Goldman Sachs

Global

Global Portfolio Analysis

November 30, 1999

Beating Benchmarks

A Stockpicker's Reality: Part II

Analysts

Steve Strongin
steve.strongin@gs.com
(New York) 1-212 357-4706

Melanie Petsch
melanie.petsch@gs.com
(New York) 1-212 357-6092

Greg Sharenow
greg.sharenow@gs.com
(New York) 1-212 902-7494

The inability of active managers to consistently outperform capitalization-weighted benchmarks can be explained by a mismatch between those benchmarks and the underlying nature of active management. We show that this mismatch cannot be effectively addressed either through macro level risk controls or through improved stock selection. However, we develop a new approach to risk management that emphasizes diversification at the individual stock level and offers significant increases in risk-return efficiency and portfolio manager consistency; it is also significantly easier to incorporate it into a bottoms-up investment process. Further, we show how plan sponsors can further improve the value of active management through combinations of new, more portfolio-manager-friendly active manager benchmarks and completion indices that move the overall allocations back to their original capitalization-weighted benchmarks.

Table of Contents

Making Skill Count

Beating benchmarks using fundamental bottoms-up stock analysis has at its core two parts: stock selection (the ordering of stocks from best to worst using fundamental analysis) and portfolio construction (the translation of that ordering into an actual portfolio). This paper focuses on portfolio construction.[1] In particular, we focus on how a portfolio manager skilled at stock selection can exploit that skill to beat a target benchmark by as much as possible and as consistently as possible.

Recent underperformance by U.S. large-cap portfolio managers has generated an intense focus on valuation methods for large-cap growth companies and on the general question of macro-level risk management for equity portfolios. We show that these efforts are likely to significantly reduce future returns without noticeably improving the quality of risk or consistency of outperformance as these efforts are based on a misunderstanding of the true nature of the recent risk management failure and on an unreasonable notion of what even the most skilled portfolio manager might be able to do in assessing the return potential of individual companies.

In particular, we find that for reasonable levels of portfolio manager skill (i.e., skill levels consistent with the level of long-run outperformance most portfolio managers would be willing to claim), it is simply impossible to improve stock valuation methods enough to solve the recent "large-cap" problem by better stock selection.

Further, our results demonstrate that the key problems that have prevented large-cap portfolio managers from generating consistent outperformance over the last decade were neither due to a macro level failure to properly control size (or any other macro risk factor) nor due to a failure of fundamentally driven stock picking strategies (either growth or value) to discriminate between high- and low-performing stocks.

Rather, almost the entirety of portfolio manager inconsistency can be explained by a failure to properly take account of a massive concentration of stock-specific risk in a small number of names at the top end of the capitalization spectrum.[2] While the difference between the macro control of size and controlling exposure to a small number of stocks at the top of the capitalization spectrum might seem a rather academic distinction, the operational implications are enormous and the resulting impact on portfolio manager performance dramatic.

In particular, we show that controlling size risk at the macro level reduces returns almost directly in proportion to the amount of benchmark tracking error eliminated (i.e., risk and return fall in equal amounts as return per unit of risk barely improves). In contrast, compensating for the concentration of stock-specific risk at the top end of the capitalization spectrum through passive individual stock positions reduces tracking error at double the rate that it reduces returns, substantially improving the quality of risk and the consistency of outperformance.

For example, we find that for a moderately skilled portfolio manager, simply market weighting the top 50 stocks will, on a pre-transactions cost basis, double their Sharpe ratio from 0.58 to 1.30 for value and from 0.62 to 1.30 for growth, increase the percentage of quarters that such portfolio managers outperform their benchmark from 58.3% to 73.8% for value and from 60.3% to 74.3% for growth, radically smooth the time series of outperformance and dramatically reduce both the size and the frequency of extreme underperformance (including transaction costs would only make the improvement more dramatic on a relative basis).

[1] Stock selection is the focus of our January 14, 1999 paper "Style, Size and Skill," Part I of *A Stockpicker's Reality*.

[2] The general presumption that indices with large numbers of stocks diversify away most idiosyncratic stock risk and that the remaining index volatility is mostly macro in character only holds if the weight on each stock is below a certain threshold. When stocks are added with weights above that threshold, more stock-specific risk is added than diversified away, creating indices with significant stock-specific volatility focused in those high-weight stocks. The mathematical details of this argument are covered in Appendix A.

More broadly, we find that the standard top-down macro risk approach fails bottoms-up portfolio managers in three ways:

1. It completely misses the need to offset the stock-specific risk embedded in large-capitalization benchmarks.
2. It misdirects stock selection toward large-cap names where active management is, in general, less effective.
3. It causes portfolio managers to concentrate stock selection risk into too small a number of positions to allow for consistency of outperformance.

Our results indicate that for risk management to actually aid portfolio manager performance, it is necessary to refocus risk control away from quantifying the macro risk characteristics of the portfolio at a point in time and toward understanding and controlling the quality (not quantity) of risk that the portfolio manager takes over time.

In particular, we find that the key risk control issue for bottoms-up fundamentally driven portfolio managers is being able to distinguish between (1) habitual concentrations of risk that arise either out of peculiarities of the benchmark or prejudices in the portfolio manager's investment process that create consistent risk positions that do not reflect current market conditions or the portfolio manager's skill,[3] and (2) those risk positions that arise naturally out of the portfolio manager's assessment of fundamentals and vary with market conditions.

Habitual risk positions create a high level of benchmark risk without any real expectation of return. Eliminating such positions, either through passive offsets or constraints on portfolio construction, can significantly improve a portfolio manager's ability to consistently outperform benchmarks.

In contrast, we find that the types of risk positions that arise naturally out of fundamental analysis by skilled portfolio managers are not only justified on a risk-return basis, but are naturally diversified. The natural diversification suggests that the best way to reduce risk without sacrificing returns unnecessarily is simply to increase the number of stock positions (both at a point-in-time and over time) in order to allow the high level of uncertainty that characterizes individual stock positions to average out as much as possible and, thus, allow the portfolio manager's skill at stock selection to dominate rather than the randomness of individual stock returns. Macro level risk controls can and often do significantly interfere with portfolio managers taking advantage of this natural diversification and actually reduce the risk-return efficiency of portfolio manager performance.

Consequently, we argue that, beyond identifying and eliminating chronic/habitual risk positions through passive offsets,[4] risk controls should be limited to (1) helping exploit the natural diversification of the portfolio manager's stock choices and (2) helping the portfolio manager modestly emphasize taking risk in categories of stocks in which they tend to be more effective and de-emphasize taking risk in areas in which they tend to be less effective.[5]

[3] Examples of habitual concentrations of risk include (1) being perpetually underweight a concentration of stock-specific risk in the high index-weight stocks discussed above or (2) a perpetual underweight in the tech sector that reflects a portfolio manager's discomfort with new technology rather than their evaluation of the actual companies.

[4] Such offsets can be accomplished either through the use of derivatives or through appropriately constructed passive portfolios either at the portfolio manager or plan sponsor level.

[5] In particular, sector controls can be useful in helping define categories of stocks in which the mapping between fundamentals and returns is more consistent. Sector controls can also provide an ability to compensate for the positive correlations between stock picks that are created by the use of common drivers of forecasted earnings within sectors.

In summary, we argue that the risk management process should be split into three distinct processes which, in order of declining importance, are:

1. Compensating for undue concentration of stock-specific risk in (large-cap) benchmarks.
2. Broadly diversifying stock-specific risk chosen by portfolio managers in order to allow the high level of uncertainty in the individual stock returns to average out.
3. Modestly concentrating stock-specific risk in areas in which the portfolio manager has demonstrated greater ability to identify higher-returning stocks.

The paper proceeds in three steps. First, we define the nature of manager skill in a way that allows us to quantify the level of skill needed to produce differing levels of long-run outperformance. We then analyze different risk control approaches, both in terms of performance and in terms of portfolio manager consistency, in order to understand how different risk management approaches impact both returns and consistency. We then enter a broader discussion of how bottoms-up managers should approach risk control, both conceptually and as a matter of practice. In particular, we look at what is necessary to tailor risk control to improve overall performance rather than simply reduce risk.

The Nature of Skill

For our purposes, portfolio manager skill is defined as the ability to rank stocks based on future fundamentals. In Part I of this series, "Style, Size and Skill," we looked at how the market pays for future fundamentals. In particular, we showed that a portfolio manager who ranks stocks based on future fundamentals using either a growth methodology (in which stocks are ranked from best to worst based on forward earnings growth) or a value methodology (in which stocks are ranked from best to worst based on a future normalized P/E ratio) can separate stocks into higher- and lower-performing groups with a high degree of consistency. (More details on the growth and value measures we use, as well as the data, can be found in Appendix B.)

A limitation of the prior analysis was that, to clearly define what the market was pricing into the market, we allowed the portfolio manager an unreasonable level of foresight into future earnings, allowing the manager to perfectly predict earnings. In the current context, we need to be able to allow the portfolio manager a fixed, but limited, ability to rank stocks based on insights into forward earnings behavior. We can then use these rankings to see how different risk management approaches will allow a portfolio manager with a given level of stock-selection skill to create portfolios capable of outperforming a given benchmark.

The way we model skill is to allow the portfolio manager to rank stocks relative to the true (i.e., perfect foresight) fundamental rankings for their style of investing with differing degrees of statistical accuracy. This allows us to hold the stock-selection skill level constant and investigate how different risk-control approaches work with different investment styles and portfolio construction approaches through large-scale simulations. In the main body of the paper, we focus on pure growth and value styles. In Appendix C, we repeat some of the key results for a hybrid valuation or growth at a reasonable price style.

To create the true rankings for pure value investing, every calendar quarter, stocks are ranked based on P/E ratios (which are based on the average realized earnings for the next four quarters) from the least to most expensive, under the expectation that less expensive stocks will outperform more expensive stocks. For pure growth investing, true rankings are based on forward earnings growth – the next four quarters for the S&P 500 and Russell 1000 (large-cap) simulations, the next two quarters for Russell 2000 (small-cap) simulations.[6]

Then we create simulated rankings in which the portfolio manager is able to approximate the true ranking more or less closely, based on their skill level. The specifics of the statistical modeling are quite simple. In the zero-skill case, the portfolio manager's stock ratings follow a uniform random distribution from 0 to 1 (think of this as the percentile rank of the stock). Thus, in the absence of stockpicking skill, each stock is equally likely to have any rating from 0 to 1, regardless of fundamentals.

To create skill, we tilt the distribution such that stocks with better fundamentals are more likely to receive higher ratings and less likely to receive low ratings. We do this simply by tilting the uniform distribution based on the true ranking of the stock and the skill of the manager. Figure 1 shows the resulting valuation rating distributions for the best, median and worst stocks, for a zero-skill, moderate-skill and max-skill portfolio manager.

For zero skill, the top stock (true rating value of 1.00 measured on a scale from 0.00 to 1.00) has roughly a 5% probability of receiving a rating between 0.95 and 1.00 and a 20% probability of getting a rating between 0.80 and 1.00 (i.e., a top quintile rating). Similarly, in the zero-skill case, the top stock also has a 20% probability of getting a rating between 0.00 and 0.20 (i.e., a bottom quintile rating).

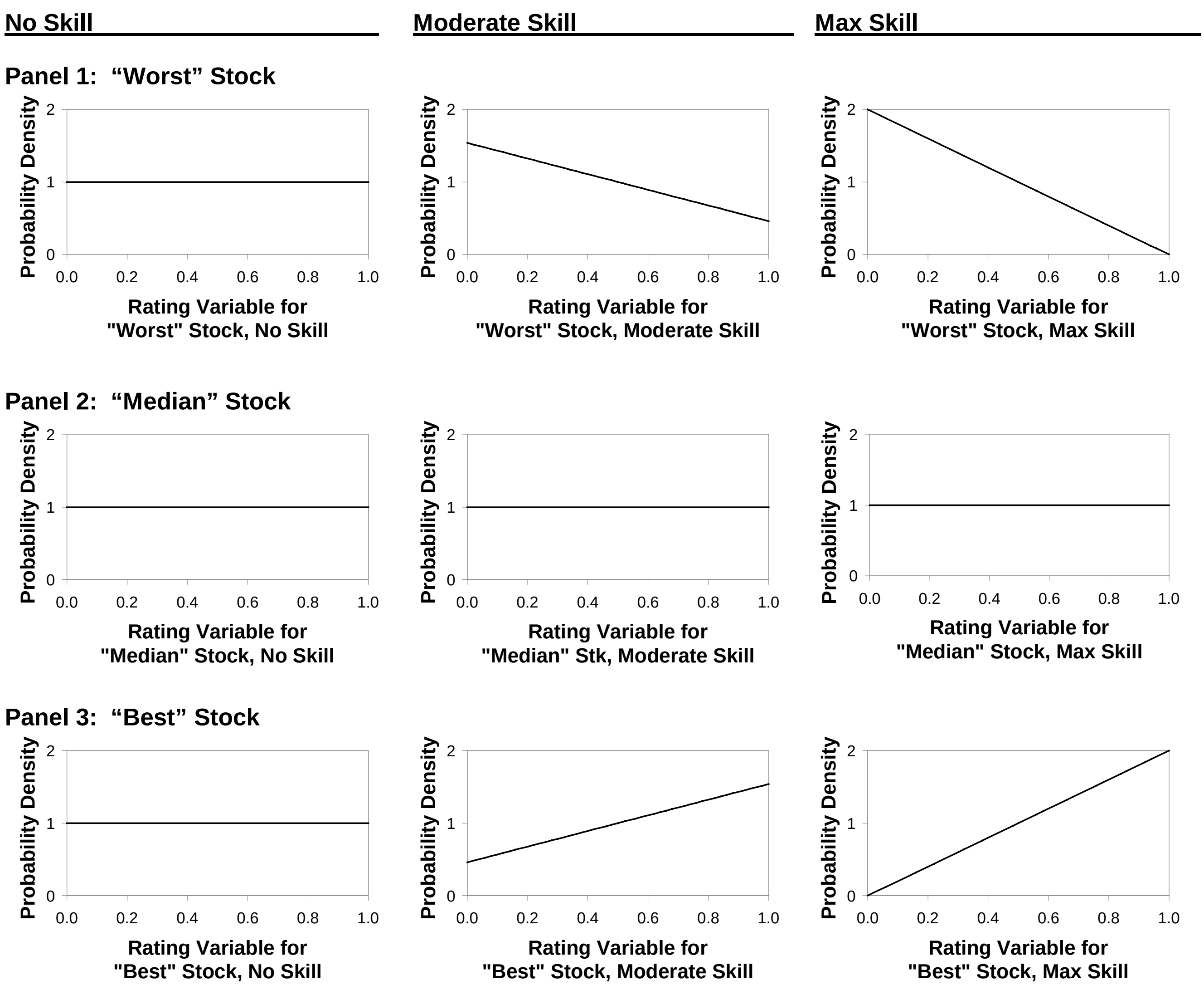
With max stock-selection skill, the top stock has a 9.75% chance of getting a rating between 0.95 and 1.00 and a 36% chance of a top quintile rating, while only a 4% chance of getting a bottom quintile rating.

[6] In "Style, Size and Skill," we showed that the optimal horizon for earnings insight for growth strategies is shorter for smaller-cap stocks. These horizons (four quarters forward for large-cap, two quarters forward for small-cap) were chosen because, of one to four quarters forward, they provide the best performance for the growth strategies for these particular data samples.

Another way of thinking about this measure of stock-selection skill, which gives some additional intuition about the level of skill implied by these tilts, is to ask how accurate the ranking is relative to the true (that is, perfect foresight) ranking. One way of doing this is to look at the quintile accuracy. That is, if the portfolio manager ranks stocks from 1 to 5 where 1 is the best quintile and 5 is the worst quintile of stocks, how likely is the portfolio manager to rank stocks in the correct quintile? Table 1 shows the map from the tilt of the skill distribution to the percentage of the stocks the portfolio manager ranks in the correct quintile bucket. No skill (0% of maximum tilt) gets it right 20% of the time. Moderate skill (54% of maximum tilt) gets it right 23% of the time, high skill (82% of maximum tilt) 25% of the time and max skill (100% of maximum tilt) 26.4% of the time.

At first glance, the implied level of accuracy appears quite low (even when the distribution is tilted as far as possible), but as we see in the next section, when we translate these skill levels into implied long-run excess returns that would be generated by a portfolio manager with these levels of stock-selection skill, the range of implied returns covers the full range of what might reasonably be expected from portfolio managers and even reaches levels well beyond what even the most skilled portfolio manager could be expected to deliver.

Figure 1: Skill



Source: Goldman Sachs Research

Mapping Skill Levels to Implied Returns

To get a solid idea of what these skill levels mean in terms of actual stock selection, we run statistical simulations in which a simulated portfolio manager of a given skill level creates 1,000 rankings, which are then translated into long-short portfolios. The portfolios are long the top 20% of the stocks (by the simulated rankings) and short the bottom 20%.[7] Returns are then calculated for 1,000 long-short portfolios. We then treat half the average of these returns as a reasonable estimate of the potential excess returns relative to market that a long-only portfolio manager of this skill level should be able to attain over time.[8]

The use of long-short portfolios in this section of the paper allows us to concentrate on the impact of skill on stock selection (i.e., the ranking of stocks) and how that relates to returns in a way that is largely independent of benchmark choice. This turns out to be important as it allows us in later sections to clearly separate which risk control problems relate to the portfolio manager's skills (and prejudices) and which relate to the particular benchmark they are attempting to beat. This, in turn, acts as a guide to which results are likely to hold for all portfolio managers and which need to be adjusted to reflect the particular skills/process of a specific portfolio manager.

The average returns and implied long-run excess returns generated by value and growth strategies at the various skill levels are reported in Table 2 (along with two rank correlations between the skilled ranks and the perfect foresight ranks). As would be expected, a portfolio manager with no skill generates no excess returns. What might be more surprising is how little of an edge in stock selection is needed to generate extraordinarily high excess returns. Over this period, the max skill level, which only has a 26.4% chance of ranking a stock in the correct quintile, produces 10.9% annualized returns for a market-neutral long-short growth manager and 9.9% for a market-neutral value manager, implying long-run long-only outperformance of 5.4% and 4.9%, respectively.

Table 1: Stockpicking Edge Associated with Skill Tilt

Percent of Maximum Skill Tilt (%)	Percent in Correct Quintile (%)	Skill Name
0	20	No Skill
20	21	
37	22	
54	23	Moderate
68	24	
82	25	High
100	26.4	Max
----	100	Perfect

Source: Goldman Sachs Research

Such return numbers would suggest that the reasonable range for portfolio manager skill would be between 23% and 25% or between 3% and 5% better than random selection, which we refer to as moderate and high skill respectively. In most cases, we also look at higher skill levels to demonstrate that our results hold even at extraordinary skill levels that are consistent with returns well beyond historical precedent.

[7] Unless otherwise indicated, all of the long-only strategies are long the top 20% of stocks and the long-short strategies are long the top 20% and short the bottom 20%. The exceptions are graphs that deal with the impact of the number of stocks in the portfolio.

[8] Transaction costs, of course, would reduce these returns, but such costs are too dependent on position size to be dealt with in a general manner. As a first approximation, reduce reported returns by 150 basis points to approximate realized returns for roughly a $1-billion portfolio. For a more complete analysis of how transactions costs would impact these results, see Appendix A of "Style, Size and Skill."

Table 2: Implied Long-Run Excess Returns and Rank Correlations for Skill Levels

(Estimated Russell 1000® Sample, 1Q1987-1Q1998)

		Average Long-Short Rtns (%)		Implied Long-Run Excess Rtns (%)			
Skill Name	**Percent in Correct Bucket (%)**	**Growth Strategy**	**Value Strategy**	**Growth Strategy**	**Value Strategy**	**Quintile Bucket Correlation**	**Rank Correlation**
No Skill	20	0.0	0.1	0.0	0.0	0.00	0.00
	21	2.3	2.2	1.1	1.1	0.07	0.07
	22	4.3	3.9	2.1	1.9	0.12	0.12
Moderate	23	6.1	5.6	3.1	2.8	0.17	0.18
	24	7.7	6.9	3.8	3.5	0.22	0.23
High	25	9.1	8.3	4.5	4.1	0.27	0.28
Max	26.4	10.9	9.9	5.4	4.9	0.32	0.33
Perfect	100	33.1	30.6	16.5	15.3	1.00	1.00

Source: Goldman Sachs Research

Getting a Realistic Picture of the Stockpicker's Edge

The thinness of the stock-picking edge described above hints that one of the core risk management problems facing portfolio managers is that they have a high probability of being wrong on individual stocks. Such randomness can only be turned into consistent short- or medium-term performance by taking a large number of positions and allowing the randomness to average out.

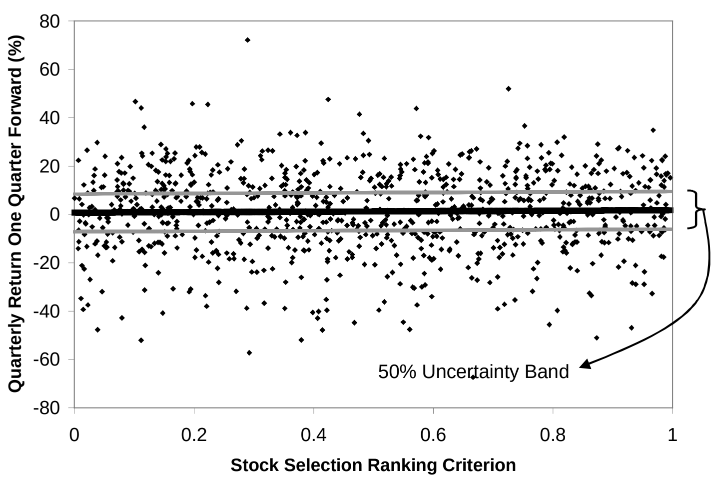
This randomness turns out to be deeply fundamental to the whole portfolio management process. Stock returns are very, very disperse and mostly random with respect to **any** particular notion of fundamentals, even for portfolio managers with very, very high levels of skill based on nearly perfect foresight of fundamentals, let alone those with more reasonable levels of skill.

To provide a baseline representation of the underlying randomness of stock returns, Figure 2 shows a representative scatter plot from a particular quarter of data of the relationship of individual stock returns to their rank for a portfolio manager with no skill, plus the regression line characterizing the relationship between rankings and returns and the limits of the 50% uncertainty band around the regression line. Because this type of graphic and the associated table are central to this paper, the contents are explained more fully in the sidebar on the next page.

Not that surprisingly, for the no-skill case, the regression line is flat and the individual equities often fall quite far from the line. Based on 1,000 simulations of the ranking process at this skill level, we calculate that 50% of the stocks will have quarterly returns within ±7.8% of portfolio manager's expectations.[9] This is a very wide band, highlighting the large extent of the underlying randomness in stock returns.

Figure 2: No Skill – Relationship of Ranking Criterion to Returns

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



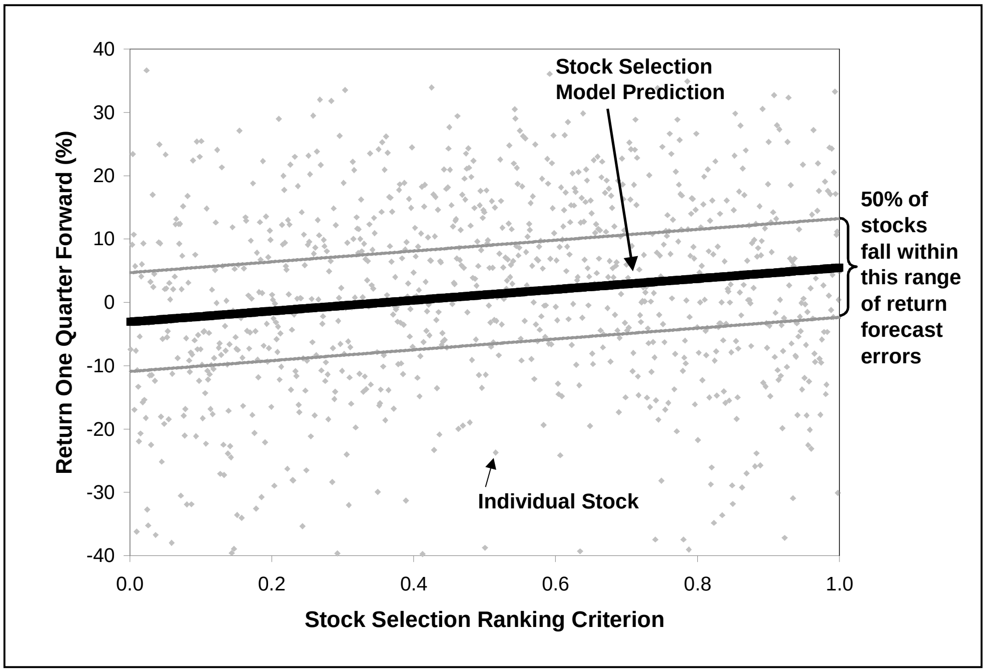
No Stock Selection Skill
(20% of Stocks in Correct Bucket)

High-Mid Spread	0.0 %
High-Low Spread	0.0 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	0.0 %
Standard Deviation	14.4 %
Average R-Squared	0.11 %
Annualized Long-Short Return	0.1 %
Annualized Implied Long-Run Excess Long-Only Return	0.0 %

Source: Goldman Sachs Research

[9] We also include an estimated standard deviation which is estimated by calculating a 95.4% confidence interval and then dividing the width of that interval by 4. This provides a more robust measure of the standard deviation than the more conventional calculation as it reduces the impact of outliers.

A Guide to the Skill Scatter Charts and Tables



The scatter plots depict one simulation of each style and skill level from the third quarter of 1997. Each dot represents an individual stock's one-quarter forward return in our estimated Russell 1000 sample. At the various levels of skill, the simulated portfolio manager ranks stocks from 0 to 1, where 0 is the worst and 1 is the best. The graphs show the relationship of this ranking variable and the subsequent return on the stocks, along with a regression line to summarize that relationship. The statistics in the accompanying tables refer to the average of 1,000 simulations for each quarter from the first quarter of 1987 to the first quarter of 1998.

The figure above illustrates how some of the key skill statistics are displayed in the scatter charts; the table below describes the statistics we report.

High-Mid Spread	The difference between the expected returns of the top-ranked stock and the median-ranked stock.
High-Low Spread	The difference between the expected returns of the top-ranked stock and the bottom-ranked stock.
50% Uncertainty Band	50% of the realized returns of the stocks falls within these bounds of the return predicted by the valuation model. That is, 50% of the realized returns are between the predicted return plus this percentage and the predicted return minus this percentage.
Average Slope	Average of the slope from the 1,000 simulated regressions of the forward return on the ranking criterion.
Standard Deviation	Standard deviation of the realized returns around the expected return line.
Average R-Squared	Average of the R-squared from the 1,000 simulated regressions of the forward return on the ranking criterion.
Annualized Long-Short Return	Annualized average of the excess returns for 1,000 simulated portfolios long the top 20% of the stocks by the ranking criteria and short the bottom 20%. That is, go long the fastest growing and short the slowest growing or go long the least expensive and short the most expensive.
Annualized Implied Long-Run Excess Long-Only Return	Annualized average excess return one might expect over the long run from a manager with this style and skill level. This number is half of the historical long-short return shown above.

To show how stock-selection skill impacts this randomness, Figures 3 and 4 show equivalent scatter plots and relationship characteristics for growth and value portfolio managers, respectively, for three skill levels:

- Moderate skill,
  - 3% edge: stocks are placed in the right quintile 3% of the time more than pure random chance, 23% vs. 20% of the time,
  - consistent with 3.1% annualized long-run long-only outperformance for growth (before transactions costs) and 2.8% for value,
- High skill,
  - 5% edge
  - consistent with 4.5% outperformance for growth and 4.1% for value
- Perfect skill,
  - 80% edge
  - ranking of stocks reflect actual future earnings with perfect foresight
  - consistent with 16.5% outperformance for growth and 15.3% for value.

The core observation is that, while increases in portfolio manager skill generate higher and higher returns, as evidenced by the slope of the regression line and the expected excess return of the top-ranked stock, **higher levels of skill do not noticeably reduce the underlying level of uncertainty at the individual stock level with respect to the link between valuation and returns**.

Even when we allow for perfect foresight of future fundamentals and extraordinary rates of implied excess returns (16.5% per year for growth and 15.3% for value), the scatter diagrams show no reduction in uncertainty visible to the naked eye and, without the regression lines and related statistical analysis, it would be impossible to assess whether the valuation method was in fact adding value.

At the individual stock level, this high level of uncertainty dominates the risk-return problem. For a highly skilled growth manager, the top-rated stock would be expected to outperform the median stock by only 1.5% in any given quarter (this is the hi-mid spread in the tables). The uncertainty around that 1.5% of outperformance is enormous. The 50% uncertainty band is ±7.8%, meaning that 25% of the time that top stock would outperform the median stock by more than 9.3% and **25% of the time the top stock would underperform the median stock by more than 6.3%. (**To ease comparison to the long-run return statistics, the hi-mid spread annualizes to 5.8% outperformance with an uncertainty band of ±31.2%.)

Even for a growth manager with perfect foresight, the 50% uncertainty band at the individual stock level is still ±7.7% or ±30.8% on an annualized basis, even though such a portfolio manager would be expected to produce long-run returns a full 16.5 percentage points above and beyond their benchmark.

The situation is even worse if we shift our focus from the top stock to what could be called key index drivers (stocks with high index weights) with middling ratings. Such stocks are clearly capable of very strong and very weak performance, but it is simply impossible to forecast those returns with any accuracy through even the most insightful fundamental analysis.

Further, these results are true no matter how skillful the portfolio manager or the type of valuation methods employed.[10] The reason for this is that the level of dispersion of stock returns is so high that any method of valuation that meaningfully reduces the uncertainty of returns at the individual stock level would generate such high returns as to defy historical reality and common sense.

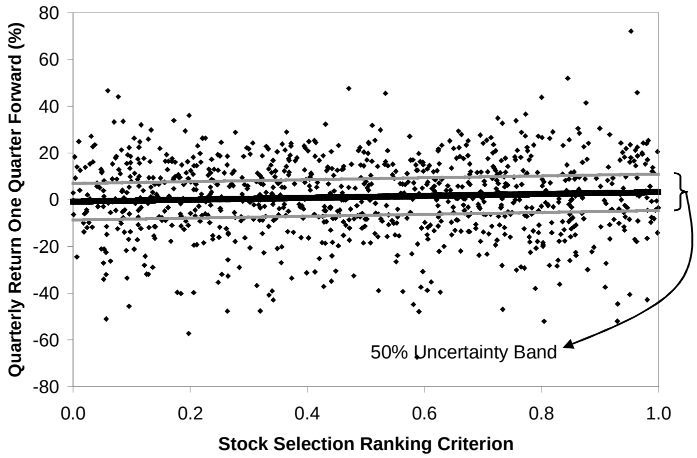
Put more simply, **even though skilled stock selection is capable of generating significant returns at the portfolio level, it is simply impractical to get individual stocks or even small groups of stocks "right."**

[10] See Appendix C to see this work repeated for a hybrid valuation method.

Figure 3: Growth – Relationship of Ranking Criterion to Returns

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

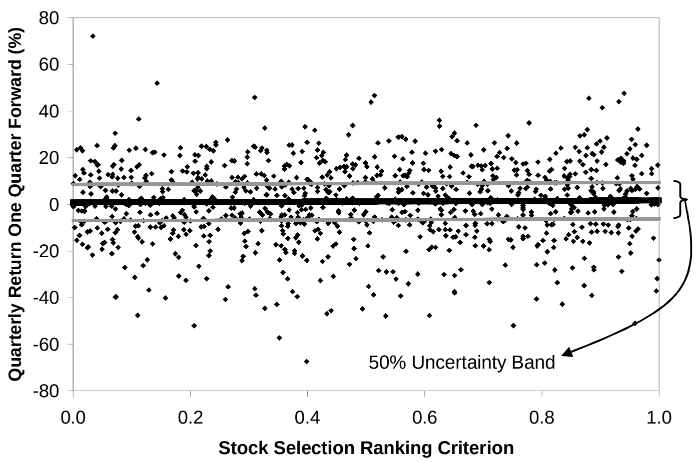
Panel 1: Portfolio Manager with Moderate Skill



Moderate Stock Selection Skill
(23% of Stocks in Correct Bucket)

High-Mid Spread	1.0 %
High-Low Spread	1.9 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	1.9 %
Standard Deviation	14.4 %
Average R-Squared	0.26 %
Annualized Long-Short Return	6.1 %
Annualized Implied Long-Run Excess Long-Only Return	3.1 %

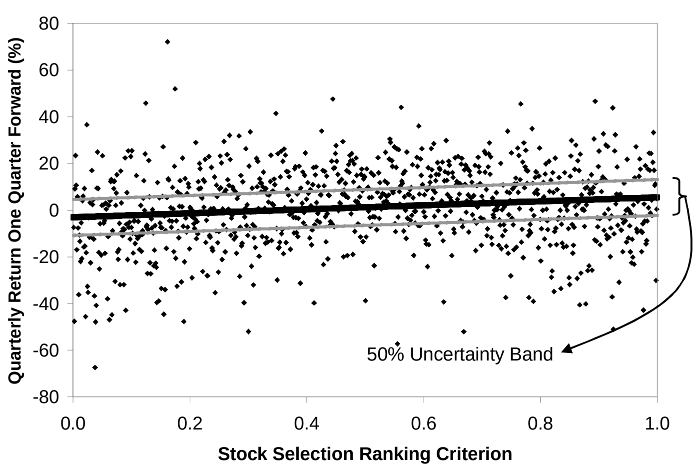
Panel 2: Portfolio Manager with High Skill



High Stock Selection Skill
(25% of Stocks in Correct Bucket)

High-Mid Spread	1.5 %
High-Low Spread	2.9 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	2.9 %
Standard Deviation	14.4 %
Average R-Squared	0.46 %
Annualized Long-Short Return	9.1 %
Annualized Implied Long-Run Excess Long-Only Return	4.5 %

Panel 3: Portfolio Manager with Perfect Skill



Perfect Stock Selection Skill
(100% of Stocks in Correct Bucket)

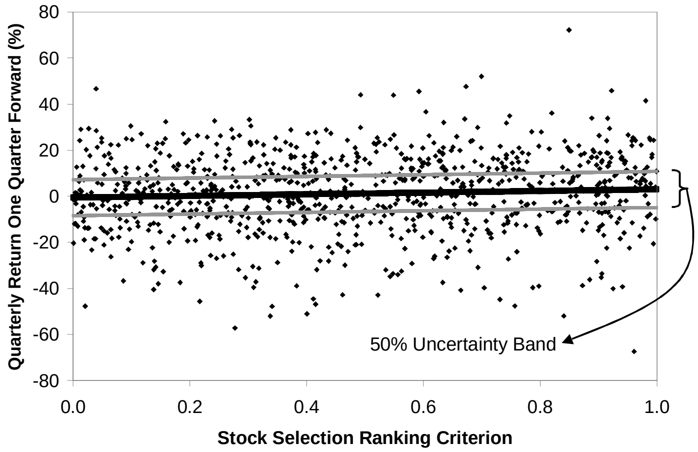
High-Mid Spread	5.1 %
High-Low Spread	10.2 %
50% Uncertainty Band	+/- 7.7 %
Average Slope	10.2 %
Standard Deviation	14.1 %
Average R-Squared	4.70 %
Annualized Long-Short Return	33.1 %
Annualized Implied Long-Run Excess Long-Only Return	16.5 %

Source: Goldman Sachs Research

Figure 4: Value – Relationship of Ranking Criterion to Returns

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

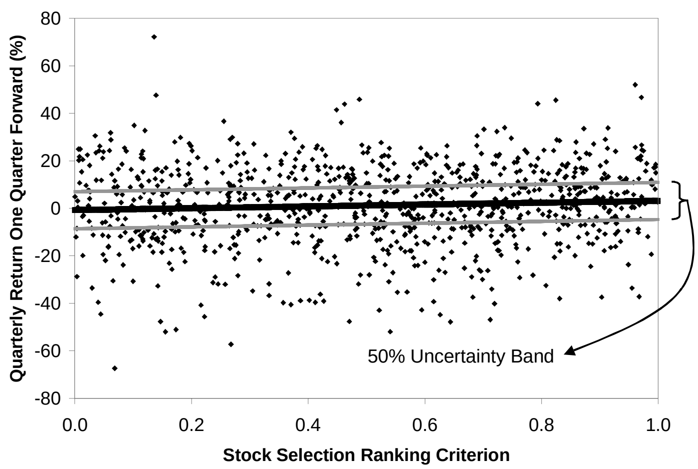
Panel 1: Portfolio Manager with Moderate Skill



Moderate Stock Selection Skill
(23% of Stocks in Correct Bucket)

High-Mid Spread	0.9 %
High-Low Spread	1.7 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	1.7 %
Standard Deviation	14.4 %
Average R-Squared	0.27 %
Annualized Long-Short Return	5.6 %
Annualized Implied Long-Run Excess Long-Only Return	2.8 %

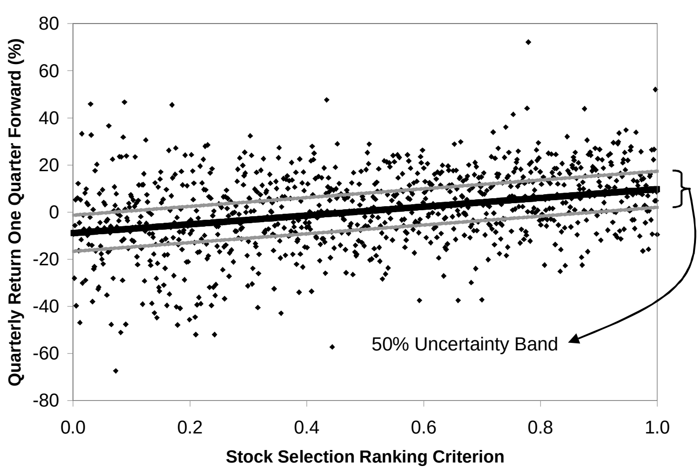
Panel 2: Portfolio Manager with High Skill



High Stock Selection Skill
(25% of Stocks in Correct Bucket)

High-Mid Spread	1.3 %
High-Low Spread	2.6 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	2.6 %
Standard Deviation	14.4 %
Average R-Squared	0.47 %
Annualized Long-Short Return	8.3 %
Annualized Implied Long-Run Excess Long-Only Return	4.1 %

Panel 3: Portfolio Manager with Perfect Skill



Perfect Stock Selection Skill
(100% of Stocks in Correct Bucket)

High-Mid Spread	4.6 %
High-Low Spread	9.3 %
50% Uncertainty Band	+/- 7.7 %
Average Slope	9.3 %
Standard Deviation	14.0 %
Average R-Squared	4.83 %
Annualized Long-Short Return	30.6 %
Annualized Implied Long-Run Excess Long-Only Return	15.3 %

Source: Goldman Sachs Research

Managing Uncertainty: Turning a Thin Edge into Consistency

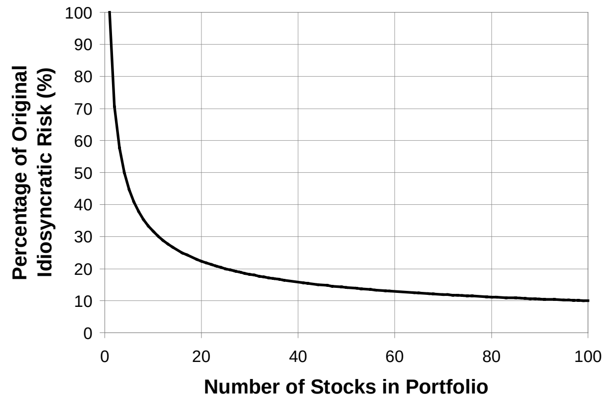
Conquering this type of randomness is easy in theory and not that difficult in practice. The solution is diversification. While individual stocks are subject to a high degree of uncertainty, as we increase the number of stocks, the randomness of the portfolio falls roughly as a function of the inverse of the square root of n $(1/\sqrt{n})$. See Figure 5, where n is the number of stocks in an equally weighted portfolio (more accurately, where n is the number of statistically independent risk positions). (Appendix A shows in detail how to calculate the level of diversification in non-equally weighted portfolios.) Thus, the standard deviation and 50% uncertainty bands fall to half their original sizes if the portfolio has 4 stocks in it and to one-tenth their original sizes if the portfolio has 100 stocks.

In practice, the effectiveness of diversification is strongly impacted by the correlations between individual stock positions – high correlations imply a smaller reduction in uncertainty, while negative correlations would imply even larger reductions in uncertainty. In fact, as will become clear, the core risk control issue for fundamentally driven portfolio managers is eliminating highly correlated risk positions so that diversification works and the underlying randomness in returns can be averaged out.

A simple way to see how effective diversification is at reducing the uncertainty of stock selection is to look at the impact of changing the number of equally weighted stocks in the portfolio on returns, tracking errors and Sharpe ratios. Returns fall as stocks with lower expected returns are added to the portfolio, but volatility (that is, tracking error) falls; thus, the risk-return efficiency as measured by the Sharpe ratio can rise as long as the resulting reduction in uncertainty more than offsets the reduction in returns.

Figure 6 shows the annualized average returns, the volatilities[11] (as measured by the standard deviations) and the Sharpe ratios for long-short portfolios with

Figure 5: Equal-Weight Portfolio Volatility Falls with the Inverse of the Square Root of n



Source: Goldman Sachs Research

different numbers of stocks.[12] We include a line of the inverse of the square root of n $(1/\sqrt{n})$[13] in the volatility graph to benchmark how well diversification is working in each case.

These graphs show that diversification works quite well without any risk control at all. The close correspondence between the square root of n baseline and the average standard deviation of the simulated portfolios, which corresponds to the tracking error of a long-short overlay portfolio, shows that the stock picks that arise from fundamental analysis are relatively uncorrelated across stocks. This is very good news in that it strongly suggests that, as we go forward to look at risk management approaches, **it will not be necessary to distort or even guide the stock selection process in any strong way as fundamentals and the underlying diversity of stocks will create all the diversification that is needed to generate much more consistent portfolio manager performance.**

The Sharpe ratio graph shows that, initially, adding stocks to the portfolio causes a rapid increase in the risk-return efficiency of the portfolio, but as the

[11] If the long-short portfolios were held as an overlay to a benchmark portfolio, these standard deviations would be the tracking errors.

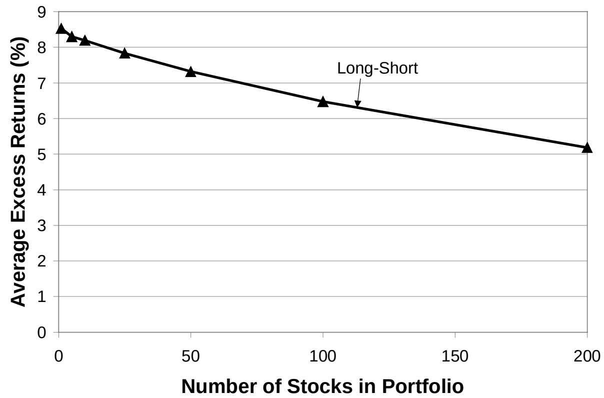
[12] Due to convergence issues for portfolios of small numbers of stocks, graphs with results from such portfolios represent 10,000 simulations rather than 1,000.

[13] Appendix D provides an explanation for why we use the square root of n in this context.

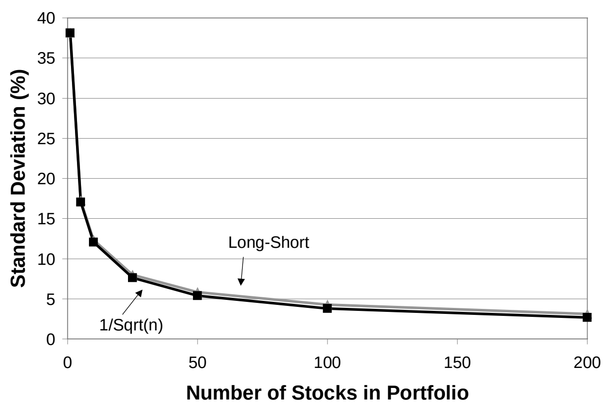
Figure 6: Impact of Increasing Number of Positions for the Average of Value and Growth, Moderate Skill, Long-Short Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

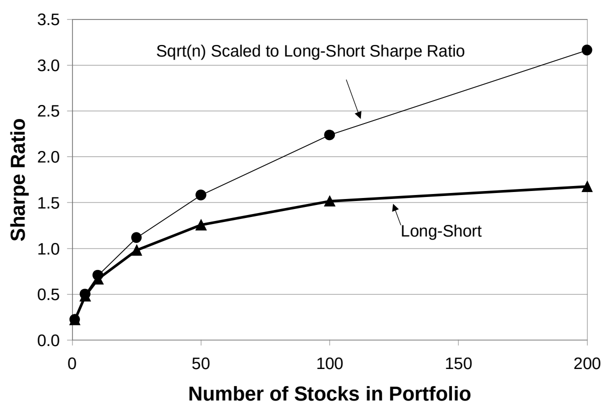
Panel 1: Annualized Returns



Panel 2: Volatility



Panel 3: Sharpe Ratios

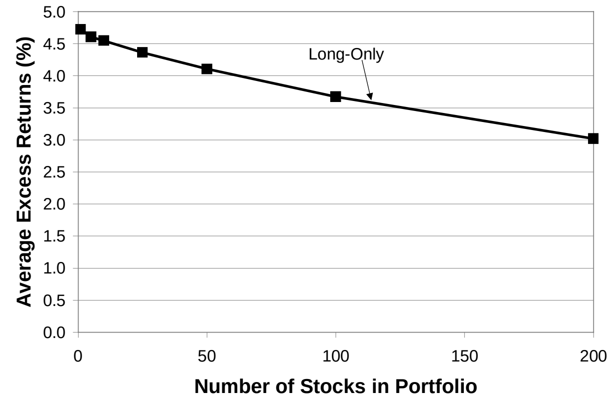


Source: Goldman Sachs Research

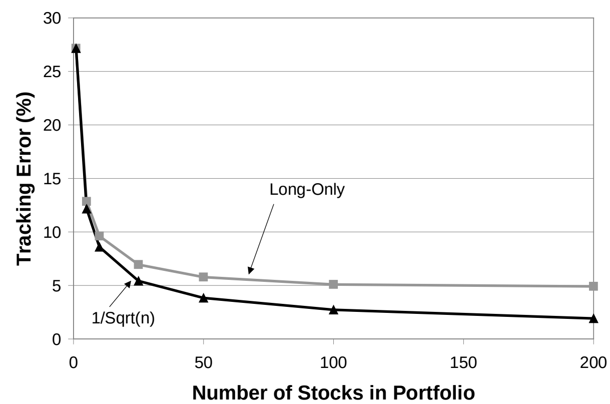
Figure 7: Impact of Increasing Number of Positions for the Average of Value and Growth, Moderate Skill, Long-Only Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

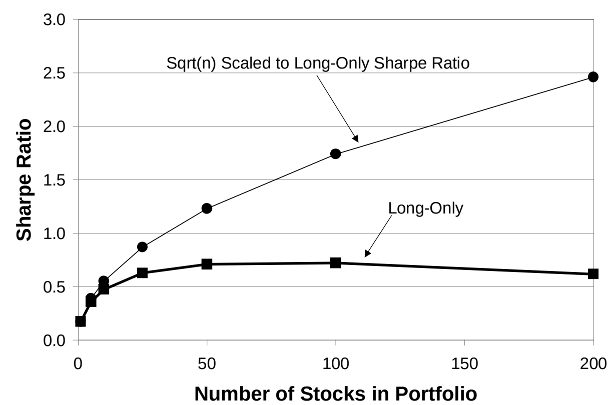
Panel 1: Annualized Returns



Panel 2: Tracking Errors



Panel 3: Sharpe Ratios



Source: Goldman Sachs Research

number of stocks exceeds 100 or roughly 10% of the stock universe, the rate of reduction in uncertainty begins to fall off, causing the decline in expected returns to take a more significant toll on the risk-return efficiency of the portfolio.

Net, the data strongly supports the notion that risk-return tradeoffs are improved by broadening rather than deepening research and stock-selection criterion. Similarly, these results suggest that extending holding periods will (by reducing the number of individual stock choices) reduce rather than improve the risk-return tradeoff.

These graphs also point to a significant trade-off between long-run returns and short-term risk-return efficiency. We will address this in more detail later, but it is already clear that consistency will have a price and that finding ways to improve this trade-off will have significant long-run benefits for portfolio managers and investors alike.

The problem facing real world portfolio managers, who cannot short stocks and thus cannot engage in long-short strategies, becomes evident if we redo this analysis using only the long portion of the portfolio and measure results against a Russell 1000 benchmark.[14] Figure 7 shows the return, tracking error and Sharpe ratio results and the line of the inverse of the square root of n for long-only portfolios. For long-only portfolios, diversification fails after the first few stocks.

In particular, the tracking error graph, which shows convergence to a higher level of nondiversifiable risk than shown by the $1/\sqrt{n}$ line rather than simply a slower convergence to a common risk level, implies some common risk position in all stock positions relative to the benchmark that diversification in the active portfolio is identifying rather than eliminating. **As we shall show, it is this common risk position and not stock selection that has made it so difficult even for skilled managers to consistently outperform benchmarks.**

[14] As described in Appendix B, the data sample we use is an approximation of the stocks in the Russell 1000 index. Because the difference in the cap-weighted mean return of our sample and the actual Russell 1000 return could bias the results, the excess returns we report are actually the excess above the cap-weighted mean of our Russell 1000 sample rather than above the index return. Unless otherwise specified, in this report, Russell 1000 and Russell 2000 refer to our estimated samples.

Solving the Risk Management Problem

Large-Cap First – The Russell 1000 Universe

So what is the common risk factor and what can portfolio managers do to eliminate it? Once the common risk factor is eliminated, what other risk control is needed/desired?

The common risk factor turns out to be stocks with large index weights. The simplest and most effective risk correction is to hold a passive position[15] in those stocks (or an equivalent derivative) to offset that concentration of stock-specific risk. Beyond that, as was implied by the long-short results, little risk management will turn out to be necessary, although as we will discuss later, some additional risk controls can help portfolio managers, but those controls need to be carefully tailored to the specific portfolio manager and depend importantly on that portfolio manager's specific skills, weaknesses and research methodologies.

These conclusions might seem surprisingly simple given the broad failure of risk models to noticeably improve portfolio manager performance over the last decade, but as we show, macro approaches do not correctly address the problems of a bottoms-up fundamentally driven portfolio manager and, once the perspective is shifted to the individual stock level, the problems become much simpler.

Put differently, **we find that the types of macro risk positions that arise naturally out of bottoms-up analysis are, in fact, justified on a risk-return basis and do not need to be controlled. The risks that prove to be both important in size and unjustified are those that arise out of mismatches between the portfolio manager's natural base portfolio and the benchmark.** Such mismatches generate persistent risk positions that do not reflect the portfolio manager's judgement about investment opportunities, and, hence, are rarely justified on a risk-return basis. **Macro risk systems indiscriminately work at reducing both types of risk and are, in general, more effective at eliminating the good risk driven by a portfolio manager's stock selection than they are at eliminating the habitual risk patterns that do not offer reasonable expectations of return.**

To show this formally and understand the key drivers of these conclusions and the real world solutions to the risk management problem, we need to analyze the match between fundamentally driven stock selection and various risk control approaches. We first look at these questions from the standpoint of an orthodox manager following a pure investment style with complete research coverage across all sectors and size groups in their investment universe and no idiosyncratic biases in accuracy across categories or types of stocks. In the final section, we re-examine the question from the perspective of a more idiosyncratic portfolio manager with research strengths and weaknesses, investment prejudices, non-standard valuation methods and correlated patterns of forecast accuracy.

In the current context, we identify three obvious potential concentrations of common risk – size, sector and individual stock positions. Some readers may wonder at the notion that benchmarks can contain large concentrations of individual stock risk. In fact, one of the key underlying assumptions in the way most portfolio managers and most macro risk models approach indices is that stock-specific risk in indices has been diversified away. In the case of large-cap indices, this assumption is patently false. (It is more reasonable for mid- and small-cap indices as we show later.)

Equal-weighted indices quickly diversify away stock-specific risk following the inverse of the square root of N $(1/\sqrt{N})$ rule discussed earlier. For cap-weighted indices, the question of diversification is much more subtle. Appendix A develops the mathematics in some detail, but the key point is quite simple – **if the weight of the stock in an index exceeds 2/(N+1) where N is the number**

[15] Because our portfolios are rebalanced every calendar quarter, the positions in the largest stocks we describe as passive are not entirely passive. Change is due to turnover in the set of the largest stocks, which, for our purposes, is more of an issue in the Russell 2000 than it is in the Russell 1000 or the S&P 500.

Table 3: Effect of Macro Risk Control Methods
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill (3% Edge)

	Value				Growth			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Unadjusted	2.8	4.73	0.58	58.3	3.0	4.85	0.62	60.3
Control for Size	1.9	2.77	0.69	57.9	2.3	2.73	0.84	59.8
Control for Sector	2.8	4.17	0.66	60.2	2.7	4.32	0.63	60.0
Control for Sector and Size	1.7	2.68	0.65	58.7	1.7	2.70	0.62	58.6
Long-Short	5.6	3.29	1.71	80.7	6.1	3.09	1.98	84.4

Panel 2: High Skill (5% Edge)

	Value				Growth			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Unadjusted	4.1	4.81	0.85	64.6	4.5	4.98	0.90	66.1
Control for Size	2.8	2.88	0.98	63.9	3.4	2.75	1.24	65.5
Control for Sector	4.1	4.21	0.99	67.1	3.9	4.40	0.89	65.9
Control for Sector and Size	2.7	2.69	1.00	63.1	2.5	2.74	0.92	62.6
Long-Short	8.3	3.70	2.23	87.5	9.1	3.28	2.77	92.3

Source: Goldman Sachs Research

of stocks in the index, then that stock adds more stock-specific risk than it diversifies away.[16]

Looking at the S&P 500[17] like this would suggest that somewhere between the 50 and 75 largest stocks are adding significant stock-specific risk. Such concentration of stock-specific risk can act as a common risk position against all of the portfolio manager's individual stock positions.

In Table 3, we show the relative effectiveness of different macro risk control approaches for our value and growth managers. For portfolio managers with moderate and high skill levels, we simulate 1,000 portfolios based on simulated rankings for different risk management approaches focused on controlling size, sector and concentrations of individual stock risk. We then report the annualized averages for the returns, tracking errors and Sharpe ratios for each approach.

The macro risk controls are imposed by stratified risk sampling methods often used in the construction of polling data. This means that stock picking is only allowed within groups of controlled categories. Thus, for the size risk control results, the stocks universe is divided into 10 decile ranges (smallest 10%, next larger 10%, ..., largest 10%) and the best stocks in each decile are chosen according to the fundamental ranking criterion. The stocks chosen in each size decile are equally weighted. Then, each decile portfolio is given a portfolio weight equal to that decile's share of the index.

For example, if we have a sample of 1,000 stocks and we want to construct a 50-stock long-only portfolio, we start by dividing the 1,000 stock universe into 10 size deciles of 100 stocks each. Then, we pick the best 5 stocks from each decile. Within each decile, the 5 stocks are equal-weighted. Then, the 10 decile portfolios are combined by weighting the portfolios by the share of the index market capitalization in that decile. As of July 30, 1999, that meant the largest decile in an estimated Russell 1000 sample was weighted at roughly 60%.

Similarly, for sector controls, we break the data into the 11 Compustat economic sectors, the best stocks

[16] Formally, Appendix A defines an effective N that takes account of the range of index weights applied across the capitalization spectrum.
[17] Appendix E provides some statistics on the concentration of market capitalization in the largest stocks in the S&P 500.

are chosen within each sector, and the sector is then weighted by the capitalization of the sector in the index. For the joint size and sector controls, the stocks are broken up into size/sector groupings, the best stocks within each size/sector group are chosen, and then the portfolio is assembled from these sub-portfolios by cap-weighting.

Using this type of risk control allows us to look at how well stock selection is working, both within the categories and how well it is working at generating cross-category risk positions. That is, we can examine whether it is better to pick stocks within sectors or it is better to allow sector overweights that arise naturally out of bottoms-up analysis.

A first pass at interpreting Table 3 would suggest some moderate gain from macro risk controls, especially controlling for size. Tracking error is reduced dramatically, but the Sharpe ratio only improves modestly as returns also fall dramatically.

Given the large drop off in returns that arise from reduced risk taking and the modest improvement in the quality of risk, it is little wonder that portfolio managers view risk control with more than modest suspicion that it is doing more harm than good over the long haul. The size "bets" would be expected to average out over time, but reduced risk taking would still impact the portfolio manager's cumulative returns exactly in proportion to the quarterly reductions in returns.

Sector controls appear to have little value in risk control as they have little impact on returns or Sharpe ratios.

Dealing with Stock-Specific Benchmark Risk

However, **because these results ignore the potential impact of concentrations of stock-specific risk, they are actually highly misleading.** As macro size risk and stock-specific risk are focused in the same large-cap stocks, it is easy to mistake one for the other. However, the operational methods of offsetting the two risks are completely different and the resulting impact on quality of risk is equally different.

Controlling concentrations of stock-specific risk is quite simple. The portfolio manager can simply market-weight the largest stocks. The downside of this approach is that every dollar used to offset these concentrations is no longer available for generating outperformance through active management, so the loss in long-run outperformance is equal to the percentage of funds used to offset stock-specific risk. (In Appendix F, we examine strategies aimed at reducing the necessary funds.)

The impact of such a risk control approach on Sharpe ratios is dramatic, especially in comparison with the modest impact of the size-based risk controls. In Table 4, we show the results of adding a market-weighting of the largest stocks (from 0 to 100) to an otherwise un-risk-controlled long-only portfolio. For comparison, we also include the unadjusted long-short returns, which can be thought of as a measure of the total unconstrained portfolio manager's potential for extracting value from their ability to rank stocks.

After offsetting the stock-specific risk of the top 50 stocks, the long-only portfolio manager has doubled their Sharpe ratio and recaptured approximately

Table 4: Effect of Offsetting Stock-Specific Risk

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill (3% Edge)

	Value				Growth			
Number of Largest Stocks Index-Weighted	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Portolios								
0	2.8	4.73	0.58	58.3	3.0	4.85	0.62	60.3
10	2.3	3.39	0.69	59.9	2.6	3.51	0.73	62.3
20	2.2	2.56	0.85	63.1	2.4	2.72	0.88	65.3
50	2.0	1.56	1.30	73.8	2.2	1.68	1.30	74.3
75	1.7	1.20	1.42	76.2	1.8	1.34	1.37	75.7
100	1.5	0.96	1.52	77.8	1.6	1.09	1.45	76.9
Long-Short Portfolios								
0	5.6	3.29	1.71	80.7	6.1	3.09	1.98	84.4

Panel 2: High Skill (5% Edge)

	Value				Growth			
Number of Largest Stocks Index-Weighted	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Portolios								
0	4.1	4.81	0.85	64.6	4.5	4.98	0.90	66.1
10	3.4	3.46	1.00	66.9	3.8	3.62	1.05	69.1
20	3.2	2.62	1.21	70.9	3.5	2.82	1.23	72.8
50	2.8	1.61	1.74	81.9	3.0	1.77	1.72	81.5
75	2.4	1.24	1.91	84.2	2.6	1.41	1.82	83.1
100	2.0	1.00	2.05	85.9	2.2	1.16	1.93	84.1
Long-Short Portfolios								
0	8.3	3.70	2.23	87.5	9.1	3.28	2.77	92.3

Source: Goldman Sachs Research

Table 5: Effect of Offsetting Stock-Specific Risk and Controlling for Macro Risk
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill (3% Edge)

Index-Weight Largest 50 Stocks	Value				Growth			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only	2.0	1.56	1.30	73.8	2.2	1.68	1.30	74.3
Control for Size	1.5	1.82	0.82	73.3	1.7	1.88	0.91	74.1
Control for Sector	2.0	1.50	1.36	75.8	2.0	1.64	1.23	74.3
Control for Sector and Size	1.4	1.95	0.72	73.1	1.4	2.03	0.67	71.9

Panel 2: High Skill (5% Edge)

Index-Weight Largest 50 Stocks	Value				Growth			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only	2.8	1.61	1.74	81.9	3.0	1.77	1.72	81.5
Control for Size	2.0	1.82	1.11	80.9	2.4	1.89	1.25	81.1
Control for Sector	2.8	1.50	1.90	83.6	2.7	1.69	1.61	80.6
Control for Sector and Size	2.0	1.94	1.01	79.3	1.8	2.05	0.90	77.6

Source: Goldman Sachs Research

three-quarters and two-thirds of the efficiency in utilizing value and growth fundamentals, respectively, to generate returns lost by being constrained to be long-only. **Clearly, the concentration of stock-specific risk has far more impact than the macro risk factors and represents the primary risk management challenge to large-cap managers.**

A simple interpretation of this result, which is shown even more clearly later, is that **the concentration of stock-specific risk in the large-capitalization indices is so large that the indices are taking more stock-specific risk than the portfolio manager.**[18] As a result, the portfolio manager's performance relative to the benchmark is driven by the index rather than the skill of the portfolio manager. Only by neutralizing the risk concentration in the index can the portfolio manager's skill show through.

The importance of this concentration of stock-specific risk in large-cap benchmarks becomes especially clear if we redo the macro risk control analysis taking account of the stock-specific risk. Table 5 repeats the analysis on macro risk controls for portfolios where the top 50 stocks are market-weighted. Note, that if these stocks are chosen as part of the active portfolio, they can have a final portfolio weight above the market weight, although they cannot be underweighted.[19] Once the stock-specific risk is offset, size controls are distinctly counterproductive, while sector controls now generate noticeable improvement for value managers in terms of higher Sharpe ratios without reducing returns.

The size controls continue to suffer from overly concentrating active management risk into a small number of stocks in the top deciles, losing efficiency (that is, increasing tracking error) as the cap-weighting of the size segments reduces the effective number of names in the portfolio (see Appendix G for the exact mathematics) and reducing returns as the effectiveness of active management falls as the capitalization of the stocks increases (see "Style, Size and Skill").

Sector controls, in contrast, eliminate the returns from over- and underweighting sectors, but appear to offset this by improving the mapping from fundamentals to returns. The gains are not strong enough to view this result as definitive without reference to the particular skill set of the portfolio

[18] For the mathematically inclined who believe that benchmarks cannot by definition take risk relative to the overall market this statement is given a precise mathematical meaning in Appendix A.

[19] In Appendix F, we look at lifting this restriction.

manager, but are sufficient to warrant careful investigation for a portfolio manager seeking to use risk control to improve performance.[20]

We would interpret these results as suggesting that **tailoring sector controls to the portfolio manager's investment process (strengths and weaknesses) is likely to offer value, but macro level size controls are unlikely to do anything but reduce returns and portfolio performance.**

Finally, if we redo the analysis on the impact of the number of stocks in the portfolio on Sharpe ratios and tracking error after controlling for the concentration of stock-specific risk in the very large-cap stocks (see Figure 8), we can see that market-weighting the top 50 and top 100 stocks does indeed remove most of the common risk factor as the tracking errors now much more closely follow the square root of *n* line in much the same way as the long-short portfolios do.

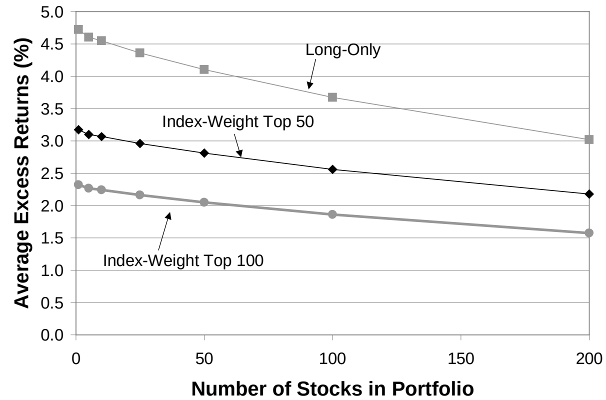
The fall-off in returns and the leveling off of the risk-return efficiency of the portfolio as the number of stock approaches 100 would suggest that portfolio managers concerned with short- and medium-term risk-return efficiency would likely want to hold between 75 and 125 stocks. The desired number of stocks would be lower the higher the portfolio manager skill level, the longer the investment horizon and the more funds under management (as a result of liquidity issues).[21]

The fall-off in returns resulting from the funds dedicated to offsetting the stock-specific risk concentration does raise some serious questions about the trade-off between consistency and long-run outperformance. This overhead cannot be eliminated as long as the benchmark contains such large concentrations of stock-specific risk.

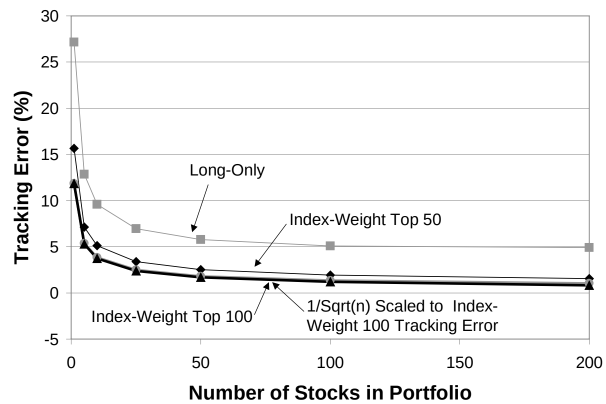
Figure 8: Impact of Increasing Number of Positions for the Average of Value and Growth, Moderate Skill, Long-Only Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

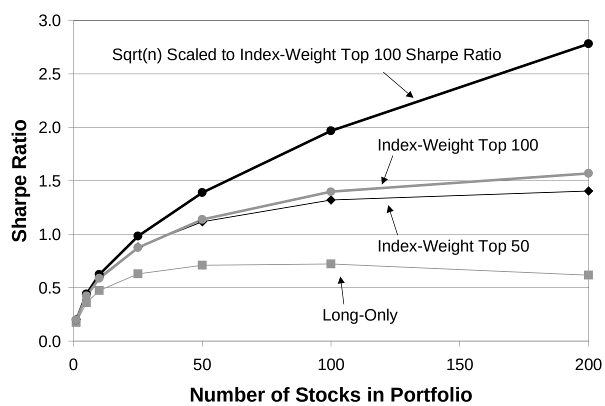
Panel 1: Annualized Returns



Panel 2: Tracking Errors



Panel 3: Sharpe Ratios



Source: Goldman Sachs Research

[20] Using risk control to improve performance in practice is discussed in the final section on optimizing risk control to take advantage of a specific portfolio manager skill set. This topic will be further developed in the next paper in the series *A Stockpicker's Reality*, which will focus on sectors from the perspective of bottoms-up portfolio managers.

[21] For portfolio managers who tend to take correlated sector positions, the desired number of stocks could be substantially higher. Some related issues are discussed in the section on how these recommendations need to be tailored to individual portfolio managers.

A first step in reducing that overhead is to allow limited short selling of the stocks in the passive portfolio. That is, the passive positions held in those stocks that would naturally qualify for short positions based on the portfolio manager's fundamental analysis could be reduced in size from the full index weight by the same dollar size as a full long position. This would both create a new active position and free up cash to be added to the active long portfolio. The two key points are that these implicit benchmark shorts need to be based on the same type of fundamental analysis as the longs and need to be limited in size by the same diversification criterion discussed earlier and developed in detail in Appendix A. That is, no short relative to benchmark should exceed twice the size of the average active portfolio position.

Appendix F looks at additional strategies to minimize the overhead. The last section of the paper, which focuses on this type of long-run outperformance trade-off, discusses how benchmarks might be revised to significantly reduce or even eliminate this problem.

Performance Over Time

Perhaps the most dramatic way of seeing the importance of the impact of the stock-specific risk embedded in large-capitalization benchmarks on portfolio manager consistency relative to their large-capitalization benchmarks is simply to graph the mean performance of simulated groups of high- and moderate-skill portfolio managers over time with and without offsetting the stock-specific risk.

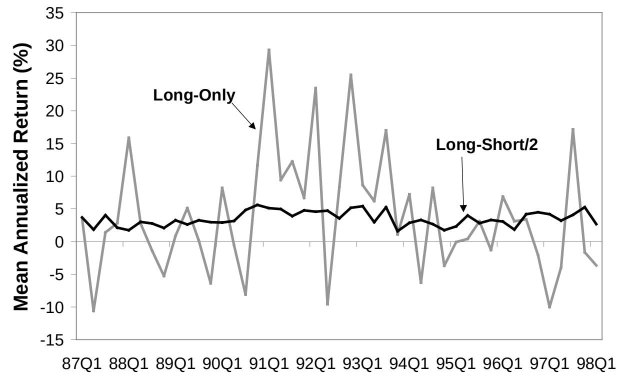
Figures 9 and 10 show the average excess return for equivalently skilled long-only portfolio managers without risk control of any type and long-short portfolio managers for moderate and high skill, respectively. The long-short returns are divided by two to bring the total returns in line with the potential for a long-only manager (much as we used one half of the long-short return to estimate the implied long-run excess returns for long-only portfolio managers) and are included to show the ability of stock selection to distinguish between high- and low-returning stocks. The returns are based on 50%-50% composites of growth and value.

The long-short returns are quite stable and show a remarkable consistency in the ability of a skilled portfolio manager to use fundamentals to distinguish between higher- and lower-returning stocks. In contrast, the long-only excess returns are far more variable and subject to massive swings that apparently have little to do with the ability of fundamentals to distinguish between higher- and lower-returning stocks.

Figures 11 and 12 repeat these graphs using the same scaling, but the long-only portfolios are supplemented with a passive market-weighted holding of the top 50 and top 100 stocks by market capitalization. The change is dramatic. The long-only portfolio is now nearly as stable and consistent as the long-short portfolio and massively more stable than the long-only without the passive stock-specific risk offset. In Figures 13 and 14, we compare passively holding the 50 and 100 largest stocks. Passively holding the largest 100 stocks noticeably smoothes the outperformance. (To facilitate a clearer comparison between the long-short and long-only with passive supplemental portfolios, Figures 11 through 14 are rescaled in Appendix H.)

Figure 9: Moderate Skill, Composite Strategy, Long-Only and Long-Short Returns

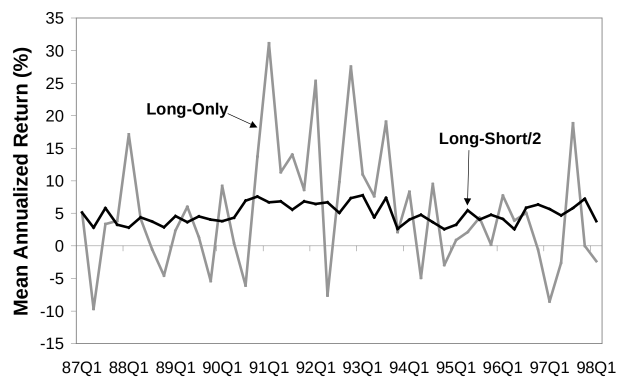
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Goldman Sachs Research

Figure 10: High Skill, Composite Strategy, Long-Only and Long-Short Returns

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Goldman Sachs Research

Thus, without any risk control beyond a market weighting of the top 50 or 100 stocks, the long-only portfolio performance relative to the benchmark is massively stabilized.

It is important to understand that passive risk control can work without any reference to the portfolio manager's portfolio because the risk problem that is being addressed is a concentration of stock-specific risk in the benchmark rather than in the portfolio manager's stock selection. Attempts to address this problem through interfering with the portfolio manager's investment process are almost certain to hinder rather than help long-run performance.

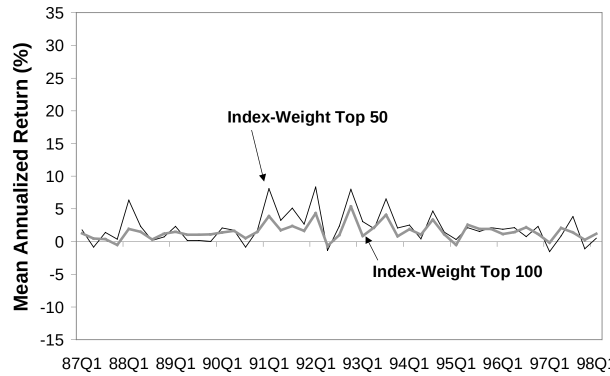
Figure 11: Moderate Skill, Composite Strategy, Long-Short and Index-Weight Top 50 Long-Only

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Source: Goldman Sachs Research

Figure 13: Moderate Skill, Composite Strategy, Index-Weight Top 50 and 100 Long-Only

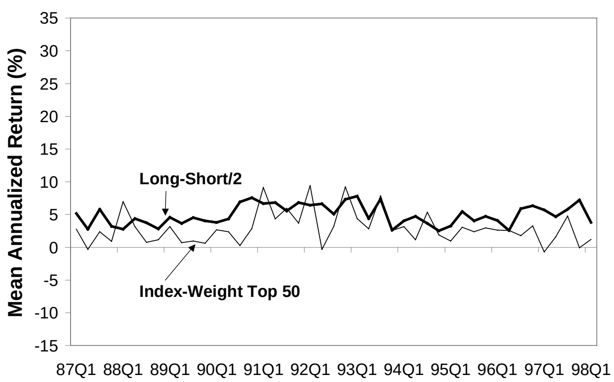
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Goldman Sachs Research

Figure 12: High Skill, Composite Strategy, Long-Short and Index-Weight Top 50 Long-Only

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Goldman Sachs Research

Figure 14: High Skill, Composite Strategy, Index-Weight Top 50 and 100 Long-Only

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Source: Goldman Sachs Research

Real World Evidence

At this point, the reader could easily be forgiven for some skepticism that such a simple solution can so dramatically solve such a large problem, after all, "Isn't the real world far more complicated than these statistical simulations suggest?"

Luckily, our results suggest some straightforward real world back-tests that can be done on actual portfolio manager performance to see if the stock-specific risk in the large-cap benchmarks is, in fact, anywhere near as important as these results seem to suggest.

Specifically, if the real risk management problem faced by large-cap portfolio managers is the concentration of stock-specific risk at the top end of the capitalization spectrum, then historical portfolio manager performance should be determined both by the overall market and the performance of this stock-specific concentration relative to an equal-weighted index (an equal-weight index minimizes the impact of stock-specific behavior).[22]

In the regressions reported in Table 6, average portfolio manager performance (for all managers in the Lipper database and for managers broken down by style[23]) is explained by market performance (as measured by the S&P 500) and by the stock-specific factor (measured as the difference between the equal-weighted performance of an estimated S&P 500 sample and capitalization-weighted performance of the S&P 500).

Table 6: Impact of Stock-Specific Risk on the Average Portfolio Manager

(1Q1987-1Q1998)

Panel 1: Average of All Managers in Lipper Database

Variable	Estimate	T-Stat
Intercept	-0.01	-0.72
Cash Drag	0.08	0.48
S&P 500	1.05	6.33
EW-CW S&P	0.70	6.54

R-Squared = 95.7%

Panel 2: Average of Growth Managers

Variable	Estimate	T-Stat
Intercept	-0.01	-0.76
Cash Drag	0.10	0.40
S&P 500	1.17	4.57
EW-CW S&P	0.65	3.91

R-Squared = 92.3%

Panel 3: Average of Value Managers

Variable	Estimate	T-Stat
Intercept	0.00	-0.02
Cash Drag	0.03	0.22
S&P 500	0.83	6.32
EW-CW S&P	0.56	6.62

R-Squared = 96.3%

Source: Lipper Funds Database and Goldman Sachs Research

---

[22] Similar regression results have been found by Joanne Hill and Bob Jones ("Domestic Equity Benchmark Underperformance," Pension & Endowment Forum, Goldman, Sachs & Co., June 1996) with the broad interpretation of estimating a size effect (these authors did not distinguish between the macro factor and the related stock-specific risk.) As our prior results make clear empirically and Appendix A makes clear mathematically, we think the key issue and appropriate interpretation of this variable is as a measure of stock-specific risk and not as a macro factor.

[23] Funds are classified into styles using a methodology akin to William Sharpe's method of clustering by performance. Funds are divided according to their correlations with the difference between our growth and value strategy returns. The existence of a third category covers cases where the correlations are not definitive. See Appendix I for more details on this methodology.

The results are dramatic. The stock-specific risk variable has coefficients between 0.56 and 0.70, implying that for every 100 basis points that the stock-specific risk concentration in the S&P 500 outperformed the equal-weight estimate of the common market factor, the average value portfolio manager underperformed their large-cap benchmark by 56 basis points, the average growth portfolio manager underperformed their large-cap benchmark by 65 basis points and the overall average portfolio manager underperformed their large-cap benchmark by 70 basis points.

To demonstrate the importance of stock-specific risk visually, Figure 15 graphs the stock-specific risk variable (i.e., the difference between the return from the capitalization-weighted and equal-weighted S&P 500 indices) and the mean portfolio manager performance over the last 12 years. The impact of the stock-specific risk is quite evident and the correlation of 0.33 is statistically significant.

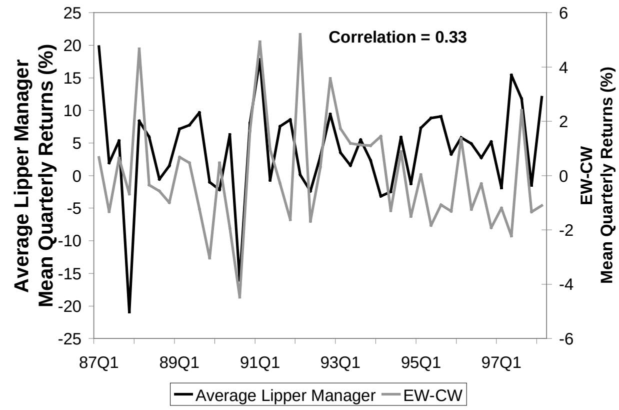
In contrast, Figure 16 compares the average portfolio manager performance to our long-short portfolio manager to see how much the ability of fundamentals to distinguish between high- and low-performing stocks is driving performance. Here there is almost no relationship and the correlation is not statistically different from zero.

The clear message of these comparisons is that **the history of short-term active manager performance relative to large-cap benchmarks has not been determined by the effectiveness of fundamental analysis, but rather it has been determined by the performance of the stock-specific risk concentration embedded in the large-cap indices.**

For portfolio managers seeking to consistently outperform a large-cap benchmark, the message is equally clear. Stock-specific risk concentration in indices must be offset with passive positions; otherwise, the concentration of stock-specific risk in the large-cap benchmark is likely to overwhelm the portfolio manager skill. **In essence, portfolio managers must make sure that the stock-specific risk they take on purpose is larger than the stock-specific risk embedded in their benchmark.**

Figure 15: Stock-Specific Risk Concentration and the Average Lipper Manager

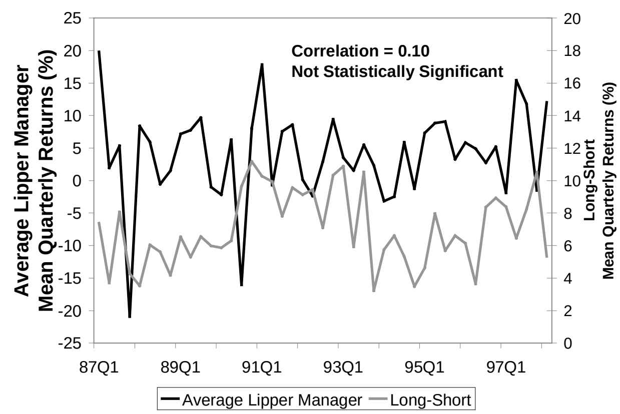
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Lipper and Goldman Sachs Research

Figure 16: Composite Long-Short Portfolio and the Average Lipper Manager

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Lipper and Goldman Sachs Research

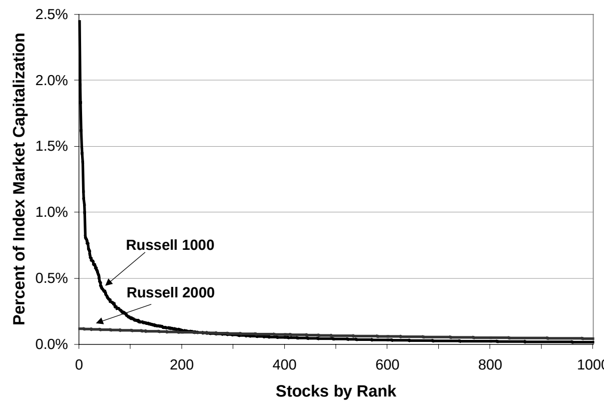
Mid- to Small-Cap, The Other Real World

Until this point, the analysis has focused on the large-cap world of the Russell 1000 and S&P 500. The Russell 2000 universe poses entirely different issues. Historically, portfolio managers who focus on this segment have had few persistent risk management problems in terms of beating benchmarks. The largest of these problems is the occasional runaway segment of stocks that outperforms the index by so much that the portfolio manager participation in that segment becomes a large, if transient, determinant of performance relative to benchmark.

Given our prior results, this is not surprising. Figure 17 shows the size distribution of the Russell 1000 and Russell 2000 samples. In the Russell 2000, few stocks cross the 2/(N+1) barrier and then only by a modest amount. These size distributions suggest that, if our focus on stock-specific risk is correct, neither size nor stock-specific risk should significantly impact portfolio manager consistency with respect to small-cap benchmarks. This, is in fact, the case.

Figure 17: Concentrations of Stock-Specific Risk in Estimated Russell 1000 and Russell 2000 Samples

(As of July 30, 1999)



Source: FactSet and Goldman Sachs Research

As Tables 7 and 8 show, for the Russell 2000, controlling for both size and stock-specific risk simply reduces returns with no compensating improvement in quality of risk. Sector controls once again add modest value.

Table 7: Controlling for Macro Risk in the Small-Cap (Estimated Russell 2000) Sample

(Estimated Russell 2000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill (3% Edge)

	Value				Growth			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Unadjusted	6.2	2.80	2.20	87.9	6.5	2.81	2.31	88.6
Control for Size	4.7	2.53	1.85	82.6	4.9	2.60	1.88	83.0
Control for Sector	6.1	2.67	2.27	88.7	5.9	2.63	2.22	87.9
Control for Sector and Size	4.5	2.59	1.75	81.3	4.3	2.66	1.62	79.2
Long-Short	10.3	3.95	2.61	90.9	10.8	3.93	2.75	92.3

Panel 2: High Skill (5% Edge)

	Value				Growth			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Unadjusted	8.7	2.93	2.96	94.9	9.0	2.98	3.04	94.7
Control for Size	7.0	2.70	2.58	90.7	7.2	2.79	2.58	90.9
Control for Sector	8.6	2.77	3.09	95.6	8.2	2.70	3.05	95.1
Control for Sector and Size	6.7	2.63	2.56	90.6	6.4	2.76	2.31	88.3
Long-Short	15.3	4.29	3.56	96.3	15.9	4.33	3.68	97.3

Source: Goldman Sachs Research

Overall, these results suggest that mid- and small-cap portfolio managers need little in the way of standard macro or stock-specific risk management, although, once again, it appears that sector controls might be useful if carefully designed to match the portfolio manager's investment process.

Perhaps most importantly, the joint absence of extremely high index-weight stocks and risk management problems in the small- and mid-cap range can be viewed as additional confirmation of the prior analysis of the importance of the concentration of stock-specific risk.

However, a very important caveat must be added. Our results as have been discussed previously use simulated benchmarks that are rebalanced quarterly. For the large-cap indices, this is a harmless approximation. In the current context, it is less innocent. In particular, the Russell 2000 is rebalanced annually. As a result, if a group of stocks near the top end of the capitalization spectrum perform strongly relative to the rest of the index, they will create a temporary, but significant concentration of stock-specific risk before they are shifted into the Russell 1000 at the reconstitution.

Table 8: Controlling for Stock-Specific Risk of the Top 50 and 100 in the Small-Cap (Estimated Russell 2000) Sample

(Estimated Russell 2000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill (3% Edge)

		Value				Growth			
	Number of Largest Stocks Index-Weighted	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only	**50**	5.4	2.48	2.19	87.6	5.7	2.44	2.35	88.9
Control for Size	**50**	4.1	2.33	1.76	81.3	4.3	2.34	1.82	82.2
Control for Sector	**50**	5.3	2.35	2.27	88.6	5.2	2.31	2.23	88.0
Control for Sector and Size	**50**	4.0	2.34	1.69	80.3	3.8	2.41	1.56	78.4
Long-Only	**100**	4.9	2.14	2.27	88.3	5.1	2.13	2.41	89.3
Control for Size	**100**	3.7	2.16	1.69	80.7	3.8	2.19	1.74	81.2
Control for Sector	**100**	4.8	2.02	2.37	89.1	4.6	2.03	2.27	88.1
Control for Sector and Size	**100**	3.5	2.15	1.64	79.8	3.4	2.27	1.48	77.4

Panel 2: High Skill (5% Edge)

		Value				Growth			
	Number of Largest Stocks Market-Weighted	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only	**50**	7.7	2.60	2.97	94.8	8.0	2.57	3.12	95.1
Control for Size	**50**	6.1	2.48	2.48	89.9	6.4	2.48	2.56	90.5
Control for Sector	**50**	7.6	2.43	3.14	95.6	7.3	2.36	3.10	95.1
Control for Sector and Size	**50**	5.9	2.35	2.52	90.2	5.6	2.48	2.27	87.8
Long-Only	**100**	6.9	2.25	3.07	95.1	7.2	2.25	3.20	95.4
Control for Size	**100**	5.5	2.28	2.41	89.9	5.7	2.31	2.46	90.5
Control for Sector	**100**	6.8	2.08	3.28	95.6	6.5	2.07	3.16	95.1
Control for Sector and Size	**100**	5.3	2.15	2.47	90.2	5.0	2.33	2.16	87.8

Source: Goldman Sachs Research

Table 9 shows the concentration that has sometimes occurred in the top ten stocks before and after rebalancing over the last four years. As this makes clear, it is quite possible that there are times when the Russell 2000 benchmarked portfolio manager may face a similar stock-specific risk concentration problem to that faced by the large-cap manager. The difference is that the problem will be more transient. As a matter of analysis, this transience reduces the available data to the point where we cannot reliably test the importance of this problem, but would strongly recommend that small-cap portfolio managers carefully monitor the top end of the capitalization of their benchmark so that, when concentration begins to noticeably exceed the 2/(N+1) rule developed in Appendix A, some passive weightings are added to the portfolio to stabilize performance.

Table 9: Concentration of Russell 2000 Before and After Reconstitution

	Portion of Russell 2000 in Largest 10 Stocks (%)		
Year	May (Before Reconstitution)	June (After Reconstitution)	If No Stock Over Approx. 2/(eff N+1) Limit
1996	3.74	1.64	1.33
1997	2.50	1.68	1.33
1998	2.96	1.73	1.33
1999	5.36	1.96	1.33

Source: Russell-Mellon and Goldman Sachs Research

Portfolio Manager Patterns, Skill Sets and Value-Enhancing Risk Management

The first key difference between applying the types of risk controls we have been discussing to real rather than simulated portfolio managers is that real portfolio managers are unlikely to be as unbiased or uncorrelated in their judgements on stocks as the simulated portfolio managers we have been studying. In particular, when portfolio managers use common forecast drivers across groups of stocks (such as oil prices for oil stocks), they are likely to create more correlated sector evaluations than our statistical simulations. This in turn suggests that, for portfolio managers whose analysis is dependent on macro level inputs, some form of sector-neutral risk control that puts groups of stocks subject to the same drivers in the same sectors is likely to be more effective than our results would indicate.

However, if the portfolio manager is skilled at cross-sector comparison, enforcing sector neutrality will unnecessarily hurt returns. A potentially better answer is to mix sector-neutral and unconstrained portfolios to reflect the portfolio manager's skill at sector comparisons.

The potential for such mixed approaches to risk control can be seen in Table 10, in which we compare the performance of a sector-neutral, an unconstrained and a mixed portfolio consisting of 50% sector-neutral and 50% unconstrained portfolio. The mixed portfolio performs better from a risk-return efficiency basis than either of the pure approaches. Unfortunately, without examining a particular portfolio manager's skill set, it is impossible to determine the optimal mix as the correct mix will be highly dependent on how effective the portfolio manager's sector over- and underweights are at generating returns and on the nature of the correlations created by the particular portfolio manager's research methodologies for forecasting fundamentals.

The key is to understand how common drivers can create correlations and then analyze the degree to which cross-sector and within-sector positions can best be mixed.

A second and related issue is that (1) the portfolio manager may be differentially effective at analyzing different sectors or (2) fundamentals may be differentially effective at forecasting returns in different sectors. Table 11, which shows the returns generated by our simulated portfolio managers on a sector-by-sector basis, shows that at least the latter is, in fact, the case. Such differences in performance suggests that performance could be improved by tilting risk taking toward areas in which the portfolio manager is more effective.

The stratified risk systems discussed in this paper provide a framework for analysis of portfolio manager skill and for breaking that skill down by areas of effectiveness (as we did for sectors in the prior section). They can also correct for correlated ratings as stratified controls will create more risk efficient portfolios if portfolio manager stock ratings across categories tend to be correlated. Having used this approach to find areas of high and low effectiveness, the portfolio manager can then tilt risk taking toward areas of high effectiveness.

We note that there is likely to be a very strict limit on how strongly a portfolio manager should concentrate risk into areas in which the portfolio manager has greater skill (or conviction). That limit arises from the math used to discuss benchmark

Table 10: Mixed Composite Strategies

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill

	Index-Weight Top 50			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Unadjusted	2.1	1.41	1.49	78.3
Sector	2.0	1.34	1.51	79.2
1/2 Unadjusted and 1/2 Sector	2.1	1.20	1.73	83.9

Panel 2: High Skill

	Index-Weight Top 50			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Unadjusted	2.9	1.45	2.02	87.3
Sector	2.8	1.36	2.05	86.6
1/2 Unadjusted and 1/2 Sector	2.9	1.22	2.33	90.9

Source: Goldman Sachs Research

diversification developed in Appendix A. **Specifically, once positions in the active segment of the portfolio are more than double-weighted, the portfolio managers are almost certainly hurting the portfolio's risk-return efficiency as they have crossed the boundary from constructing an optimized portfolio to taking specific single stock risk positions (see pages 36-37 in Appendix A).**

While this result is not an absolute certainty, the evidence presented earlier in the paper on the nature of skill and the underlying uncertainty in returns, even for a portfolio manager with perfect foresight of fundamentals, strongly suggests that it would be unwise to take specific positions in individual stocks at the expense of optimizing the performance of the entire portfolio.

An intriguing possibility raised by this analysis is that even greater portfolio manager effectiveness could be achieved if we took a more flexible approach to defining sectors. In particular, we suggest looking for categories of stocks in which the portfolio manager's effectiveness is more clearly differentiated and, thus, would be more suitable for under- and overweighting, or which provides the best within-group comparability across stocks and, thus, improves the effectiveness of stock selection. Such an approach is quite feasible, but requires us to develop a new set of tools to look at categories more dynamically; these tools are developed in the next paper in the series *A Stockpicker's Reality*.

Risk Control and Long-Run Portfolio Manager Performance

The prior analysis focuses on risk-return efficiency and takes portfolio manager consistency as a primary goal. In reality, there is a strong trade-off between such efficiency and long-run returns. For large-cap portfolio managers, gains in efficiency could be had from holding a market-weighted passive position in the top 50 stocks, but this strategy averaged 80 b.p. lower returns per year for both moderately skilled value and growth managers as a result of diverting funds under management away from actively selected stocks to passive risk management positions. (Market-weighting the top 100 stocks, which takes even more assets away from active management, averaged 130 b.p. lower returns per year for moderately skilled value managers and 140 b.p. for moderately skilled growth managers.)

The potential for using short positions against the passive portfolio (which are limited in size to match the active long risk positions in dollar size) modestly reduces this overhead, but does not fundamentally change the question.

Does it make sense to lower expected long-run returns in order to create more consistent quarterly/annual performance relative to benchmark?

The answer clearly depends on investment goals and restrictions on leverage. With leverage, the portfolio

Table 11: Composite Long-Short Portfolio Returns by Sector – Moderate Skill
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Sector	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Consumer Cyclicals	7.2	5.53	1.30	74.4
Technology	11.7	10.07	1.16	72.3
Capital Goods	6.2	5.96	1.05	69.9
Financials	3.4	3.75	0.91	67.6
Consumer Staples	4.7	5.59	0.84	66.5
Basic Materials	4.9	5.98	0.82	66.2
Health Care	7.4	10.01	0.74	64.4
Utilities	1.6	3.41	0.46	59.1
Energy	2.8	8.62	0.33	56.3
Transportation	3.1	13.00	0.24	55.1
Communication Services	2.7	11.84	0.23	55.0
Cap-Weight Sectors Long-Short	5.4	1.88	2.88	91.0

Source: Goldman Sachs Research

manager would simply lever the more efficient portfolio to bring risk back up to the original level, thereby translating the gain in efficiency into increased returns rather than reduced risk. Equivalently, a plan sponsor/investor could adjust their active-to-passive management mix to a higher active percentage and bring total active management risk back up to the desired level and achieve the same translation of risk control into increased returns.

An alternative solution that does not use leverage is to split the benchmark into (1) a portfolio manager's benchmark, designed to minimize risk management overhead and maximize the potential for portfolio manager outperformance, and (2) a completion benchmark that tracks the difference between the portfolio manager's benchmark and the original capitalization-weighted benchmark. These two benchmarks could then be used in combination by investors to recreate a total portfolio that would still be expected to track the original capitalization benchmark, but individual portfolio managers would no longer have to make the trade-off between consistency of outperformance and long-run returns.

The broader point is that capitalization-weighted benchmarks are designed to track the market without reference to their impact on active portfolio managers. The unintended impact of their use in benchmarking active portfolio managers has been to distort the investment process and to create an unwanted and unintended conflict between tracking the benchmark and generating long-run returns. **Both portfolio managers and investors would be better off if the benchmarks were redesigned to promote the success of active portfolio managers while still allowing investors to construct overall portfolios that would track their desired asset allocation benchmarks.**

The mismatch between portfolio manager behavior and benchmarks is a long-standing irritant to both portfolio managers and plan sponsors. In fact, style based indices were developed as an attempt to address this conflict. The problem, however, is that these indices are capitalization weighted and still possess the same (and, in some cases, higher) levels of stock-specific risk in the high index weight stocks, and, thus, do little if anything to either reduce risk management overhead or create a better match between actual active portfolios and the benchmark. In fact, the artificial constraints such indices (and the underlying partitioning of stocks into value and growth categories) put on stock selection can act as a drag on performance if enforced too rigorously (see our January 15, 1998 paper "Making the Most of Value and Growth Investing"). This is not to say that style based diversification is not useful, as our prior results on style strongly show major benefits from diversification across styles, but simply that cap-weighted style indices do not address the portfolio manager's actual risk control problem of needing to offset high concentrations of stock-specific risk.

As long as benchmarks contain high concentrations of stock-specific risk, portfolio managers will be forced to choose between consistency of outperformance and long-run returns. While managing to a benchmark is usually viewed as a portfolio manager problem, the resulting distortion of the active management process is the investor's problem. The types of combination active/passive benchmarks described above would eliminate the conflict between consistency and long-run returns without impacting the overall portfolio benchmarks.

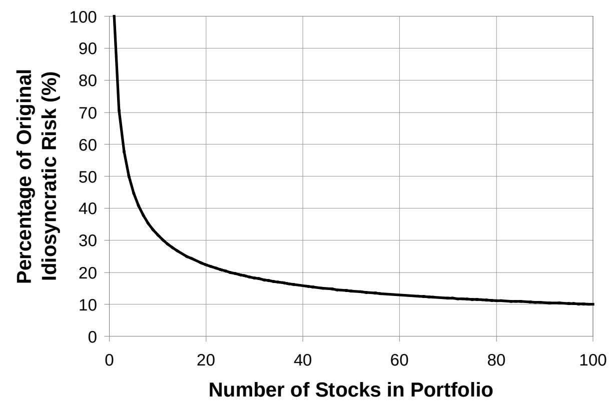
The conflict between risk control and long run returns is a result of the intense concentration of stock-specific risk at the top end of the capitalization spectrum and should not be viewed as a general statement about the use of risk control. In fact, we find noticeable evidence that sector and tailored sector type controls, used either standalone or in mixed forms, have the potential for creating more risk efficient portfolios without long term return losses. The key to making risk control an aid rather than a drag on performance is to match those risk controls against the portfolios manager's skill set.

Appendix A:
The Mathematics of Diversification

In the standard treatment of diversification in portfolio theory, as stocks are added to a portfolio, the idiosyncratic or stock-specific risks diversify away, leaving only the common or market risk. Further, in the standard treatment, the stock-specific risk diversifies away at a rate of $1/\sqrt{n}$, where n is the number of stocks in the portfolio. Thus, the textbooks show a graph of $1/\sqrt{n}$ like the one in Figure A1 and conclude that, in equal-weighted portfolios, only a relatively small number of stocks are required to get most of the benefits of diversification.

The key problem with this analysis is that most portfolios and benchmarks are not fully equal-weighted. If we add a stock with sufficient weight (or capitalization) to an equal-weighted portfolio, the portfolio actually becomes *less* diversified, with a higher concentration of stock-specific risk. A stock added at any weight over $\frac{2}{n+1}$, where n is the number of stocks in the portfolio, adds to the concentration of stock-specific risk and decreases the diversification of the portfolio. This formula is generalized below for non-equally weighted portfolios using the concept of an effective number of stocks in a portfolio, also derived below.

Figure A1: Textbook Example of Diversification



Source: Goldman Sachs Research

In terms of understanding which stocks are adding stock-specific risk rather than diversifying it away and how to handle concentrations of stock-specific risk, it is necessary to develop more sophisticated models of the relationship between index weights and diversification in index construction. The necessary mathematics are developed below.

Derivation of the Diversification Results

The risk of each stock return is divided into common factor risk and idiosyncratic (or stock-specific) risk. By definition, these two components of risk are not correlated. Each stock is assumed to have the same amount of common factor risk and the same amount of idiosyncratic risk. The stock-specific risks for all of the stocks are taken to be independently and identically distributed.

In symbols, let the return from the common risk factor be R_{common}, the return from the stock-specific risk be R_i and the total return for the individual stock be $R_{i,total}$. The variance of the common factor is σ^2_{common} and the variance of each of the stock-specific risks is $\sigma^2_{stock\_specific}$.

The total return for each stock is[24]

$$R_{i,total} = R_{common} + R_i .$$

Thus, the variance for each stock is

$$\sigma^2_{i,total} = \sigma^2_{common} + \sigma^2_{stock\_specific} .$$

If we construct a portfolio of n stocks, each with weight w_i, the resulting portfolio return is

$$R_{port\_of\_n\_stocks} = R_{common} + \sum_{i=1}^{n} w_i R_i$$

and the variance of the portfolio of N stocks is

$$\sigma^2_{port\_of\_n\_stocks} = \sigma^2_{common} + \sigma^2_{stock\_specific} \sum_{i=1}^{n} w_i^2 .$$

If we just look at the stock-specific portion of the portfolio variance, calling it θ^2, we have

$$\theta^2_{port\_of\_n\_stocks} = \sigma^2_{stock\_specific} \sum_{i=1}^{n} w_i^2$$

and the stock-specific portion of the portfolio risk (in terms of standard deviation or the square-root of the variance), θ, is

$$\theta_{port\_of\_n\_stocks} = \sigma_{stock\_specific} \sqrt{\sum_{i=1}^{n} w_i^2} .$$

[24] This equation abstracts both differentiation in betas across stocks and potential correlations across related groups of stocks. In practice, this does not appear to be a significant assumption but does radically simplify the mathematics.

Effective n

The *effective* number of stocks in a portfolio, $\tilde{n}$, can be defined as the number of equal-weighted stocks that create a portfolio with the same stock-specific risk as the portfolio we are trying to characterize.

To define the effective n, we need the stock-specific portion of the variance of a portfolio of n equal-weighted stocks. If the n stocks were equal-weighted, (that is, if $w_i = \frac{1}{n}$ for i=1 to n), the variance of the portfolio is

$$\sigma^2_{equal-wt\_port\_of\_n\_stocks} = \sigma^2_{common} + \frac{\sigma^2_{stock\_specific}}{n}.$$

Thus, the stock-specific portion of the variance is

$$\theta^2_{equal-wt\_port\_of\_n\_stocks} = \frac{\sigma^2_{stock\_specific}}{n}$$

and the stock-specific risk is

$$\theta_{equal-wt\_port\_of\_n\_stocks} = \frac{\sigma_{stock\_specific}}{\sqrt{n}}.$$

As we said above, the effective n, $\tilde{n}$, is the number of equal-weighted stocks that have the same stock-specific risk as the original portfolio. That is, we define the effective n, $\tilde{n}$, to be the $\tilde{n}$ that makes this equality true:

$$\theta_{equal-wt\_port\_of\_\tilde{n}\_stocks} = \theta_{port\_of\_n\_stocks}.$$

Substituting in our formulas for these two stock-specific risks, we get

$$\frac{\sigma_{stock\_specific}}{\sqrt{\tilde{n}}} = \sigma_{stock\_specific}\sqrt{\sum_{i=1}^{n} w_i^2}.$$

Solving for the effective number of stocks, we get

$$\tilde{n} = \frac{1}{\sum_{i=1}^{n} w_i^2}.$$

The largest effective number of stocks a portfolio can have is the actual number of stocks (that is, $\tilde{n} \le n$) and the effective number of stocks equals the actual number of stocks only when the portfolio is equal-weighted. (To see that $\tilde{n} = n$ for an equal-weight portfolio, substitute equal weights, $w_i = \frac{1}{n}$ for i=1 to n, into the equation above.)

To illustrate this point in more concrete terms, we use a two-stock portfolio as an example. If we have an equal-weight portfolio of two stocks (so each stock's weight is ½), the effective n is 2:

$$\tilde{n} = \frac{1}{\sum_{i=1}^{2} w_i^2} = \frac{1}{w_1^2 + w_2^2} = \frac{1}{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{1}{\frac{1}{4} + \frac{1}{4}} = \frac{1}{\frac{1}{2}} = 2.$$

If we do not equally weight the two stocks, the effective n of the portfolio drops below 2. For example, if the weight of the first stock is twice as much as the weight of the second stock, then w_1=2/3 and w_2=1/3, and the effective n is 1.8:

$$\tilde{n} = \frac{1}{\sum_{i=1}^{2} w_i^2} = \frac{1}{w_1^2 + w_2^2} = \frac{1}{\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = \frac{1}{\frac{4}{9} + \frac{1}{9}} = \frac{1}{\frac{5}{9}} = \frac{9}{5} = 1.8.$$

The large-cap benchmarks provide more relevant examples. As of July 30, 1999, the S&P 500 had approximately the same stock-specific risk as an equally weighted index of 107 stocks, while the Russell 1000 had approximately the same stock-specific risk as an equally weighted index of 145 stocks. In contrast, the more equally weighted Russell 2000 had approximately the same stock-specific risk as an equally weighted index of 1,627 stocks.

The 2/(n+1) Rule

When we reexamine the stock-specific risk of a portfolio in terms of the effective n, $\tilde{n}$, we get the more general formula for the rate at which adding stocks increases diversification, which is $1/\sqrt{\tilde{n}}$.

The key issue in portfolio construction is that adding stocks at weights above a certain threshold adds, rather than diversifies away, stock-specific risk. For example, if we add one more stock (with return $R_{a,total}$ and weight *a*) to our portfolio of n stocks, the resulting portfolio return is

$$R_{port\_of\_n\_stocks\_plus\_a} = R_{common} + aR_a + \sum_{i=1}^{n}(1-a)w_i R_i.$$

(The weights of the original portfolio are scaled down by (1-*a*) to adjust them for adding the additional stock.)

The portfolio variance is

$$\sigma^2_{port\_of\_n\_stocks\_plus\_a} = \sigma^2_{common} + \sigma^2_{stock\_specific}\left[a^2 + \sum_{i=1}^{n}(1-a)^2 w_i^2\right].$$

And the stock-specific portion of the risk of the new portfolio is

$$\theta_{port\_of\_n\_stocks\_plus\_a} = \sigma_{stock\_specific}\sqrt{a^2 + \sum_{i=1}^{n}(1-a)^2 w_i^2}.$$

The reduction in the stock-specific risk is the difference between the stock-specific risk of the original portfolio of n stocks and the stock-specific risk of the portfolio of n stocks plus the additional stock:

$$\theta_{reduction} = \theta_{port\_of\_n\_stocks} - \theta_{port\_of\_n\_stocks\_plus\_a} = \sigma_{stock\_specific}\left[\sqrt{\sum_{i=1}^{n} w_i^2} - \sqrt{a^2 + \sum_{i=1}^{n}(1-a)^2 w_i^2}\right].$$

In Figure A2, we graph this reduction in stock-specific risk as a function of the weight of the additional stock using index weights from the Russell 1000 (as of July 30, 1999), which has an effective N of 145. Of course, adding the additional stock at a weight of 0 (or not adding the stock at all) leaves the stock-specific risk unchanged. Then, as the weight of the additional stock is increased, the amount of stock-specific risk reduction increases, hits a maximum, declines and falls below zero, meaning that, at lower weights, adding the stock reduces stock-specific risk, but at higher weights, adding the stock adds to the stock-specific risk rather than reducing it.

To find the highest weight *a* of the additional stock that does not increase the stock-specific risk, we set the reduction in stock-specific risk to zero

$$\theta_{reduction} = 0.$$

$$\sigma_{stock\_specific}\left[\sqrt{\sum_{i=1}^{n} w_i^2} - \sqrt{a^2 + \sum_{i=1}^{n}(1-a)^2 w_i^2}\right] = 0.$$

Solving for the weight *a*, we get two solutions:

$$a = 0 \quad \text{and} \quad a = \frac{2}{\frac{1}{\sum_{i=1}^{n} w_i^2} + 1}.$$

The second solution is more interesting. If we substitute in the formula for $\tilde{n}$, it has the interpretation that **adding stocks with weights less than $a = \frac{2}{\tilde{n}+1}$, where $\tilde{n}$ is the effective number of stocks already in the portfolio (or benchmark), keeps the additional stock from adding to the stock-specific risk of the portfolio (or benchmark).** Figure A3 shows these critical weights as a function of the effective number of stocks in the portfolio.

The implication for portfolio managers is that over- and underweighting stocks in their portfolios up to the $\frac{2}{\tilde{n}+1}$ weight limit can be an expression of their investment strategy. However, over- or underweights larger than this size become bets on a single stock as the stock-specific risk of that stock increases the concentration of stock-specific risk in the portfolio.

Further, the weight for the additional stock that maximizes the reduction in stock-specific risk is half of this limit or $\frac{1}{\tilde{n}+1}$. For portfolio managers, the implication is that, to deal with the stock-specific risk in their portfolios (as opposed to their benchmarks, which are treated in Appendix F), a sensible strategy is roughly equal-weighting most stock picks, occasionally adding stocks at up to, but not beyond, double their typical position size.

Finally, if we consider adding more than one stock, we first note that it is optimal to add all of the additional stocks at the same weight. That is, each additional stock is added at the same new weight a. Then, the generalized form of the critical portfolio weight is that the largest weight a for adding m stocks without adding stock-specific risk is

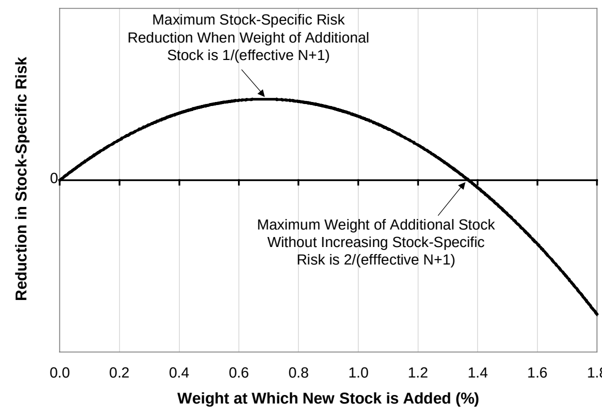
$$a = \frac{2}{\tilde{n}+m}$$

and the weight for the additional stocks that maximizes the reduction in stock-specific risk is

$$a = \frac{1}{\tilde{n}+m}.$$

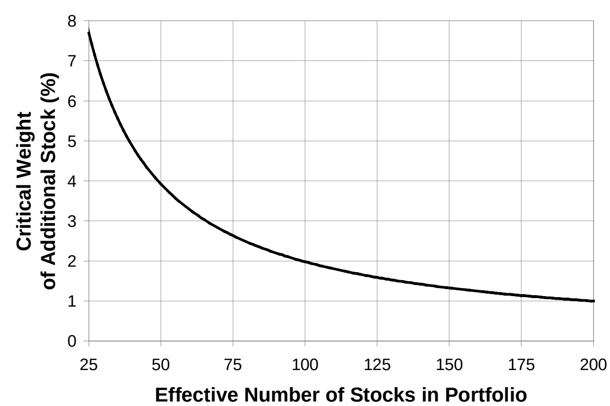
Figure A2: Reduction in Stock-Specific Risk as a Function of the Weight of One Additional Stock

(Russell 1000 Sample as of 7/30/99, Effective N = 145.)



Source: Goldman Sachs Research

Figure A3: Largest One-Stock Portfolio Weights that Do Not Increase the Concentration of Stock-Specific Risk



Source: Goldman Sachs Research

Tracking Error

A simple way of thinking about the benchmark and portfolio tracking problem can be derived from the theoretic formula for tracking error for a random active portfolio, which can be derived from the prior analysis. In particular, the tracking error (TE) equals

$$TE = \left(\sigma^2_{portfolio} + \sigma^2_{benchmark} - 2\sigma_{benchmark,portfolio}\right)^{1/2},$$

where $\sigma_{benchmark,portfolio}$ is the covariance of the benchmark return and the portfolio return.

Rewriting in terms of common market risk and stock-specific risk generates the formula

$$TE = \sigma_{stock\_specific}\left(\frac{1}{n} + \frac{1}{\tilde{N}}\right)^{1/2},$$

where n is the number of stock in the active portfolio and $\tilde{N}$ is the effective number of stocks in the benchmark. (This abstracts from any correlation between the active portfolio's stock-specific risk and the benchmark index's stock-specific risk. This correlation is quite small as long as either n or $\tilde{N}$ is noticeably smaller than N, where N is the number of stocks in the portfolio manager stock universe.)

This formula for tracking error makes clear a number of things about why increasing the number of effective stocks in the benchmark reduces tracking error and can help a portfolio manager's risk efficiency. In particular, it shows why increasing $\tilde{N}$ is more effective when $\tilde{N}$ is small (the Russell 1000, $\tilde{N}$ =145) than when it is large (the Russell 2000, $\tilde{N}$ =1,627) and also why risk management is less important if n is low and more important when n is high.

The point is that whichever of the n's (n, $\tilde{N}$) is smaller is creating the most tracking error. If they are of similar size, both will matter, but if one is much larger than the other, only the smaller will really matter. In a very real way, the effective n of a portfolio (active or benchmark) is a measure of how active the portfolio is in terms of stock-specific risk. Thus, a benchmark index with a low effective N is an active portfolio and an index like the S&P 500, with an $\tilde{N}$ of 107 is a very active portfolio, a kind of closet hedge fund. In order to have their own skill dominate the comparison of portfolio to benchmark performance, it is imperative that the portfolio manager take more stock-specific risk than the index they are measured against. The portfolio manager can either do this by taking very concentrated positions in their own active portfolio or by using passive offsets to match the concentrations of risk in the benchmark, effectively creating a new benchmark that is better diversified and, thus, easier to manage against. Using macro controls to reduce the risk the portfolio manager takes makes this problem worse not better.

Appendix B: The Data and Strategies

For this paper, we start with the Compustat universe of U.S. companies, include companies that are no longer active in order to mitigate survivorship bias, remove secondary and tertiary issues, and remove companies and data points for which the data appears to be seriously flawed. Companies without basic price and earnings data are also excluded.

The earnings and returns data covers the period from the third quarter of 1985 to the third quarter of 1998 on a calendar quarterly basis. Because some strategies need four quarters each of backward-looking and forward-looking earnings information plus a quarter to be sure the earnings would have been reported, most of our results are based on returns from the first quarter of 1987 to the first quarter of 1998.

Table B1 shows summary statistics for our base sample for this paper, which is an estimated Russell 1000 sample. We also examine estimated S&P 500 and the Russell 2000 samples. To determine which stocks are in the S&P 500 sample, we use Compustat's monthly indicator. To determine which stocks are in the Russell 1000 and 2000 samples, we use market capitalization cutoffs to first construct a Russell 3000 sample. Then, the companies in the Russell 3000 sample are divided into the top one-third by market capitalization, which becomes the Russell 1000 sample, and the bottom two-thirds by market capitalization, which becomes the Russell 2000 sample.

For the years (1992 to 1998) for which we have the actual market-cap cutoffs used in the annual Russell index reconstititutions, we use those cutoffs. For earlier years (1987 to 1991), we use our estimates of the market-cap cutoffs based on index constituent lists from Russell-Mellon and data from Compustat and FactSet. For the earliest years (1985 to 1986), we have neither the actual market-cap cutoffs nor index constituent lists, so we propagate the 1987 bottom market-cap of the Russell 2000 back in time using the average return on the Russell 2000.

Table B1: Estimated Russell 1000 Sample Summary Statistics
(1Q1987 - 1Q1997)

	Annualized Mean Return (%)
Cap-Weight Russell 1000 Sample	16.8
Equal-Weight Russell 1000 Sample	16.6
Cap-Weight Top 50	17.8
Equal-Weight Top 50	18.2
Active Benchmark if Index-Weight Top 50	16.1
Active Benchmark if Modified Index-Weight Top 50	16.6

Source: Goldman Sachs Research

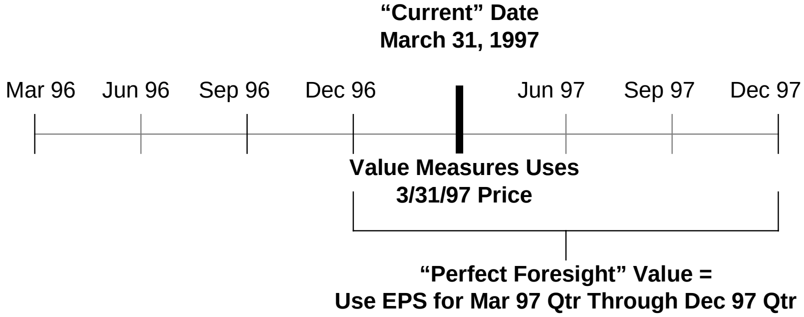
The Strategies

The style investment strategies we use for stock selection in this paper are based on forward-looking earnings, "predicted" with greater or lesser skill as described in the section of the paper on simulating skill. For value, we use a P/E ratio based on four quarters ahead cumulative (smoothed) earnings[25] (see Figure B1). For growth, we use a four-quarter earnings growth rate. Both are based on primary earnings per share excluding extraordinary items. We also construct a hybrid or growth at a reasonable price strategy that uses both the value and growth measures.

In "Style, Size and Skill," we show that the horizon of earnings insight that is most useful for growth strategies is longer for larger-cap stocks. Thus, in this paper, we use four-quarter forward earnings growth rates for the perfect foresight growth strategy for the Russell 1000 and S&P 500 samples, and two-quarter forward earnings growth for the Russell 2000 sample (see Figure B2), the horizon of growth rates between one and four quarters forward that produces the highest Sharpe ratios.

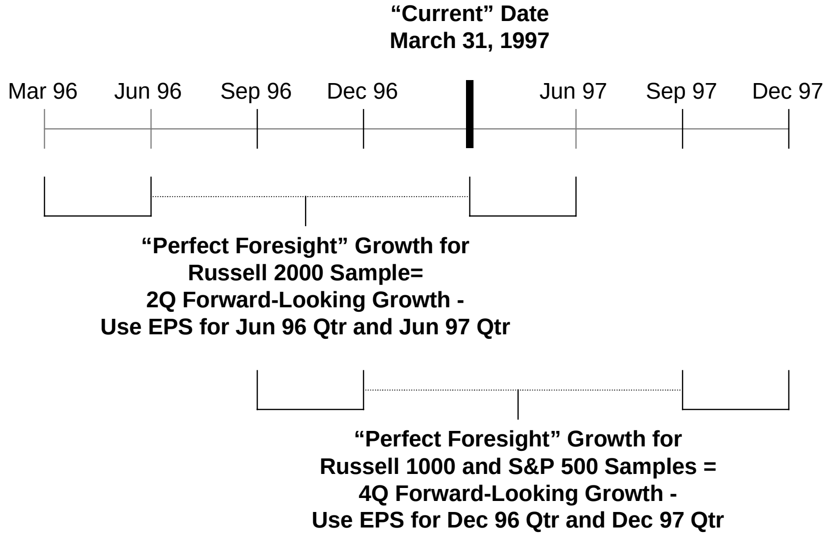
[25]To handle negative P/Es well and have a smooth transition from a small positive value to a small negative value when earnings vary from a small positive number to a small negative number, we actually use E/P.

Figure B1: Earnings Timing Conventions for Value Strategies



Source: Goldman Sachs Research

Figure B2: Earnings Timing Conventions for Growth Strategies



Source: Goldman Sachs Research

Portfolio Construction

The basic long-only portfolios are equal-weight portfolios of the top 20% of the stocks based on the ranking criterion, which is based on the value, growth or hybrid measure. For some of the graphics, we also form long-only portfolios of a specific number of stocks rather than the top 20% (i.e., we form an equal-weight portfolio of the top 5 or 50 stocks by the ranking criterion). The returns we show are in excess of a cap-weighted index of all of the stocks in the sample. Some of the long-only portfolios are modified to market weight some of the largest stocks. The portfolios labeled "Index-Weight Top 50" are long-only portfolios for which the largest 50 stocks have been index- or market-weighted, which also means that assets were removed from active management in sufficient quantity to create the market-weighting.

The long-short portfolios are formed by going long the top 20% of the stocks and short the bottom 20% of the stocks according to the ranking criterion. The long-short returns can be thought of as a measure of the ability of a portfolio manager of a given skill level using that underlying style strategy to distinguish between high- and low-returning stocks. Both the returns and Sharpe ratios from the long-short portfolios are measures of results a portfolio manager might produce without the widespread constraints of being long-only and leverage-free.

Adjusting for Size and Sector

To adjust the style strategies for characteristics like size and sector, we adapt the value and growth ranking strategies to rank within a size or sector category rather than across the whole universe of stocks. In particular, in this paper, each Compustat economic sector is a sector category and each of the ten size deciles is a size category. To adjust for both size and sector at the same time, we use size-sector categories like largest-decile energy. Compustat has 11 economic sectors – basic materials, consumer cyclicals, consumer staples, health care, energy, financials, capital goods, technology, communication services, utilities and transportation.

For the unadjusted strategies, the stocks are ranked from 1 to N (where N is the number of stocks in the target investment universe) and for the size- and sector-adjusted strategies, the stocks are ranked from 1 to $N_{category}$ (where $N_{category}$ is the number of stocks in that size, sector or size-sector category). Then equal-weight portfolios of the stocks ranked within the top 20% of the category and of the stocks ranked within the bottom 20% of the category.[26] Finally, the returns from these categories are cap-weighted (i.e., weighted by the market capitalization of that category) to form the final size-, sector or size-sector-adjusted portfolios.

[26] At least one stock from each category is chosen for each of the long and short portfolios. This condition is rarely binding.

Appendix C: Key Results for a Hybrid or Growth at a Reasonable Price Strategy

Portfolio Construction

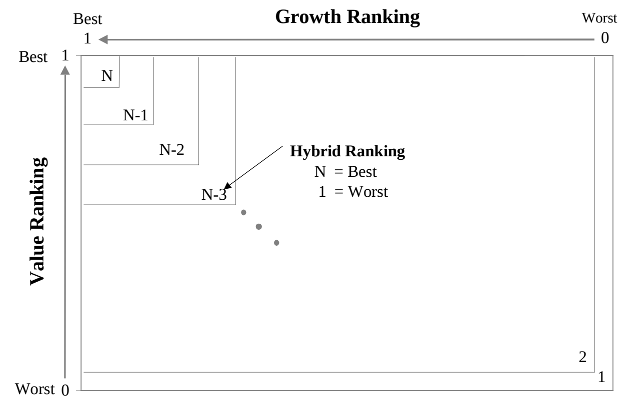
The hybrid ranking methodology combines the value and growth characteristics of the stocks into one ranking criteria for a given skill level. Using the rating methodology discussed in the paper and in Appendix B, we rate the stocks from 0 to 1 for both value and growth. We assign a value rating of 1 for the stock with the highest value (lowest P/E ratio) and a value rating of 0 for the stock with the lowest value (highest P/E ratio). Similarly, we assign a growth rating of 1 for the fastest growing stock and a growth rating of 0 for the slowest growing stock.

Using these value and growth ratings, we create a hybrid rating by assigning an individual stock the lesser of their value or growth rating. That is, the hybrid rating is the minimum of the value rating and growth rating assigned to an individual stock. The set of stocks are then ranked by this hybrid rating, with the highest-rated hybrid stock given a hybrid ranking of N (the number of stocks in the portfolio), the second highest stock N-1, down to the lowest hybrid rated stock receiving a hybrid ranking of 1. Figure C1 illustrates this methodology.

Results

The key results for our hybrid strategy are consistent with the results for our value and growth strategies. The average returns and implied long-run excess returns generated by the hybrid strategy at various skill levels are reported in Table C1. As is consistent with value and growth, slight improvements in stock selection ability lead to substantial outperformance. The hybrid returns are slightly higher than either the pure value or pure growth returns for a given skill level, which is consistent with using more information (that is, both growth and value characteristics) in forming the hybrid rankings.

Figure C1: Illustration of Hybrid Ranking Method



Source: Goldman Sachs Research

Table C1: Implied Hybrid Long-Run Excess Returns
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Skill Name	Percent In Correct Bucket (%)	Average Long-Short Returns	Implied Long-Run Excess Rtns (%)
No Skill	20	0.0	0.0
	21	2.6	1.3
	22	4.8	2.4
Moderate	23	7.0	3.5
	24	8.7	4.3
High	25	10.4	5.2
Max	26 .4	12.3	6.2
Perfect	100	35.3	17.7

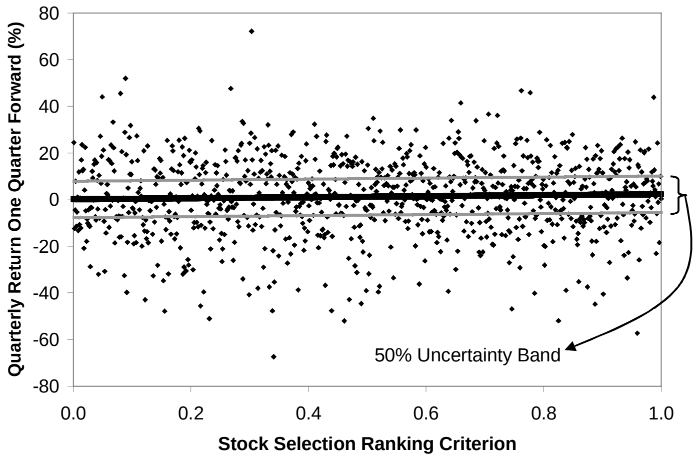
Source: Goldman Sachs Research

Although the hybrid strategy uses more information and improves the expected returns, using the hybrid valuation strategy does not overcome the basic underlying randomness of individual stock returns as can be seen in Figure C2. As it was for the value and growth strategies, and would be for any strategy that did not produce unbelievably large returns, the core observation from these graphs is that, even with more information or with higher skill levels, the randomness of individual stock returns is pervasive.

Figure C2: Hybrid – Relationship of Ranking Criterion to Returns

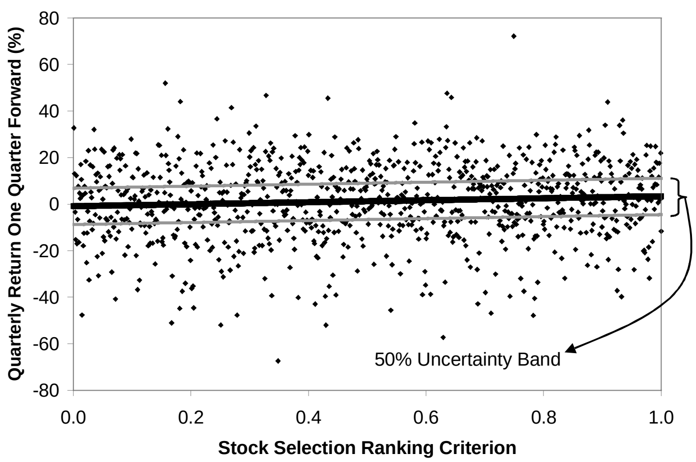
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Portfolio Manager with Moderate Skill



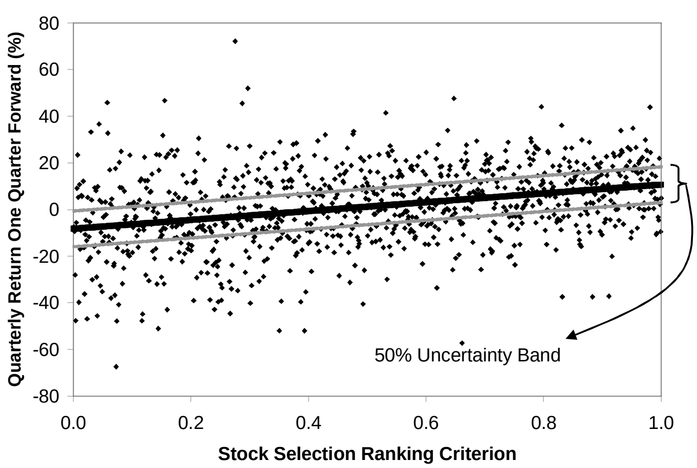
Moderate Stock Selection Skill (23% of Stocks in Correct Bucket)	
High-Mid Spread	1.1 %
High-Low Spread	2.1 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	2.1 %
Standard Deviation	14.4 %
Average R-Squared	0.30 %
Annualized Long-Short Return	7.0 %
Annualized Implied Long-Run Excess Long-Only Return	3.5 %

Panel 2: Portfolio Manager with High Skill



High Stock Selection Skill (25% of Stocks in Correct Bucket)	
High-Mid Spread	1.6 %
High-Low Spread	3.2 %
50% Uncertainty Band	+/- 7.8 %
Average Slope	3.2 %
Standard Deviation	14.4 %
Average R-Squared	0.55 %
Annualized Long-Short Return	10.4 %
Annualized Implied Long-Run Excess Long-Only Return	5.2 %

Panel 3: Portfolio Manager with Perfect Skill



Perfect Stock Selection Skill (100% of Stocks in Correct Bucket)	
High-Mid Spread	5.4 %
High-Low Spread	10.9 %
50% Uncertainty Band	+/- 7.7 %
Average Slope	10.9 %
Standard Deviation	14.0 %
Average R-Squared	5.45 %
Annualized Long-Short Return	35.3 %
Annualized Implied Long-Run Excess Long-Only Return	17.7 %

Source: Goldman Sachs Research

Managing Uncertainty

Figure C3 shows the annualized returns, the volatility and Sharpe ratios for simulated long-short hybrid portfolios as the number of stocks increase, with a line for the inverse of the square root of n in the volatility graph as a benchmark for how well diversification is working. Figure C4 reproduces Figure C3 for long-only portfolios.

As was the case for the average of value and growth shown in the body of the paper, the long-short graphs show that diversification works quite well without any risk control. Further, for hybrid, the Shape ratio for the long-only portfolios continues to increase even out to portfolios of 200 stocks, while for the average of value and growth, the loss of return from adding more stocks to the portfolio causes the Sharpe ratio to start declining after 100 stocks.

Solving the Risk Management Problem

Table C2 shows the relative effectiveness of different macro risk control approaches for the hybrid managers before controlling the stock-specific risk in the large-cap (Russell 1000) benchmark. In contrast to the value and growth strategies, for which these types of risk control provided some modest, if misleading, improvement, for hybrid, these risk control measures uniformly hurt the risk-return trade-off (that is, these risk control measures lower the Sharpe ratio).

However, as can be seen in Table C3, offsetting the stock-specific risk by index-weighting the largest stocks in the benchmark improves the Sharpe ratio for the hybrid strategy, at least up until the largest 50 stocks are passively held.

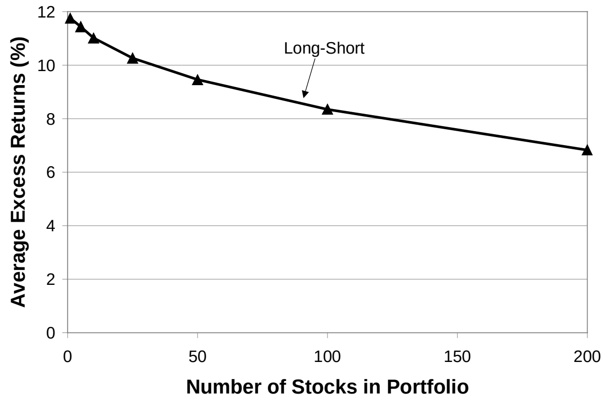
After offsetting the stock-specific benchmark risk with passive positions in the largest stocks, the impact of the risk control approaches on the hybrid portfolios (see Table C4) is consistent with the impact on the value and growth portfolios – size controls only hurt, while sector control shows sufficient promise to warrant careful investigation in light of a manager’s particular process and skill set.

If we redo the analysis of the impact of the number of stocks in the hybrid portfolio on the Sharpe ratios and tracking error after controlling for the concentration of stock-specific risk in the very large-cap stocks (see Figure C5), we see that market-weighting the largest 50 and 100 stocks removes most of the common risk factor, much as it did for the average of the value and growth strategies.

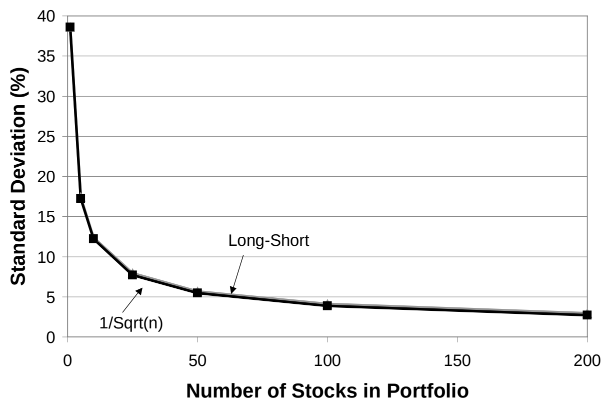
Figure C3: Impact of Increasing Number of Positions for Hybrid, Moderate Skill, Long-Short Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

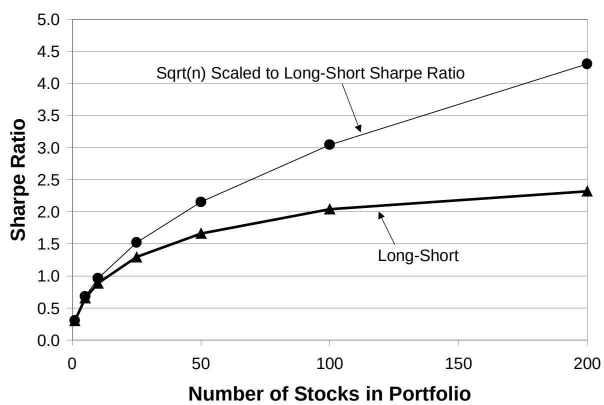
Panel 1: Annualized Returns



Panel 2: Volatility



Panel 3: Sharpe Ratios

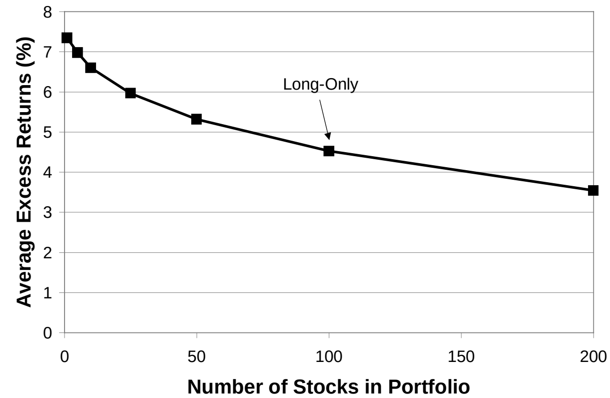


Source: Goldman Sachs Research

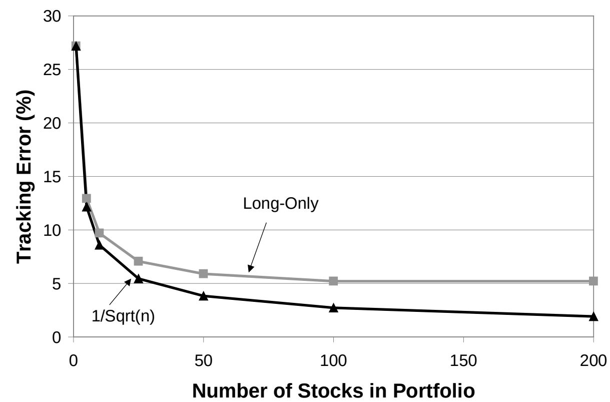
Figure C4: Impact of Increasing Number of Positions for Hybrid, Moderate Skill, Long-Only Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

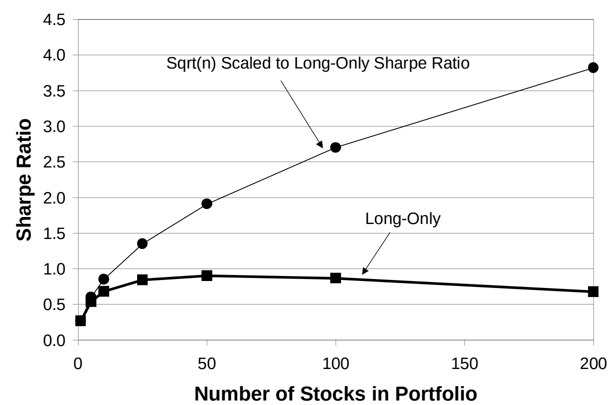
Panel 1: Annualized Returns



Panel 2: Tracking Errors



Panel 3: Sharpe Ratios



Source: Goldman Sachs Research

Table C2: Effect of Macro Risk Control Methods – Hybrid Strategy
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

	Moderate Skill (3% Edge)				High Skill (5% Edge)			
	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Unadjusted	3.6	4.8	0.8	62.3	5.7	4.9	1.1	70.8
Control for Size	2.6	2.7	1.0	68.7	4.1	2.8	1.5	77.4
Control for Sector	3.5	4.3	0.8	63.8	5.4	4.3	1.2	72.9
Control for Sector and Size	2.1	2.7	0.8	65.7	3.4	2.7	1.3	74.3
Long-Short	7.0	3.1	2.3	87.8	10.4	3.3	3.2	95.3

Source: Goldman Sachs Research

Table C3: Effect of Offsetting Stock-Specific Risk – Hybrid Strategy
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

	Moderate Skill (3% Edge)				High Skill (5% Edge)			
Number of Largest Stocks Index-Weighted	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Portfolios								
0	3.6	4.85	0.75	62.3	5.7	4.95	1.14	70.8
10	3.1	3.49	0.88	65.0	4.8	3.57	1.33	74.2
20	2.8	2.67	1.06	68.4	4.3	2.74	1.59	78.8
50	2.5	1.64	1.56	78.8	3.7	1.69	2.21	88.9
75	2.1	1.27	1.68	80.8	3.2	1.32	2.40	90.4
100	1.9	1.03	1.80	82.5	2.7	1.06	2.58	91.6
Long-Short Portfolios								
0	7.0	3.08	2.27	87.8	10.4	3.25	3.18	95.3

Source: Goldman Sachs Research

Table C4: Effect of Offsetting Stock-Specific Risk and Controlling for Macro Risk – Hybrid Strategy
(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

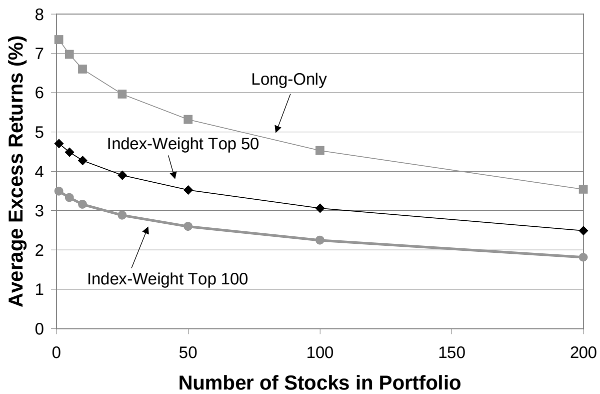
	Moderate Skill (3% Edge)				High Skill (5% Edge)			
Index-Weight Largest 50 Stocks	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)	Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Percent Positive (%)
Long-Only Unadjusted	2.5	1.6	1.6	78.8	3.7	1.7	2.2	88.9
Control for Size	1.9	1.8	1.1	71.5	2.8	1.8	1.5	79.2
Control for Sector	2.4	1.6	1.6	79.0	3.6	1.6	2.3	88.3
Control for Sector and Size	1.6	2.0	0.8	67.2	2.4	2.0	1.2	73.9

Source: Goldman Sachs Research

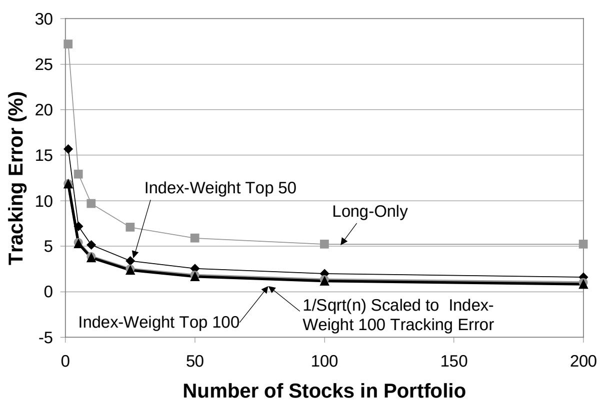
Figure C5: Impact of Increasing Number of Positions for Hybrid, Moderate Skill, Long-Only Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

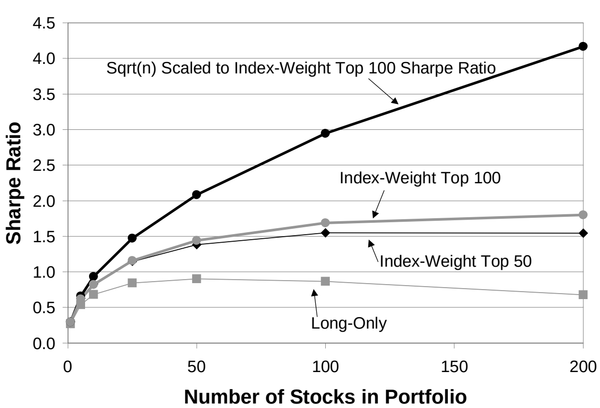
Panel 1: Annualized Returns



Panel 2: Tracking Errors



Panel 3: Sharpe Ratios



Source: Goldman Sachs Research

Appendix D: Volatility and the Square Root of N

In the main body of the paper, we use the square root of n ($\sqrt{n}$) as the limit of how fast the portfolio volatility (or tracking error) can decrease and the portfolio Sharpe ratio (that is, the average return divided by the volatility) can increase. The purpose of this appendix is to show where the importance of the square root of n comes from. The key is how tracking error decreases as the number of stocks in the portfolio increases.

If we take the same model we used in Appendix A, each stock's return can be split into the common portion (R_{common}) and the stock-specific portion (R_i). The stock-specific portions are independent and identically distributed and each has the variance $\sigma^2_{stock\_specific}$.

Thus, each stock's return is

$$R_{i,total} = R_{common} + R_i$$

and each stock's variance is

$$\sigma^2_{i,total} = \sigma^2_{common} + \sigma^2_{stock\_specific}$$

If we construct an equal-weighted portfolio of n stocks, each stock's weight in the portfolio is 1/n and the portfolio return is

$$\begin{aligned} R_{equal\_wt\_port\_of\_n\_stocks} &= \sum_{i=1}^{n}\left(\frac{R_{common}+R_i}{n}\right) \\ &= n\left(\frac{R_{common}}{n}\right)+\sum_{i=1}^{n}\frac{R_i}{n} \\ &= R_{common}+\frac{1}{n}\sum_{i=1}^{n}R_i. \end{aligned}$$

Then, the excess return is the difference between the portfolio return and the common (or market) return:

$$\begin{aligned} R_{excess} &= R_{equal\_wt\_port\_of\_n\_stocks} - R_{common} \\ &= R_{common}+\frac{1}{n}\sum_{i=1}^{n}R_i - R_{common} \\ &= \frac{1}{n}\sum_{i=1}^{n}R_i. \end{aligned}$$

The volatility of the excess return is

$$\begin{aligned} \sigma^2_{excess} &= vol\left(\frac{1}{n}\sum_{i=1}^{n}R_i\right) \\ &= \left(\frac{1}{n}\right)^2\sum_{i=1}^{n}vol(R_i) \\ &= \left(\frac{1}{n}\right)^2\sum_{i=1}^{n}\sigma^2_{stock\_specific} \\ &= \left(\frac{1}{n}\right)^2 n\sigma^2_{stock\_specific} \\ &= \frac{\sigma^2_{stock\_specific}}{n}. \end{aligned}$$

Then, the tracking error (TE) is the square-root of this volatility of the excess return:

$$\begin{aligned} TE &= \sqrt{\sigma^2_{excess}} \\ &= \sqrt{\frac{\sigma^2_{stock\_specific}}{n}} \\ &= \frac{\sigma_{stock\_specific}}{\sqrt{n}}. \end{aligned}$$

Thus, if a portfolio is constructed of independent, identically distributed and equally weighted stock picks, the tracking error declines at a rate of $1/\sqrt{n}$ as the number of stocks in the portfolio (n) is increased. If the stock-specific portions of the returns were positively correlated (as they might be if they were in the same sector), the tracking error would increase and, if the stock-specific returns were negatively correlated, the tracking error would decrease.

Appendix E: S&P 500 Index Concentration

The S&P 500 typically has a significant concentration of stock-specific risk. Figure E1 shows the cumulative market capitalization for the largest 10, 50 and 100 stocks, which shows that there is a disproportionate concentration of weight on the largest stocks in the index and this has been a persistent characteristic in the S&P 500 over the past 11 years.

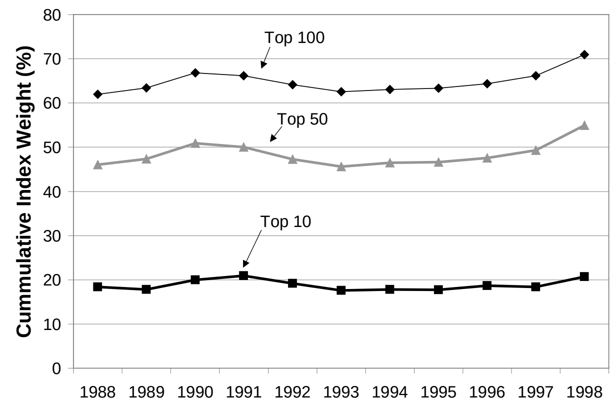
According to the Standard and Poor's Index Focus, the top 100 stocks (the largest 20% of the stocks in the index) account for close to 71% of the total index market value as of December 31, 1998. In addition, the top 100 stocks have accounted for at least 62% of the S&P 500 total market capitalization since 1988. Similarly, the top 50 stocks (the largest 10% of the stocks in the index) have had a cumulative market capitalization between 45% and 55% of the total market value since December 31, 1998.

Much of this concentration occurs in the largest 10 stocks, which are only 2% of the stocks in the index, but comprise between 17% and 21% of the total index market value. Figure E2 graphs the size distribution of the S&P 500 index as of (July 30, 1999), again showing that there is a disproportionately high concentration of market value in the largest stocks.

Figure E3 compares the concentration of the market capitalization of the S&P 500 and the 500 largest U.S. stocks. Figure E3 demonstrates that most of the high concentration of market capitalization in the largest stocks in the S&P 500 is due to the high market capitalization concentration in the actual 500 largest U.S. companies. However, relative to these actual 500 largest companies, the S&P 500 further overweights the largest stocks. These overweights are a result of the managing of the constituents of the index. In the S&P 500, some middle-sized stocks have been excluded and some stocks that are smaller than the 500 largest stocks have been included.

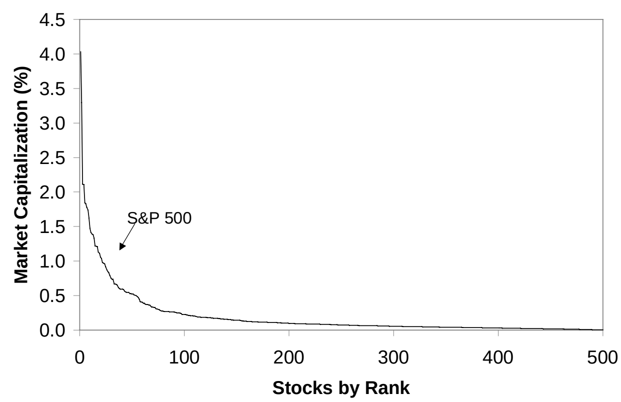
For reference, Table E1 on the following page provides a list of the largest 100 companies in the S&P 500 index as of December 31, 1998, ranked according to market value and cumulative index weight.

Figure E1: Concentration of Market Capitalization in the S&P 500 Index



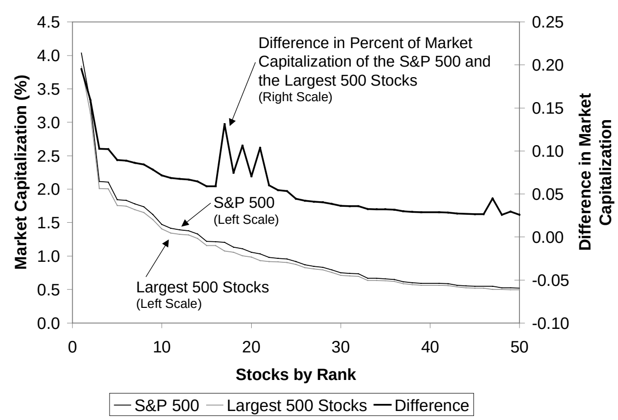
Source: Standard and Poor's Index Focus

Figure E2: Size Distribution of the S&P 500



Source: FactSet and Goldman Sachs Research

Figure E3: Size Distribution of the S&P 500 and the 500 Largest U.S. Stocks



Source: FactSet and Goldman Sachs Research

Table E1: S&P 500 Largest 100 Market Value Rankings as of December 31, 1998

Market Value Rank	Ticker	Company	Index Weight (%)	Cummulative Index Weight (%)	Market Value Rank	Ticker	Company	Index Weight (%)	Cummulative Index Weight (%)
1	MSFT	Microsoft	3.48	3.5	51	MCD	McDonald's	0.52	55.4
2	GE	General Electric	3.36	6.8	52	TYC	Tyco International	0.49	55.9
3	INTC	Intel	1.99	8.8	53	GM	General Motors	0.47	56.4
4	WMT	Wal-Mart	1.85	10.7	54	AXP	American Express	0.46	56.9
5	XON	Exxon	1.79	12.5	55	FRE	Federal Home Loan Mortgage	0.44	57.3
6	MRK	Merck	1.77	14.2	56	EMC	EMC	0.43	57.7
7	IBM	IBM	1.73	16.0	57	ORCL	Oracle	0.42	58.2
8	KO	Coca Cola	1.66	17.6	58	ATI	AirTouch Communications	0.41	58.6
9	PFE	Pfizer	1.64	19.3	59	MWD	Morgan Stanley Dean Witter	0.41	59.0
10	CSCO	Cisco Systems	1.47	20.7	60	XRX	Xerox	0.39	59.4
11	LU	Lucent Technologies	1.45	22.2	61	MOT	Motorola	0.37	59.7
12	T	AT&T	1.37	23.6	62	MDT	Medtronic	0.37	60.1
13	BMY	Bristol-Myers Squibb	1.34	24.9	63	FON	Sprint Corp. FON Group	0.36	60.5
14	WCOM	MCI WorldCom	1.32	26.2	64	TXN	Texas Instruments	0.34	60.8
15	MO	Philip Morris	1.31	27.5	65	NT	Northern Telecom	0.33	61.1
16	PG	Procter & Gamble	1.22	28.8	66	BA	Boeing	0.33	61.5
17	JNJ	Johnson & Johnson	1.13	29.9	67	GPS	The Gap	0.33	61.8
18	C	Citigroup	1.13	31.0	68	SUNW	Sun Microsystems	0.33	62.1
19	SBC	SBC Communications	1.05	32.1	69	USW	US West	0.33	62.5
20	BAC	BankAmerica	1.05	33.1	70	ALL	Allstate	0.32	62.8
21	RD	Royal Dutch Petroleum	1.03	34.1	71	BUD	Anheuser-Busch	0.32	63.1
22	AIG	AIG	1.02	35.2	72	AFS	Associates First Capital	0.31	63.4
23	LLY	Eli Lilly	0.98	36.1	73	BK	Bank of New York	0.31	63.7
24	BLS	BellSouth	0.98	37.1	74	MTC	Monsanto	0.30	64.0
25	HD	Home Depot	0.96	38.1	75	TCOMA	Tele-Communications	0.30	64.3
26	DELL	Dell Computer	0.94	39.0	76	SWY	Safeway	0.30	64.6
27	BEL	Bell Atlantic	0.83	39.9	77	KMB	Kimberly-Clark	0.30	64.9
28	SGP	Schering-Plough	0.82	40.7	78	WAG	Walgreen	0.29	65.2
29	FNM	Fannie Mae	0.77	41.4	79	PNU	Pharmacia & Upjohn	0.29	65.5
30	TWX	Time Warner	0.77	42.2	80	CCL	Carnival	0.29	65.8
31	ABT	Abbott Labs	0.75	43.0	81	MMM	3M	0.29	66.1
32	AHP	American Home Products	0.75	43.7	82	UMG	Media One Group	0.29	66.4
33	CPQ	COMPAQ Computer	0.72	44.4	83	TX	Texaco	0.28	66.6
34	F	Ford Motor	0.71	45.1	84	CL	Colgate-Palmolive	0.27	66.9
35	HWP	Hewlett-Packard	0.71	45.9	85	WMI	Waste Management	0.27	67.2
36	AIT	Ameritech	0.70	46.6	86	EMR	Emerson Electric	0.27	67.5
37	MOB	Mobil	0.68	47.2	87	AMGN	Amgen	0.27	67.7
38	WFC	Wells Fargo	0.65	47.9	88	USB	U.S. Bancorp	0.26	68.0
39	GTE	GTE	0.63	48.5	89	SLE	Sara Lee	0.26	68.2
40	WLA	Warner-Lambert	0.62	49.1	90	VIA.B	Viacom	0.26	68.5
41	DIS	Walt Disney	0.62	49.8	91	FLT	Fleet Financial Group	0.26	68.8
42	FTU	First Union	0.61	50.4	92	SLB	Schlumberger Ltd.	0.25	69.0
43	PEP	PepsiCo	0.60	51.0	93	ALD	Allied Signal	0.25	69.3
44	ONE	Bank One	0.60	51.6	94	EDS	Electronic Data Systems	0.25	69.5
45	DD	Du Pont (E.I.)	0.60	52.2	95	CPB	Campbell Soup	0.25	69.8
46	CMB	Chase Manhattan	0.58	52.7	96	UTX	United Technologies	0.25	70.0
47	AN	Amoco	0.57	53.3	97	AUD	Automatic Data Processing	0.24	70.3
48	CHV	Chevron	0.54	53.9	98	NCC	National City	0.24	70.5
49	G	Gillette	0.54	54.4	99	DH	Dayton Hudson	0.24	70.7
50	UN	Unilever N.V.	0.53	54.9	100	MER	Merrill Lynch	0.24	71.0

Source: Standard and Poor's Index Focus

Appendix F: Managing the Stock-Specific Risk in Large-Cap Benchmarks

As we saw, large-cap equity benchmarks have high concentrations of stock weight and stock-specific risk in the largest few stocks. In the main body of the paper, the risk control method we used to deal with this risk is to index weight the largest 50 or 100 stocks, which is quite effective at managing the risk, but takes a significant portion of assets away from active management to fund the passive index-weight positions.

A potentially more efficient way to manage the stock-specific risk in the benchmark would be to market weight the largest stocks only partially. In particular, we would want the solution that offsets the most stock-specific risk for the least cost of assets taken away from active management, given the constraint against short-selling under which most portfolio managers operate.[27]

The way to develop this optimal passive holding strategy is to ask how might my marginal dollar of passive assets best be held? The optimal solution is to use the next dollar of passive holdings to hold the stocks that have the highest remaining weights in the effective active management benchmark (the original benchmark minus the passive holdings).

If we use the Russell 1000 as our example benchmark (see Table F1 for the largest 10 stocks and their index weights as of July 30, 1999), the first dollar of assets held passively to mitigate stock-specific risk should be used to purchase the largest stock in the benchmark, which is Microsoft.

In fact, if we are managing a $1-billion portfolio, the next several million dollars of passive holdings should be used to buy Microsoft. If we are managing a $1-billion portfolio, the amount of Microsoft in our benchmark is $34.6 million (Microsoft's portfolio weight (3.46%) in the benchmark times our $1 billion) and the amount of the next biggest stock, General Electric, in our benchmark is $28.2 million (see Table F1).

The core of the modified index-weight strategy is to spend the first dollars of passive holdings to add passive holdings of the largest stock in the portfolio until the effective benchmark position of the largest stock equals the effective benchmark position of the second-largest stock. Then, add passive holdings of both the first and second-largest stock until their effective benchmark positions equal the effective benchmark position of the third largest stock and so on. Figure F1 illustrates the effective benchmark positions of the largest stocks in our Russell 1000 example for passive holdings of the largest one, two and five stocks.

Thus, the first $6.254 million (the difference between the benchmark positions of Microsoft and General Electric or $34.6 million - $28.2 million) in passive holdings for our $1-billion portfolio go toward passively holding Microsoft. Then, the next $20.224 million (twice the difference between the benchmark positions of General Electric and Intel or 2*($28.2 million - $18.1 million)) would be split evenly between passive holdings of General Electric and additional passive holdings of Microsoft.

The key point is that after decreasing the weight of the largest stock until it equals the weight of the second-largest stock, the next dollar of passive holdings is more effective if it is used to decrease the weights of both stocks rather than just continuing to reduce the weight of the largest stock. The goal of the modified index strategy for offsetting the stock-specific benchmark risk is to raise the effective N of the benchmark index as much as possible using as little funds as possible.

The key to understanding why the modified rule is potentially more efficient is simply observing that, under the modified rule, each dollar is applied to the stock that is currently adding the most stock-specific risk. The index weights at the top of the index are slowly being moved toward, but not reaching the $1/\left(\tilde{N}+1\right)$ optimum weight (this optimum is derived in Appendix A). In the full market weight rule, some of the passive funds actually reduce the diversification of the index as, in the effective active management benchmark, the weights of the top

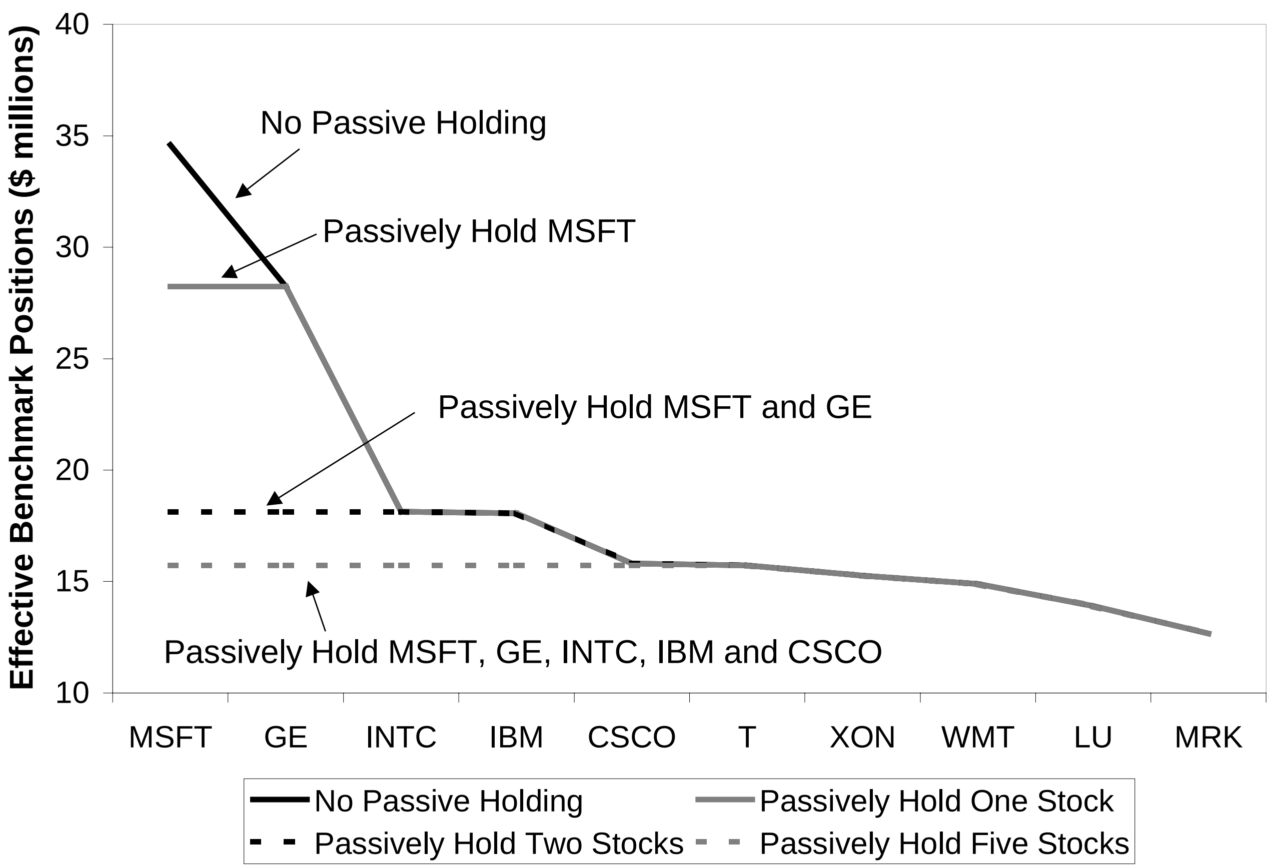
[27] If short-selling is allowed, the optimal solution for managing stock-specific risk would be to transform the benchmark into an equally weighted portfolio with long passive positions in the large-cap stocks and short passive positions in the small-cap stocks.

Table F1: Largest 10 Stocks in the Russell 1000

	Ticker	Company	Weight in Benchmark (%)	Position in a $1 billion Portfolio ($ millions)	Diff. From Next-Largest Position ($ thousands)
1	MSFT	Microsoft	3.46	34.6	6,354
2	GE	General Electric	2.82	28.2	10,112
3	INTC	Intel	1.81	18.1	67
4	IBM	IBM	1.81	18.1	2,251
5	CSCO	Cisco	1.58	15.8	95
6	T	AT&T	1.57	15.7	455
7	XON	Exxon	1.53	15.3	373
8	WMT	Wal-Mart	1.49	14.9	988
9	LU	Lucent Technologies	1.39	13.9	1,242
10	MRK	Merck	1.27	12.7	530

Source: FactSet and Goldman Sachs Research

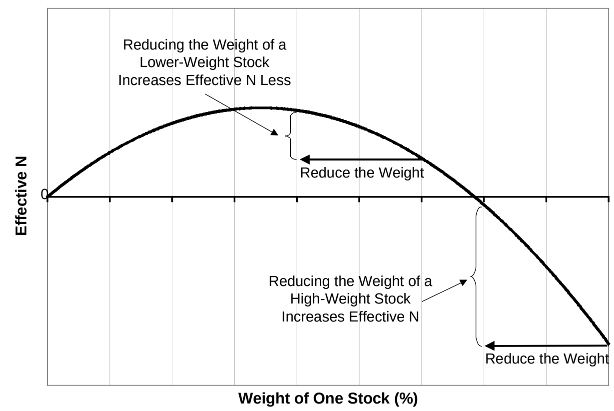
Figure F1: Effective Benchmark Positions with 0, 1, 2 and 5 Stocks Held Passively



Source: FactSet and Goldman Sachs Research

stocks are actually pushed down below the $1/(\tilde{N}+1)$ maximum diversification point to 0. Figure F2 shows, more generally, how varying index weights impacts the effective N of the benchmark index. Reducing the weight of a high-weight stock can substantially increase the effective N of the benchmark. Reducing the weight of a lower-weight stock increases the effective N of the benchmark less.

Figure F2: Reducing the Weight of High Weight Stocks Has Greater Impact on $\tilde{N}$



Source: FactSet and Goldman Sachs Research

Passive Allocation Impact on Returns

The comparison of modified index-weighted portfolios and full index-weighted portfolios raises two issues that we ignored in the earlier discussion.

The passive portfolio, while eliminating stock-specific risk, also contains a subtle macro strategy risk position that will have modest, but noticeable impact on realized returns. In particular, the passive portfolio is unlikely to produce the same returns as the overall benchmark over any specific interval of time. This difference will create some short-run distortions in returns that will either artificially add to or reduce the apparent returns from active management.

Secondly, if there is a true macro risk factor that differentiates the top 50 to 100 stocks from the rest of the stock universe, such as international exposure or liquidity, the full index-weight passive position will eliminate that macro factor at the same time that it eliminates the stock specific risk. In contrast, the modified index-weights will likely not completely eliminate such macro factors and, thus, may fail to fully eliminate the common risk factors so central to making diversification based risk control work.

Although these differences were clearly second order in the comparisons of the different risk control strategies in the main body of the paper, in the current context of comparing two very similar risk control approaches that deal with all the first order risk issues in equally effective ways, these secondary differences become much more important.

The reality is that the return differences are small and the evidence for the common risk factor is mixed. The modified index weights do not work as well as the full index weights on a number of stocks basis, but are roughly equivalent on a dollar spent basis (see Table F2).

Table F2: Effect of Index-Weighting and Modified Index-Weighting the Largest Stocks

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate (23%) Skill

			Value						Growth				
	Number of Stocks Passively Weighted	Percent Passively Weighted (%)	Mean Returns (%)	Adjusted Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Adjusted Returns Sharpe Ratio	Mean Returns (%)	Adjusted Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Adjusted Returns Sharpe Ratio
Index-Weight	50	41.5	2.0	1.8	1.56	1.30	1.15	2.2	2.0	1.68	1.30	1.17
Modified Index-Weight	50	34.2	2.1	----	2.00	1.06	------	2.3	----	2.13	1.08	------
Modified Index-Weight	75	41.6	1.9	----	1.61	1.17	------	2.0	----	1.75	1.16	------
Index-Weight	100	55.7	1.5	1.4	0.96	1.52	1.42	1.6	1.5	1.09	1.45	1.36
Modified Index-Weight	100	47.1	1.7	----	1.35	1.25	------	1.8	----	1.49	1.24	------
Modified Index-Weight	152	55.7	1.4	----	1.01	1.41	------	1.5	----	1.11	1.38	------

Panel 2: High (25%) Skill

			Value						Growth				
	Number of Stocks Passively Weighted	Percent Passively Weighted (%)	Mean Returns (%)	Adjusted Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Adjusted Returns Sharpe Ratio	Mean Returns (%)	Adjusted Mean Returns (%)	Tracking Error (%)	Sharpe Ratio	Adjusted Returns Sharpe Ratio
Index-Weight	50	41.5	2.8	2.6	1.61	1.74	1.60	3.0	2.8	1.77	1.72	1.60
Modified Index-Weight	50	34.2	3.0	----	2.06	1.46	------	3.3	----	2.22	1.47	------
Modified Index-Weight	75	41.6	2.6	----	1.66	1.59	------	2.9	----	1.84	1.58	------
Index-Weight	100	55.7	2.0	1.9	1.00	2.05	1.95	2.2	2.1	1.16	1.93	1.84
Modified Index-Weight	100	47.1	2.4	----	1.39	1.71	------	2.6	----	1.57	1.67	------
Modified Index-Weight	152	55.7	2.0	----	1.05	1.90	------	2.2	----	1.18	1.86	------

For simplicity, for the modified index-weight strategies, the top stocks are weighted at the optimally diversifying weight of 1/(effective N + number of stocks held passively), where the effective N is based on the remaining smaller stocks in the benchmark and is based on the estimated index weights as of 7/30/99. This means that the effective N used is 400 for modified index-weighting the top 50 stocks, 471 for the top 75 stocks, 512 for the top 100 stocks and 566 for the top 152 stocks. The adjusted mean returns and the adjusted returns Sharpe ratios are calculated by reducing the returns of the index-weight strategy by the difference between the passive returns of the modified index-weight and index-weight strategies. These adjusted returns are one way to control for the difference in returns due to the different implied strategy positions in the two risk control approaches.

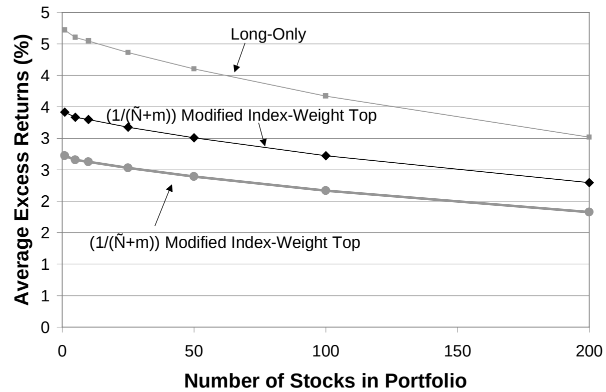
Source: Goldman Sachs Research

There appears to be some evidence that there is a macro factor that is specific to the top 50 stocks. The tracking error graphs (see Figure F3) do not fully converge, but the remaining common factor is of relatively small size and appears to be fully isolated within the top capitalization stocks (quite different from the normal size factor). These results raise the possibility that, if we fully understood this top cap stock phenomenon, it might be possible to create a risk system that would be modestly more efficient than the full index weights. In general, we argue that the real efficiency gains occur from shorts against the passive portfolio (limited by the 2/(n+1) cap rule) or from splitting the original benchmark into an active management index and a passive completion index.

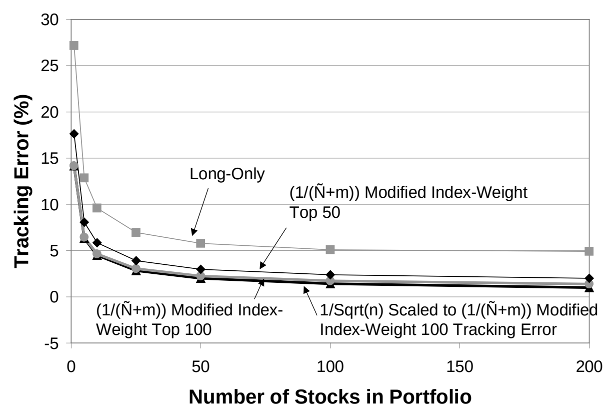
Figure F3: Impact of Increasing Number of Positions for the Average of Value and Growth, Moderate Skill, Long-Only Portfolios

(Estimated Russell 1000 Sample, 1Q1987-1Q1998, 10,000 Simulations)

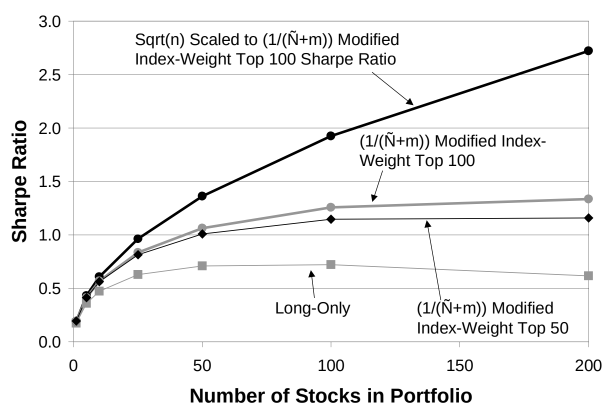
Panel 1: Annualized Returns



Panel 2: Tracking Errors



Panel 3: Sharpe Ratios



Source: Goldman Sachs Research

Appendix G: How Controlling for Size Decreases Risk-Return Efficiency

In the main body of the paper, we suggest that the size control strategy suffers from overly concentrating active management risk into a small number of stocks in the top deciles, losing efficiency (that is, increasing tracking error) as the cap-weighting of the size segments reduces the effective number of names in the portfolio.

The size control strategy is based on equally weighting 20% of the stocks in each size decile of stocks and then weighting the portfolio from each decile by the market-cap of that decile. For example, in the Russell 1000, we would first divide the 1000 stocks into 10 deciles of 100 stocks each. Then, we would use our fundamental stock-selection strategy to pick 20 stocks in each decile. The 20 stocks in each decile would be equally weighted and the resulting 10 portfolios would be weighted by the capitalization weights of the deciles. If we use the decile weights from the Russell 1000 as of July 30, 1999 (see Table G1), over 60% of the weight goes on the largest decile and the resulting effective number of stocks , $\tilde{n}$, is 49.

If, instead, we had not controlled for size, we would have used our fundamental stock-selection strategy to pick 200 stocks and equal-weighted them, resulting in an effective number of stocks of 200. Since the concentration of stock-specific risk decreases with $1/\sqrt{\tilde{n}}$, the concentration of stock-specific risk is twice ($\frac{\sqrt{200}}{\sqrt{49}} = 2.0$) as high in the size-controlled portfolio than in the un-risk-controlled portfolio. Thus, the tracking error of the size-controlled strategy higher than it is for the un-risk-controlled portfolio.

The size-controlled strategies also suffer from lower expected returns from two sources. First, the size controlled strategies lose a small amount of return from picking stocks within the size deciles rather than in an unconstrained manner. This small loss is the difference between the long-only unadjusted and the control for size, equal-weight across deciles results in Table G2, which show that a moderately skilled portfolio manager loses approximately 10 basis points of expected outperformance for choosing stocks within deciles.

The second source of return loss, which is more dramatic, comes from cap-weighting across the deciles. As Table G2 shows, a moderately skilled portfolio manager loses more like 80 basis points in a value strategy and 60 basis points in a growth

Table G1: Estimated Russell 1000 Decile Weights

(as of July 30, 1999)

Decile	Weight in Benchmark (%)	Composite Strategy Mean Long-Short Excess Returns (%)
Largest	61.5	3.6
2	13.3	4.7
3	7.4	3.7
4	4.9	5.6
5	3.6	5.7
6	2.7	5.9
7	2.2	6.3
8	1.8	6.6
9	1.5	7.7
Smallest	1.2	7.4

Source: FactSet and Goldman Sachs Research

Table G2: Returns from Size-Controlled and Un-Risk-Controlled Strategies

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)

Panel 1: Moderate Skill

	Mean Returns (%)	
	Value	Growth
Long-Only Unadjusted	2.8	3.0
Control for Size, Equal-Weight Across Deciles	2.7	2.9
Control for Size, Cap-Weight Across Deciles	1.9	2.3

Panel 2: High Skill

	Mean Returns (%)	
	Value	Growth
Long-Only Unadjusted	4.1	4.5
Control for Size, Equal-Weight Across Deciles	3.9	4.4
Control for Size, Cap-Weight Across Deciles	2.8	3.4

Source: Goldman Sachs Research

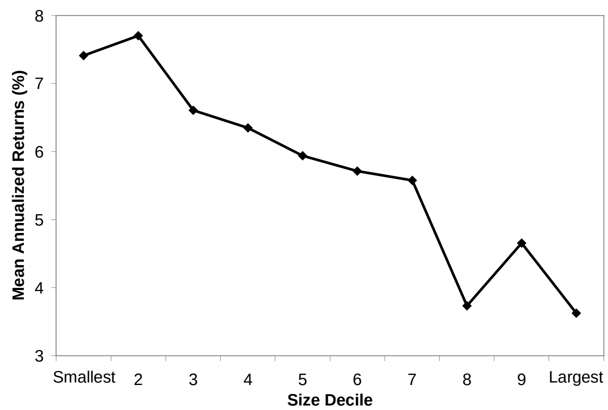
strategy (the differences in returns from cap-weighting and equal-weighting across the deciles).

The reason the cap-weighting across the deciles has such a big impact on expected returns is that cap-weighting puts most of the portfolio weight on the largest stocks – 61.5% on the largest decile, 13.3% on the second-largest – and stockpicking is less effective in the largest deciles. This result, that stockpicking is more effective in smaller-cap stocks, can be seen in the average returns by decile in Table G1 and in a graph of these average returns in Figure G1. By most heavily weighting the deciles in which stockpicking works least well, controlling for size by picking within deciles and cap-weighting across deciles reduces expected return substantially.

Thus, the size-controlled strategy is losing risk-return efficiency on both the risk and return fronts – risk (tracking error) is higher and expected return is lower.

Figure G1: Composite Strategy Mean Long-Short Excess Returns by Decile

(Estimated Russell 1000 Sample, 1Q1987-1Q1998)



Source: Goldman Sachs Research

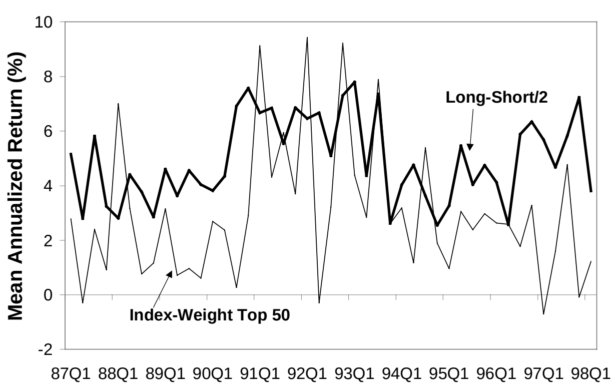
Appendix H: Rescaled Time Series Graphs

In the main body of the paper, we scale these graphs to facilitate comparison with the graphs of the long-only and long-short/2 time series. In Figures H1 through H4, we rescale them so more of the detail of these time series can be seen.

Figure H1: Moderate Skill, Long-Short and Index-Weight Top 50 Long-Only

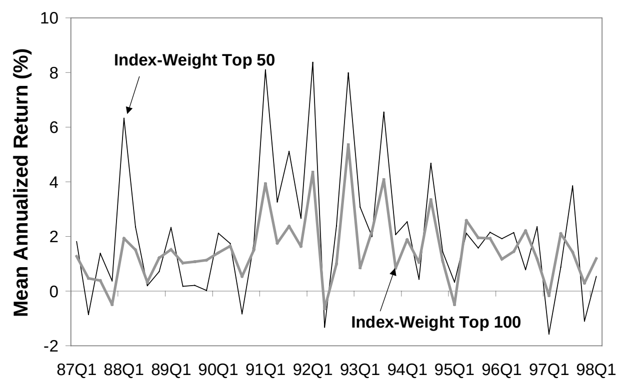
Source: Goldman Sachs Research

Figure H2: High Skill, Long-Short and Index-Weight Top 50 Long-Only



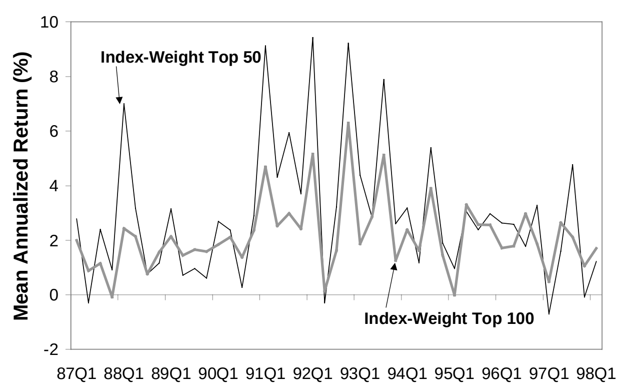
Source: Goldman Sachs Research

Figure H3: Moderate Skill, Index-Weight Top 50 and 100 Long-Only



Source: Goldman Sachs Research

Figure H4: High Skill, Index-Weight Top 50 and 100 Long-Only



Source: Goldman Sachs Research

Appendix I: Data and Methodology for Lipper Analysis

To examine real world portfolio manager returns, we use the Lipper database of monthly returns of 3897 funds from February 28, 1963 through February 28, 1999. We restrict the data set to months for which there are at least 10 funds reporting returns. We then convert the monthly returns into quarterly returns and limit the data set to the first quarter of 1987 through the first quarter of 1998 so that we can compare the Lipper returns to our value and growth strategy returns. The resulting database has 3814 funds across all fund objectives.

To classify the Lipper funds into growth or value funds, we follow a classification methodology akin to William Sharpe's method for clustering funds by their performance.[28] We eliminate all funds with less than 12 consecutive quarters of returns, leaving us with 1976 funds. Then, we regress the returns for each fund on three common factors: cash drag, S&P 500 returns and the stock-specific risk factor (measured by the difference between the equal-weighted and capitalization-weighted performance of the S&P 500).

We regress the part of the fund returns not explained by these common factors on the difference between our value and growth strategy returns. If we find that the residuals from the earlier regression are positively related to the difference between our value and growth returns, we classify the fund as value. Similarly, if the portion of the fund returns not explained by the common factors is negatively related to the difference between our value and growth returns, the fund is classified as growth. All other funds are grouped into a separate category that is neither value nor growth. We use a t-stat cutoff of ±1 to determine the significance of the relationships (significance levels tend to be low given the limited number of quarters in many of the fund return series).

668 or 33.8% of the funds are classified as growth, 377 or 19.1% are classified as value and 931 or 47.1% of the funds are classified as neither value nor growth.

[28] William Sharpe, "Asset Allocation: Management Style and Performance Measurement," *Journal of Portfolio Management*, Winter 1994.

Goldman Sachs Global Research Centers

New York
Goldman, Sachs & Co.
One New York Plaza
New York, New York 10004
United States
Tel: (1) 212-902-1000
Fax: (1) 212-902-2145

London
Goldman Sachs International
Peterborough Court
133 Fleet Street
London, EC4A 2BB, England
Tel: (44) 171-774-1000
Fax: (44) 171-774-1146

Hong Kong
Goldman Sachs (Asia) L.L.C.
Asia Pacific Finance Tower
Citibank Plaza, 3 Garden Road
Central, Hong Kong
Tel: (852) 2978-1000
Fax: (852) 2978-0479

Tokyo
Goldman Sachs (Japan) Ltd.
ARK Mori Building
12-32, Akasaka 1-chome
Minato-ku, Tokyo 107, Japan
Tel: (81) 3-3589-7000
Fax: (81) 3-3587-9263

Korea
Goldman, Sachs & Co.
Dong Ah Life Insurance Building
33 Da-Dong, Chung-Ku
Seoul, South Korea
Tel: (822) 3788-1000
Fax: (822) 3788-1001

Frankfurt
Goldman, Sachs & Co. oHG
MesseTurm
60308 Frankfurt am Main
Germany
Tel: (49) 69-7532-1000
Fax: (49) 69-7532-2800

Paris
Goldman Sachs Paris Inc. et Cie
2, Rue de Thann
75017 Paris
France
Tel: (33) 1-4212-1000
Fax: (33) 1-4212-1499

Singapore
Goldman Sachs (Singapore) Pte.
50 Raffles Place
#29-01 Singapore Land Tower
Singapore 048623
Tel: (65) 228-8100
Fax: (65) 228-8474

Stock Ratings

RL =	Recommended List	LL =	Latin America Recommended List	TB =	Trading Buy
MO =	Market Outperformer	MP =	Market Performer	MU =	Market Underperformer