part of NBIM's fundamental analysis. These factors also have very long histories and can be done globally. The Factor Benchmark Development team would also be charged with creating appropriate "backfill" data if not available from tradable returns to facilitate risk-return analysis. A next step is then to add dimensions of size, value-growth, and momentum beyond just the equities portfolio and construct global portfolios that implement these strategies within and across asset classes.

While this process is underway, an intermediate step towards implementation is for the Factor Benchmark Development division to deliver to the Ministry of Finance appropriate portfolios that can be used for simulations, distribution analysis, and other exercises to calibrate the asset owner's preferences and risk tolerances with respect to each factor. There may be some factors which the asset owner judges as inappropriate to hold.

A next step, if not already capable of being implemented by NBIM, is to add derivatives to the factor portfolios. This enables the construction of volatility factors. The use of derivatives may require enhancements in NBIM's risk management capabilities. Derivatives could also be used

to enhance the returns of all factor portfolios. Management of derivative and dynamically leveraged portfolios may require upgrades in NBIM's risk budgeting and risk management process.

While this is being done, creation of new, market-weighted capstone portfolios should commence. These portfolios could take the existing Barcap fixed income and FTSE equity portfolios as a foundation and then use alternative data and pricing sources to expand these benchmarks. Liquidity screens are very important and the market-weighted foundations should comprise only liquid, traded assets. We comment on an alternative vehicle for holding illiquid assets below.

The final step, which can only be done after the creation of the factor benchmarks, is to create a process for using these factor benchmarks in selecting and evaluating alternative asset classes. In this final step, the "benchmark of the benchmark" would also be done. This would take the new market-weighted portfolios as the baseline case, and the decisions by the asset owner with respect to factor risk can be measured and tracked relative to the baseline case. An important part of the final step is to set rebalancing rules to maintain factor exposure over the long run, analogous to the rebalancing rules done at the equity-bond level for asset classes.

The Factor Benchmark Development division could be responsible for the presentation of a "business plan" for factor benchmark management to the Ministry of Finance that explains fully the systematic factor exposures contemplated and the measurement and control process for these factor exposures. NBIM, through this division, would further have the task of communicating these processes to the Ministry of Finance so that the development of the factors is fully explained to the client and constituents of the Fund. The value of this communication process is that it may help ensure a broad consensus on the decision about exposure to factor risk premiums and prevent a disjunction between risk expectations and realizations as in 2008. Once the factor benchmarks are built, NBIM should also re-formulate a "statement of purpose" or "mission statement" on the underlying philosophy of how active management can add value in excess of taking on systematic factor risk.

Finally, we recommend that the process of creating and maintaining the factor benchmarks by the new division be externally reviewed by an outside panel of experts appointed by the Ministry of Finance. This should occur both early in the transition process, like a year after the division has been created and the definitions and implementation of the factors have commenced, and also later when the transition process is almost finished, perhaps after three years, to review, among other things, the entire factor benchmark process and to ensure the integration of the factor benchmarks into NBIM's other active risk activities.

IIIC 10. Summary

The factor benchmarks we advocate mimic active strategies currently undertaken by NBIM and many other active managers. Bringing these factors, and their exposures, into the benchmark of the Fund allows the benchmark to reflect risk premiums recognized by both theory and practice, creates more robust portfolios, and most importantly allows the investor to determine the appropriate amount of each factor exposure.

IIID: Horizon Categorization

A horizon-based differentiation of trading strategies can help moderate any tendencies towards "short-termism."

IIID 1. Creating Horizon Buckets

As the recent financial crisis has made all too evident, some risk factors, in particular liquidity default risk, and volatility risk, are episodic in nature and the period over which it is reasonable to try to measure returns will be correspondingly long term. This perspective helps understand why active management of the NBIM portfolio first added value regularly until the crisis, then appeared to fail during the crisis, and, recently, after the crisis active, once again added value as markets stabilized. Our analysis shows that the positive performance was due in large part to factor exposures. Over the long term, theory suggests that this strategy should continue to be profitable. However, if the process of fund evaluation is conditional upon periods of poor performance, then there is the risk that such strategies might be discontinued exactly at the point when they are potentially most valuable. Financial theory suggests that if your portfolio does not lose value in a crisis, then it does not have a long-term expected return above the riskless rate. However, inherently cautious human nature makes it difficult to remember this equilibrium argument during the crisis.

The Shliefer-Vishny theory about the limits of arbitrage, reviewed in Section I, provides a natural motivation for classification by the horizon over which equilibrium price adjustment occurs and the credit co-factors influencing this price adjustment. The theory tells us that, not only should longer-horizon strategies for return be understood as such, but that there are moments when the investment in such a strategy will look much worse due to the forced selling by other arbitrageurs.

We recommend that factor benchmarks be classified by horizon, or cycle over which the strategies are likely to pay off. Strategies such as real estate development, for example, will be placed in a long horizon category. This classification could either be used as a guide for the

frequency and nature of formal review (see also below), or at the very least be useful as a guideline for interpretation of measured performance at higher frequencies. The classification applies to both asset classes as well as to factor strategies. This categorization will frame the evaluation of performance for that strategy appropriately. Strategies that blend long and short term factors will also be characterized as such, and expectations about their performance during periods of negative factor realizations can be adjusted accordingly. In fact, during the crisis NBIM took actions along these lines by separating certain illiquid fixed income instruments into a separate portfolio as stated in NBIM's 2008 annual report.

We propose three horizon categories:

1. Cash

   This category includes only instruments like U.S. T-bills and other very short-term claims on developed sovereign nations. Besides being necessary for rebalancing, having cash will be necessary to meet payouts when cash inflows to the fund cease. Having sufficient cash reserves on hand is necessary to act as a counter-cyclical investor and to provide liquidity.

2. Short Term

   All listed equities fall into this category. Other instruments with typical large daily volumes and daily pricing like foreign exchange, long-term sovereign bonds, agency bonds, and liquid corporate bonds would also be included. These are securities for which market prices are reliable and readily available.

3. Long Term

   This category is designed for any asset where market prices are not available at high frequencies due to lack of counterparties and for which predominantly "mark to model" or appraised values are used. This includes real estate, infrastructure, private equity, venture capital, and other "alternative" asset classes. Certain credit strategies, like illiquid securitized products would also fall into this category. The typical payout horizon for these assets is well over five years. Infrequent opportunistic, long-term trades over three-five years or even longer, which are within managerial discretion should go into the long-horizon bucket

We advocate that returns and market values for each horizon bucket be reported separately. Similarly, we also recommend that the simulation and risk-return analysis take into account horizon categories. Because some of the reported values in the long-term bucket are not market values, but mark-to-model and appraisal values, we also recommend that performance

reports of the long-term category emphasize metrics other than just returns: cashflow generating measures such as dividend yields, capitalization rates for real estate, and other cashflow distributions are appropriate.

An immediate implication of horizon bucketing is horizon allocation. Horizon buckets have the additional advantage of forcing the Fund to be aware of the issue of horizon diversification, and monitoring the temporal allocation of strategies. The lesson learned from the performance of endowments in 2008 was that the long horizon tilt of many institutional portfolios led to unanticipated management difficulties. While the Fund has a different mandate from an endowment, it is still useful to classify investments by horizon and to monitor that classification for potentially adverse concentrations. This will be more important when cash inflows to the Fund become very small relative to the size of the Fund or when cash inflows cease.

The factor benchmark portfolios are designed to apply only to the short-term horizon category. However, the factor benchmarks should still be used in evaluating and benchmarking the long term category. We now discuss this issue.

IIID 2. Evaluating Long-Term Assets

Recommendation: Evaluate alternatives to the best extent possible in a factor-based framework consistent with short-term assets. Use scale to buy the best expertise or operations available.

As discussed in Section I assets in the long-term category are generally considered attractive to institutional investor for three reasons. First, they provide diversification to the portfolio. Second they offer a chance to access potentially mispriced assets, since they are thought to be traded in less efficient markets and third, it is thought that they come with a liquidity premium since many investors need the benefit of being able to quickly liquidate securities.

The recent crash confirmed the prior belief that alternatives are exposed to a liquidity factor. We believe that it is important to consider the factor exposures of alternative assets in the context of the overall portfolio, even when they are not well measured by historical data. The classical example of this problem is venture capital. Because one cannot measure the monthly or annual returns to venture capital, one might presume that correlations to other asset classes are zero, or rather that returns are not generated by known factors.

Finding factor exposures of expected returns on assets that do not have good return data is difficult. Nevertheless, this problem can be addressed in a framework that allows reasonable priors about factor exposures and uses performance to update these beliefs. The base assumption would assume that the expected returns to any asset class – regardless of the observability of returns – is a function of underlying factor exposures, and that historical returns, as unreliable as they might be, provide some information about these unobserved

betas. It presumes that a historically successful strategy is compensation for systematic risk exposure rather than a function of manager skill. As such it conservatively budgets for risk when that strategy is included in the portfolio.

IIIE: Conclusion of Section III

In Section III we have been asked to evaluate the Fund's comparative strengths and to assess the extent to which they support active management. We have also been asked to describe a strategic approach that utilizes these strengths. In addressing this mandate, we have taken the liberty to propose quantitative, conceptual and procedural changes that we believe build directly on NBIM's particular strengths and address some of its distinctive needs.

Our main conclusions and recommendations are:

1. Comparative Advantage and Active Management

   We recognize several distinctive features of the fund that contribute to its ability to fulfill its mission, ranging from governance structure to market experience. Throughout the report we have also highlighted the extent to which the active component of the Fund's returns has been generated largely as a result of a rational, equilibrium-based investment policy implicitly based on capturing factor premiums, controlled tracking error, and judicious reactions during a period of crisis. These capabilities are well-suited to the needs of the fund going forward.

   We believe that significant resources should be focused on the development and further refinement of NBIM's factor investment and management capabilities, given the role that risk premiums play in excess return strategies and the relative paucity and limitations of existing investable passive global benchmark portfolios. We have noted elsewhere in the report that NBIM may be underutilized with respect to its capacity for active management.

2. Implementation: Existing and New Asset Classes

   The risk-based approach to active management can and should be extended uniformly across existing and new asset classes. The main challenge with respect to new asset classes is the lack of data for benchmarking. We have proposed two means to address this issue. First, a horizon-bucketing approach that would treat alternative investments that are illiquid or strategies in liquid instruments with long verification horizons differently with respect to performance evaluation. Second, we have recommended the

adoption of an approach that would seek to attribute historical alternative asset returns to a parsimonious set of tradable factors as a first approximation and would recognize that the residual may also reflect systematic risk exposure.

3. Strategies that Utilize Comparative Advantages

   We have proposed an investment approach that builds on NBIM's comparative advantage in benchmarking. Our proposal for the development of a factor allocation model is effectively an extension of current NBIM practice, although it de-emphasizes the quest for mispriced securities and emphasizes investment to capture risk premiums.

4. Identify and Construct Factors

   The foundation of our proposal is the identification and construction of factors that bear risk premiums. The need to do this in-house is driven by the lack of consistent, investable benchmarks for equity and fixed income factors around the world, as well as the lack of factors that capture meaningful sources of risk and returns, such as liquidity. We point out that the factor portfolios are dynamic and thus requiring rebalancing. We argue that they provide clarity with respect to the source of return and the exposure to risk. We have shown how the risk-return profile of the factor-based allocation strategy can be described in terms of a probability distribution. Similar tools can be used to gauge the public's risk aversion with respect to different factors and after identifying the amount of desired exposure to each factor, the factor exposure is brought explicitly into the fund benchmark.

5. Risk Categorization by Horizon

   We recommend that different strategies be classified according to the horizon appropriate to the cycle of equilibration. For any strategy, theory may suggest long-run profitability while in the short run there may be extended periods of underperformance. We propose three "horizon buckets" that reflect relative liquidity and different verification periods of the various strategies. Reporting and performance for each class of investment should be segregated, although the management of the portfolio would use the factor model as a framework for integrating all three. We point out that the lack of such a differentiation is a disincentive for NBIM to take active risks that generate positive long-horizon rewards.

6. Implementation via a Transition Process

   We recommend that a new division within NBIM be responsible for constructing, evaluating, and communicating the factors to the Ministry of Finance. Term structure

and credit risk in the fixed income portfolio and size, value-growth, and momentum in the equities portfolio can be started immediately and have long histories. Expansion of these factors to other asset classes and markets, and construction of other factors like volatility risk, should be taken as a second step. The division would also be responsible for creating customized, market-weighted portfolios.

[This page is intentionally left blank]

Appendix A: Analysis with Fama-French Factors

This appendix reports results with the Fama-French HML and SMB factors instead of VALGRTH and SMLG constructed from the MSCI universe. There are differences between these series because the Fama-French factors are constructed only from U.S. stocks, whereas the MSCI indices include global stocks, the Fama-French factors use many small and micro-cap stocks, whereas only mid to large firms are included in MSCI indices, and the Fama-French factors have high turnover and a rebalancing strategy that would not suit a large investor with low tracking error relative to a market-weighted benchmark index, whereas the MSCI indices are deliberately built to be easily mimicked by index funds.

As an example of these differences, we plot cumulated returns on the size factors SMB and SMLG in Panel A of Figure A-1, which have a correlation of 0.75. In Panel B we plot HML and VALGRTH, which have a correlation of 0.73. Thus, the Fama-French factors are similar to the MSCI factors we employ.

If the HML and SMB factors are used, translated to NOK, instead of VALGRTH and SMLG, the partial correlations of the overall Fund's active returns with the systematic factors do change, which we report in Table A-1. The partial correlations with respect to the other factors change because the partial correlations take into account the dynamics of the other factors and the time series variation of HML and SMB is different, as Figure A-1 illustrates, to VALGRTH and SMLG. Table A-1 reports that partial correlations of active Fund returns to HML and SMB are insignificantly different from zero. In the presence of HML and SMB there is still a significantly positive partial correlation with respect to CREDITAa, a negative partial correlation with CREDITBaa, and large positive correlations with LIQUIDITY and VOL. Whereas TERM had a borderline significant partial correlation of -0.17 with a p-value of 0.05 in Table 3, Table A-1 reports that with HML and SMB the partial correlation with respect to TERM is now statistically significant at the 95% level, but its magnitude is almost the same at -0.18.

Including HML and SMB factors make very little difference to the fitted versus actual active returns in the full sample and rolling regressions. Figure A-2 is the corresponding graph to Figure 4, except with HML and SMB replacing the systematic factors VALGRTH and SMLG. The whole sample fitted regression has a R^2 of 0.62 with HML and SMB, approximately the same as the full sample regression R^2 of 0.68 with VALGRTH and SMLG. The fitted active returns from the rolling regressions are extremely similar. VALGRTH and SMLG do not play large roles in explaining the 2008 negative performance; the large downturn in October 2008 is almost entirely due to LIQUIDITY and VOL exposure.

Appendix B: Regime Analysis

In this appendix we report some results on structural breaks through the sample. The most notable feature of Figure 2 is the very negative returns experienced in 2008 after a strong upward trend in active returns since 1998. Formal sequential break tests allowing for breaks in volatility dynamics, such as Andrews (1993) and those used by Andreou and Ghysels (2002), reveal the presence of additional structural breaks. Thus, a natural model to investigate changes in conditional means and volatility dynamics of the Fund's active returns is a regime-switching model following Hamilton (1989).

We estimate the following model by maximum likelihood on the overall Fund active returns:

$$ActRet_t = \mu(s_t) + \sigma(s_t)\varepsilon_t, \quad \text{(B1)}$$

where ε_t is IID N(0,1) and the conditional mean and conditional volatility change across regimes s_t. The regimes s_t take on the value 1 or 2 and follow a Markov chain with the probabilities

$$P = \Pr(s_t = 1 \mid s_{t-1} = 1) \text{ and } Q = \Pr(s_t = 2 \mid s_{t-1} = 2). \quad \text{(B2)}$$

The estimation reveals that both regimes are persistent and there is one regime with a slightly negative mean and very large volatility and the other regime has a significantly positive mean and low volatility:

	Estimate	Std Error
P	0.900	0.028
Q	0.986	0.008
$\mu(s_t=1)$	-0.037	0.085
$\mu(s_t=2)$	0.037	0.009
$\sigma(s_t=1)$	0.480	0.086
$\sigma(s_t=2)$	0.064	0.007

This pattern of high volatility occurring with negative conditional means also appears in global equity markets as shown by Ang and Bekaert (2002) and many others.

Figure A-3 plots regime probabilities smoothing over the full sample of being in the high volatility regime. Transitions to this regime occur in late 1999, just prior to 2006, and just prior to 2008. Figure A-3 shows that despite the stabilization of the global financial sector towards the end of our sample, the estimation still finds the Fund's active return volatility high enough to classify the entire period post-2007 as the high volatility regime.

The transition to the high volatility regime at the end of the sample occurs in August 2007 as there is a relatively large shock that month. However, the extremely large negative losses for the Fund do not start until 2008. The relatively large losses in August 2007 and also in November 2007. Although sub-prime mortgage securities started declining earlier in 2007, the losses on sub-prime and other mortgage securities started accelerating in August and during that month, AXA, BNP Paribas, and other investment managers announced support for their mortgage funds or suspended redemptions.

The large August and November 2007 losses are the same order of magnitude as relatively large positive gains experienced earlier in the sample, particularly the positive gains causing the switches to the first regime in late 1999 and late 2005. Thus, although the econometric identification chooses August 2007 as a break date, we use a pre-2008 and post-2007 sample period in our analysis. The December 2007 break date also has the advantage of being a complete calendar year and coincides with the same annual calendar reporting periods used by the Ministry of Finance and NBIM. Note that we also estimate rolling regressions beginning from early 2007 which has the advantage of not having to pre-specify different subsamples.

Appendix C: Rolling Partial Correlations

This appendix reports the evolution of selected rolling partial correlations for the active returns of the overall Fund estimated through the sample. These are analogous to the rolling fitted regressions, shown in Figure 4 for the overall Fund, and represent one way of normalizing the coefficients with respect to the factors in the regressions. We consider the rolling partial correlations with respect to LIQUIDITY and VOL, the two factors that contributed most to the Fund's active losses in 2008.

Panel A of Figure A-4 plots rolling partial correlations computed from the beginning of the sample to the current month for the overall Fund. The rolling partial correlation with respect to LIQUIDITY is relatively stable across 2007 to the end of the sample. We previously noted that in Figure 1, Panel B the large spike in illiquidity in late 2008 coincided with some large active losses of the Fund. Interestingly, Figure A-4 does not show a marked increase in the LIQUIDITY rolling partial correlation during this period. This is because the Fund also experienced some large, negative returns in the late 1990's when returns to LIQUIDITY were negative as a result of the default of LTCM and the emerging market crisis (see Figure 1, Panel B and Figure 3). Thus, the partial correlation with respect to LIQUIDITY remains high over the entire sample.

In contrast, the rolling partial correlation with respect to VOL increases from below 0.2 to close to 0.5 in September 2008 and then further increases the following month. During September 2008, VOL returned -12% and the Fund endured its largest loss in active returns that it had experienced to date. October 2008 saw a return in VOL near -50% and the Fund also had a very large drawdown that month. These are the first dramatic reversals in VOL's returns in the sample, which prior to this period had enjoyed generally steady gains (see Figure 1, Panel C). Thus, September and October 2008 are very important for identification and there is a large increase in the partial correlation with respect to VOL over these months.

Section IIE presents evidence that is consistent with NBIM having successfully changed its exposure to VOL in equities to avoid a much more catastrophic loss than actually occurred during the second half of 2008. Panel B of Figure A-4 shows the rolling partial correlations with respect to VOL, controlling for only the equity factors (as in Figure 7) and this increases from around 0.2 to above 0.4 in September 2008, a similar increase to that for the VOL factor for the overall Fund in Panel A. However, unlike the overall Fund, the partial correlation with VOL for the equity active return decreases in October 2008 to 0.3 when the worst negative return to VOL occurs. Although we have no direct evidence on the actions taken by the manager, this is consistent with the VOL exposure being managed in equities, but not in fixed income.

Appendix D: Analysis of Fixed Income Autocorrelations

Our analysis revealed strong positive autocorrelation of residual returns for the fixed income area reported in Table 4. Given the evidence in the academic literature about autocorrelations as evidence of valuation challenges, we tested some additional hypotheses about the cause of the trends in the fixed income sample.

For example, it is possible that end-of-quarter estimates of these illiquid instruments are closer to true market values than estimates at other end-of-month periods. Such reporting would imply that the implied monthly autocorrelation from end-of-quarter returns would be lower than the autocorrelation computed directly from monthly-frequency returns. We find some confirmatory evidence of this. The implied monthly autocorrelation from quarterly-frequency returns is 0.68, which is lower than the direct monthly autocorrelation of 0.72. Another possibility is that the high autocorrelation over the full sample is caused by the negative endpoint at the end of the sample. The mean of the sample is defined ex post. The negative active returns towards the end of the sample pull down the mean making the earlier observations appear to have high realizations relative to the mean. Computing rolling estimates of the mean through the sample would mitigate this problem. We do not find that this is causing the high autocorrelations through the full sample. The autocorrelation computed with "up-to-date" estimates of the mean is still very high at 0.73 and almost identical to the standard autocorrelation of 0.72.

We believe it is most likely that the high fixed income autocorrelation stemming from the period after 2008 is being caused by smoothed valuations of fixed income instruments for which there is little information on market values. To confirm this hypothesis requires position-level data, which was not provided. Such investments may be entirely appropriate for a long-term investor like the Fund. But, the 2007-2008 financial crisis has highlighted one challenge with holding such illiquid investments and determining appropriate transparent reporting schemes. We address how such investments could be actively managed in Section III.

Appendix E: Common Factor Model of External Active Returns

This appendix describes the estimation of the common factor model fitted to external active returns. For convenience, we restate the model in equation (7) here:

$$ActRet_{it}^{*} = \lambda_i F_t + \sigma_i u_{it}, \quad \text{(D1)}$$

with the common factor *F* following

$$F_t = \phi F_{t-1} + \sigma_F v_t, \quad \text{(D2)}$$

where *ActRet** is the de-meaned active return for external fund *i*, *F* is a latent common factor, λ is the loading of fund *i* on the common factor, and the residuals *u* are N(0,1) and independent across time and across funds, and the shock to the common factor *v* is also N(0,1), orthogonal to all *u*, and also IID across time.

To estimate this model we employ a Gibbs-sampler Bayesian algorithm. This involves iteratively drawing the parameters and the latent factor from a series of conditional distributions, which in steady state yields the distributions of the parameters, the latent factor, and all missing returns of the funds. Since each fund enters and exits at different points in time, the algorithm draws a return distribution for those periods where the fund is not present, and this distribution represents the returns that would have occurred according to equation (D1) if the fund had been held by NBIM. A notable feature of this model and algorithm is that it accounts for the entry and exit of funds in estimating the common factor. We denote the set of unobserved active returns as R^{unobs} and the observed returns as R^{obs}. Drawing the unobserved returns is a procedure known as "data augmentation". Similar estimations to equation (D1) are done by Stock and Watson (2002), who develop a principal components algorithm allowing for missing observations, and Ang and Chen (2007), who estimate latent persistent factors extracted returns, among others. A textbook exposition of Gibbs sampling procedures is provided in Robert and Casella (1999).

We iterate over the following draws in the Gibbs sampler:

1) $F \mid \lambda_i, \phi, \sigma_i, \sigma_F, R^{unobs}, R^{obs}$

We use the forward-backward algorithm of Carter and Kohn (1994) to draw *F*. Equation (D2) represents a state equation while equation (D1) is a series of measurement equations in a Kalman filter system. The algorithm works by running the Kalman filter forward through the sample and then sampling backwards following Carter and Kohn.

2) σ_i, σ_F, | F, λ_i, ϕ, R^{unobs}, R^{obs}

Given the latent factor *F*, equation (D2) is simply a regression and we can draw these parameters using a conjugate normal-inverse gamma distribution. Specifically, we assume a diffuse normal prior for ϕ yielding a normal posterior and an uninformative inverse gamma prior for σ_F yielding a inverse gamma posterior.

3) λ_i, σ_i | F, ϕ, σ_F, R^{unobs}, R^{obs}

Equation (D1) is a regression of demeaned fund returns on observable factors *F*. This is also a conjugate normal-inverse gamma draw. However, we place an informative prior on λ_i so the factor loadings are centered at 1 with a prior standard deviation of 1. To motivate this prior, a simple estimate of *F* would take an equal-weighted average of active returns across funds and first principal components of returns usually return a factor with close to equal weights. The prior standard deviation of 1 allows a reasonable dispersion of the λ_i factor loadings. The posterior standard deviation of lambdas is around 1.3.

4) R^{unobs} | F, λ_i, ϕ, σ_i, σ_F,, R^{obs}

Equation (D1) states that all returns are normally distributed given F, λ_i, and σ_i.

In estimation, we use a burn in period of 20,000 followed by 100,000 simulations. The following table reports the estimated coefficients for equation (D2):

	Fixed Income		Equity	
	Posterior Mean	Posterior Stdev	Posterior Mean	Posterior Stdev
ϕ	0.638	0.081	0.434	0.100
σ_F	0.644	0.075	0.445	0.055

Note that both fixed income and equity external common active return factors exhibit relatively high persistence. The posterior standard deviations are small so these parameters are identified by the data very precisely.

Appendix F: Factor Benchmarks and the Fama-French Model

This appendix shows the relation between the factor benchmarks proposed in Section III with the Fama-French three-factor model. The Fama-French model decomposes a managed portfolio return, *r* , into a systematic component dependent on market (MKT), size (SMB), and value-growth (HML) factors and an orthogonal residual component, $\alpha + \varepsilon$, where ε has a mean of zero:

$$r_t = \alpha + \beta \cdot MKT_t + s \cdot SMB_t + h \cdot HML_t + \varepsilon_t, \tag{E1}$$

where β, *s*, and *h* are the factor loadings on the MKT, SMB and HML factors, respectively. As explained in Section I, the philosophy behind this portfolio evaluation method is that the systematic component, $\beta \cdot MKT + s \cdot SMB + h \cdot HML$, can be replicated by the investor and represents exposure to risk which must carry a commensurate reward. The return in excess of the systematic component is alpha.

The MKT, SMB, and HML factors are portfolios and have constituent weights, which we denote by w_i^{MKT} , w_i^{SMB} , and w_i^{HML} , respectively for stock i. Thus, denoting individual stock returns as r_i, the return on the managed portfolio, *r*, in equation (E1) can be written in terms of individual stock weights and returns as:

$$\begin{aligned} r_t &= \left(\sum_i \left(\beta \cdot w_{i,t-1}^{MKT} + s \cdot w_{i,t-1}^{SMB} + h \cdot w_{i,t-1}^{HML} \right) r_{i,t} \right) + \alpha + \varepsilon_t \\ &= Bmk_t + ActRet_t. \end{aligned} \tag{E2}$$

The benchmark return, *Bmk*, is the systematic component and the active return, *ActRet* = $\alpha + \varepsilon$, is the return of the managed fund in excess of the systematic component. The benchmark weights adjust the market weights of each stock by the size and value-growth deviation implied by the SMB and HML factors. The size of the exposures are the Fama-French factor loadings, which are assumed to be the choice of the investor. Note that the weights for SMB and HML sum to zero, i.e.

$$\sum_i w_{i,t-1}^{SMB} = \sum_i w_{i,t-1}^{HML} = 0, \tag{E3}$$

and are positive (negative) for small (large) stocks and positive (negative) for stocks with high (low) book-to-market ratios. These weights change when the Fama-French factors are rebalanced.

Appendix G: Empirical Overview of Candidate Factors

This appendix provides a brief overview of the candidate factors suggested for the Fund to include in its benchmark. All of these factors would be constructed to be zero-cost, i.e. have long and short positions that sum to zero, that would be overlaid on the market benchmarks. The portfolios should be constructed to minimize turnover and have the lowest trading costs possible. In many cases, the short side may not be vital to portfolio construction.

1. Term Structure Risk

   Long-term government bonds have historically provided higher yields than short-term bonds and this difference is regarded as a compensation for the exposure to the risk in variation in the future short-term rate, although several theories of the yield curve propose additional reasons for this yield gap, including variation in demand for money at different maturities (see most recently Vayanos and Vila, 2009). Embedded in the long-term rates are also expectations about inflation and inflation risk premiums, since long-term bonds are nominal securities (cf. Ang and Piazzesi, 2003). The long-term bond rate for many countries can be traced back well more than a century through various macroeconomic regimes (see, e.g., Ibbotson and Sinquefield, 1976; Dimson, Marsh and Staunton, 2002). Thus, its range of realizations can be well understood with sufficient historical analysis. In addition, this long-term data over many political economic regimes also allows a good understanding of the co-movement of the factor with other historically documented premiums such as the equity risk premium. This factor would take long positions in long-term government bonds and short positions in short-term Treasury securities, similar to the TERM factor of Section II.

2. Credit Risk

   This captures the compensation for the risk of default on debt instruments. For risky corporate securities this is effectively a conditional equity premium and thus is likely to be correlated to the equity premium, although it is conditional upon economic events, since defaults tend to be clustered in time. Like long-term government debt, corporate and risky government or agency debt is documented over long periods of time (see, e.g., Ibbotson and Sinquefield, 1976; Dimson, Marsh and Staunton, 2002), although the cross-sectional quality is limited. At present, an academic attempt is being pursued to extend substantially the historical period over which default risk at the individual bond level can be measured. This factor would take long positions in long-term corporate debt and short positions in corresponding long-term government bonds of the same maturity or duration. Corporate bonds of various credit qualities could be used as default rates of different classes of credit risk are imperfectly correlated, but in practice

bonds in lower risk credit classes tend to be more illiquid. This factor is similar to the CREDITAa, CREDITBaa, and CREDITHY factors used in Section II.

3. FX Carry Risk

   This factor captures the return to lending in high-interest currencies and borrowing in low-interest currencies. It corresponds to the FXCARRY factor used in Section II. This strategy has an implicit premium due to the risk of interest rate convergence, but its use has been documented only over the modern era for which currencies have traded in the capital markets, which is the post-1970's periods after the breakdown of Bretton Woods. Jurek (2007) and many other authors document large gains over multiple year horizons for carry trade strategies, but also point out that they are significantly negatively skewed, indicative of an insurance-like payoff. Given that the strategy is specific to the era of floating exchange rates, it may be difficult to assess the risk characteristics of the factor over periods of global segmentation. This factor could be folded into the next value-growth factor because it is a form of value-growth: it goes long currencies with high yields and short currencies with low yields.

4. Value-Growth Risk

   This is typically constructed from a long position in stocks with an unusually high book-to-market ratio and a short position in stocks with an unusually low book-to-market ratio. In practice, many other indications of "value" would be used including prices relative to other accounting variables (such as earnings, sales, forecasted and realized earnings, etc.) A conservative view of this factor is that it represents compensation for risk, however researchers lack a good understanding of the nature of that risk. Current theories for the latent risk in the value factor abound and include the hypothesis that the value premium is compensation for low-growth options by inflexible firms with assets in place during periods of distress (cf. Zhang , 2005), or time-varying sensitivities of value stocks that manifest themselves as changing betas in macroeconomic states (cf. Ang and Chen, 2007; Goetzmann, Watanabe and Watanabe, 2008). The standard behavioral explanation for the value premium is the over-extrapolation of past growth rates into the future (cf. Lakonishok, Shleifer and Vishny, 1994). Section I provides further references on the value effect. Close to a century of data exists for the value-growth premium (see Davis, Fama and French, 2000) via the CRSP database of U.S. equities. This factor corresponds to the VALGRTH factor of Section II. Besides being constructed more efficiently, one important difference between the MSCI data used in VALGRTH and an optimal factor to be constructed by NBIM is that value-growth would be pursued in fixed income, commodities, potentially foreign exchange (see above), and other asset classes in addition to international equities.

5. Small-Large Risk

   This is typically constructed from a long position in small-cap stocks and a short position in large cap stocks. The challenge for an investor the size of the Fund is that it may require holding large relative ownership stakes in the small companies of the world, which has already been partially addressed by the increase in the limit to 10% of the Fund's holdings in any single company. Section I provides references to the size effect. This factor corresponds to SMLG in Section II. Unlike the Value-Growth factor, this would be primarily restricted to equities.

6. Momentum Risk

   As with the value premium, this factor has a very strong historical premium but no clearly articulated risk. Jegadeesh and Titman (1993) documented positive returns to buying past winners and selling past losers over the post-1926 period. Rouwenhorst (1998) observed profitable momentum returns in international equity portfolios as well. Recent research by Chabot, Ghysels and Jagannathan (2008) demonstrate that momentum existed in the Victorian era, indicating that it is not limited to a recent window in capital market history. The scale of momentum returns adjusted for other factor exposures but before transactions costs is on the order of 90 basis points per month, although the returns from the 19th century are lower. These returns are not well-explained by value and size factors, nor are they related to booms and recessions in the macro-economy. To date, the most compelling explanations for momentum are behavioral and are based on investors under-reacting to news (cf. Barberis, Shleifer and Vishny, 1998). Cooper, Gutierrez and Hameed (2004) and Chabot, Ghysels and Jagannathan (2008) note that momentum profits depend on whether the stock market itself is in a bull or bear market. Momentum profits turn negative during an extended bear market – the implication being that bull markets attract naïve investors who slow price equilibration. This factor corresponds with MOM in Section II. However, an important difference is that the momentum factor constructed by NBIM would take advantage of momentum in every asset class, within each asset classes, and even across asset classes (cf Asness, Moskowitz and Pedersen, 2008). Most momentum factors, including MOM, have relatively high turnover. The momentum factor would also need to be constructed carefully to minimize turnover and transactions costs.

7. Volatility Risk

   The VOL factor used in Section II took only the difference between realized and implied volatility of the S&P500. This premium arises, among other reasons, because agents are averse to periods of increased volatility and are willing to pay high prices to hedge

against significant increases in market volatility – which typically also coincide with downward market moves (cf. Bakshi and Kapadia, 2003). A volatility factor should also trade the cross section of options: out-of-the money options are expensive compared to at-the-money options (cf. Coval and Shumway, 2001), there are differences in prices between options on indices and individual options on index components (called correlation trades, cf. Driessen, Maenhout and Vilkov, 2007), and exploit this relation in all asset classes (fixed income, currencies, commodities, etc.). This factor is not restricted to just derivatives as any relation between volatility and returns should be captured by a volatility risk factor. For example, Ang et al. (2009) show that stocks with low volatility have high returns in the global cross section of stock returns. It is also possible to include other higher moment trades like skewness and downside risk (cf. Harvey and Siddique, 2000; Ang, Chen and Xing, 2006). Since derivatives are leveraged, many derivatives have relatively large transactions costs, and counterparty risk is an issue for realized vs. implied variance swaps, this factor presents more challenges than the other factors.

As a guide to the risk-return premiums of some of these factors, the table below is taken from Morningstar's Ibbotson Associates, which has been documenting risk premiums with U.S. data for more than 25 years. It provides historical evidence on some of these premiums for U.S. markets from 1926 through 2008, as well as a t-value for testing whether a premium of zero can be rejected.

U.S. Data Series 1926-2008	Geometric Mean (%)	Arithmetic Mean (%)	Standard Deviation (%)	T Statistic
Inflation	3.01	3.10	4.23	6.68
US 30 Day TBill Inflation Adjusted	0.68	0.75	3.92	1.75
US Bond Horizon Premium	1.92	2.28	8.75	2.37
US Bond Default Premium	0.18	0.24	3.50	0.63
US Equity Risk Premium	5.70	7.72	20.28	3.47
SMB	1.52	2.28	12.76	1.62
HML	5.56	6.87	19.00	3.27
MOM	8.09	9.34	14.82	5.71

The second table is taken from the Dimson Marsh and Staunton [DMS] database which reports the performance of major asset classes for a wide range of global markets from 1900. DMS calculate premiums related to major factors including real interest rates, bond premiums maturity premiums and equity premiums.

	Geometric Mean (%)	Arithmetic Mean (%)	Standard Deviation (%)	T Statistic
DMS World Inflation	2.98	3.09	4.89	6.61
DMS World Real Bill TR	0.97	1.08	4.68	2.41
DMS World Real Bond TR	1.80	2.30	10.28	2.34
DMS World Maturity Premium	0.82	1.14	8.26	1.44
DMS World Equity Premium vs. Bills	4.22	5.65	17.06	3.46
DMS World Equity Premium vs. Bonds	3.37	4.64	15.64	3.10

These historical averages can likely be improved upon for purposes of forecasting expected returns and risk-return trade-offs of the factors because the factors ultimately constructed will be global and take advantage of further diversification across many different asset classes. The portfolios in both tables also do not take advantage of efficient portfolio construction techniques and use simple breakpoints to construct portfolios, rebalancing only at specified points in time, etc. More efficient portfolio techniques are available to the Fund. Also, it is important to point out that these risk premiums are not “additive.”

[This page is intentionally left blank]

References

ACKERMANN, C., R. McENALLY and D. RAVENSCRAFT, 1999. The Performance of Hedge Funds: Risk, Return and Incentives, Journal of Finance 54, 833-874.

AGARWAL, V., N. D. DANIEL and N. Y NAIK, 2004. Flows, Performance, and Managerial Incentives in Hedge Funds, SSRN Working Paper.

AGARWAL, V. and J. R. KALE, 2007. On the Relative Performance of Multi-Strategy and Funds of Hedge Funds, Journal of Investment Management, 5, 41-63.

AGARWAL, V., W. FUNG, Y. C. LOON and N. Y. NAIK, 2009. Risk and Return in Convertible Arbitrage: Evidence from the Convertible Bond Market, SSRN Working Paper.

AGARWAL, V. and N. Y. NAIK, 2000. Multi-Period Performance Persistence Analysis of Hedge Funds, Journal of Financial and Quantitative Analysis 35, 327-342.

AMIHUD, Y. and H. MENDELSON, 1991. Liquidity, Maturity, and the Yields on U.S. Treasury Securities, Journal of Finance 46, 1411-1425.

ANDREOU, E. and G. GHYSELS, 2002. Detecting Multiple Breaks in Financial Market Volatility Dynamics, Journal of Applied Econometrics 17, 579-600.

ANDREWS, D. W. K., 1993. Tests for Parameter Instability and Structural Change with Unknown Change Point, Econometrica 61, 821-856.

ANG, A. and G. BEKAERT, 2002. International Asset Allocation with Regime Shifts, Review of Financial Studies 15, 1137-1187.

ANG, A., and G. BEKAERT, 2007. Stock Return Predictability: Is it There? Review of Financial Studies 20, 651-707.

ANG, A., G. BEKAERT and M. WEI, 2008. The Term Structure of Real Rates and Expected Inflation, Journal of Finance 63, 797-849.

ANG, A. and J. CHEN, 2002. Asymmetric Correlations of Equity Portfolios, Journal of Financial Economics 63, 443-494.

ANG, A. and J. CHEN, 2007. CAPM Over the Long Run: 1926-2001, Journal of Empirical Finance 14, 1-40.

ANG, A., J. CHEN and Y. XING, 2006. Downside Risk, Review of Financial Studies 19, 1191-1239.

ANG, A., R. J. HODRICK, Y. XING and X. ZHANG, 2006. The Cross Section of Volatility and Expected Returns, Journal of Finance 61, 259-299.

ANG, A., R. J. HODRICK, Y. XING and X. ZHANG, 2009. High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence, Journal of Financial Economics 91, 1-23.

ANG, A. and M. PIAZZESI, 2003. A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, Journal of Monetary Economics 50, 745-787.

ARNOTT, R. D., J. HSU and P. MOORE, 2005. Fundamental Indexation, Financial Analysts Journal 61, March/April, 83-99.

ASNESS, C. S., T. J. MOSKOWITZ and L. H. PEDERSEN, 2008. Value and Momentum Everywhere, Working Paper, NYU.

AVRAMOV, D., L. BARRAS and R. KOSOWSKI, 2009. Understanding Hedge Fund Return Predictability: A Comprehensive Outlook Using a Fund by Fund Analysis, SSRN Working Paper.

BAILEY, W. B., H. LI and X. ZHANG, 2004. Hedge Fund Performance Evaluation: A Stochastic Discount Factor Approach, SSRN Working Paper.

BAKER, M. and J. WURGLER, 2006. Investor Sentiment and the Cross-Section of Stock Returns, Journal of Finance 61, 1645-1680.

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.

BAKSHI, G. and N. KAPADIA, 2003. Delta-Hedged Gains and the Negative Market Volatility Risk Premium, Review of Financial Studies 16, 527-566.

BANZ, R. W., 1981. The Relationship Between Return and Market Value of Common Stocks, Journal of Financial Economics 9, 3-18.

BARBERIS, N., A. SHLEIFER and R. W. VISHNY, 1998. A Model of Investor Sentiment, Journal of Financial Economics 49, 307-343.

BARDHAN, A., R. H. EDELSTEIN and D. TSANG, 2008. Global Economic-Financial Integration and Returns of Publicly Traded Real Estate Firms, Real Estate Economics 36, 285-311.

BARRAS, L., D. AVRAMOV and R. KOSOWSKI, 2009. Understanding Return Predictability: A Comprehensive Outlook Using a Fund by Fund Analysis, SSRN Working Paper.

BORTOLOTTI, B. V. FOTAK, W. L. MEGGINSON and W. MIRACKY, 2008. Sovereign Wealth Fund Investment Patterns and Performance, EFA 2009 Bergen Meetings Paper.

BASU, S., 1983. The Relationship Between Earnings Yield, Market Value and Return for NYSE Common Stocks, Journal of Financial Economics 12, 129-156.

BAUER, R., M. CREMERS and R. FREHEN, 2009. The Performance of U.S. Pension Funds, SSRN Working Paper.

BECHT, M., J. FRANKS, C. MAYER and S. ROSSI, 2009, Returns to Shareholder Activism: Evidence from a Clinical Study of the Hermes UK Focus Fund, Review of Financial Studies 22, 3093-3129.

BERK, J. and R. C. GREEN, 2004. Mutual Fund Flows and Performance in Rational Markets, Journal of Political Economy 112, 1269-1295.

BERNARD, V. L. and J. K. THOMAS, 1989. Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium? Journal of Accounting Research 27, 1-36.

BLACK, F. and M. SCHOLES, 1973. The Pricing of Options and Corporate Liabilities, Journal of Political Economy 81, 637-654.

BOLLEN, N. P. B. and V. K. POOL, 2008. Conditional Return Smoothing in the Hedge Fund Industry, Journal of Financial Economics 43, 267-298.

BOND, S. A., A. G. KAROLYI and A. B. SAUNDERS, 2003. International Real Estate Returns: A Multi-Factor, Multi-Country Approach, Real Estate Economics 31, 481-420.

BOUDOUKH, J. and R. WHITELAW, 1993. Liquidity as a Choice Variable: A Lesson from the Japanese Government Bond Market, Review of Financial Studies 6, 265-292.

BRAV, A., W. JIANG, F. PARTNOY and R. THOMAS, 2008. Hedge Fund Activism, Corporate Governance, and Firm Performance, Journal of Finance 63, 1729-1775.

BRINSON, G. P., L. R. HOOD and G. L. BEEBOWER, 1986. Determinants of Portfolio Performance, Financial Analysts Journal 42, 39-48.

BROWN, K. C. and C. TIU, 2009. Do Endowment Funds Select the Optimal Mix of Active and Passive Risk? Working Paper, UT Austin.

BROWN, K. C., L. GARLAPPI and C. TIU, 2009. Asset Allocation and Portfolio Performance: Evidence from University Endowment Funds, Working Paper, UT Austin.

BROWN, S. J. and W. N. GOETZMANN, 1995. Performance Persistence, Journal of Finance 50, 679-698.

BROWN, S. J., W. N. GOETZMANN and R. G. IBBOTSON, 1999. Offshore Hedge Funds: Survival and Performance 1989-1995, Journal of Business 72, 91-117.

BRUEGGEMAN W. B., A. H. CHEN and T. G. THIHODEAU, 1984. Real Estate Investment Funds: Performance and Portfolio Considerations, Real Estate Economics 12, 333-354.

BUSSE, J., A. GOYAL and S. WAHAL, 2009. Performance Persistence in Institutional Investment Management, forthcoming Journal of Finance.

CAMPBELL, J. Y. and J. H. COCHRANE, 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, Journal of Political Economy 107, 205-250.

CAMPBELL, J. Y., J. HILSCHER and J. SZILAGYI, 2008. In Search of Distress Risk. Journal of Finance 63, 2899-2939.

CAMPBELL, J. Y. and R. J. SHILLER, 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, Review of Financial Studies 1, 195-228.

CAMPBELL, J. Y. and L. VICEIRA, 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors, Oxford University Press, Oxford.

CAPOZZA D. R. and P. J. SEGUIN, 1996. Expectations, Efficiency, and Euphoria in the Housing Market, Regional Science and Urban Economics 26, 369-386.

CARHART, M. M., 1997. On Persistence in Mutual Fund Performance, Journal of Finance 52, 57-82.

CARTER, C. K. and R. KOHN, 1994. On Gibbs Sampling for State Space Models, Biometrika 81, 541-553.

CASE, K. E. and R. J. SHILLER, 1989. The Efficiency of the Market for Single-Family Homes, American Economic Review 79, 125-137.

CHABOT, B., E. GHYSELS, and R. JAGANNATHAN, 2008. Price Momentum in Stocks; Insights from Victorian Age Data, NBER Working Paper 14500.

CHEN, Y., W. E. FERSON and H. PETERS, 2009. Measuring the Timing Ability and Performance of Bond Mutual Funds, NBER Working Paper 15318.

CHEN, H., N. JU and J. MIAO, 2009. Dynamic Asset Allocation with Ambiguous Return Predictability, Working Paper, MIT.

CHEN, L., D. A. LESMOND and J. WEI, 2007. Corporate Yield Spreads and Bond Liquidity, Journal of Finance 62, 119-149.

CHEN, N. F., R. ROLL and S. A. ROSS, 1986. Economic Forces and the Stock Market, Journal of Business 59, 383-403.

CHEN, Z., W. STANZL and M. WATANABE, 2002. Price Impact Costs and the Limit of Arbitrage, Yale ICF Working Paper No. 00-66.

CHEVALIER, J. and G. ELLISON, 1997. Risk Taking by Mutual Funds as a Response to Incentives, Journal of Political Economy 105, 1167-1200.

CHRISTOPHERSON, J. A., W. E. FERSON and D. A. GLASSMAN, 1998. Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance, Review of Financial Studies 11, 111-42.

CIOCHETTI, B. and J. FISHER, 2002. The Characteristics of Commercial Real Estate Holding Period Returns (IRRs), Real Estate Research Institute Working Paper.

COCHRANE, J. H., 1999. Portfolio Advice for a Multifactor World, Economic Perspectives Federal Reserve Bank of Chicago 23, 59-78.

COCHRANE, J. H., 2005. The Risk and Return of Venture Capital, Journal of Financial Economics 75, 3-52.

COHEN, L. and A. FRAZZINI, 2008. Economic Links and Predictable Returns, Journal of Finance 63, 1977-2011.

COLLIN-DUFRESNE, P, GOLDSTEIN, R. S. and J. S. MARTIN, 2001. The Determinants of Credit Spread Changes, Journal of Finance 56, 2177-2207.

CONROY, R. M. and R. S. HARRIS. 2007. How Good are Private Equity Returns? Journal of Applied Corporate Finance 19, 96-108.

COOPER, M. and D. H. DOWNS, 1999. Real Estate Securities and a Filter-based, Short-term Trading Strategy, Journal of Real Estate Research 18, 313-334.

COOPER, M. J., R. C. GUTIERREZ and A. HAMEED, 2004. Market States and Momentum, Journal of Finance 59, 1345-1365.

COVAL, J. D. and T. SHUMWAY, 2001. Expected Option Returns, Journal of Finance 56, 983-1009.

COWLES, A., 1933. Can Stock Market Forecasters Forecast? Econometrica 1, 309-324.

COWLES, A. and H. E. JONES, 1937. Some A Posteriori Probabilities in Stock Market Action, Econometrica 5, 280–294.

CREMERS, M., A. PETAJISTO and E. ZITZEWITZ, 2008. Should Benchmark Indices Have Alpha? Revisitng Performance Evaluation? SSRN Working Paper.

DANIEL, K. and S. TITMAN, 2006. Market Reactions to Tangible and Intangible Information. Journal of Finance 61, 1605-1643.

DARRAT, A and J. GLASCOCK, 1993. On the Real Estate Market Efficiency. Journal of Real Estate Finance and Economics 7, 55-72.

DAVIS, J. L., E. F. FAMA and K. R. FRENCH, 2000. Characteristics, Covariances, and Average Returns: 1929 to 1997, Journal of Finance 55, 389-406.

DE BONDT, W. F. M. and R. THALER, 1985. Does the Stock Market Overreact? Journal of Finance 40, 793-805.

DELLAVIGNA, S. and J. M. POLLET, 2009. Investor Inattention and Friday Earnings Announcements, Journal of Finance 64, 709-749.

DEWENTER, K. L., X. HAN and P. H. MALATESTA, 2009. Firm Values and Sovereign Wealth Fund Investments, SSRN Working Paper.

DIMSON, E., P. MARSH and M. STAUNTON, 2002. Triumph of the Optimists: 101 years of Global Investment Returns, Princeton University Press, Princeton.

DRIESSEN, J., P. MAENHOUT and G. VILKOV, 2007. The Price of Correlation Risk: Evidence from Equity Options, forthcoming Journal of Finance.

DUARTE, J., F. A. LONGSTAFF and F. YU, 2005. Risk and Return in Fixed-Income Arbitrage: Nickels in Front of a Steamroller? Review of Financial Studies 20, 769-811.

DYBVIG, P. H. and S. A. ROSS, 1985. Differential Information and Performance Measurement Using a Security Market Line, Journal of Finance 40, 383-399.

EDELSTEIN, R. H., A. BARDHAN and D. TSANG , 2007. Global Financial Integration and Real Estate Security Returns, SSRN Working Paper.

EDMANS, A., D. GARCIA and O. NORLI, 2007. Sports Sentiment and Stock Returns. Journal of Finance 62, 1967-1998.

ELTON, E., M. GRUBER and C. BLAKE, 1993. The Performance of Bond Mutual Funds, Journal of Business 66, 371-403.

ELTON, E., M. GRUBER and C. BLAKE, 2003. Incentive Fees and Mutual Funds, Journal of Finance 58, 779-804.

FAMA, E. F., 1965. The Behavior of Stock-Market Prices, Journal of Business 38, 34–105.

FAMA, E. F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work, Journal of Finance 25, 383–41.

FAMA, E. F., 1991. Efficient Capital Markets: II, Journal of Finance 46, 1575–1617.

FAMA, E. F. and R. R. BLISS, 1987. The Information in Long-Maturity Forward Rates, American Economic Review 77, 680-692.

FAMA, E. F. and K. R. FRENCH, 1988a. Permanent and Temporary Components of Stock Prices, Journal of Political Economy 96, 246–273.

FAMA, E. F. and K. R. FRENCH, 1988b. Dividend Yields and Expected Stock Returns, Journal of Financial Economics 22, 3-25.

FAMA, E. F. and K. R. FRENCH, 2008. Mutual Fund Performance, SSRN Working Paper.

FERNANDES, N. G., 2009. Sovereign Wealth Funds: Investment Choices and Implications Around the World, EFA 2009 Bergen Meetings Paper.

FERSON, W. E. and C. R. HARVEY, 1991. The Variation of Economic Risk Premiums, Journal of Political Economy 99, 385-415.

FERSON, W. E., T. R. HENRY and D. KISGEN, 2006. Evaluating Government Bond Fund Performance with Stochastic Discount Factors, Review of Financial Studies 19, 423-456.

FERSON, W. E. and K. KANG, 2002. Conditional Performance Measurement Using Portfolio Weights: Evidence for Pension Funds, Journal of Financial Economics 65, 249-282.

FERSON, W. E. and R. W. SCHADT, 1996. Measuring Fund Strategy and Performance in Changing Economic Conditions, Journal of Finance 51, 425-461.

FRANZONI, F., E. NOWAK and L. PHALIPPOU, 2009. Private Equity and Liquidity Risk, SSRN Working Paper.

FRENCH, K. R., 1988. Crash-Testing the Efficient Market Hypothesis, NBER Macroeconomics Annual 3, 277-285.

FRENCH, K. R., 2008. Presidential Address: The Cost of Active Investing, Journal of Finance 63, 1537-1573.

FUNG, W. and D. A. HSIEH, 1997. Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds, Review of Financial Studies 10, 275-302.

FUNG, W. and D. A. HSIEH, 2002. Risk in Fixed-Income Hedge Fund Styles, Journal of Fixed Income 12, 6-27.

GALLO, J. G., L. J. LOCKWOOD and M. RODRIGUEZ , 2006. Differentiating CREF Performance. Real Estate Economics 34, 173-209.

GALLO, J. G., L. J. LOCKWOOD and R. C. RUTHERFORD, 2000. Asset Allocation and the Performance of Real Estate Mutual Funds, Real Estate Economics 28, 165-185.

GATEV, E., W. N. GOETZMANN and K. G. ROUWENHORST, 2006. Pairs Trading: Performance of a Relative-Value Arbitrage Rule, Review of Financial Studies 19, 797-827.

GATZLAFF, D. H. and D. TITIROGLU, 1995. Real Estate Market Efficiency: Issues and Evidence, Journal of Real Estate Literature 3, 157-189.

GETMANSKY, M., A. W. LO and I. MAKAROV, 2004. An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns, Journal of Financial Economics 74, 529-609.

GIBSON, R. and S. WANG, 2009. Hedge Fund Alphas: Do They Reflect Managerial Skills or Mere Compensation for Liquidity Risk-Bearing? SSRN Working Paper.

GOMPERS, P. A. and A. METRICK, 2001. Institutional Investors and Equity Prices, Quarterly Journal of Economics 116, 229-259.

GOETZMANN, W. N., 1993. The Single Family Home in the Investment Portfolio, Journal of Real Estate Finance and Economics 6, 201-222.

GOETZMANN, W. N. and J. D. FISHER, 2005. Performance of Real Estate Portfolios: A Simulation Approach, Journal of Portfolio Management 31, 32-45.

GOETZMANN, W. N. and R. G. IBBOTSON, 1990. The Performance of Real Estate as an Asset Class, Journal of Applied Corporate Finance 3, 65-76.

GOETZMANN, W. N., J. INGERSOLL, M. SPIEGEL and I. WELCH, 2008. Portfolio Performance Manipulation and Manipulation-Proof Performance Measures, Review of Financial Studies 20, 1503-1546.

GOETZMANN, W. N. and P. JORION, 1993. Testing the Predictive Power of Dividend Yields, Journal of Finance 48, 663-679.

GOETZMANN, W. N. and P. JORION, 1995. A Longer Look at Dividend Yields, Journal of Business 68, 483-508.

GOETZMANN, W. N. and M. MASSA, 2003. Index Funds and Stock Market Growth, Journal of Business 76, 1-28.

GOETZMANN, W. N., A. WATANABE and M. WATANABE, 2009. Investor Expectations, Business Conditions, and the Pricing of Beta-Instabiliaty Risk, SSRN Working Paper.

GOMPERS, P. A. and A. METRICK, 2001. Institutional Investors and Equity Prices, Quarterly Journal of Economics 116, 229-259.

GOYAL, A. and S. WAHAL, 2008. The Selection and Termination of Investment Management Firms by Plan Sponsors, Journal of Finance 63, 1805-1847.

GRIFFIN, J. M. and J. XU, 2009. How Smart are the Smart Guys? A Unique View from Hedge Fund Stock Holdings, Review of Financial Studies 22, 2531-2570.

GRINBLATT, M., S. TITMAN and R. WERMERS, 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior, American Economic Review 85, 1088-1105.

GRINOLD, R. C. and R. KAHN, 2000. Active Portfolio Management: A Quantitative Approach for Providing Superior Returns and Controlling Risk, McGraw Hill, New York.

GROSSMAN, S. J. and J. E. STIGLITZ, 1976. Information and Competitive Price Systems, American Economic Review 66, 246–253.

GRUBER, M. J., 1996. Another Puzzle: The Growth in Actively Managed Mutual Funds, Journal of Finance 51, 783-810.

GURKAYNAK, R. S., B. SACK and J. H. WRIGHT, 2007. The U.S. Treasury Yield Cruve: 1961 to the Present, Journal of Monetary Economics 54, 2291-2304.

GUTIERREZ, R. C. and E. K. KELLEY, 2008. The Long-Lasting Momentum in Weekly Returns. Journal of Finance 63, 415-447.

HAMILTON, J. D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica 57, 357-384.

HAND, J. R., 1990. A Test of the Extended Functional Fixation Hypothesis, Accounting Review 65, 740-763.

HANSEN, L. P. and R. JAGANNATHAN, 1997. Assessing Specification Errors in Stochastic Discount Factor Models, Journal of Finance 52, 557-590.

HARVEY, C. R. and A. SIDDIQUE, 2000. Conditional Skewness in Asset Pricing Tests, Journal of Finance 55, 1263-1295.

HUANG, M., J. M. QUIGLEY and S. E. WOODWARD, 2006. An Index For Venture Capital, 1987-2003, Contributions to Economic Analysis and Policy 4, article 13.

HUBERMAN, G. and S. KANDEL, 1987. Mean-Variance Spanning. Journal of Finance 42, 873-888.

HUIJ, J. and J. DERWALL, 2008. 'Hot Hands' in Bond Funds, Journal of Banking and Finance 32, 559-572.

IBBOTSON, R. G. and P. CHEN, 2003. Long-Run Stock Returns: Participating in the Real Economy, Financial Analysts Journal 59, Jan/Feb, 88-98.

IBBOTSON, R. G. and R. A. SINQUEFIELD, 1976. Stocks, Bonds, Bills and Inflation: Year-by-Year Historical Returns (1926-1974), Journal of Business 49, 11-47.

JAGANNATHAN, R., A. MALAKHOV and D. NOVIKOV, 2006. Do Hot Hands Exist Among Hedge Fund Managers? An Empirical Evaluation, NBER Working Paper.

JAGANNATHAN, R. and Z. WANG, 1996. The Conditional CAPM and the Cross Section of Expected Returns, Journal of Finance 51, 3-53.

JEGADEESH, N. and S. TITMAN, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, Journal of Finance 48, 65-91.

JENSEN, M. C., 1968. The Performance of Mutual Funds in the Period 1945-1964, Journal of Finance 23, 389–416.

JUREK, J. W., 2007. Crash-Neutral Currency Carry Trades, Working Paper, Princeton University.

KAN, R. and G. ZHOU, 2008. Tests of Mean-Variance Spanning, Working Paper Washington University in St. Louis.

KANDEL, E. and R. F. STAMBAUGH, 1987. On Correlations and Inferences about Mean-Variance Efficiency, Journal of Financial Economics 18, 61-90.

KANIEL, R., G. SAAR and S. TITMAN, 2008. Individual Investor Trading and Stock Returns, Journal of Finance 63, 273-310.

KAPLAN, S. N. and A. SCHOAR, 2005. Private Equity Performance: Returns, Persistence, and Capital Flows, Journal of Finance 60, 1791-1823.

KEIM, D. B., 1983. Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence, Journal of Financial Economics 12, 13-32.

KHORANA, A., H. SERVAES and L. WEDGE, 2007. Portfolio Manager Ownership and Fund Performance, Journal of Financial Economics 85,179-204.

KNILL, A. M., B. S. LEE and N. MAUCK, 2009. 'Sleeping with the Enemy' or 'An Ounce of Prevention': Sovereign Wealth Fund Investments and Market Destabilization, SSRN Working Paper.

KORAJCYZK, R. A. and R. SADKA, 2004. Are Momentum Profits Robust to Trading Costs? Journal of Finance 59, 1039-1082.

KORAJCYZK, R. A. and R. SADKA, 2008. Pricing the Commonality Across Alternative Measure of Liquidity, Journal of Financial Economics 87, 45-72.

KORTEWEG, A. and M. SORENSEN, 2008. Estimating Risk and Return of Infrequently-Traded Assets: A Bayesian Selection Model of Venture Capital, Working Paper, Columbia University.

KOSOWSKI, R., A. G. TIMMERMANN, R. WERMERS, and H. WHITE, 2006. Can Mutual Fund 'Stars' Really Pick Stocks? New Evidence from a Bootstrap Analysis, Journal of Finance 61, 2551-2595.

KOSOWSKI, R., N. Y. NAIK and M. TEOH, 2007. Do Hedge Funds Deliver Alpha? A Bayesian and Bootstrap Analysis, SSRN Working Paper.

KOTTER, J. and U. LEL, 2009. Friends or Foes? Target Selection Decisions and Performance Effects of Sovereign Wealth Funds, SSRN Working Paper.

KRISHNAMURTHY, A., 2002. The New Bond/Old Bond Spread, Journal of Financial Economics 66, 463-506.

KUMAR, A. and C. M. C. LEE, 2006. Retail Investor Sentiment and the Cross-Section of Stock Returns, Journal of Finance 61, 2451-2486.

KUO, C. L., 1996. Serial Correlation and Seasonality in the Real Estate Market, Journal of Real Estate Finance and Economics 12, 139-62.

LAKONISHOK, J., A. SHLEIFER and R. W. VISHNY, 1994. Contrarian Investment, Extrapolation, and Risk, Journal of Finance 49, 1541-1578.

LERNER, J., A. SCHOAR and J. WANG, 2008. Secrets of the Academy: The Drivers of University Endowment Success, Journal of Economic Perspectives 22, 207-222.

LERNER, J., A. SCHOAR and W. WONGSUNWAI, 2007. Smart Institutions, Foolish Choice: The Limited Partner Performance Puzzle, Journal of Finance 62, 731-764.

LIN, C. Y. and K. YUNG, K., 2004. Real Estate Mutual Funds: Performance and Persistence. Journal of Real Estate Research 26, 69-93.

LJUNGQVIST, A., Y. V. HOCHBERG and A. VISSING-JORGENSEN, 2009. Informational Hold-Up and Performance Persistence in Venture Capital, NYU Working Paper.

LO, A., 2007. What Happened to the Quants in August 2007? Journal of Investment Management 5, 5–54.

MANDELBROT, B., 1963. The Variation of Certain Speculative Prices, Journal of Business 36, 394-419.

MALKIEL, B. G. and A. SAHA, 2005. Hedge Funds: Risks and Return, Financial Analysts Journal 61, 80-88.

MAMAYSY, H. and M. SPIEGEL, 2001. A Theory of Mutual Funds: Optimal Fund Objectives and Industry Organization, SSRN Working Paper.

MAMAYSKY, H., M. SPIEGEL and H. ZHANG, 2008. Estimating the Dynamics of Mutual Fund Alphas and Betas, Review of Financial Studies 21, 233-264.

MARSH, T. A. and R. C. MERTON, 1986. Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices, American Economic Review 76, 483–498.

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, 867-887.

METRICK, A. and A. YASUDA, 2009. The Economics of Private Equity Funds, forthcoming Review of Financial Studies.

MITCHELL, M. and T. PULVINO, 2001. Characteristics of Risk and Return in Risk Arbitrage, Journal of Finance 56, 2135-2175.

MONETA, F., 2009. Measuring Bond Mutual Fund Performance with Portfolio Characteristics, Boston College Working Paper.

MOSKOWITZ, T. J. and A. VISSING-JORGENSEN, 2002. The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle? American Economic Review 92, 745-778.

OKUNEV, J. and D. R. WHITE, 2002. Hedge Fund Risk Factors and Value at Risk of Credit Trading Strategies, SSRN Working Paper.

PAGLIARI, J. L., K. A. SCHERER and R. T. MONOPOLI, 2005. Public versus Private Real Estate Equities: A More Refined, Long-Term Comparison, Real Estate Economics 33, 147-187.

PAN, J. and K. J. SINGLETON, 2008. Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads, Journal of Finance 63, 2345-2384.

PASTOR, L. and R. F. STAMBAUGH, 2002. Investing in Equity Mutual Funds, Journal of Financial Economics 63, 351-380.

PASTOR, L. and R. F. STAMBAUGH, 2003. Liquidity Risk and Expected Stock Returns, Journal of Political Economy 111, 642-685.

PHALIPPOU, L. and O. GOTTSCHALG, 2009. The Performance of Private Equity Funds, Review of Financial Studies 22, 1747-1776.

POTERBA, J. M. and L. H. SUMMERS, 1988. Mean Reversion in Stock Prices: Evidence and Implications, Journal of Financial Economics 22, 27–59.

RAYMOND, K. and G. ZHOU, 2008. Tests of Mean-Variance Spanning, SSRN Working Paper.

REINGANUM, M. R., 1983. The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects, Journal of Financial Economics 12, 89-104.

RICHARDSON, M., 1993. Temporary Components of Stock Prices: A Skeptic's View, Journal of Business and Economic Statistics 11, 199-207.

ROBERT, C. P. and G. CASELLA, 1999. Monte Carlo Statistical Methods. Springer-Verlag.

ROBERTS, H., 1967. Statistical versus Clinical Prediction of the Stock Market, Unpublished manuscript.

ROLL, R., 1977. A Critique of the Asset Pricing Theory Tests, Journal of Financial Economics 4, 129-176.

ROMANO, R., 2001. Less is More: Making Institutional Investor Activism a Valuable Mechanism of Corporate Governance, Yale Journal of Regulation 18, 174-251.

RODRIGUEZ, J., 2007. A Critical Look at the Forecasting Ability of Real Estate Mutual Fund Managers, Journal of Real Estate Portfolio Management 13, 99-106.

ROSENBERG, B., K. REID and R. LANSTEIN, 1985. Pervasive Evidence of Market Inefficiency, Journal of Portfolio Management 11, 9-16.

ROSENTHAL, R., 1979. The “File Drawer Problem” and Tolerance for Null Results, Psychological Bulletin 86, 638-641.

ROSS, S. A., 1973. The Economic Theory of Agency: The Principal's Problem, American Economic Review 63, 134-139.

ROSS, S. A., 1976. The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory, 13, 341-360.

ROUWENHORST, K. G., 1998. International Momentum Strategies, Journal of Finance 53, 267-284.

SAMUELSON, P. A., 1965. Proof that Properly Anticipated Prices Fluctuate Randomly, Industrial Management Review 6, 41–49.

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

SCHWARZ, K., 2009. Mind the Gap: Disentangling Credit and Liquidity in Risk Spreads. Working Paper, Wharton.

SHARPE, W. F., 1966. Mutual Fund Performance, Journal of Business 39, 119-138.

SHARPE, W. F., 1992. Asset Allocation: Management Style and Performance Analysis, Journal of Portfolio Management, Winter, 7-19.

SHILLER, R. J., 1981, Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? American Economic Review 71, 421-436.

SHLEIFER, A., 1986. Do Demand Curves for Stocks Slope Down? Journal of Finance 41, 579–590.

SHLEIFER, A. and R. W. VISHNY, 1997. The Limits of Arbitrage, Journal of Finance 52, 35 -55.

SIAS, R. W., L. T. STARKS and S. TITMAN, 2006. Changes in Institutional Ownership and Stock Returns: Assessment and Methodology, Journal of Business 79, 2869-2910.

SIRRI, E. R. and P. TUFANO, 1998. Costly Search and Mutual Fund Flows, Journal of Finance 53, 1589-1622.

SLOAN, R. G., 1996. Do Stock Prices Reflect Information in Accruals and Cashflows about Future Earnings? Accounting Review 71, 289-315.

SPIEGEL, M. I. and H. MAMAYSKY, 2001. A Theory of Mutual Funds: Optimal Fund Objectives and Industry Organization, Yale ICF Working Paper No. 00-50.

STATTMAN, D., 1980. Book Values and Stock Returns, The Chicago MBA: A Journal of Selected Papers 4, 25-45.

STOCK, J. H. and M. W. WATSON, 2002. Macroeconomic Forecasting Using Diffusion Indices, Journal of Business and Economic Statistics 20, 147-162.

STEWART, S., J. NEUMAN, C. KNITTEL and J. HEISLER, 2009. Absence of Value: An Analysis of Investment Allocation Decisions by Institutional Plan Sponsors, Financial Analysts Journal 65, 1-17.

SUMMERS, L. H., 1986. Does the Stock Market Rationally Reflect Fundamental Values? Journal of Finance 41, 591–601

SWENSEN, D. F., 2000. Pioneering Portfolio Management: An Unconventional Approach to Institutional Investment, The Free Press, New York.

TITMAN, S. and A. WARGA, 1986. Risk and the Performance of Real Estate Investment Trusts: A Multiple Index Approach, Real Estate Economics 14, 414-431.

TEOH, M, 2009. Does Size Matter in the Hedge Fund Industry? SSRN Working Paper.

TONKS, I., 2005. Performance Persistence of Pension-Fund Managers, Journal of Business 78, 1917-1942.

VAYANOS, D. and J. L. VILA, 2009. A Preferred-Habitat Model of the Term Structure of Interest Rates, Working Paper, LSE.

WACHTER, J. A. and M. WARUSAWITHARNA, 2009. Predictable Returns and Asset Allocation: Should a Skeptical Investor Time the Market? Journal of Econometrics 148, 162-178.

WELCH, I. and A. GOYAL, 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, Review of Financial Studies 4, 1455-1508.

WERMERS, R. R., 2000. Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses. Journal of Finance 55, 1655-1703.

WURGLER, J. and E. ZHURAVSKAYA, 2002. Does Arbitrage Flatten Demand Curves for Stocks? Journal of Business 75, 583-608.

ZHANG, L., 2005. The Value Premium, Journal of Finance 60, 67-103.

[This page is intentionally left blank]

Tables[37]

[37] Returns for the overall Fund (Table 3), the fixed income and equity subclasses (Table 4), and internal NBIM returns (Tables 5 and 6) are gross returns. Returns for external funds (Table 7) are net of fees.

Table 1: Definitions

	Notation	Definition	Source
Panel A: Returns			
Fund Return	Ret	Total return of the fund. These are monthly gross returns for the overall Fund, the fixed income and equity subclass, and internal NBIM returns. These are net returns for external funds.	NBIM
Benchmark Return	Bmk	Return of the fund benchmark.	NBIM
Active Return	Ret - Bmk	Difference between the fund return and the benchmark return. Also called an excess return or relative return.	NBIM or own calculations
Alpha	Alpha	Mean of the active return.	Own calculations
Residual Return	Ret - β Bmk	Difference between the fund return and a beta-adjusted (or leveraged) return on the benchmark. These are the residuals in a regression of the fund return on the benchmark return.	Own calculations

Table 1 Continued: Definitions

	Notation	Definition	Source
Panel B: Systematic Factors			
Term	TERM	Difference in returns on the total return BarCap US Treasury 20+ yr index and the total return BarCap US Treasury Bill 1-3 mth index	Morningstar Encorr
Credit Aa	CREDITAa	Difference in returns on the total return BarCap US Corporate Aa Long Maturity index and the	Morningstar Encorr
Credit Baa	CREDITBaa	Difference in returns on the total return US Corporate Baa Long Maturity index and the total return BarCap US Corporate Aa Long Maturity	Morningstar Encorr
Credit High Yield	CREDITHY	Difference in returns on the total return BarCap US Corporate High Yield Caa index and the total return BarCap US Corporate Baa Long Maturity Baa index	Morningstar Encorr
FX Carry	FXCARRY	Difference in returns between currency returns on the top three G10 currencies with the highest short-term yields and the bottom three G10 currencies with the lowest short-term yields	Bloomberg (FXFB)
Liquidity	LIQUIDITY	The negative of innovations in the on-the-run/off-the-run spread on 10-year US Treasury bonds	Federal Reserve Board
Value/Growth	VALGRTH	Difference in returns between global "value" stocks and global "growth" stocks computed using MSCI world indices	Datastream (MSVWLD$ and MSGWLD$)
Small/Large	SMLG	Difference in returns between global small cap stodcks and global large cap stocks computed using MSCI all country indices	Datastream (MSSAWF$ and MSLAWF$)
Momentum	MOM	Difference in returns between US stocks with past high returns and US stocks with past low returns	Kenneth French at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
Volatiltiy	VOL	Return on a variance swap between implied and realized volatility on the S&P500 in excess of LIBOR	Bloomberg (MLHFEV1E)

Table 2: Systematic Factor Correlations

	TERM	CREDITAa	CREDITBaa	CREDITHY	FXCARRY	LIQUIDITY	VALGRTH	SMLG	MOM	VOL
TERM	1.00	-0.49	-0.45	-0.63	0.08	0.06	0.02	0.04	0.16	-0.11
CREDITAa	-0.49	1.00	0.34	0.45	0.29	0.18	0.00	0.20	-0.35	0.42
CREDITBaa	-0.45	0.34	1.00	0.56	0.19	0.42	-0.05	0.23	-0.25	0.58
CREDITHY	-0.63	0.45	0.56	1.00	0.23	0.30	0.09	0.23	-0.37	0.34
FXCARRY	0.08	0.29	0.19	0.23	1.00	0.50	0.07	0.22	-0.14	0.49
LIQUIDITY	0.06	0.18	0.42	0.30	0.50	1.00	0.01	0.37	-0.25	0.60
VALGRTH	0.02	0.00	-0.05	0.09	0.07	0.01	1.00	0.09	-0.38	-0.06
SMLG	0.04	0.20	0.23	0.23	0.22	0.37	0.09	1.00	-0.04	0.25
MOM	0.16	-0.35	-0.25	-0.37	-0.14	-0.25	-0.38	-0.04	1.00	-0.21
VOL	-0.11	0.42	0.58	0.34	0.49	0.60	-0.06	0.25	-0.21	1.00

Table 3: Overall Fund Returns

(Monthly gross returns)

Panel A: Variance Attribution

	Full Sample	Pre-2008
Benchmark Return	99.1%	99.7%
Active Return	0.9%	0.3%
Total Return	100.0%	100.0%

Panel B: Active and Residual Returns

		Active Returns				Residual Returns			
		Mean	Autocorr	Skew	Reward-to-Risk Ratio	Alpha	Beta	Skew	Reward-to-Risk Ratio
Full Sample	Coeff	0.02	0.55	-2.41	0.07	0.01	1.02	-2.57	0.04
	P-value	0.56				0.72			
Pre-2008	Coeff	0.03	0.16	-0.30	0.15	0.03	1.01	-0.40	0.25
	P-value	0.01				0.01			

Panel C: Partial Correlations of Active Returns with Systematic Factors

	Full Sample		Pre-2008	
	Partial Corr	P-value	Partial Corr	P-value
TERM	-0.17	0.05	-0.25	0.01
CREDITAa	0.40	0.00	0.02	0.81
CREDITBaa	-0.40	0.00	-0.16	0.09
CREDITHY	0.02	0.83	-0.01	0.90
FXCARRY	0.07	0.44	0.13	0.16
LIQUIDITY	0.31	0.00	0.25	0.01
VALGRTH	-0.35	0.00	-0.45	0.00
SMLG	0.21	0.01	0.44	0.00
MOM	-0.03	0.76	0.07	0.48
VOL	0.39	0.00	0.24	0.01

Table 3 Continued: Overall Fund Returns

Panel D: Correlations of Active Returns with Hedge Fund Index Returns

		Corr	P-value
Relative Value	RV Fixed Income Asset Backed	-0.21	0.26
	RV Fixed Income Convertible Arbitrage	0.25	0.05
	RV Fixed Income Corporate	0.04	0.65
	RV Multistrategy	-0.02	0.83
	RV Yield Alternatives	-0.01	0.89
	RV Index	-0.04	0.67
Event Driven	ED Distressed/Restructuring	0.05	0.55
	ED Merger Arbitrage	-0.21	0.27
	ED Index	0.08	0.38
Equity Hedge	EH Equity Market Neutral	-0.28	0.15
	EH Quantitative Directional	0.20	0.00
	EH Index	0.21	0.00
Macro	Macro Systematic Diversified	-0.25	0.29
	Macro Index	-0.23	0.32

Table 4: Fixed Income and Equity Returns

(Monthly gross returns)

Panel A: Variance Attribution

	Fixed Income		Equity	
	Full Sample	Pre-2008	Full Sample	Pre-2008
Benchmark Return	97.1%	99.8%	99.7%	99.7%
Active Return	2.9%	0.2%	0.3%	0.3%
Total Return	100.0%	100.0%	100.0%	100.0%

Panel B: Active and Residual Returns

		Active Returns				Residual Returns			
		Mean	Autocorr	Skew	Reward-to-Risk Ratio	Alpha	Beta	Skew	Reward-to-Risk Ratio
Fixed Income									
Full Sample	Coeff	0.00	0.72	-0.80	0.00	0.02	0.95	-0.10	0.06
	P-value	0.98				0.64			
Pre-2008	Coeff	0.01	0.11	-2.51	0.08	0.01	0.99	-2.29	0.10
	P-value	0.45				0.29			
Equity									
Full Sample	Coeff	0.05	0.17	-0.83	0.17	0.04	1.02	-0.88	0.16
	P-value	0.08				0.09			
Pre-2008	Coeff	0.06	0.10	0.52	0.24	0.06	1.01	0.37	0.22
	P-value	0.02				0.02			

Panel C: Partial Correlations of Active Returns with Systematic Factors

	Fixed Income		Equity	
	Partial Corr	P-value	Partial Corr	P-value
TERM	-0.21	0.01		
CREDITAa	0.35	0.00		
CREDITBaa	-0.33	0.00		
CREDITHY	-0.01	0.92		
FXCARRY	-0.04	0.64		
LIQUIDITY	0.35	0.00		
VALGRTH			-0.56	0.00
SMLG			0.41	0.00
MOM			0.02	0.80
VOL	0.37	0.00	0.28	0.00

Table 4 Continued: Fixed Income and Equity Returns

Panel D: Correlations of Active Returns with Hedge Fund Index Returns

		Fixed Income		Equity	
		Corr	P-value	Corr	P-value
Relative Value	RV Fixed Income Asset Backed	-0.21	0.23	-0.12	0.31
	RV Fixed Income Convertible Arbitrage	0.22	0.09	0.19	0.06
	RV Fixed Income Corporate	0.02	0.81	0.03	0.66
	RV Multistrategy	-0.08	0.46	0.05	0.50
	RV Yield Alternatives	-0.05	0.53	0.03	0.68
	RV Index	-0.09	0.36	0.04	0.56
Event Driven	ED Distressed/Restructuring	0.00	1.00	0.09	0.28
	ED Merger Arbitrage	-0.28	0.12	-0.01	0.96
	ED Index	-0.02	0.76	0.16	0.08
Equity Hedge	EH Equity Market Neutral	-0.33	0.09	-0.07	0.57
	EH Quantitative Directional	0.01	0.90	0.35	0.00
	EH Index	0.04	0.54	0.34	0.00
Macro	Macro Systematic Diversified	-0.37	0.08	0.06	0.74
	Macro Index	-0.35	0.09	0.03	0.87

Table 5: Internal Fixed Income Returns

(Monthly gross returns)

Panel A: Variance Attribution

	Full Sample	Pre-2008
Benchmark Return	98.4%	99.8%
Active Return	1.6%	0.2%
Total Return	100.0%	100.0%

Panel B: Active and Residual Returns

		Active Returns				Residual Returns			
		Mean	Autocorr	Skew	Reward-to-Risk Ratio	Alpha	Beta	Skew	Reward-to-Risk Ratio
Full Sample	Coeff	0.02	0.67	0.22	0.05	0.04	0.96	0.60	0.11
	P-value	0.78				0.51			
Pre-2008	Coeff	0.00	-0.19	-1.24	0.03	0.01	0.99	-0.92	0.06
	P-value	0.80				0.54			

Panel C: Partial Correlations of Active Returns with Systematic Factors

	Partial Corr	P-value
TERM	-0.08	0.36
CREDITAa	0.38	0.00
CREDITBaa	-0.38	0.00
CREDITHY	0.01	0.93
FXCARRY	-0.08	0.34
LIQUIDITY	0.34	0.00
VOL	0.32	0.00

Table 5 Continued: Internal Fixed Income Returns

Panel D: Correlations of Active Returns with Hedge Fund Index Returns

		Corr	P-value
Relative Value	RV Fixed Income Asset Backed	-0.21	0.20
	RV Fixed Income Convertible Arbitrage	0.22	0.12
	RV Fixed Income Corporate	0.00	0.96
	RV Multistrategy	-0.08	0.35
	RV Yield Alternatives	-0.10	0.19
	RV Index	-0.10	0.27
Event Driven	ED Distressed/Restructuring	-0.02	0.79
	ED Merger Arbitrage	-0.27	0.09
	ED Index	-0.03	0.62
Equity Hedge	EH Equity Market Neutral	-0.32	0.07
	EH Quantitative Directional	0.00	0.91
	EH Index	0.02	0.62
Macro	Macro Systematic Diversified	-0.35	0.06
	Macro Index	-0.34	0.07

Table 6: Internal Equity Returns

(Monthly gross returns)

Panel A: Variance Attribution

	AS	
	Full Sample	Pre-2008
Benchmark Return	100.0%	100.0%
Active Return	0.0%	0.0%
Total Return	100.0%	100.0%

Panel B: Active Returns

		Active Returns				Residual Returns			
		Mean	Autocorr	Skew	Reward-to-Risk Ratio	Alpha	Beta	Skew	Reward-to-Risk Ratio
Internal									
Full Sample	Coeff	0.05	0.25	-3.35	0.20				
	P-value	0.16							
Pre-2008	Coeff	0.09	0.08	0.26	0.58				
	P-value	0.00							
AS									
Full Sample	Coeff	0.00	0.28	-3.05	0.05	0.00	1.01	-2.16	0.04
	P-value	0.70				0.76			
Pre-2008	Coeff	0.02	0.14	-0.45	0.35	0.01	1.01	-0.24	0.29
	P-value	0.02				0.11			

Panel C: Partial Correlations of Active Returns with Systematic Factors

	Internal		AS	
	Partial Corr	P-value	Partial Corr	P-value
VALGRTH	-0.33	0.01	-0.29	0.02
SMLG	0.26	0.04	0.05	0.68
MOM	-0.25	0.05	-0.10	0.41
VOL	0.26	0.03	0.53	0.00

Table 6 Continued: Internal Equity Returns

Panel D: Correlations of Active Returns with Hedge Fund Index Returns

		Internal		AS	
		Corr	P-value	Corr	P-value
Relative Value	RV Fixed Income Asset Backed	-0.27	0.22	-0.44	0.11
	RV Fixed Income Convertible Arbitrage	0.36	0.05	0.30	0.14
	RV Fixed Income Corporate	0.01	0.96	-0.10	0.47
	RV Multistrategy	0.00	0.98	-0.13	0.41
	RV Yield Alternatives	0.06	0.64	-0.14	0.50
	RV Index	0.00	1.00	-0.15	0.39
Event Driven	ED Distressed/Restructuring	0.04	0.79	-0.13	0.51
	ED Merger Arbitrage	-0.16	0.49	-0.30	0.27
	ED Index	0.10	0.46	-0.05	0.77
Equity Hedge	EH Equity Market Neutral	-0.25	0.22	-0.45	0.09
	EH Quantitative Directional	0.30	0.00	0.08	0.48
	EH Index	0.29	0.00	0.05	0.64
Macro	Macro Systematic Diversified	-0.22	0.37	-0.43	0.13
	Macro Index	-0.26	0.27	-0.47	0.07

Table 7: External Active Returns

(Net monthly returns)

Panel A: Active Returns and Residual Returns (wrt Mandate Benchmark)

	Active Returns		Residual Returns	
	Mean	Skew	Alpha	Skew
Fixed Income External Returns (wrt Mandate Benchmark)				
Mean	-0.45	-0.89	-0.53	-1.12
Median	-0.06	-0.67	-0.07	-1.19
Cross-Sectional Stdev	0.94	1.42	1.03	1.49
Equity External Returns (wrt Mandate Benchmark)				
Mean	0.09	0.00	0.05	0.03
Median	0.05	0.01	0.05	0.02
Cross-Sectional Stdev	0.73	0.69	0.67	0.70

Panel B: Partial Correlations of External Active Return Factor with Systematic Factors

	Fixed Income		Equity	
	Mean	p-value	Mean	p-value
TERM	-0.38	0.00		
CREDITAa	0.08	0.43		
CREDITBaa	-0.06	0.52		
CREDITHY	0.13	0.21		
FXCARRY	0.17	0.09		
LIQUIDITY	0.01	0.88		
VALGRTH			-0.68	0.00
SMLG			0.39	0.00
MOM			0.27	0.00
VOL	0.36	0.00	0.38	0.00

Table 7 Continued: External Active Returns

Panel C: Correlations of External Active Return Common Factor with Hedge Fund Index Returns

		Fixed Income		Equity	
		Partial Corr	P-value	Partial Corr	P-value
Relative Value	RV Fixed Income Asset Backed	-0.11	0.46	-0.12	0.35
	RV Fixed Income Convertible Arbitrage	0.21	0.04	0.15	0.12
	RV Fixed Income Corporate	0.13	0.16	0.02	0.81
	RV Multistrategy	0.00	1.00	0.05	0.65
	RV Yield Alternatives	0.10	0.31	0.00	0.96
	RV Index	-0.02	0.89	0.03	0.79
Event Driven	ED Distressed/Restructuring	0.14	0.11	0.06	0.57
	ED Merger Arbitrage	-0.17	0.31	-0.05	0.72
	ED Index	0.09	0.32	0.13	0.29
Equity Hedge	EH Equity Market Neutral	-0.22	0.17	-0.09	0.56
	EH Quantitative Directional	0.11	0.16	0.37	0.01
	EH Index	0.13	0.14	0.37	0.03
Macro	Macro Systematic Diversified	-0.25	0.14	0.06	0.78
	Macro Index	-0.27	0.09	0.05	0.82

Panel D: Gross vs Net Returns

	FIBA	EBL
Expense Ratios Across Funds		
Mean	0.87%	0.76%
Median	0.04%	0.14%
Cross-Sectional Stdev	5.08%	5.04%
Performance Fee Ratios Across Funds		
Mean	0.17%	0.29%
Median	0.01%	0.01%
Cross-Sectional Stdev	0.64%	0.77%
Fees Aggregated Across All External Funds		
Mean Expense Ratio	0.06%	0.17%
Mean Performance Fee Expense Ratio	0.07%	0.17%
Correlation between Gross and Net Returns		
Correlation between Gross and Net Mean Active Returns	85.8%	98.2%
Correlation between Gross and Net Alphas	99.8%	97.4%

Table 8: Summary of Active Management Scope and Advantages

Activities to be Included in Active Management

Time variation in exposure to risk factors included in the benchmark	There is some limited evidence that such factor timing skill may add value. The potential to add value cannot be ruled out and thus this should be a dimension of active management, constrained by tracking error targets.
Exposure to less established, new risk factors, and possibly new segments and markets outside current benchmark.	As the world economy evolves, there may be new risk factors available for investment. Some of these will be re-packaged, known factors, some may not. NBIM should be charged with identification and evaluation of potentially new, priced factors, and with determining how to capture them in an efficient manner.
Activist investing	NBIM currently expresses the intent of the asset owner with respect to socially responsible investing. It also seeks to add value through strategic exercise of its ownership stakes in corporations. While the evidence for excess returns to the latter are limited, NBIM should continue to exercise this power.
Some scope for single-company fundamental value based stock-picking.	Theory supports some level of active management of this type, despite the weak evidence in favor of it significantly contributing to overall value in section II. Thus, it this should be a limited dimension of active management, depending upon the cost. There may be relevant managerial reasons to maintain a component of fundamental investing. The practice of selectivity may include taking advantage of brand and size to access pre-IPOs and managers with evidence of skill.

Activities NOT to be Included in Active Management

NBIM should continue to avoid investing in below-market opportunities with potential social benefits.	In line with the present goals of the Fund, careful stewardship to invest wealth for the Norwegian public should be the aim.
NBIM should not initiate a Swenson-model strategy of seeking alpha in illiquid alternative asset classes unless or until the Fund has a model that integrates these into a factor framework and has in place a top management team.	Before investing in illiquid alternative asset classes the Fund should determine the extent to which performance of these alternatives is based on factor exposure vs. asset selection. The Fund must have a team with demonstrable capability to replicate the success of the top endowment managers who have succeeded with the strategy.

Table A-1: Overall Fund Active Return Correlations with Fama-French Factors

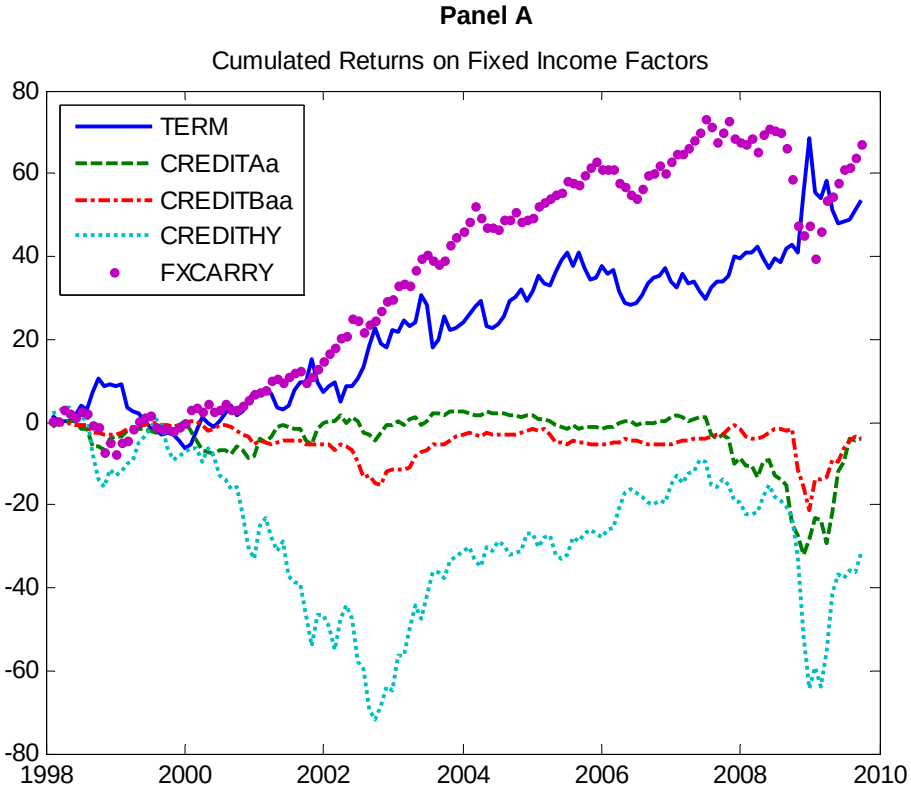
	Partial Corr	P-value
TERM	-0.18	0.04
CREDITAa	0.40	0.00
CREDITBaa	-0.43	0.00
CREDITHY	0.02	0.84
FXCARRY	0.02	0.81
LIQUIDITY	0.37	0.00
Fama-French HML	-0.14	0.12
Fama-French SMB	0.11	0.22
MOM	0.08	0.39
VOL	0.42	0.00

Figures[38]

[38] Returns for the overall Fund (Figures 2-4), the fixed income and equity subclasses (Figures 5-7), and internal NBIM returns (Figures 8 and 9) are monthly gross returns. Returns for external funds (Figure 11) are net of fees.

Figure 1: Factors

Panel A



Panel B

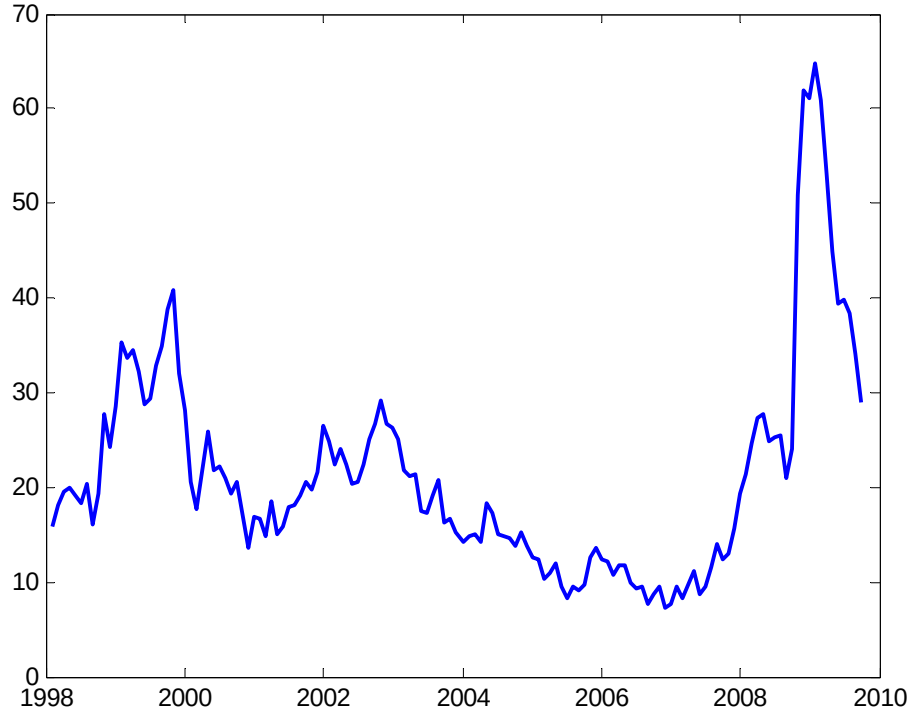
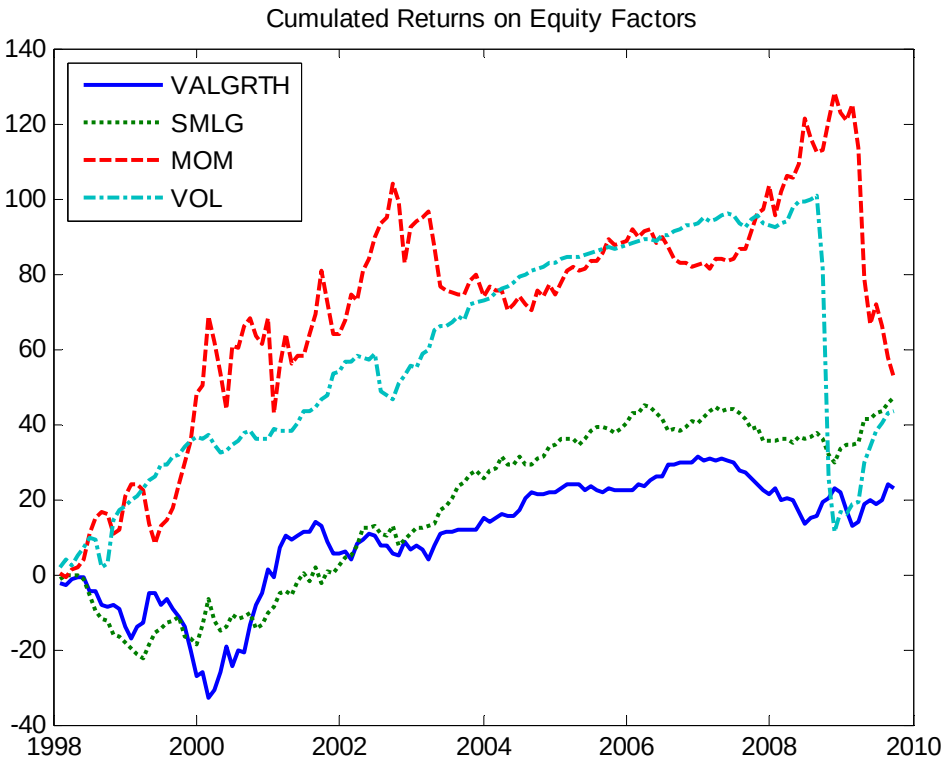


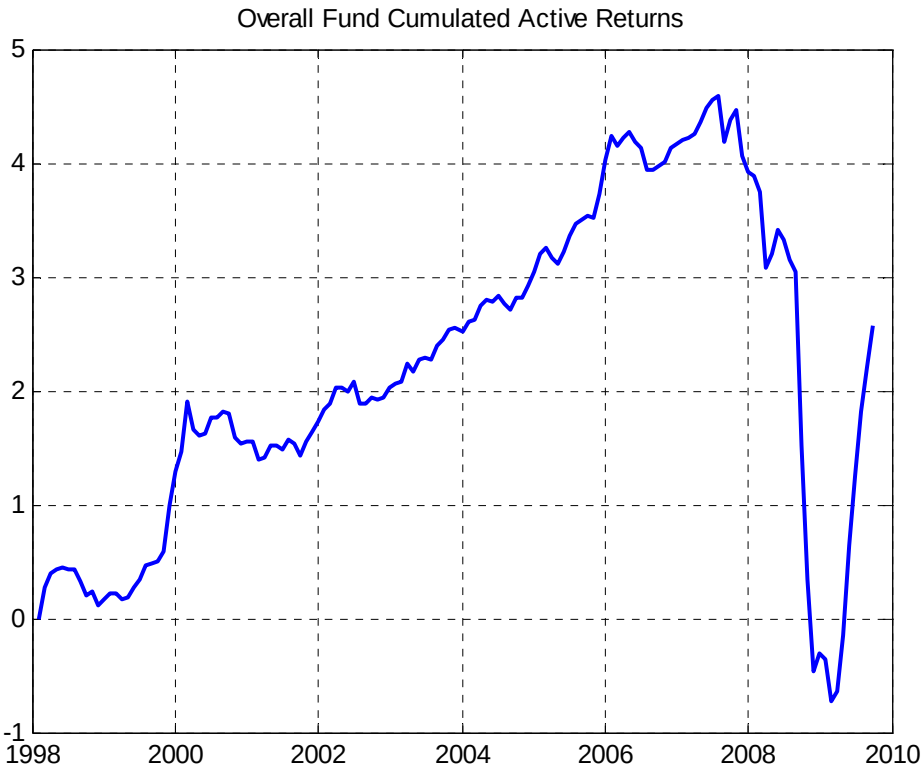
Figure 1 Continued: Factors

Panel C



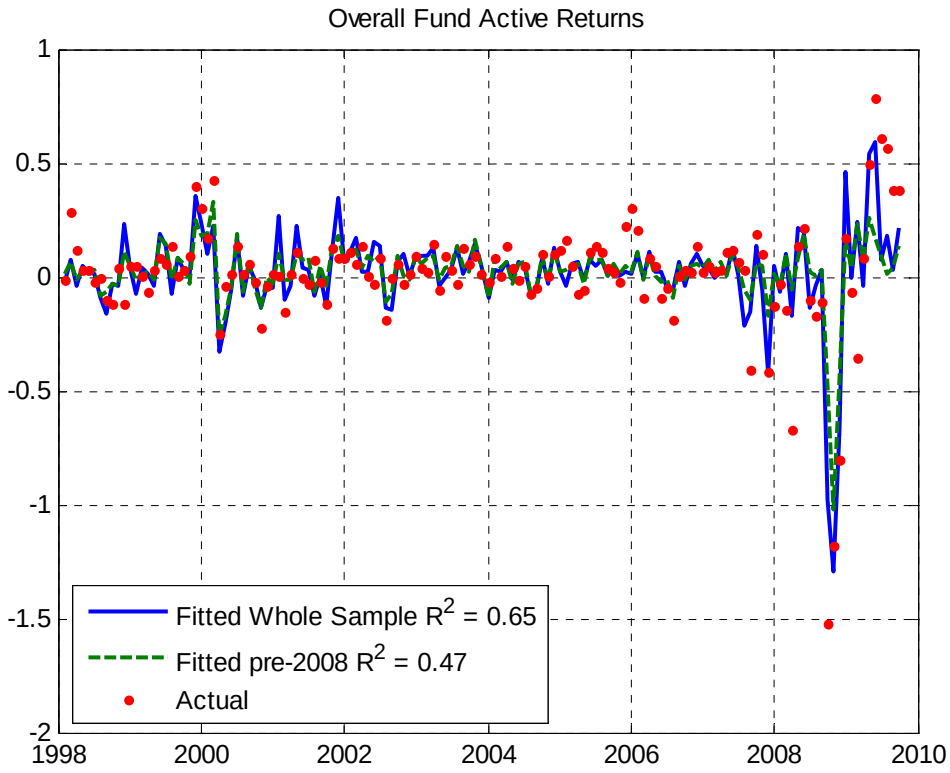
Panel A plots cumulated returns, which are cumulated sums of the monthly returns from the begining of the sample, on the fixed income factors. Units on the y-axis are cumulated monthly percentage returns. Panel B plots the on-the-run/off-the-run spread in basis points on 10-year US Treasury bonds. The liquidity factor is the negative of the first difference of this spread. Panel C plots cumulated returns on the equity factors. Units on the y-axis are cumulated monthly percentage returns.

Figure 2: Overall Fund Cumulated Active Returns



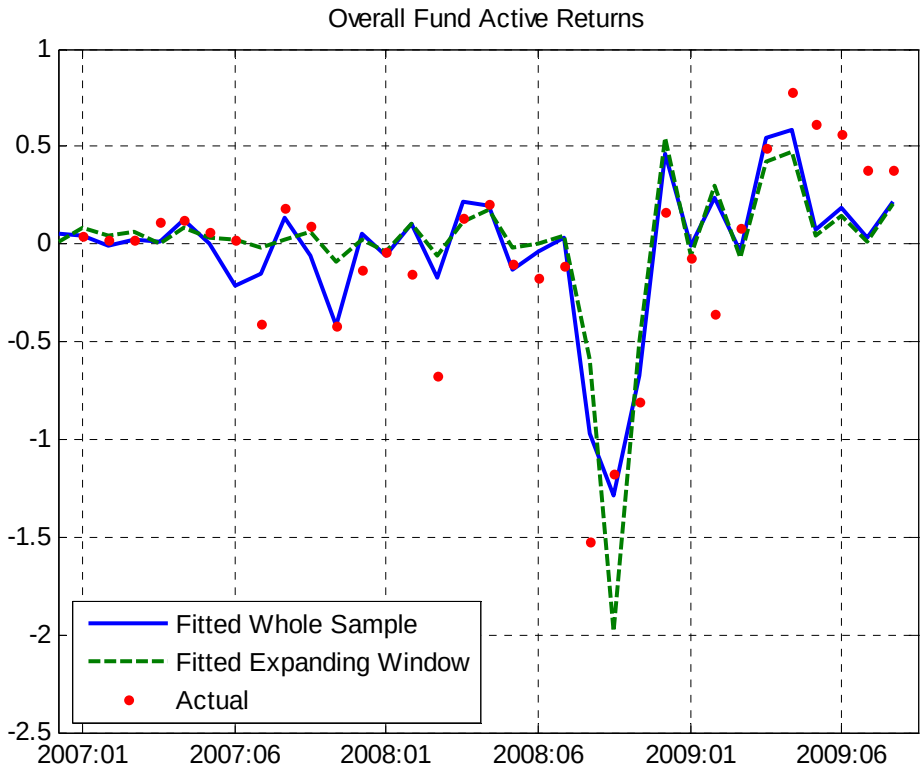
The figure plots cumulated active returns, which are cumulated sums of the monthly active returns from the beginning of the sample, on the overall Fund. Units on the y-axis are cumulated monthly percentage returns.

Figure 3: Overall Fund Active Returns



We plot realized active returns on the overall Fund in red dots, the fitted active return where the regression coefficients are computed over the whole sample in the blue solid line, and the fitted active return where the regression coefficients are computed over the pre-2008 sample in the green dashed line. The regressions are computed using all systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

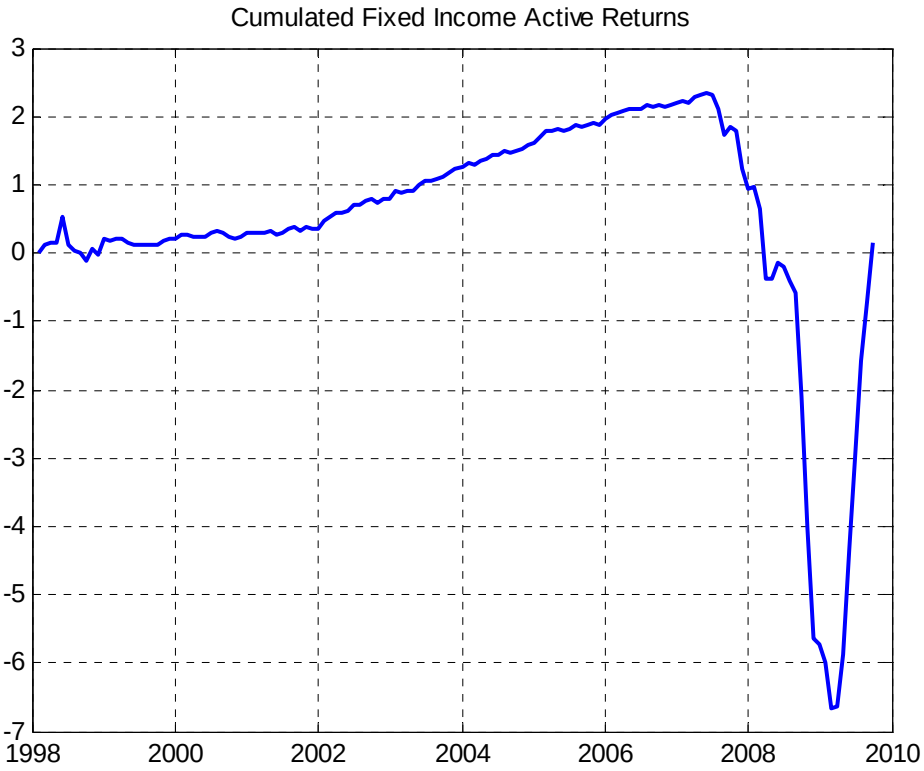
Figure 4: Overall Fund Active Returns Post-2007



We plot realized active returns on the overall Fund in red dots, the fitted active return where the regression coefficients are computed over the whole sample in the blue solid line, and the fitted active return where the regression coefficients are computed over an expanding window using data up to the previous month in the green dashed line. For the latter the coefficients are estimated using data up to the previous month and the fitted active return for the current month is produced using that month's realized factors. The regressions are computed using all systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

Figure 5: Cumulated Fixed Income and Equity Active Returns

Panel A

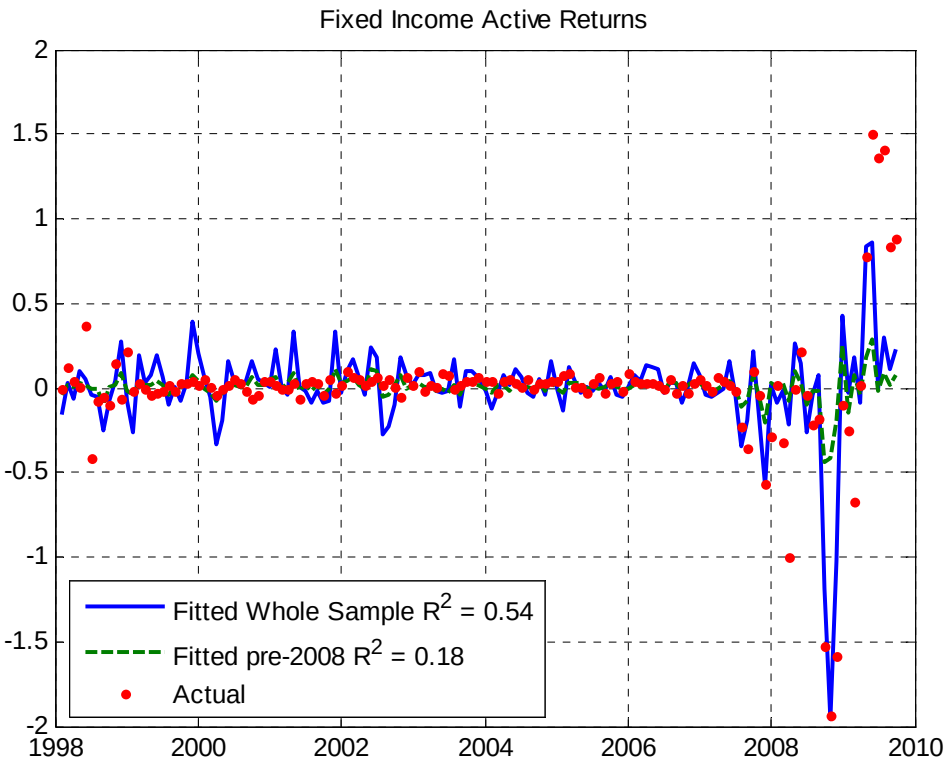


Panel B

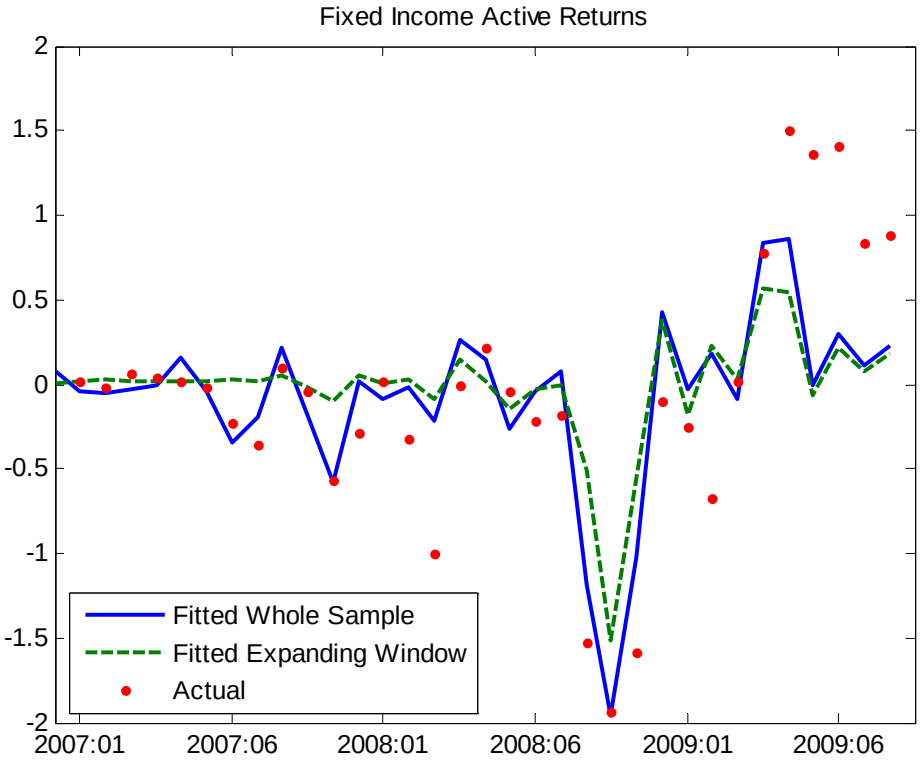
The figure plots cumulated active returns, which are cumulated sums of the monthly active returns from the beginning of the sample, on the fixed income (Panel A) and equity (Panel B) portions of the fund. Units on the y-axis are cumulated monthly percentage returns.

Figure 6: Fixed Income Active Returns

Panel A



Panel B

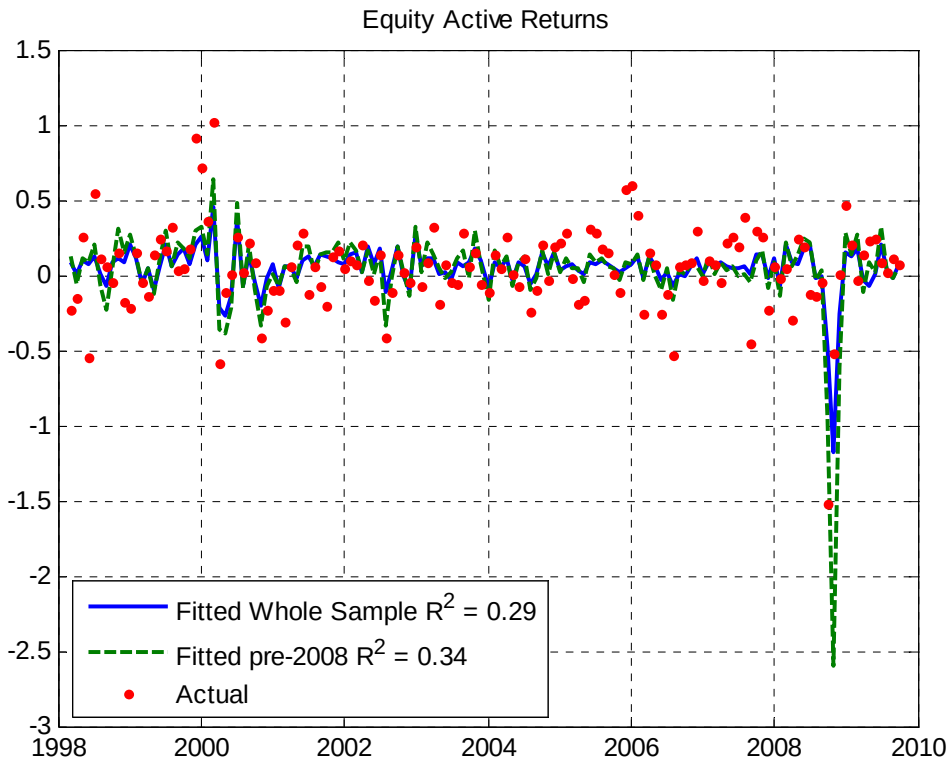


Note To Figure 6

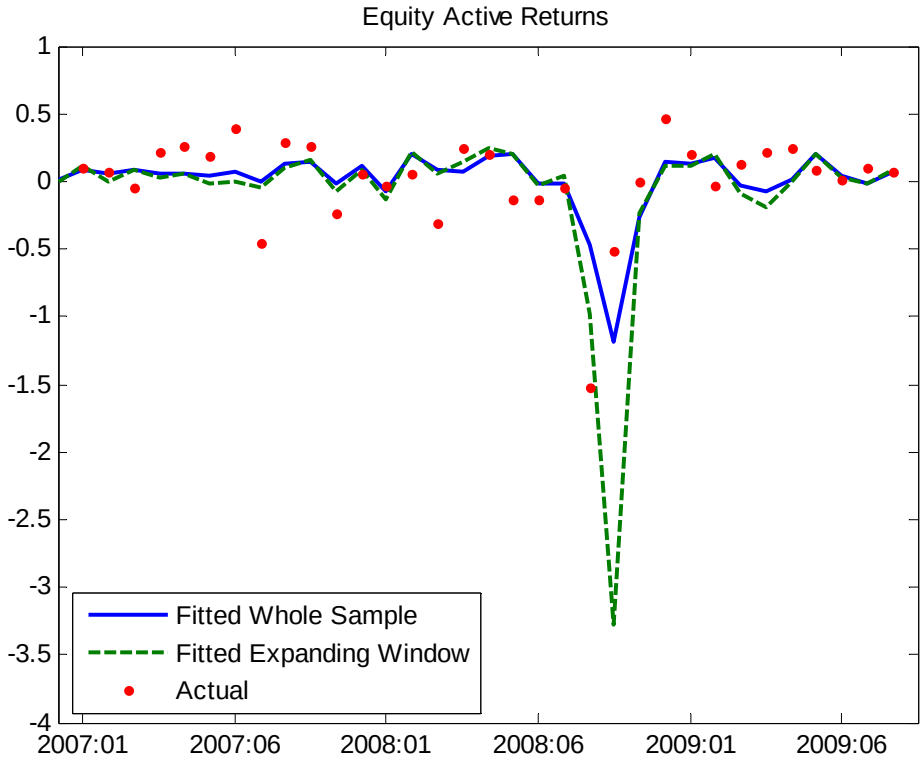
In Panels A and B we plot realized active returns on the fixed income portion of the fund in red dots and the fitted active return where the regression coefficients are computed over the whole sample in the blue solid line. In Panel A we plot the fitted active return where the regression coefficients are computed over the pre-2008 sample in the green dashed line. In Panel B the green dashed line corresponds to the fitted active return where the regression coefficients are computed over an expanding window using data up to the previous month in the green dashed line. For the latter the coefficients are estimated using data up to the previous month and the fitted active return for the current month is produced using that month's realized factors. The regressions are computed using the fixed income and volatility systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

Figure 7: Equity Active Returns

Panel A



Panel B

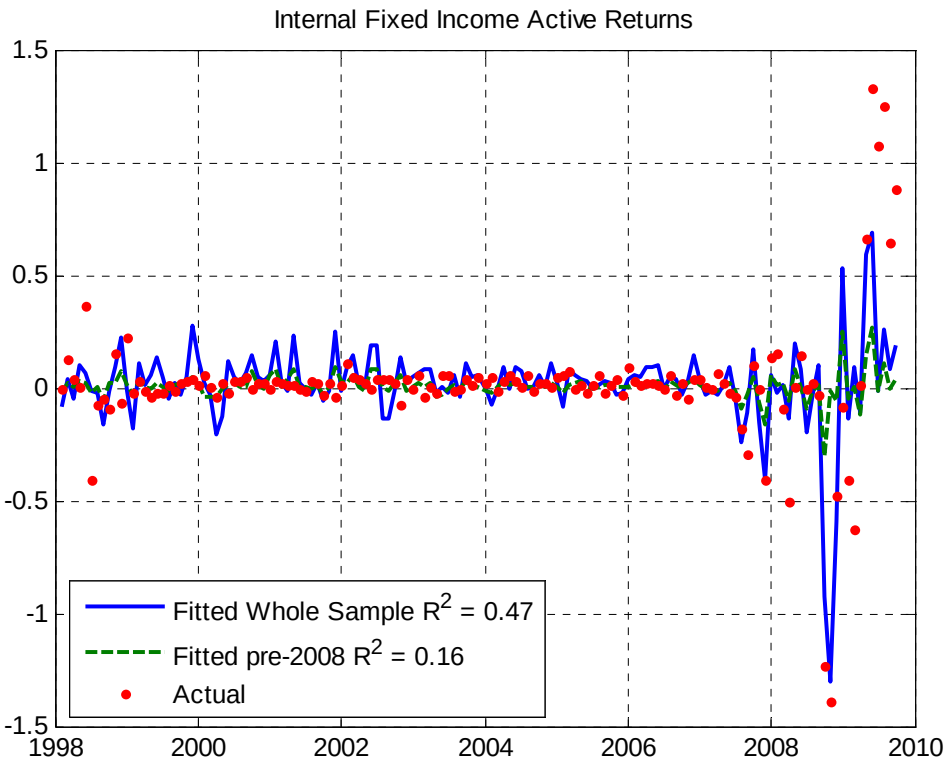


Note To Figure 7

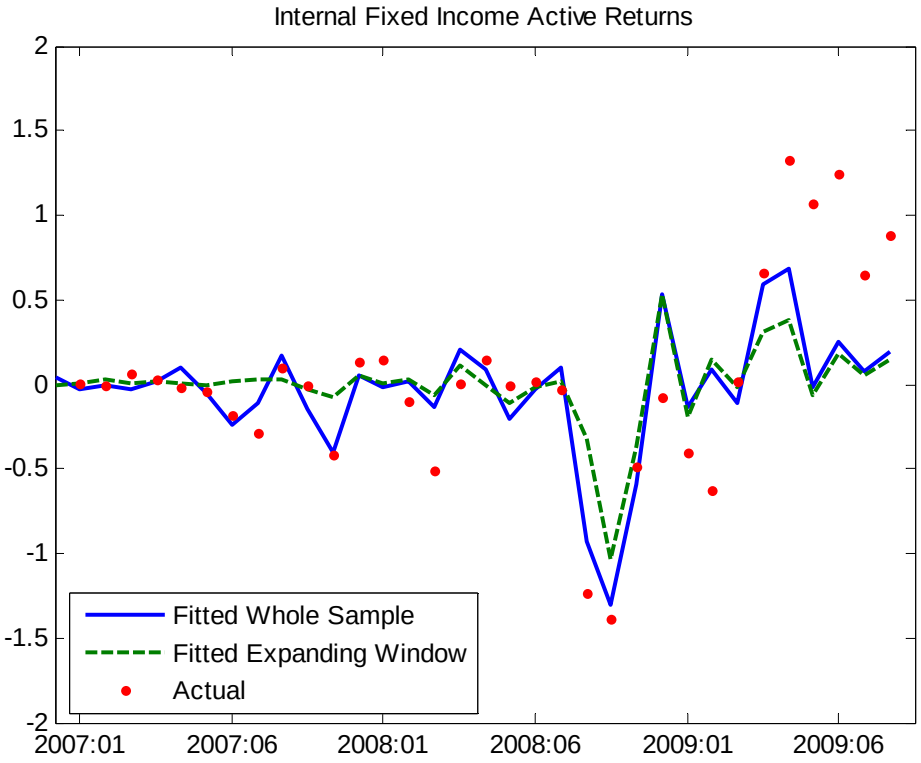
In Panels A and B we plot realized active returns on the equity portion of the fund in red dots and the fitted active return where the regression coefficients are computed over the whole sample in the blue solid line. In Panel A we plot the fitted active return where the regression coefficients are computed over the pre-2008 sample in the green dashed line. In Panel B the green dashed line corresponds to the fitted active return where the regression coefficients are computed over an expanding window using data up to the previous month in the green dashed line. For the latter the coefficients are estimated using data up to the previous month and the fitted active return for the current month is produced using that month's realized factors. The regressions are computed using the equity systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

Figure 8: Internal Fixed Income Active Returns

Panel A



Panel B

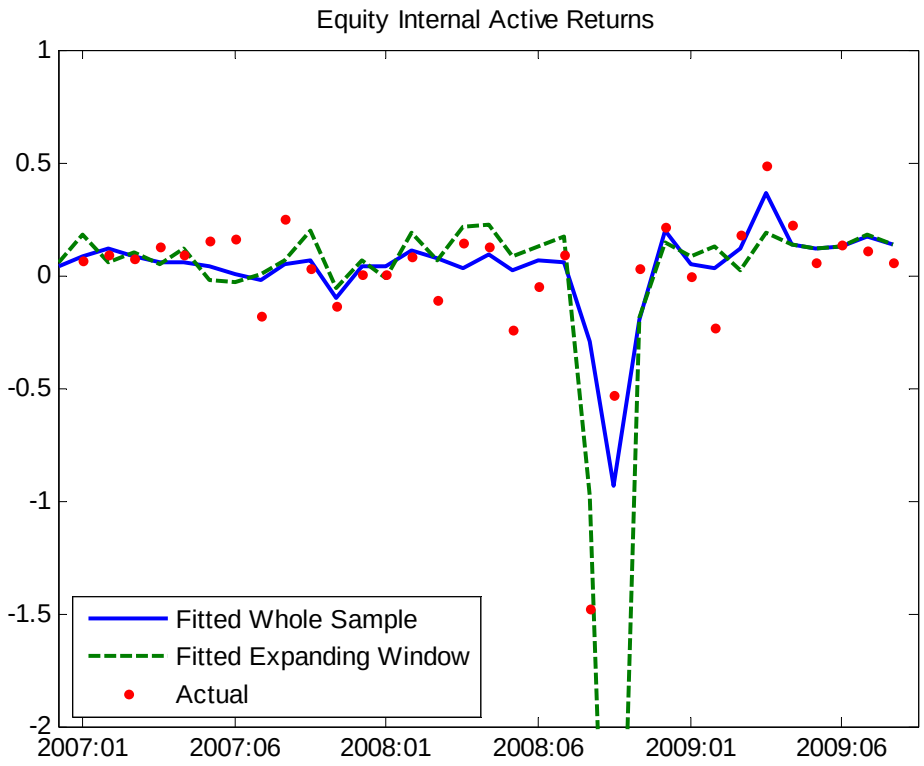


Note to Figure 8

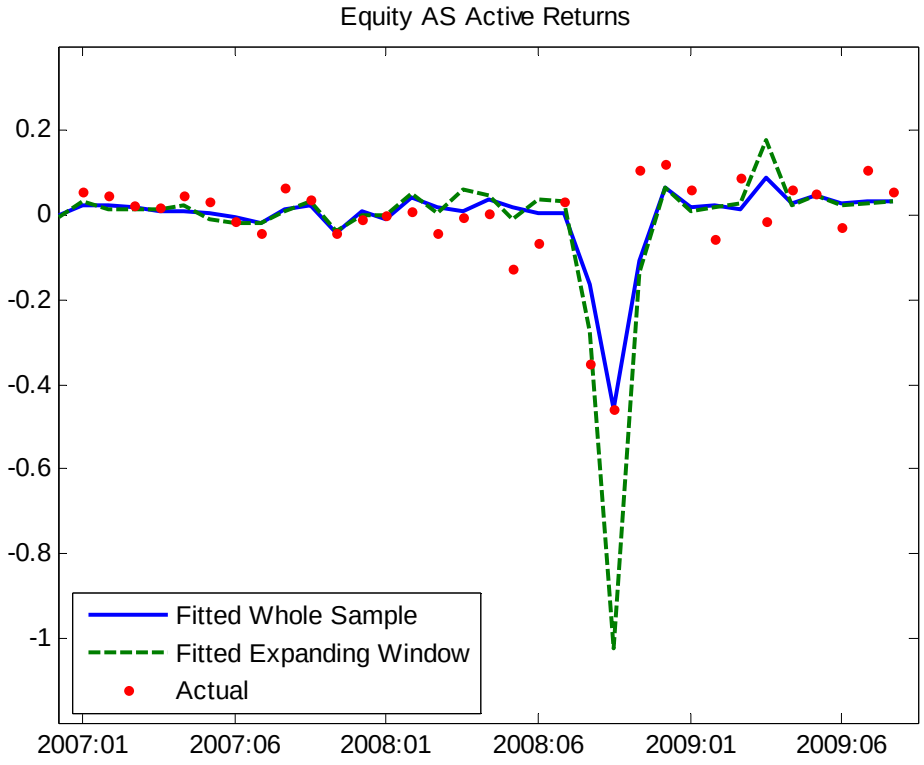
In Panels A and B we plot realized active returns on the internal fixed income portion of the fund in red dots and the fitted active return where the regression coefficients are computed over the whole sample in the blue solid line. In Panel A we plot the fitted active return where the regression coefficients are computed over the pre-2008 sample in the green dashed line. In Panel B the green dashed line corresponds to the fitted active return where the regression coefficients are computed over an expanding window using data up to the previous month in the green dashed line. For the latter the coefficients are estimated using data up to the previous month and the fitted active return for the current month is produced using that month's realized factors. The regressions are computed using the fixed income and volatility systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

Figure 9: Internal Equity Active Returns

Panel A



Panel B

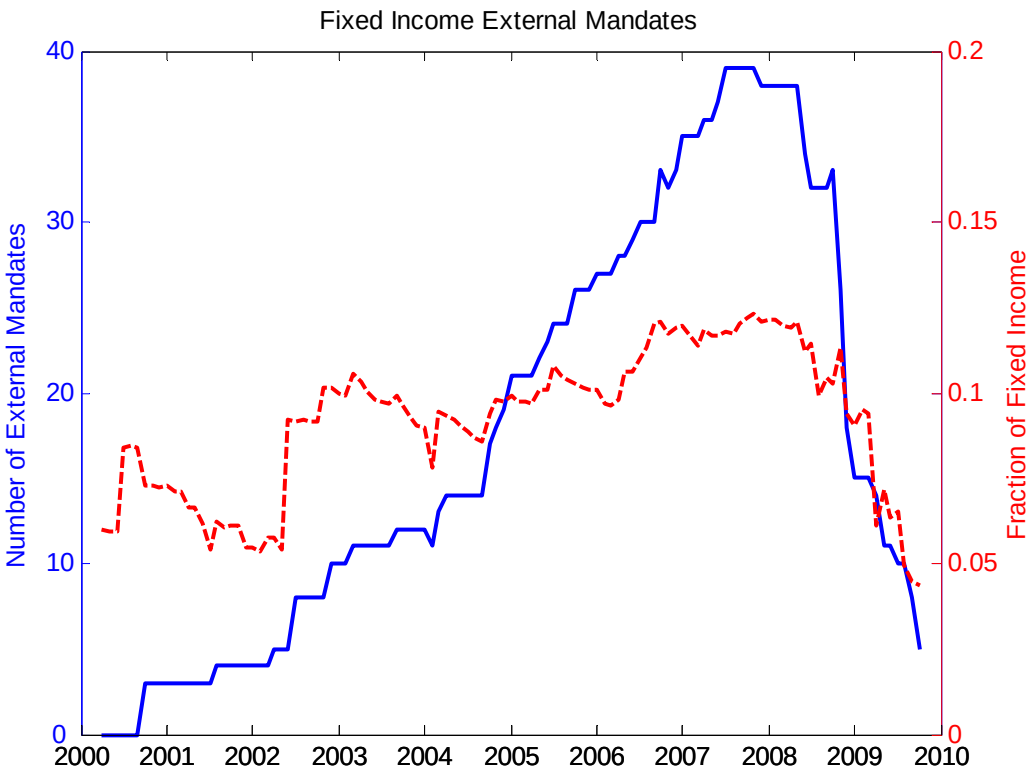


Note to Figure 9

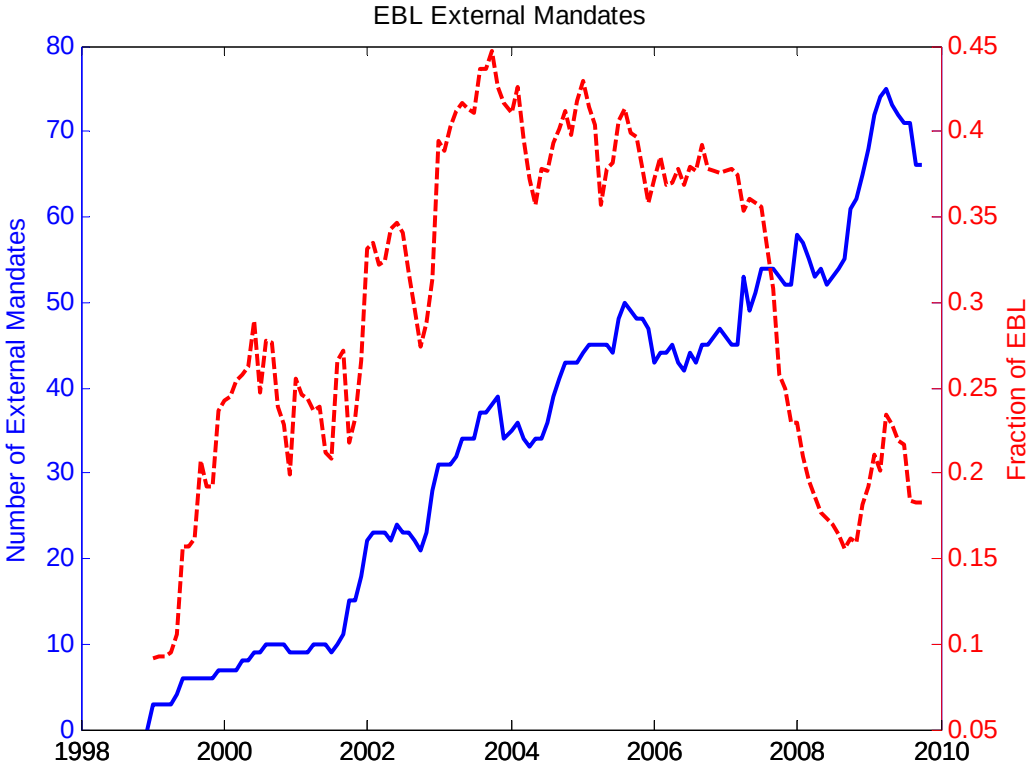
We plot realized active returns on the internal equity portions of the fund in red dots and the fitted active return where the regression coefficients are computed over an expanding window using data up to the previous month in the green dashed line. For the latter the coefficients are estimated using data up to the previous month and the fitted active return for the current month is produced using that month's realized factors. We plot active returns for the internal and AS strategies in Panels A and B, respectively. The regressions are computed using the equity systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

Figure 10: External Funds

Panel A



Panel B

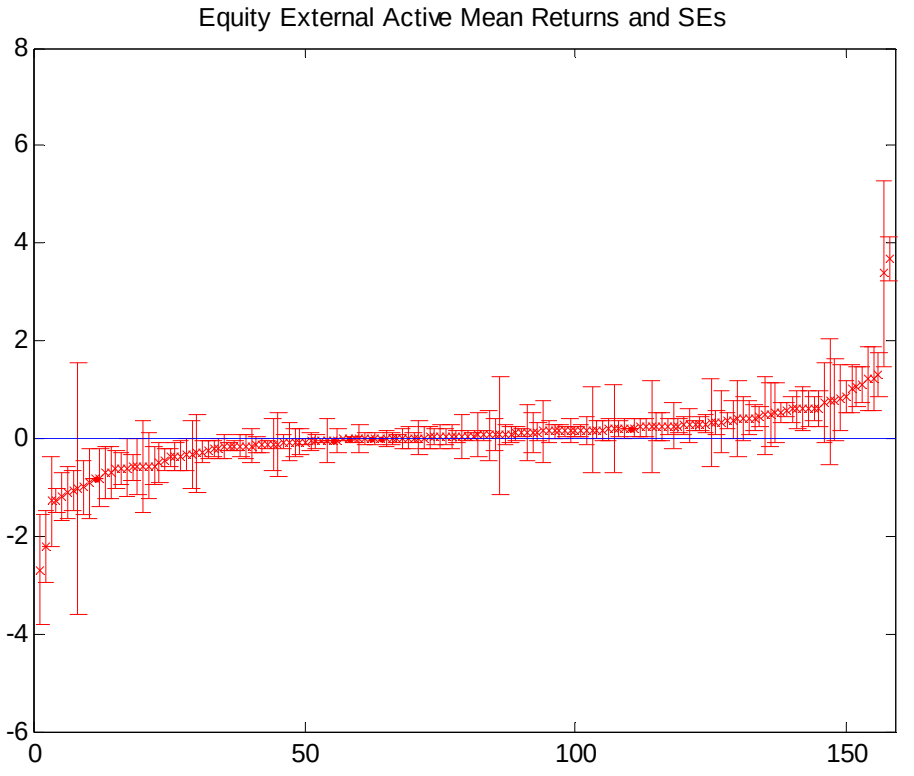


We plot the number of external mandates in the solid blue line on the LHS axis and the proportion of the portfolio externally managed in the dashed red line on the RHS axis for fixed income (Panel A) and equities (Panel B).

Figure 11: External Active Returns

Panel A

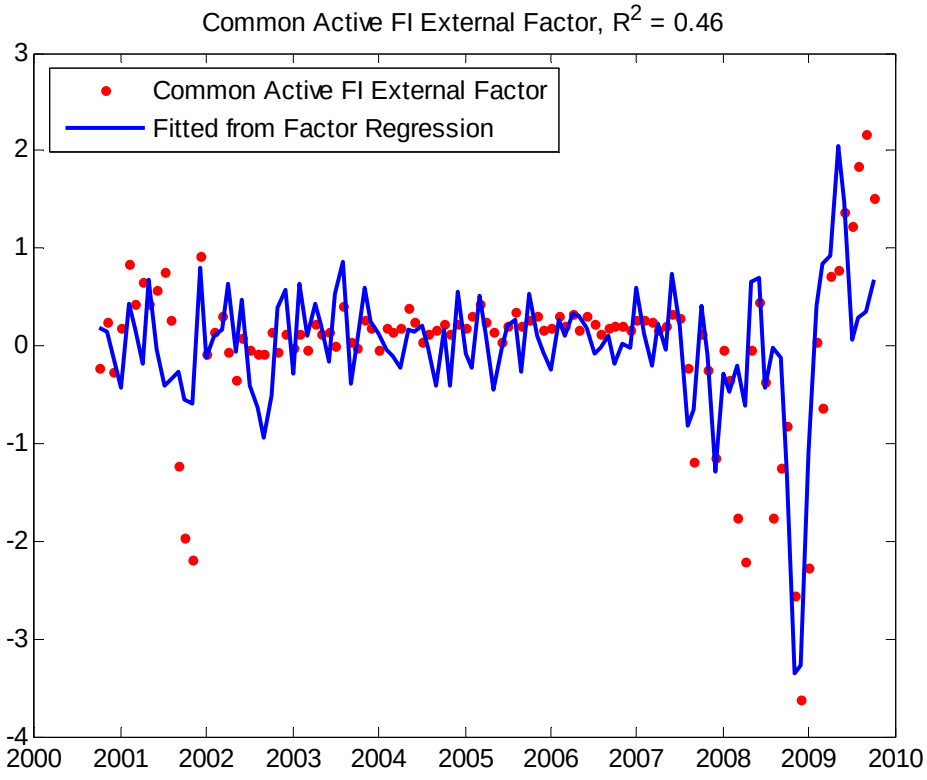
Panel B



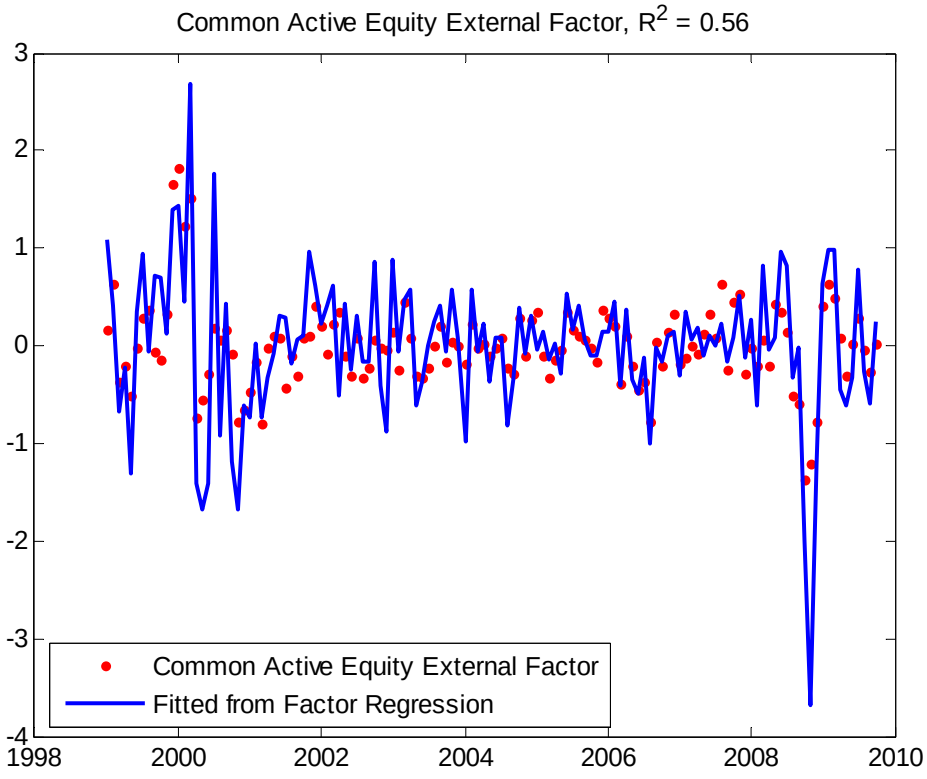
We plot the mean external active return for fixed income (Panel A) and equity (Panel B) external funds. The 'x' represent the mean return ordered in increasing mean active return and robust +/- 2 standard error bounds are shown in the error bars.

Figure 12: External Active Returns

Panel A



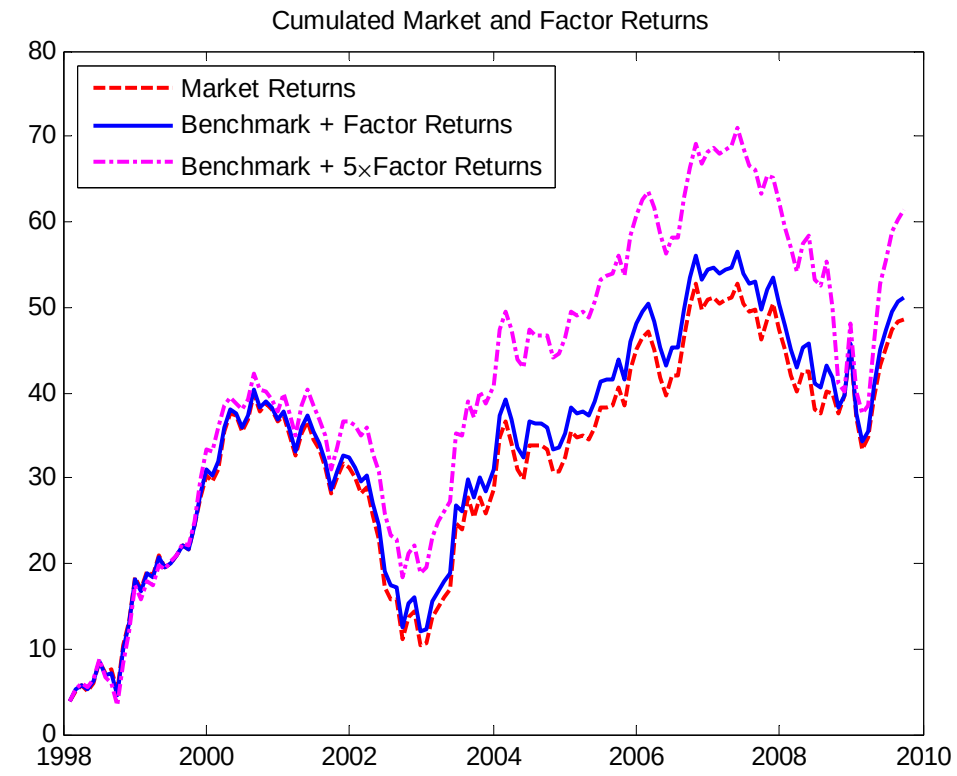
Panel B



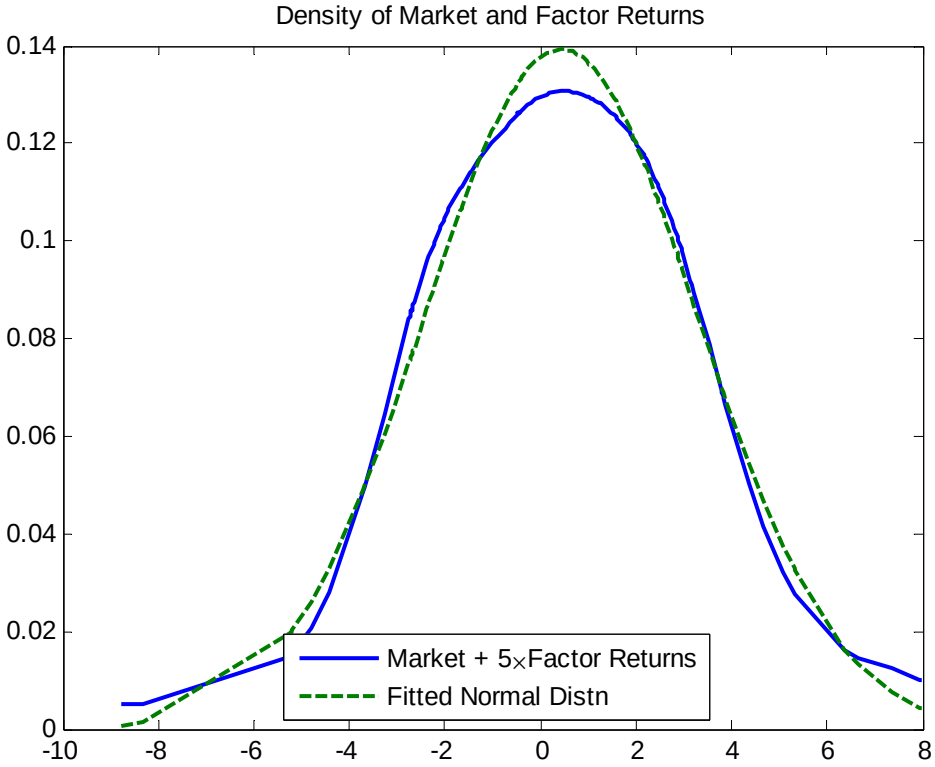
We plot estimates of the common active fixed income (Panel A) and equity (Panel B) external common active factor in the red dots. We overlay the fitted values from the factor regression in the solid lines. The factors used for fixed income are the systematic fixed income and volatility factors. The factors for equity are the systematic equity factors.

Figure 13: Market and Factor Returns

Panel A



Panel B

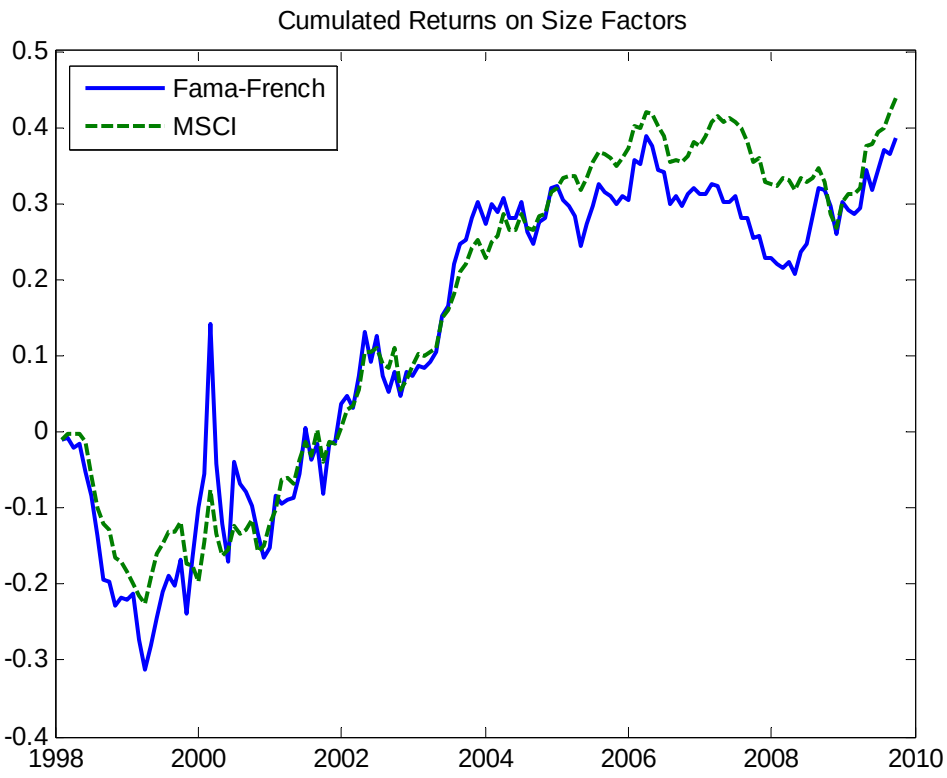


Note to Figure 13

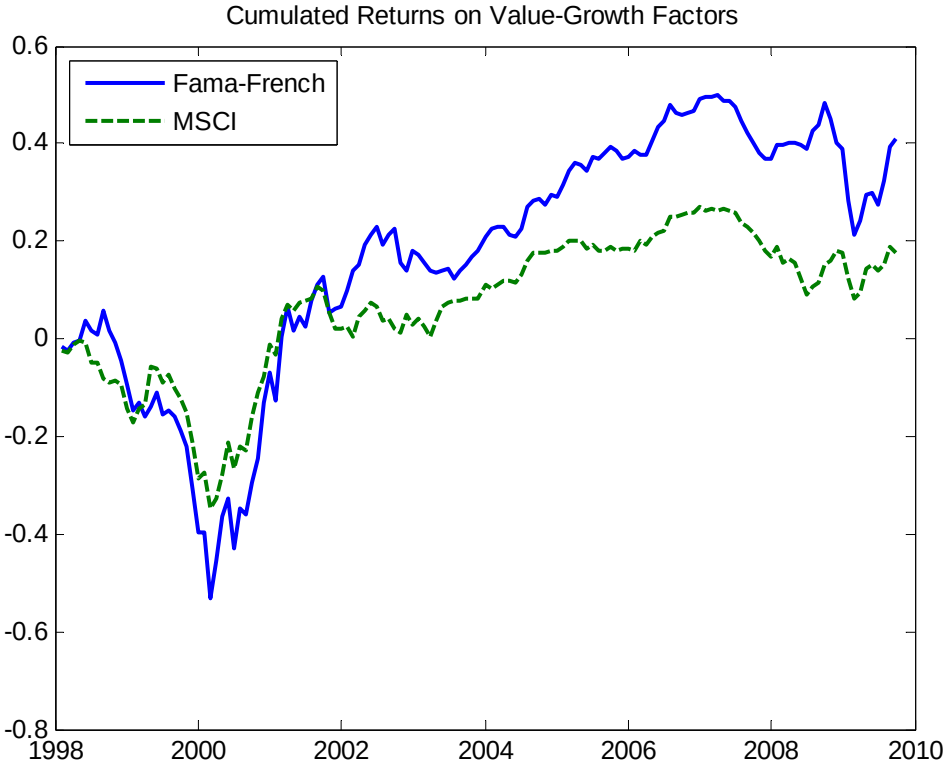
The market return is the overall Fund benchmark and the factor return is the fitted portion of the active overall Fund return regressed onto the systematic factors using a regression over the whole sample. Panel A graphs cumulated market and factor returns together with actual fund returns. Panel B plots the fitted probability density function produced by a kernel estimation method of market and factor returns together with a normal distribution with the same mean and variance.

Figure A-1: Cumulated Returns on Size Factors

Panel A

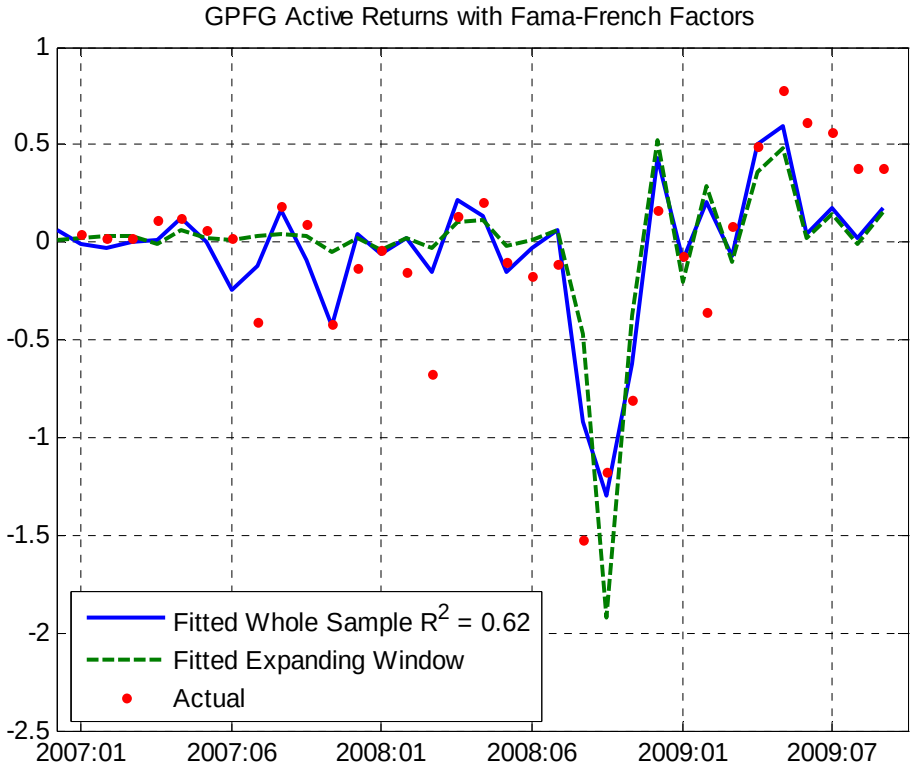


Panel B



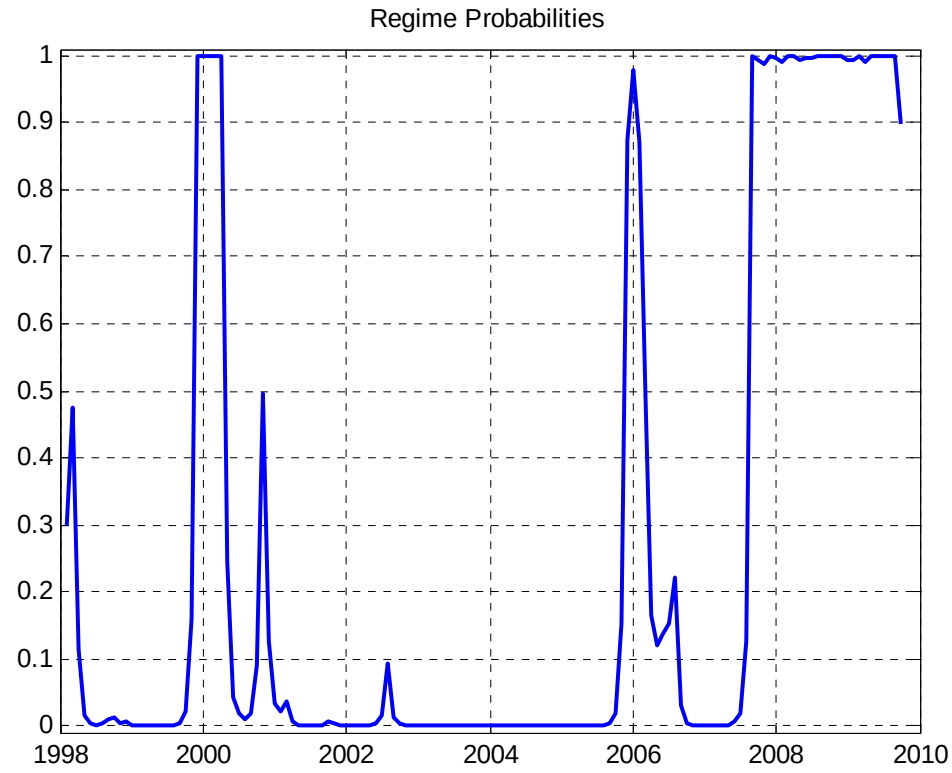
We compare cumulated returns on size factors SMLG and SMB constructed by Fama-French in Panel A and cumulated returns on value-growth factors VALGRTH and HML constructed by Fama-French in Panel B. Units on the y-axis are cumulated monthly percentage returns.

Figure A-2: Overall Fund Active Returns Using Fama-French Factors



This is the corresponding plot to Figure 4 estimated using Fama-French SMB and HML factors as systematic factors. We plot realized active returns on the overall Fund in red dots, the fitted active return where the regression coefficients are computed over the whole sample in the blue solid line, and the fitted active return where the regression coefficients are computed over an expanding window using data up to the previous month in the green dashed line. For the latter the coefficients are estimated using data up to the previous month and the fitted active return for the current month is produced using that month's realized factors. The regressions are computed using Fama-French SMB and HML instead of the SMLG and VALGRTH systematic factors as regressors. All fitted returns are computed using the realized factors at each date.

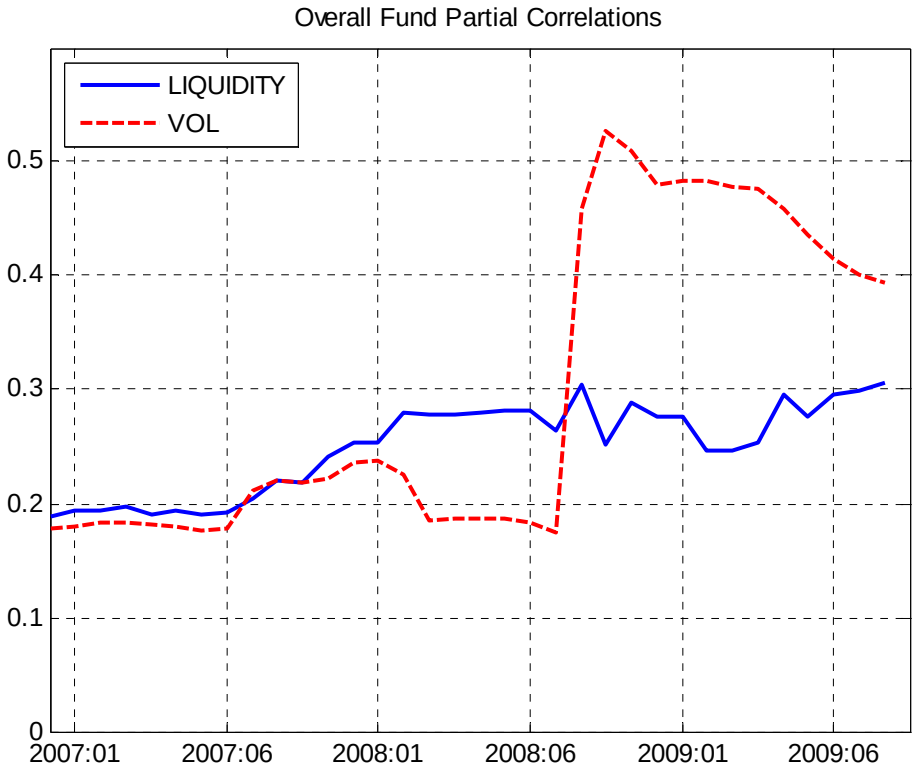
Figure A-3: Overall Fund Regime Probabilities



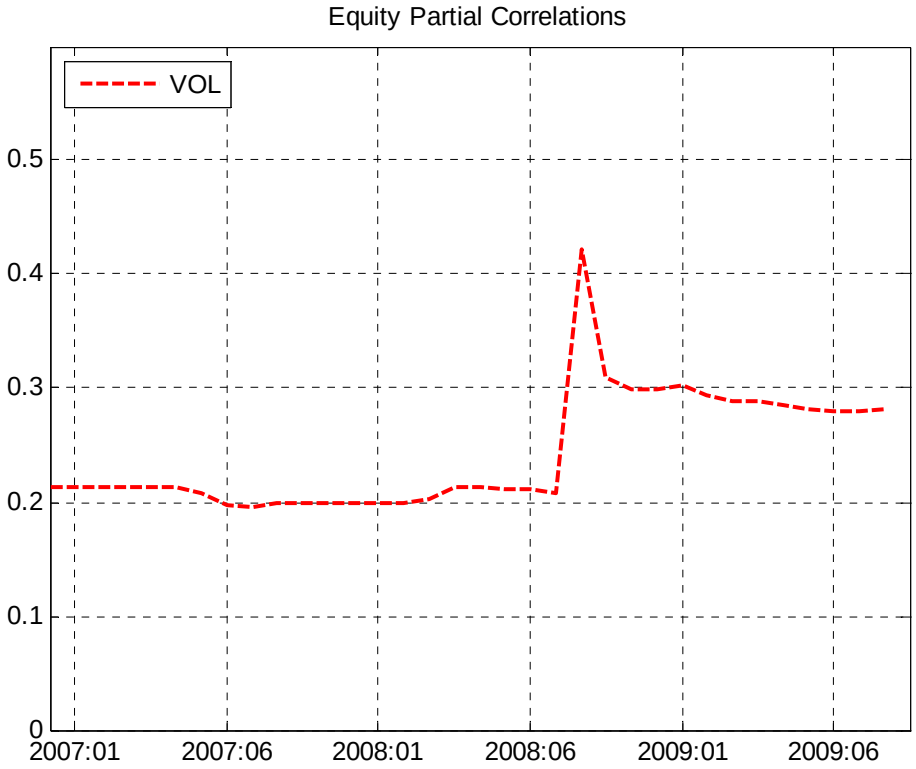
The figure plots regime probabilities from a regime-switching model estimated on the active returns of the Fund smoothed over the whole sample for being in a regime with a small, negative mean and very large volatility. The other regime has a small, positive mean and low volatility.

Figure A-4: Rolling Partial Correlations

Panel A



Panel B



The figure plots rolling partial correlations of the overall Fund (Panel A) and the equity asset class (Panel B) active returns with LIQUIDITY and VOL factors. The partial correlations are computed over an expanding window using data up to the current month from the beginning of the sample.