Goldman Sachs

A Stockpicker's Reality: Part III
Global Portfolio Analysis

Sector strategies for maximizing returns to stockpicking

January 22, 2002

Part III of a series examining portfolio management from the portfolio manager's perspective. Part I was *Style, Size, and Skill*. Part II was *Beating Benchmarks*.

Optimal risk management differs for growth and value managers

The third paper of the series *A Stockpicker's Reality* examines the degree to which risk management can be used to increase returns to stockpicking. For value investing styles, the results indicate that risk controls based on choosing stocks within groups of comparable stocks (sector controls) can substantially add to returns. For growth styles, group-against-group risk positions play such an important role in generating returns that sector risk controls can hurt performance substantially.

Strategies that focus risk taking on sectors in which a portfolio manager's style is most effective (technology and healthcare for growth and technology, consumer cyclicals and transportation for value) can increase returns, although the overall portfolio efficiency is enhanced if all sectors are actively managed.

Steven Strongin
steve.strongin@gs.com
New York: 1-212-357-4706

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

Lewis Segal
lewis.segal@gs.com
New York: 1-212-357-4322

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

Goldman Sachs
Global Equity Research

Table of contents

Overview

This paper (third in the series, *A Stockpicker's Reality*) examines the degree to which risk management strategies can be used to increase returns. This question differs substantially from the normal application of risk control, which is focused on tracking error rather than returns. There is, of course, a natural tension between reducing tracking error and increasing returns. However, in the current context, our goal is to understand where that tension is greatest (which risk controls hurt returns the most) and where the tension is least or even reversed such that risk controls can actually help portfolio managers increase returns by focusing on taking the risks with the highest expected returns.[1]

To that end, this paper, using the simulation techniques developed in *Beating Benchmarks* (November 1999), looks at risk control in the context of equal-weighted benchmarks and equal-weighted portfolios (i.e., all we want to know is if we can use risk management techniques to pick a better group of stocks; the question of portfolio construction to beat a particular benchmark was dealt with in *Beating Benchmarks*).

Not surprisingly, our results indicate that different risk management approaches work for different styles of investing. In particular, for value investing styles, the results indicate that risk controls based on choosing stocks within groups of comparable stocks (sector controls) can substantially add to returns.[2] In contrast, for growth-based styles, under most conditions, group-against-group risk positions play such an important role in generating returns that forcing managers toward sector neutral weighting significantly hurts performance.

These results suggest that much of the fundamental insights of growth managers (whether derived top down or bottoms up) have to do with predicting the common movement of one group of stocks versus another. In contrast, the insights of value managers appear to be more company against company in nature and are more accurate the more similar the companies.

Size controls, in contrast, appear to have little impact on returns on equal-weighted portfolios regardless of style. This might seem surprising giving the emphasis on size as a risk problem in recent years, but as shown in *Beating Benchmarks*, the core of the recent "size" risk management problem was the concentration of stock-specific risk in the top 50 names in large-cap US benchmarks rather than a macro size factor. In particular, *Beating Benchmarks* showed that far better results could be attained by

[1] The ability of risk control systems to increase returns might appear contrary to the normal use of risk systems and optimizers. Appendix A explains the apparent contradiction.

[2] Most of the benefits to "sector controls" are derived from breaking the stock universe into three to five comparable groups (sectors), although we find little damage from having more sectors.

carefully limiting deviations from the benchmark in the top 50 stocks[3] than by controlling size as a risk factor.

The simulations also indicate that strategies that focus portfolio manager risk taking on sectors in which the portfolio manager's style is most effective (technology and healthcare for growth and technology, consumer cyclicals and transportation for value) can increase returns, although overall portfolio efficiency is higher if all sectors are actively managed.

These results also suggest the following:

- **Value-driven methods** are most compatible with quantitative risk management of benchmark-driven portfolios.
- **Growth-driven methods** are far less compatible with strict quantitative risk limits and are more effective in relatively more concentrated, less risk-controlled portfolio construction applications where risk management is handled at the asset allocation level by diversifying across managers.

[3] This number varies with the concentration of size within an equity market and thus varies by benchmark and market.

Stock drivers: Commonality and comparability

The core pattern of behavior in stock performance that we are trying to understand in the context of this paper is the degree to which fundamentals can predict performance and the degree to which that prediction can be more profitably utilized within groups of stocks versus between groups of stocks.

For instance, we can think of the value manager assessing two similar companies and determining that company A is inexpensive relative to B based on some forward-looking valuation criterion—a typical **within-group** comparison. In contrast, we can think of a growth manager saying that recent events will speed group A's earnings growth and slow group B's and, thus, investors should overweight group A relative to group B—a typical comparison **across groups**.

In truth, both managers are doing both types of comparisons (both implicitly and explicitly), but they will not necessarily be equally effective at them. Two concepts define the basic trade-offs: comparability and commonality.

- **Comparability** is the notion that company A's stock will outperform company B's stock if some fundamentally based forward-looking criterion indicates that company A is better than company B. In terms of comparability, we can think of risk management as constructing groupings within which comparability is improved and, thus, fundamental analysis is more effective. Further, we need to look at the question, "is comparability greater in some groups than others," indicating a reason to focus stockpicking on groups in which fundamentals drive relative returns.

- **Commonality** is the notion that groups of stocks will move relative to each other based on fundamentals but that comparisons within the groups may be more difficult. For commonality, we look for groupings in which company fundamentals help drive the groups relative to other groups, but we are less concerned about the relative performance of the stocks within the groups. Commonality can be thought of as the macro drivers, while comparability is the stockpicker's arena.

More broadly, in terms of real world portfolio manager behavior, we want to understand the conflicts and synergies between these views of the world and how they relate to particular styles of stockpicking so that we can tune risk management approaches to maximize the value of portfolio manager insight.

In particular, commonality sits at the heart of risk control. If a group of stocks move according to some common driver away from stocks not in that group, then those common movements either represent the core of the stockpicker's insight or the macro winds that buffet stocks away from the company-specific fundamentals at which the stockpicker is looking, destroying comparability.

If there is no commonality, then there is little reason for any risk control, as stock picks would be naturally independent. With commonality, the question becomes whether the commonality is the insight or the problem. To that end, we start with a quantification of commonality and will then proceed to investigate the extent to which it is a problem and to what extent it is insight by looking at how sector controls remove returns from cross-sector risk positions while improving returns from better comparability.

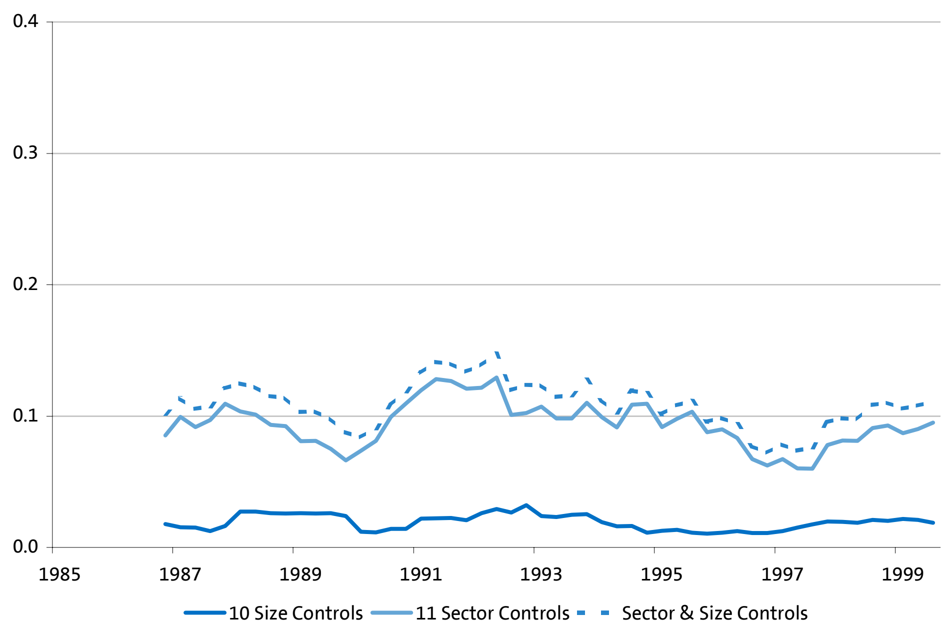
Finally, we will we look carefully at how dependent these results are on a particular set of group definitions by re-categorizing stocks into customized sectors that are optimized (using backward-looking techniques) for specific investment styles. These optimized categories will allow us to cross check that the insights we have gained along the way are not artifacts arising from the particular sector and size definitions used in the initial analysis, and also to understand the magnitude of the potential gain from fully optimized forward-looking risk systems, thus providing a benchmark against which to assess specific risk systems.

Commonality

Exhibit 1 shows the ability of size and sector macro factors to explain individual stock returns quarter by quarter. In particular, the exhibit shows the R-Squareds of quarterly regressions of stock returns on 11 Compustat sector dummies and 10 size dummies. Our data sample is described in detail in Appendix B.

The graph shows that, at this level of disaggregation, sector can explain about 9% of individual stock returns, while size can only explain about 2%. In terms of commonality, about 9% of stock movement is due to between-sector movement (commonality), while 91% of the movement is within sector or stock specific. The explanatory power varies over time but is fairly constant once the results are averaged over two-year intervals, as is done in the exhibit.

Exhibit 1: Ability of size and sector to explain returns
R-Squared, eight-quarter moving average

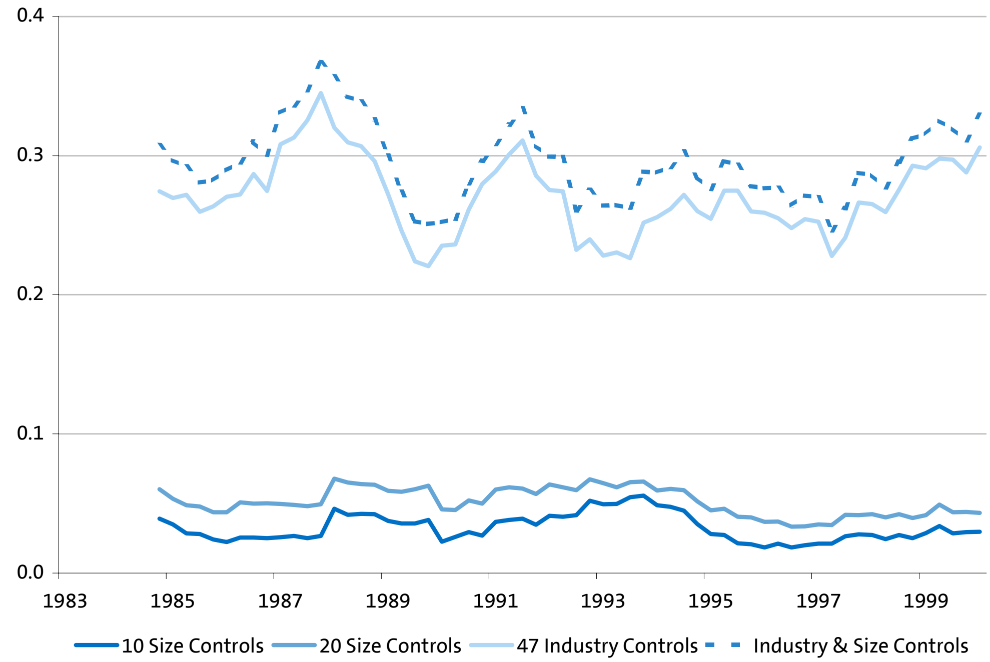


Source: Goldman Sachs Research.

If we increase the level of disaggregation to include more sectors (47 Goldman Sachs industry groupings and 20 size categories), as in Exhibit 2, we can increase sector explanatory power to about 27% but the explanatory power of size remains trivial at 5%.[4]

[4] This sample, which is restricted to stocks assigned to Goldman Sachs industries, differs from the broad sample in Exhibit 1. If we ran the 11 sector regressions on this more restricted sample, sector would explain about 16% of the variation. Thus, increasing the number of industries to 47 only explains about 11% more of the variation in returns.

Exhibit 2: Ability of finer industry definitions to explain returns
R-Squared, eight-quarter moving average



Source: Goldman Sachs Research.

A number of conclusions can be drawn from these results. First, common sector movements play a very important role in driving returns at the individual stock level, and removing the stockpicker's ability to exploit that volatility could significantly reduce returns. In contrast, size drives virtually no commonality and thus can almost certainly be safely ignored from the standpoint of stockpicking.

Although this result may seem highly inconsistent with recent history and our previous paper on beating benchmarks, it is, in fact, quite consistent. In *Beating Benchmarks*, we showed that the "size" risk that drove portfolio manager relative to benchmark performance was, in fact, not "common" size risk but stock-specific risk in the largest stocks in the benchmark. In essence, it was not size as a macro factor that was the problem; rather, it was the extreme weights placed on the largest stocks that drove benchmark performance. Hence, we recommended nearly passive positions in all 50 of the largest stocks rather than simply controlling size as a macro factor.

As a consequence, in the current context, sector controls matter not size. (Size-controlled results are included in Appendix C for completeness.) In particular, we show that investment methods that predict relative sector movements will be severely hampered by sector controls, while methods that do well at comparing similar companies but do not do well at predicting relative sector movements will be greatly helped by tight sector disciplines. The key is to understand how various investment styles differ in their relative ability to compare across sectors and their ability to compare stocks within sectors.

Comparability and style

The overall question of sector and comparability can be split into three types:

- In which sectors do fundamentals drive results? (Do particular styles have natural advantages in different sectors, and should we invest differently in different sectors?)
- Do sector controls help stockpickers? That is, can we increase returns by improving the efficiency with which fundamentals predict returns by forming comparable groups of stocks and then focusing on within-group stockpicking (disciplined sector controls)? Or, are the key fundamental insights strongly related to common sector performance (making sectors controls counter-productive)?
- Can the stock universe be segmented into meta-groups where stock selection works well and where it does not, and can results be improved by excluding those stocks from consideration for active portfolios (focus strategies)?

Comparability by sector

Exhibit 3 shows the excess return produced by a broad range of investment styles for each of the 11 Compustat sectors.[5] The five investment styles we show here are growth, value, shorter-horizon growth, standardized unanticipated earnings, and momentum (change in consensus earnings).

A style-free pure return result is also included as a reference point to show the difference in dispersion of returns available for a stockpicker to exploit. In particular, the pure return strategy shows the relative dispersion of returns within each sector and thus the potential returns from stockpicking in that sector.

The relative returns of fundamentals-based styles to the pure returns baseline indicates the degree to which fundamentals are driving returns in that sector. Actual realized returns are a function of both the dispersion of returns and the correlation of those returns to fundamentals. (See Appendix E for further explanation of the mathematics and statistics underlying these statements.)

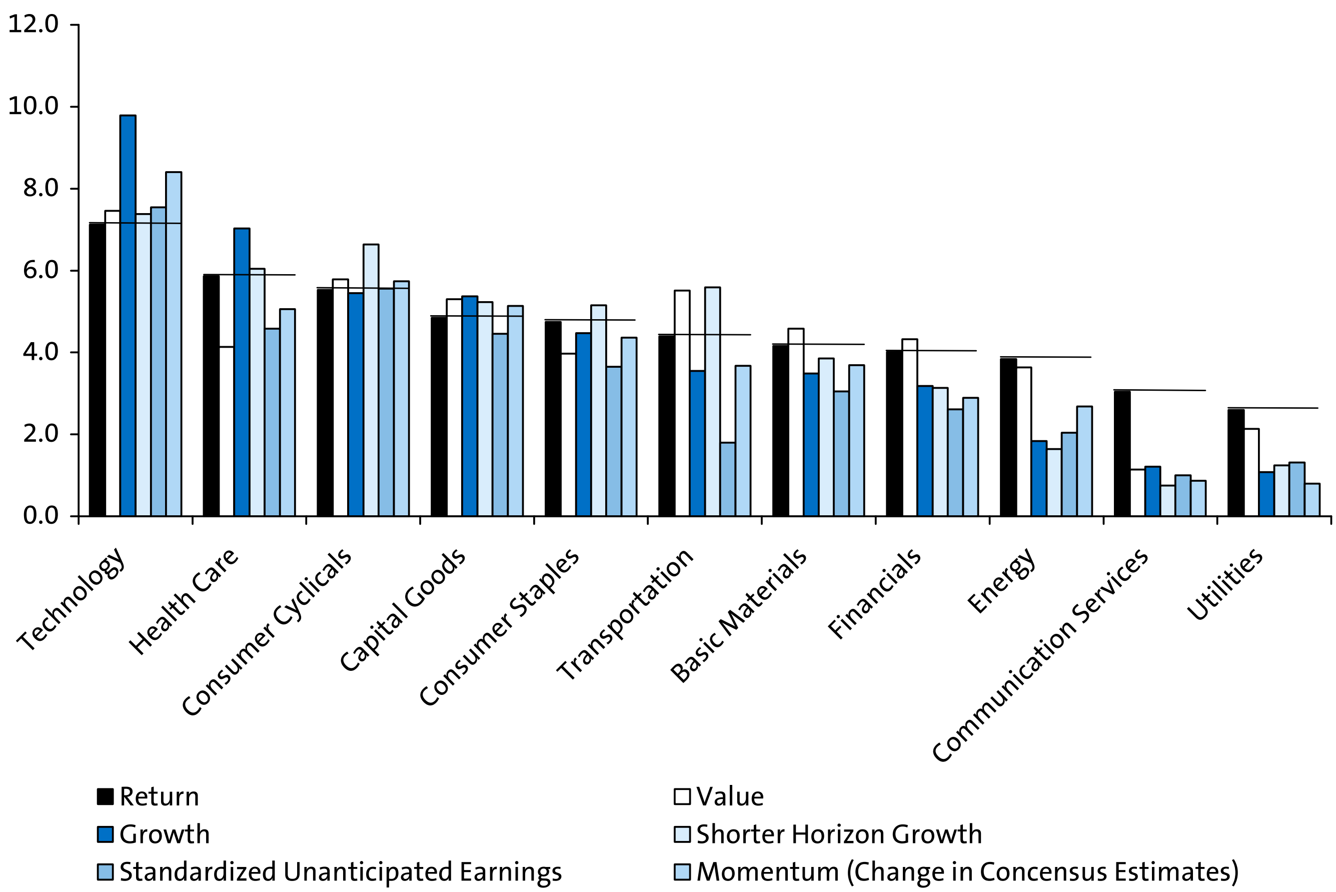
Skill levels in the simulations have been set to produce 5% excess return in a non-sector-controlled portfolio to eliminate any differences due to relative effectiveness of different styles. Thus, the relative height of bars measures relative effectiveness of the style across sectors rather than the effectiveness of that particular style. Essentially,

[5] Results are based on equally weighted portfolios of the 20% most highly rated stocks from simulations calibrated to generate a 5% return from a non-sector-controlled portfolio of all 11 sectors measured relative to an equally weighted index of all of the stocks. The presented results are excess to an equally weighted index of all of the stocks in that sector.

Please see Appendix B for more details on the investment style strategies. The notion of stockpicking skill was developed in some detail in *Beating Benchmarks*. The relevant material is repeated in Appendix D for the convenience of the reader.

Exhibit 3 shows the types of relative sector returns we would expect highly skilled managers using each particular style to achieve in each sector.

Exhibit 3: Returns from style by Compustat sector
average return excess to equal-weight sector mean %



Source: Goldman Sachs Research.

We find that there are very large differences in the dispersion of returns across sectors and thus large difference in the potential returns to stockpicking in different sectors, which suggests that sector risk budgeting may be an effective strategy. Further, it is clear that the bulk of the differences between sectors is the size of potential returns as measured by the pure returns style rather than the difference between styles.

The only notable style-generated results are the inability of fundamentals, in general, to predict returns in the communications and utilities sectors, the strong impact of fundamentals in technology, the dominance of value-based analysis over growth in the energy sector, and the dominance of growth-based analysis over value in technology and healthcare.

However, in the context of the wide dispersion of potential returns across sectors, the relatively small difference between styles does not seem sufficient to justify attempting to use different methods in different sectors, particularly if we assume a manager's style reflects a natural comparative advantage in a particular form of analysis. In that case, the gains from emphasizing the strategy at which the portfolio manager excels would overwhelm these relatively small differences in relative style effectiveness.

Further, as will be discussed later, the broad similarity in effectiveness of fundamentals to exploit the returns in each sector and the wide dispersion of returns available across sectors suggest that focus strategies that limit which sectors are invested in will push up returns but not portfolio efficiency, unless the investor is willing to take long-term sector bets. However, it may make sense to take fewer, more concentrated positions in the high-return sectors and more, better diversified positions in the lower-active-return sectors.

To get a sense of the deeper implications of these differences between sectors, we split all the dispersion (i.e., potential returns) in the data sample into categories. Exhibit 4 shows that the equal-weight benchmark (potential gains from market timings) explains 24% of the volatility in returns, common sector movements explain 7% (sector picks), and comparability within sector explains the remainder. The sectors with the highest potential returns to stockpicking are technology and consumer cyclicals, which account for nearly 33% of the total potential returns to stockpicking (43% if we insist on being fully invested and eliminate the potential returns from market timing).

Exhibit 4: Decomposition of quarterly stock return variance
pooled cross-sectional and time series variance

	Percent of Variance Explained			
Source of Variation	**Quarterly Stock Returns**		**Quarterly Sector Index Returns**	
Equal-weight benchmark		24%		69%
Commonality (between sector variation)		7%		
Sector decomposition	Comparability (within sector variation)		Commonality (between sector variation)	
Technology	18%		5%	
Health Care	6%		4%	
Consumer Cyclicals	15%		1%	
Capital Goods	7%		1%	
Consumer Staples	6%		1%	
Transportation	2%		2%	
Basic Materials	4%		1%	
Financials	8%		2%	
Energy	2%		8%	
Communications Services	1%		3%	
Utilities	1%		4%	
Total	69%	69%	31%	31%
Total variation		100%		100%

Source: Goldman Sachs Research.

We can perform a similar decomposition from the perspective of the strategist attempting to make a sector call, as in the second column of Exhibit 4. The thought experiment here is about selecting sectors not stocks within sectors; hence, the object of study is the sector return. Market timing, fluctuations over time, and permanent differences across sectors account for 69% of the sector return variation. The remaining 31% is spread unevenly across sectors.

Energy stands out as having high commonality (more than double the average) and low comparability, indicating a relatively greater importance of getting the sector call right (at least in terms of quarterly performance). Consumer cyclicals, in contrast, has very low commonality and high comparability, indicating a nearly complete dominance of stockpicking. Technology has both high commonality and high comparability, indicating both a strong sector call and a strong return to pure stockpicking within sector.

The impact of sector controls

In this section, we compare relatively complex risk-control strategies applied to simulated portfolio manager stock selections. To evaluate such comparisons, it is necessary to have as much data as possible, both in terms of numbers of stocks and quarters of data. As a result, we discontinue the use of styles that involve consensus earnings estimates and focus on pure growth and value styles. Given the results in the prior section and other work we have done in the past, we expect these results to be broadly indicative of more complex value and growth styles.

Again, we focus on comparisons of the performance of highly skilled growth and value managers, whose skill levels have been normalized to a 5% return for an equally weighted active portfolio of 20% of the stocks measured relative to an equally weighted benchmark of the entire sample regardless of style.[6] Thus, the comparisons within a style and relative movements across styles are comparable as we shift risk controls, but we cannot draw any valid conclusions based on the level of excess return across styles.

Understanding the balance between commonality and comparability in the context of a particular sector definition is actually quite easy; all we have to do is apply the sector controls and observe what happens. (As we discuss later, it is much more difficult to determine how much those results reflect the particular sector definition rather than the styles of investing we are studying. However, we will show that the following results are in fact quite robust to changes in sector definition.) If we apply sector controls (using the 11 Compustat sectors) in Exhibit 5, we see a strong result that tracking errors decline modestly regardless of style and in roughly similar amounts, but the returns respond quite differently.[7] For value, returns climb 50 basis points (bp) as comparability

---

[6] In the main body of the paper, we focus on our primary data sample and the relatively broad portfolios of 20% of the stocks in the universe. The results on a large-cap universe and more concentrated (4% of the stocks in the universe) portfolios, shown in Appendices F and G, respectively, are qualitatively the same as those for the broader universe and portfolio. The large-cap results, as with many stockpicking results on large-cap stocks, are somewhat muted relative to those of the broader sample.

[7] For portfolios with sector control, the style characteristics are ranked within sector and the top 20% of each sector (by the style characteristic) are equally weighted. Sector/quarter combinations with fewer than five stocks are removed, and at least one stock is picked from all remaining sectors. These sector portfolios are then weighted by the number of stocks in each sector.

is improved, while for growth, returns drop 50 bp as it becomes impossible to exploit commonality of movements (i.e., implicit sector calls).

Exhibit 5: Effect of sector controls on growth and value

Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Growth			
Without Sector Control	5.0	2.0	2.44
With Sector Control	4.5	1.9	2.38
Difference	**-0.5**	**-0.1**	**-0.06**
Value			
Without Sector Control	5.0	2.3	2.23
With Sector Control	5.5	1.9	2.83
Difference	**0.5**	**-0.3**	**0.60**

Source: Goldman Sachs Research.

In Exhibit 6, we break down the change in Sharpe ratio (return per unit of risk) between the strategies without sector control and with sector control into two components: one due to the change in return and one due to the change in tracking error. To get the change in Sharpe ratio due to the change in excess return, we calculate the difference between the new Sharpe ratio, which uses the new (sector controlled) excess return and the base (without sector control) tracking error, and the base (without sector control) Sharpe ratio.

As Exhibit 6 shows, for growth, the change in Sharpe ratio due to the excess return is negative, which is another way of seeing that controlling for the 11 Compustat sectors drops the excess return for the growth strategy. In contrast, for value, the change in Sharpe ratio due to the excess return is positive, as is the increase in excess return from controlling for sector in value.

The change in Sharpe ratio due to tracking error is the rest of the change in the Sharpe ratio between sector-controlled and non-sector controlled (i.e., this change is the difference between the Sharpe ratio based on the actual tracking error from the sector-controlled strategy and the Sharpe ratio just calculated with the sector-controlled return and the base tracking error). For both growth and value, the change in Sharpe ratio due to tracking error is positive.

Exhibit 6: Sharpe ratio decomposition

Investment Style	Sharpe Ratio Without Sector Control	Sharpe Ratio With Sector Control	Change in Sharpe Ratio	Change Due to Excess Return	Change Due to Tracking Error
Growth	2.44	2.38	-0.06	-0.23	0.17
Value	2.23	2.83	0.60	0.21	0.39

Source: Goldman Sachs Research.

Focus strategies

Exhibit 7 shows the returns, tracking error, and Sharpe ratios for the growth and value strategies if they were applied as though each individual sector was the portfolio manager's complete universe. The returns are excess relative to an equally weighted benchmark of only the stocks in that sector.

Exhibit 7: Sector breakdown
returns excess to equal-weight sector mean

	Growth			Value		
Sector	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Basic Materials	3.2	5.1	0.61	5.1	5.2	0.97
Consumer Cyclicals	4.9	4.4	1.12	6.3	4.5	1.42
Consumer Staples	4.0	5.0	0.81	4.3	5.0	0.87
Health Care	6.3	8.8	0.72	4.6	8.7	0.53
Energy	1.6	9.1	0.18	4.0	9.0	0.45
Financials	2.9	3.3	0.87	4.8	3.4	1.40
Capital Goods	4.9	4.8	1.02	5.8	4.6	1.25
Technology	8.9	7.4	1.19	8.2	7.3	1.13
Communication Services	1.0	11.4	0.09	1.4	10.7	0.13
Utilities	1.0	3.4	0.28	2.4	3.4	0.70
Transportation	3.1	10.9	0.29	6.1	10.8	0.57
Equal weight average	3.8	6.7	0.7	4.8	6.6	0.9

Source: Goldman Sachs Research.

If we look at the sector results, the high returns in the best sector make it tempting to suggest that we should simply focus on the best sectors. However, the very low sector Sharpe ratios suggest that this level of focus probably both exceeds the risk tolerance of most investors and makes it all but impossible to evaluate manager skill. However, it is possible that a more limited reduction in the stock universe would be of value.

In Exhibit 8, we show what happens if we limit the stock universe by reducing the number of sectors. Here, our new, more focused universe consists of the sectors in which the individual sector excess return (in Exhibit 7) is greater than or equal to the median excess return. That is, we focus on the 6 of the 11 sectors in which stockpicking was most effective in terms of generating excess return.

Exhibit 8: Focused strategies
returns excess to equal-weight mean of whole or focused sample

Investment Style	Growth* Average Return to Skill (%)	Growth* Average Tracking Error (%)	Growth* Sharpe Ratio	Value** Average Return to Skill (%)	Value** Average Tracking Error (%)	Value** Sharpe Ratio
11 Sectors						
Without Sector Controls	5.0	2.0	2.44	5.0	2.3	2.23
With Sector Controls	4.5	1.9	2.38	5.5	1.9	2.83
6 Highest Excess Return Sectors						
Without Sector Controls	5.9	2.6	2.26	5.5	2.6	2.10
With Sector Controls	5.6	2.5	2.22	6.1	2.3	2.62

*Growth sectors include Basic Materials, Consumer Cyclicals, Consumer Staples, Health Care, Capital Goods, Technology
**Value sectors include Basic Materials, Consumer Cyclicals, Financials, Capital Goods, Technology, Transportation

Source: Goldman Sachs Research.

The results show some modest improvement in returns but a reduction in overall risk efficiency as the portfolio Sharpe ratios decline. Thus, although the returns of focus strategies may be attractive, unless the investor is willing to take the benchmark risk from simply not investing in the less-active, management-friendly sectors, that investor is still better off taking active management risk across the entire stock universe.

Further, as noted earlier, these results also suggest that it may make more sense to take larger, more concentrated positions in the focus sectors and smaller, better diversified positions in the non-focus sectors as a hybrid strategy.

This result is not an argument against specialty funds that focus skill and research on a smaller universe of stocks. It is simply a caution that in constructing a portfolio of active sector managers, it still may be efficient to include managers for sectors in which active management has historically generate poor results, because the overall portfolio may still be more efficient than a true focus strategy, which has zero weight in some sectors. This caution is especially necessary given the backward-looking nature of these results and the possibility that a "cold" sector for active management might turn "hot."

Alternative sector definitions

Until now, we have focused on classical sector definitions. It is reasonable to ask whether there is any reason to expect that "economic sectors" applied to companies that do not always neatly fit such categories will lead to the best results, and whether the prior results are sensitive to the way the sectors are constructed.

Optimized groups

There are nearly an infinite number of possible sector definitions, and to examine a sufficient sampling of those definitions to claim our conclusions are robust is not feasible. However, we can clearly define the extent of the potential problem by creating a set of best sectors that set an upper bound on how good sector risk controls could possibly be. To do this, we created optimized groups (performance-defined sectors) that are chosen to optimize the performance of each investment style over the sample periods. We also look at how many categories there should be.

We do not argue that such backward-looking groups would work in the future (although we find strong evidence of stability for the value groups when we test out of sample). These groups simply provide an idea of how much a best split between groups of stocks could do for stock selection and whether further research into such categories might add significantly to the performance of managers. Optimized growth groups also help determine whether the negative impact of risk controls on growth managers was the result of badly implemented sectors or was fundamental to the growth style of investment.

The intuition behind the creation of optimized groups is simple. As mentioned earlier, the ability of active stock selection based on fundamental analysis to create returns is dependent on two factors:

- **Comparability**—the correlation between the fundamental measure and future return performance.
- **Dispersion**—the greater the potential difference in performance between stocks, the greater the value of being able to discriminate between them.

We start with a fundamental measure (growth or value) and allocate each stock in the universe into two groups so that the count-weighted returns to stockpicking within those groups are maximized. We then add another possible group and reallocate the stocks among the three groups, and so forth. We show results through ten growth or value groups. This process creates groups that maximize the returns to "sector-constrained" stockpicking by maximizing the comparability of the stocks within each group. (See Appendix H for a more detailed and rigorous description of this process.)

Once we have the optimized groups, we simulate highly skilled portfolio managers controlling for these new sector groups. Exhibit 9 shows the results for three and five of the optimized groups, along with the base (without sector control) and sector-controlled (i.e., controlled for the 11 Compustat sectors) for comparison. Note that the groups optimized for growth are used for the growth style, and the value groups, which are different, are used for the value style.

Exhibit 9: Optimal groups
returns excess to equal-weight mean of whole sample

	Growth			Value		
Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Without Sector Controls	5.0	2.0	2.44	5.0	2.3	2.23
11 Sectors, Sector Controls	4.5	1.9	2.38	5.5	1.9	2.83
3 Optimal Groups, Sector Controls	5.3	2.1	2.53	8.1	2.1	3.92
5 Optimal Groups, Sector Controls	5.5	2.1	2.60	8.8	2.1	4.18

Source: Goldman Sachs Research.

The key result is that both returns and Sharpe ratios climb considerably for value managers as the groups are optimized and as the number of groups is increased. In contrast, for growth, the impact is less dramatic, although both returns and Sharpe ratios increase. Given the backward-looking nature of these optimized groups, this result suggests that a significant share of a growth manager's performance arises from identifying differences in group performance (commonality); thus, sector over/underweights, even when fully derived from bottoms-up analysis, are still fundamental to the performance of a growth portfolio. Even optimized group restrictions would likely hurt performance.[8]

These results become even more clear in the final section of this paper, in which we examine the stability of these optimized groups and adjust for the in-sample biases of these procedures. Specifically, what little positive impact we find for optimized groups for growth managers can be attributed to the in-sample biases of the way in which we construct the optimized groups and that, once this bias is removed, sector restrictions provide no incremental returns for growth styles. In contrast, the net impact remains strongly positive for value managers.

We are not saying that growth managers should run their portfolios without thought of risk. Rather, their risk control should be thought of as informational rather than performance enhancing. It is clearly in growth managers' best interest to know what risks they are taking to be sure they are in fact taking the risks they want. However, we would be very wary of controls that sought to systematically limit the growth manager's ability to overweight one group of stocks against another. In contrast, in value investing, we think that performance can be enhanced by such restrictions, as the valuation comparisons can become more accurate and more indicative of future performance.

Hence, we argue that value managers have much to gain from segmenting stocks into comparable groupings and then choosing stocks within those groups, while growth managers needs to exploit the between-group commonality and, thus, need more sector choice. We also note that a modified value method could provide greater insight into cross-sector performance and could push these results toward relaxing sector controls. Broadly, as we have said before, the clear implication is that risk systems need to be

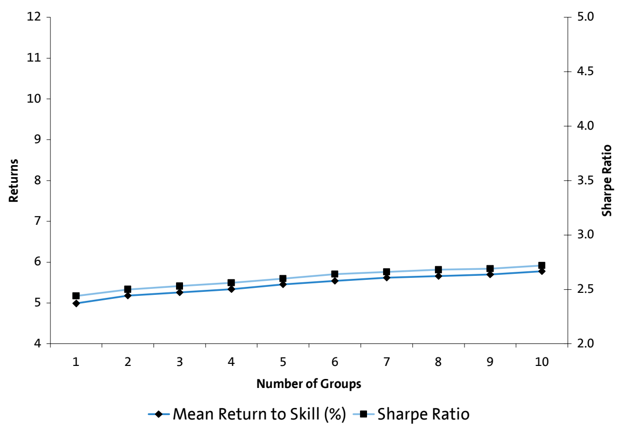
[8] Later results that make statistical adjustments for the backward-looking bias provide additional evidence for this intuition.

tailored to the specific investment style of the manager and that "one-size-fits-all" approaches to risk management are almost certain to be less than optimal.

How many optimized groups?

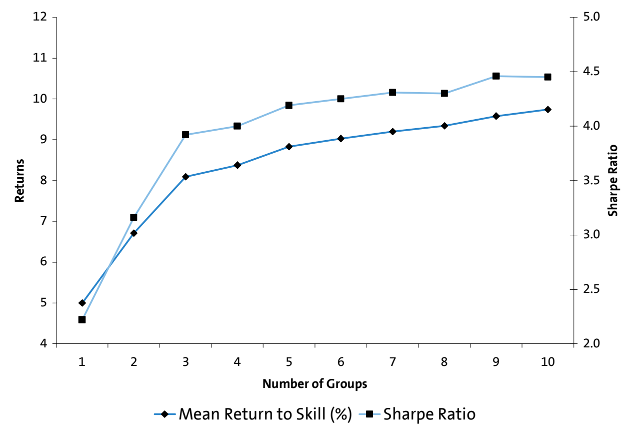
Exhibits 10 and 11 show performance as the number of groups is increased over a broad range (one to ten). We see that performance gains are most dramatic for the first three value groups and not particularly dramatic at any number of growth groups. These results suggest that there is little gain from going beyond three to five groups—far fewer groups than most risk systems or research departments use.

Exhibit 10: Returns as number of groups increase
growth



Source: Goldman Sachs Research.

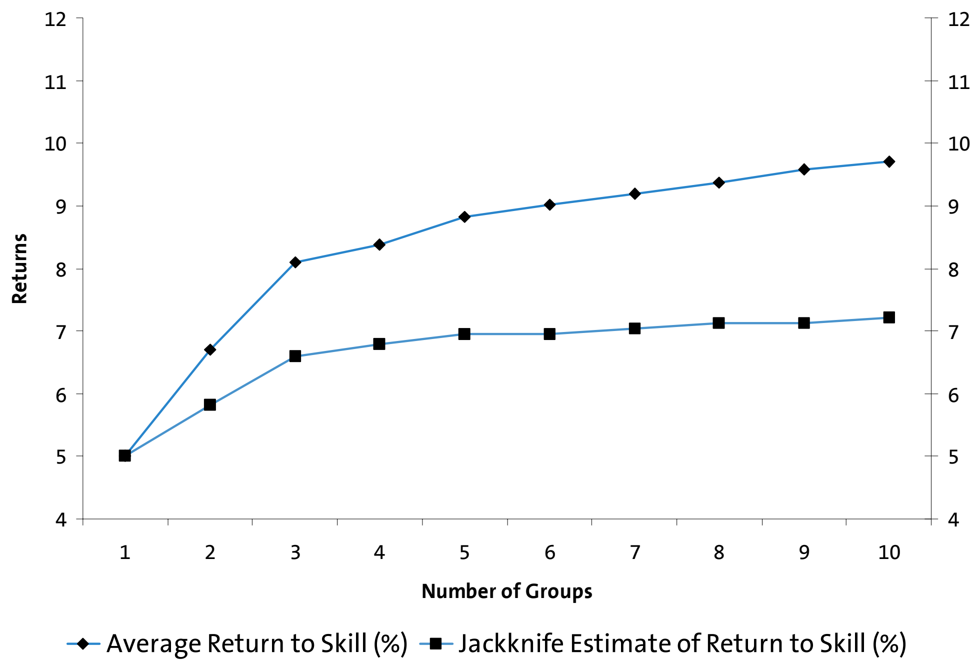
Exhibit 11: Returns as number of groups increase
value



Source: Goldman Sachs Research.

The lack of benefits beyond five optimal groups becomes more apparent when the procedure is adjusted for in-sample bias via a jackknife procedure. The procedure estimates the return to skill for each individual quarter based on clusters constructed without the data from the specific quarter. Thus, separate data enters the clustering and return calculations. The resulting jackknife estimates range from 87% of the in-sample counterpart for two groups to 74% for ten groups, indicating that the bias increases with the number of groups. Eliminating the bias reveals that the incremental return of ten groups relative to five groups is on the order of 25 bp. This finding, coupled with the stability results in the final section of the paper, suggest that one growth group and three to five value groups are sufficient levels of partitioning.

Exhibit 12: In-sample and jackknife estimates of returns as number of groups increase
value



Source: Goldman Sachs Research.

We note that with only three to five groups, value managers would still be allowed significant sector discretion by the standards of most risk-control systems. As we also do not find significant losses from over-specifying the number of groups, more finely delineated sector definitions do not seem to entail serious loss and may improve either research discipline or tracking error.

Focusing within optimized groups

Exhibit 13 shows the results from focus strategies within the optimized groups. Again, we take the groups with excess return over their individual group equal-weight benchmarks greater than or equal to the median excess return. Thus, we take the best two out of three and the best three out of five optimized groups.

Exhibit 13: Optimal groups, focused strategies
returns excess to equal-weight mean of whole or focused sample

	Growth			Value		
Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
With Sector Controls						
3 Optimal Groups	5.3	2.1	2.53	8.1	2.1	3.92
Focused - 2 of 3 Highest Excess Return Groups	6.2	2.5	2.50	9.0	2.7	3.31
5 Optimal Groups	5.5	2.1	2.60	8.8	2.1	4.18
Focused - 3 of 5 Highest Excess Return Groups	6.7	2.8	2.38	9.9	2.3	4.39

Source: Goldman Sachs Research.

These results show that, for growth, focusing can improve excess return at the expense of a slight drop in efficiency, as the portfolio effects of diversification slightly outweigh the gains from focusing on higher-return areas. For value, focusing can improve excess return and may or may not increase efficiency. Note, however, that the performance of the focused strategies is measured against the benchmark for only the focus groups. That is, stocks that are ignored for stockpicking are also ignored for the benchmark.

A benchmark-sensitive portfolio manager with a full universe benchmark would still be better off taking active stockpicking risk in all sectors, provided the manager had positive stockpicking efficiency in every sector. From a risk budgeting perspective, the manager might want to take slightly more risk in groups in which the stockpicking was most efficient.

Stability of the optimized groups

As we have noted a number of times, our optimized group analysis is biased by its backward-looking structure (our focus strategies suffer from the same type of backward-looking bias)—i.e., we used returns and fundamental data from the entire period to decide in which optimized groups each stock should belong. Thus, we know our optimized groups are effective over the universe and time period in our sample, but we know less about how effective they would be over the next five years.

Optimal group membership based on the growth strategy is considerably less stable than membership based on the value strategy. By stable we mean that a stock is more likely to retain its group membership as the sample changes over time. This is easily demonstrated in the jackknife procedure by examining the transition rates between groups for adjacent quarters. For example, we compute the fraction of observations that remain in the same group between the first and second quarters of 1990. Averaging this value across all the quarters yields the aggregate measure of instability summarized in Exhibit 14. The increase in instability as the number of groups grows is further reason to prefer solutions based on a relatively low number of groups.

Exhibit 14: Group instability
average misclassification rate

Number of groups	Value	Growth	Growth/Value
2	5%	23%	5.0x
3	5%	27%	5.4x
4	10%	39%	3.8x
5	20%	46%	2.3x

Source: Goldman Sachs Research.

For the main results of this paper, backward-looking bias in unimportant. Our main point regarding the optimized clusters is that it is possible to do a much better job of grouping stocks into "sectors" than is currently done, particularly for value-based investment strategies. As in much of our research, we are not trying to show how to build a better mousetrap, only the importance of some characteristics of the perfect mousetrap.

However, if one actually wanted to build better stockpicking groups, it would be important that the groups be stable over some horizon so they could be exploited to gain future performance. Although we do not attempt to maximize the stability of our optimized groups, we examine their stability with an out-of-sample test. In Exhibit 15, we optimize the growth and value groups over the first half of the sample and simulate portfolios over the second half of the sample.

Exhibit 15: Out-of-sample group stability
returns excess to equal-weight mean of whole or focused sample

	Growth			Value		
Investment Style	**Average Return to Skill (%)**	**Average Tracking Error (%)**	**Sharpe Ratio**	**Average Return to Skill (%)**	**Average Tracking Error (%)**	**Sharpe Ratio**
Without Sector Control	5.0	2.0	2.58	5.0	2.4	2.06
With Sector Control						
11 Compustat Sectors	4.5	1.8	2.42	5.2	2.1	2.53
3 Optimal Groups	5.0	2.0	2.55	5.6	2.2	2.52
Focused - 2 of 3 Optimal Groups with Highest Excess Returns	5.0	2.2	2.30	6.0	2.3	2.62
5 Optimal Groups	5.0	2.0	2.53	5.9	2.2	2.69
Focused - 3 of 5 Optimal Groups with Highest Excess Returns	5.3	2.6	2.05	6.3	2.5	2.50

Create optimal groups from 12/31/1984 - 3/30/1990
Run portfolios from 3/30/1990 - 9/30/1999

Source: Goldman Sachs Research.

We find reasonable stability in the value groups. The out-of-sample value results show the same pattern as the in-sample results, with an expected reduction in size of effect. That is, we still find that both value returns and value Sharpe ratios go up when we use the first-half value groups. For growth, however, the optimized groups perform slightly worse than the uncontrolled group, indicating that the modest positives in the prior results were likely due to the in-sample biases of the group construction methods.

Quick recap

Sector controls restrict the fundamental insight of the form group X will outperform group Y that sits at the heart of growth styles. For value styles, there seems to be significant gains from restricting stockpicking to groups of comparable stocks, but those groups appear to be far broader than would normally be implicit in most sector-neutral risk systems. More broadly, we find that effectiveness of style varies considerably from sector to sector, and the importance of the sector call versus the stock call varies as well. Net, we find strong gains from tailoring risk controls and risk allocations to the style of the manager and find considerable evidence that one-size-fits-all risk approaches are likely to noticeably impair manager performance.

Appendices

Appendix A: Why optimizers might not optimize

From an academic perspective, it might be argued that the appropriate use of an optimizer to form portfolios eliminates the issues discussed in this paper. In reality, the relationship between our results and the structure of most optimizers is quite subtle. In the standard optimizer structure, the portfolio manager generates predicted alphas, and then the optimizer creates a portfolio that maximizes expected alpha for a given tracking error.

A core assumption of such a process is that the predicted alphas are independent of the risk management process. In the current paper, we find that for some styles of portfolio management, the risk management filters through which the portfolio managers' judgments are passed actually impact the anticipated alphas.

The reason this is possible is that the core bit of information the portfolio manager generates in our models is not expected returns (or price targets) but rankings of stocks. The filters (such as take the top-50 stocks in the universe or take the top-10 stocks from five specified groups) then generate portfolios that have anticipated returns based on the accuracy of the rankings. From a technical perspective, the sector filters or size filter are actually being used to change the models that predict alpha rather than to control tracking error. In a pure modeling context, this is much like asking if sector sub-models work better than full universe models and whether they have the ability to predict sector returns.

Viewed from this perspective, the paper finds that value-based models work better as sub-sector models and have little ability to pick sector weightings, while growth models work better as full universe models and excel at sector overweighting/underweighting prediction.

From the perspective of non-quantitative managers, the results are more usefully interpreted as simply saying that some types of risks have higher payoffs than others and that managers should focus risk-taking on areas in which their style is most effective. We also note that for those attempting to blend qualitative judgments with quantitative risk control, our results indicate that simply translating qualitative judgment into estimated alphas stock by stock and applying an optimizer will not in general produce the best results.

One simple way of understanding this is to think of three stocks: Exxon, Shell, and Microsoft. At a particular moment in time, the manager might have anticipated alphas of 100 bp for Exxon and –50 bp for both Shell and Microsoft. A standard optimizer would view Exxon as being 150 bp better than Shell and Microsoft and overweight Exxon and underweight Shell and Microsoft proportionally based on their relative contribution to tracking error. In reality, however, the portfolio manager might be better at comparing within sector, implying that overweighting Exxon and underweighting Shell should be the dominant strategy. Alternatively, that manager might be better at sector calls, in which case, overweighting Exxon and underweighting Microsoft should be dominant. The optimizer assumes that the expected alphas hold all of the information about the portfolio managers' skill and judgment, when in fact they may not.

Appendix B: Data, portfolio construction, and investment style

In this appendix, we describe the data, portfolio construction, and investment style strategies.

Primary data sample

For this paper, we start with the Compustat universe of US companies. We include companies that are no longer active to mitigate survivorship bias. We remove secondary and tertiary issues and companies and data points for which the data appears to be seriously flawed.

We remove micro-cap companies by requiring that market caps be greater than what would be historically comparable to more than $500 million at the end of 1999. We calculate the market cap cutoff for each quarter by decrementing that $500 million by the return on a broad market measure, the Russell 3000. Thus, in the primary sample, the smallest market cap ranges from $41 million on December 31, 1984, to $461 million on June 30, 1999.

Our primary sample is used to simulate returns from value and growth investment strategies. Thus, for our primary sample, we remove stock/quarter combinations that lack value or growth numbers, where value and growth are defined below. We also require that Compustat has assigned an economic sector to each stock.

Our pricing data and some of our earnings data are from Compustat. The rest of our earnings data and our earnings estimate data is from I/B/E/S.

Timing of data

Our data is organized according to calendar quarters. We require that all stocks have a market cap at the beginning of the quarter and a return over the next quarter. The returns in our data sample run from December 31, 1984, to September 30, 1999. Fundamental data can precede the returns by up to two quarters and extend beyond the returns by up to four quarters.

To allow time for a reporting lag, we assume that December 1995 earnings were reported by March 1996, the end of the next quarter. Thus, when we form a portfolio on March 30, 1996, the current fundamentals are from the December 1995 quarter and the current one-quarter forward estimates are for the March 1996 quarter.

Summary statistics

The primary sample consists of 3,951 stocks that enter the sample for one or more periods. Several other samples are described below. Exhibit 16 provides some descriptive statistics about each of the samples.

Exhibit 16: Descriptive statistics by sample

Sample	Number of Unique Stocks	Number of Stock/Qtr Data Points	Number of Quarters	Return Start Date	Return End Date	Min Mkt Cap at Start Date ($ million)	Min Mkt Cap at End Date ($ million)
Primary	3,951	82,612	59	12/31/84	9/30/99	41	461
Style Samples							
Growth	3,951	82,612	59	12/31/84	9/30/99	41	461
Value	3,951	82,612	59	12/31/84	9/30/99	41	461
Return	3,951	82,612	59	12/31/84	9/30/99	41	461
Shorter-Horizon Growth	3,710	71,631	59	12/31/84	9/30/99	41	461
Standardized Unanticipated Earnings	2,937	29,756	59	12/31/84	9/30/99	52	461
Momentum (Change in Consensus Estimates)	3,544	53,415	59	12/31/84	9/30/99	42	461
Largest 500 Stocks	1,168	29,500	59	12/31/84	9/30/99	277	2,937
Broad Sample	6,054	146,430	69	12/31/82	3/31/00	32	500
Goldman Sachs Industry Sample	832	38,220	69	12/31/82	3/31/00	33	507

Source: Goldman Sachs Research.

Secondary data samples

Samples for other style strategies

To simulate other style strategies, the style indicator variables must be present. For the pure return-based strategy, this does not force us to trim the primary sample, as returns were needed in the value and growth strategies as well. The indicator variables for each of the other three styles of investment strategies are less broadly available, reducing the sample on which those strategies can be run. For shorter-horizon growth, the overall number of stock/quarter data points is reduced from 82,612 to 71,631. For standardized unanticipated earnings and momentum (change in consensus earnings estimates), the effect is more dramatic, reducing the overall number of stock/quarter data points to 29,756 and 53,415, respectively.

Although we only use these other four measures to show the robustness of the results to variations in the value or growth metric, we suggest caution in directly comparing the absolute levels of the results from the different strategies given the different samples.

Largest 500 stocks

To check the robustness of the results in a more purely large-cap universe, we further restrict the primary sample to include only the largest 500 stocks in each period.

Broad sample

For the size versus sector analysis, we wanted the broadest sample possible. Thus, although we start from the cleansed Compustat and I/B/E/S data, we only require the existence of a return, a market value, and a Compustat sector.

Goldman Sachs industry sample

We use the Goldman Sachs industry classification as a finer sector grid to check the robustness of the size versus sector analysis. For this sample, we start with the broad sample described above and then require a Goldman Sachs industry classification and at least five stocks within that classification per quarter. That is, if there are only four stocks in a particular industry for the third quarter of 1985, then all four of those stocks are dropped for that quarter, but only for that quarter. The idea is to keep only sectors in which a stockpicker aiming for a 20% portfolio would keep one of the stocks in that industry for that quarter. There are 47 Goldman Sachs industries in this sample.

Portfolio construction

Each quarter, we rank the stocks by their style characteristics (described in detail later in this appendix). For the results in the body of the paper, the portfolios without sector control are formed by equally weighting the 20% of the stocks that have the best style characteristics (e.g., highest growth rate and lowest P/E) for that quarter. That portfolio is held for one quarter, and the ranking and portfolio formation process is repeated at the beginning of the next quarter.

For portfolios with sector control, the style characteristics are ranked within sector and the top 20% of each sector (by the style characteristic) are equally weighted. Sector/quarter combinations with fewer than five stocks are removed, and at least one stock is picked from all remaining sectors. These sector portfolios are then weighted by the number of stocks in that sector.

The investment strategies we examine in this paper utilize forward-looking information (i.e., return, earnings, and estimate data that is not yet reported) as proxies for the insights of fundamental analysts and portfolio managers. Of course, if we gave managers perfect foresight, their portfolio returns would be incredibly high. Thus, we simulate a skill level more consistent with real world returns. The details of that skill simulation can be found in Appendix D.

Here, the point is that our investment strategies are not designed to be the optimal indicator for a quantitative stockpicking model. Our investment strategies are designed to be transparent, clear, useful examples that typify the performance of their style class.

The results in this paper also ignore transaction costs, which we discuss in Appendix A of our January 1999 report in this series, entitled *Style, Size, and Skill*.

Investment strategy styles

Growth

Our growth measure is the percentage change (on a log or continuously compounded basis) in I/B/E/S one-quarter actual primary earnings per share (excluding extraordinary items) over four quarters. This means comparing, for example, December 1996 earnings with December 1995 earnings.

In equation form, the growth calculation is as follows:

$$\text{growth measure } = \ln(e_4/e_0),$$

where e_4 is the earnings four quarters forward and
e_0 is the earnings for the quarter just reported.

We use a four-quarter change because a long-horizon growth rate is more effective than a shorter-horizon growth rate when investing in large-cap stocks (see *Style, Size, and Skill*). We also use a shorter-horizon growth strategy, which is described below.

Value

We use P/E based on I/B/E/S actual earnings per share (excluding extraordinary items) as the value measure. 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. In particular, we sum four forward quarters of actual earnings from I/B/E/S for this value measure.

In equation form, the value calculation is as follows:

$$\text{value measure } = \frac{e_1 + e_2 + e_3 + e_4}{p},$$

where e_n is the earnings n quarters forward.

Return

For a more agnostic style, we simply use the one-quarter forward return as an indicator of future performance. Because this measure is perfectly correlated with the stock performance we are measuring, we calibrate the simulated portfolio manager skill to an equivalent portfolio return, which requires a significantly lower level of skill tilt in our methodology.

Shorter-horizon growth

As we show in *Style, Size, and Skill*, a shorter-horizon measure of growth can be more effective for stockpicking in some samples, particularly samples of smaller-cap stocks. This growth measure is the percentage change (on a log basis) in Compustat one-quarter primary earnings per share (excluding extraordinary items) over two quarters. This means comparing, for example, June 1996 earnings with December 1995 earnings.

In equation form, the shorter-horizon growth calculation is as follows:

$$\text{shorter-horizon growth measure } = \ln(e_2/e_0),$$

where e_2 is the earnings two quarters forward and
e_0 is the earnings for the quarter just reported.

Standardized unanticipated earnings

For this measure of earnings surprise, we use four-quarter forward standardized unanticipated earnings from I/B/E/S data. That is, for the numerator of this ratio, we add four forward quarters of actual earnings and subtract four forward quarters of estimates. For the denominator of this ratio, we use the standard deviation of the four-quarter estimate of earnings.

In equation form, the standardized unanticipated earnings (sue) calculation is as follows:

$$\text{sue measure} = \frac{e_1 + e_2 + e_3 + e_4 - \left(e_1^c + e_2^c + e_3^c + e_4^c\right)}{\sqrt{\sigma^2_{e_1^c} + \sigma^2_{e_2^c} + \sigma^2_{e_3^c} + \sigma^2_{e_4^c}}},$$

where e_n is the earnings n quarters forward

e_n^c is the consensus earnings estimate for the earnings n quarters forward at the time those earnings are reported and

$\sigma^2_{e_n^c}$ is the variance of the individual analysts' estimates around the consensus at the time the earnings are reported.

Momentum (change in consensus earnings)

For this measure of momentum, we use a one-quarter change in the consensus earnings estimates for the second, third, and fourth quarters forward from I/B/E/S data. (The change in the first-quarter forward surprise is not used because it does not typically exist. One quarter forward, the earnings have typically already been reported at some point during the quarter.)

That is, we calculate the percentage change (on a log basis) between the sum of today's consensus estimates for earnings two, three, and four quarters and the consensus estimates for those same quarters one quarter later.

In equation form, the momentum calculation is as follows:

$$\text{momentum measure} = \ln\left(\frac{e^c_{2,1Q\,fwd} + e^c_{3,1Q\,fwd} + e^c_{4,1Q\,fwd}}{e^c_{2,\,today} + e^c_{3,\,today} + e^c_{4,\,today}}\right),$$

where $e^c_{n,\,today}$ is today's consensus estimate for the earnings n quarters forward and

$e^c_{n,\,1Q\,fwd}$ is the consensus estimate for the same earnings one quarter later.

Appendix C: Size controls

In this appendix, we demonstrate the ineffectiveness of size controls, especially after neutralizing the benchmark risk of the largest stocks. Size controls, regardless of investment strategy, decrease return at a faster rate than they decrease tracking error with a net result of lower Sharpe ratios.

Exhibit 17 summarizes the returns to skill, tracking error, and Sharpe ratio of applying size controls to value and growth strategies. The size controls are based on deciles of the size distribution each quarter.

Exhibit 17: Effect of size controls on growth and value

			Value Strategy			Growth Strategy		
Risk Control	Size Decile Weighting	Benchmark	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
None	Equal weight	Equal weight	5.0	2.3	2.22	5.0	2.0	2.45
10 Size controls	Equal weight	Equal weight	4.9	2.2	2.20	4.9	2.0	2.41
Difference			**-0.1**	**0.0**	**0.0**	**-0.1**	**0.0**	**-0.04**
None	Equal weight	Cap weight	4.9	6.9	0.71	5.0	6.7	0.74
10 Size controls	Cap weight	Cap weight	2.0	3.3	0.61	3.2	3.1	1.06
Difference			**-2.9**	**-3.6**	**-0.1**	**-1.7**	**-3.7**	**0.32**
Market cap weight largest 50 stocks	Cap weight	Cap weight	3.5	2.4	1.45	3.5	2.6	1.35
Market cap weight largest 50 stocks + 10 size controls	Cap weight	Cap weight	1.7	1.9	0.92	2.4	2.1	1.16
Difference			**-1.8**	**-0.6**	**-0.5**	**-1.1**	**-0.5**	**-0.19**

Source: Goldman Sachs Research.

If we were in an equal-weight world (i.e., if both our portfolio and our benchmark were equal weighted), size controls would be roughly neutral. The top panel of Exhibit 17 shows that, in this equal-weight world, size controls reduce the returns for both value and growth styles by approximately 10 bp with a corresponding decline in Sharpe ratio.

In sharp contrast, in a normal cap-weight world, size controls dramatically lower the returns to skill. The center panel of Exhibit 17 shows that, in the cap-weight world, size controls reduce the value return more than half—by 2.9 percentage points (pp)—and the growth return by 1.7 pp. This drop in return is accompanied by a drop in tracking error. Thus, in the absence of other risk control, size controls can either hurt or help the Sharpe ratio. Here, size controls decrease the value Sharpe ratio and increase the growth Sharpe ratio.

However, once we offset the stock-specific risk in the benchmark, size controls only hurt both returns and Sharpe ratios. As we suggest in our prior work, *Beating Benchmarks, A Stockpicker's Reality Part II*, holding a passive benchmark-weighted position in the largest 50 companies can offset the concentration of risk in the largest stocks in the US large-cap benchmarks. On net, passive holding of the largest stocks doubled the Sharpe ratio for both value and growth. The bottom panel of Exhibit 17 shows that after controlling the benchmark risk, size controls reduce the return dramatically (by 1.8 pp for value and by 1.1 pp for growth) without a sufficient decrease in tracking error.

We find that size controls have little impact so long as they do not artificially force portfolio managers to concentrate stockpicking risk in the largest stocks.

Appendix D: The nature of skill

For our purposes, portfolio manager skill is defined as the ability to rank stocks based on future fundamentals. 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.

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 fourth-quarter forward earnings growth.

We then 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 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 lower ratings. We do this simply by tilting the uniform distribution based on the true ranking of the stock and the skill of the manager. Exhibit 18 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).

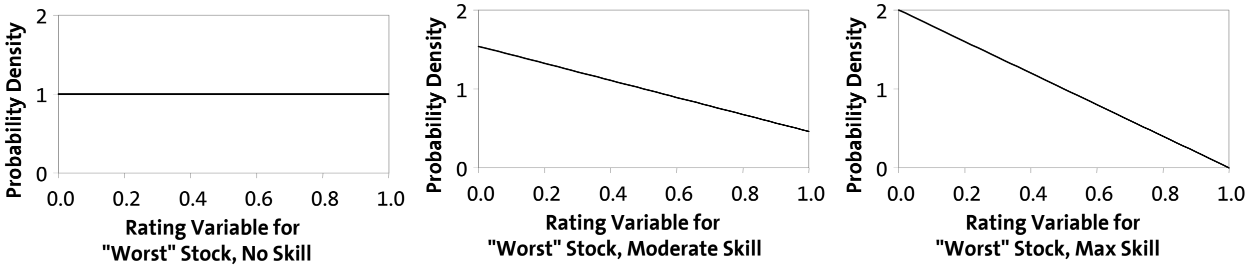
For max skill, the top stock has a 9.75% chance of getting a rating between 0.95 and 1.00 and a 36% chance of a top-quintile rating while only a 4% chance of getting a bottom-quintile rating.

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, how likely is the portfolio manager to rank stocks in the correct quintile? No skill (0% of maximum tilt) gets it right 20% of the time, moderate skill (54% of maximum tilt) 23%, high skill (82% of maximum tilt) 25%, and max skill (100% of maximum tilt) 26.4%. We focus on comparisons of the performance of highly skilled growth and value managers, whose skill levels have been normalized to a 5% return for an equally weighted active portfolio of 20% of the stocks measured relative to an equally weighted benchmark of the entire sample regardless of style.

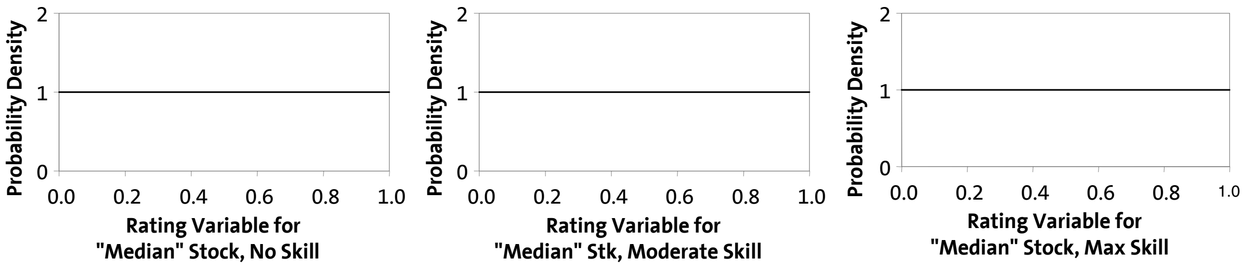
Exhibit 18: Skill

No Skill **Moderate Skill** **Max Skill**

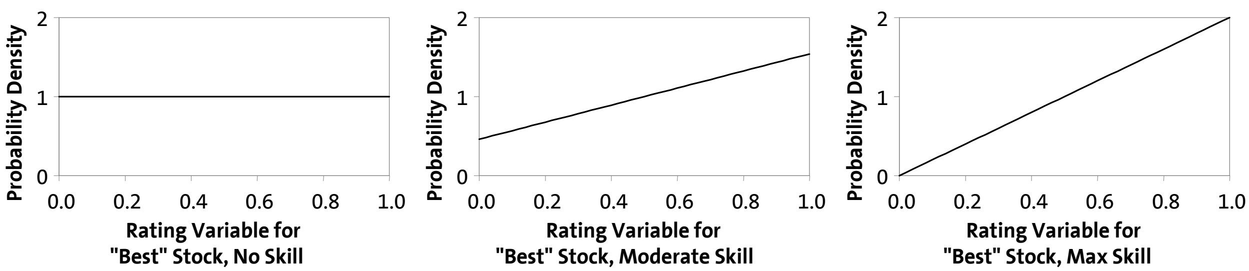
Panel 1: "Worst" Stock



Panel 2: "Median" Stock



Panel 3: "Best" Stock



Source: Goldman Sachs Research.

Appendix E: Model of returns to stockpicking

This appendix develops a statistical framework for understanding the returns to stockpicking skill.

Conceptual model of returns to stockpicking skill

At the conceptual level, returns to stockpicking skill are driven by three forces other than the skill of the analyst: returns of the entire sector, the dispersion of returns within the sector, and the strength of the relationship between returns and the stock rating variable (e.g., growth, value, and momentum). The first driver is obvious: Selecting stocks in sectors with higher average returns produces a higher return than selecting stocks in sectors with lower average returns relative to a common benchmark. Sectors with greater dispersion of returns will also generate higher returns to stockpicking, all else equal. Stockpicking involves selecting the stocks with the highest returns, or phrased differently, selecting stocks with returns further into the positive tail of the return distribution. Sectors with larger dispersion have stocks further into the tails and therefore greater returns to stockpicking ability. Analysts select stocks based on style characteristics, such as growth of earnings per share or P/E; the more closely these measures proxy returns, the more successful the stockpicker is using the selection criteria. Stockpicking skill can be conceived of as part of the relationship between fundamentals and returns.

Each of the three forces underlying the returns to skill has a simple statistical corollary. Mean return of all stocks in the sector is a measure of the overall sector, the standard deviation of returns is a measure of dispersion, and the correlation between returns and style characteristic assesses the usefulness of an investment style. Returns to stockpicking skill increase in all three of these measures. The following section derives an explicit form of the relationship.

A statistical model of the returns to stockpicking skill

The formal model begins by postulating that the analyst observes or generates a style characteristic (V) for each stock in the sector and infers the corresponding return (R). Investing in all stocks with $V \geq v$ is a reasonable investment rule in this context.

The returns from selecting a single stock are the expected value of R given $V \geq v$. Adding a distributional assumption, we can derive a closed form expression for the expectation. Assume that R, V are bivariate normal random variables with mean (μ_R, μ_V), standard deviation (σ_R, σ_V), and correlation $\rho_{R,V}$. The expression for the expected returns from a single stock is as follows:

$$E(R|V > v) = \mu_R + \sigma_R \rho_{R,V} \lambda\left(\frac{v - \mu_V}{\sigma_V}\right),$$

where $\lambda(a)$ is the inverse Mills ratio formed from the ratio of the normal density, $\phi(\cdot)$, to the normal cumulative distribution, $\Phi(\cdot)$ as in $\lambda(a)=\dfrac{\phi(a)}{(1-\Phi(a))}$.[9]

The simulations in the main text are based on a long-only portfolio of the best 20% of the stocks. Selecting this fraction of the stocks in the sector corresponds to a cutoff threshold of $v=\mu_V+0.84\,\sigma_V$ and an inverse Mills ratio of approximately 1.4. Incorporating this value into our formula yields

$$E(R|V>v)=\mu_R+1.4\,\sigma_R\;\rho_{R,V}\ \text{ for }\ v=\mu_V+0.84\,\sigma_V\,.$$

Empirical evidence

Actual returns and style characteristics are not bivariate normal. However, the above analysis suggests that a linear additive function of μ_R and $\sigma_R\;\rho_{R,V}$ may be a reasonable proxy to returns to skill. Using the primary data sample for value and growth, we compute sample estimates of the mean, standard deviation, and correlation for each of the 11 economic sectors across the 59 quarters from December 31, 1984, through September 30, 1999. We also compute the return to perfect skill, corresponding to an analyst accurately selecting the 20% of the stocks with the highest style characteristic. We then fit linear regressions of the form

$$\text{Return to perfect skill }=\alpha+\beta\,\mu_R+\gamma\,\sigma_R\;\rho_{R,V}+\varepsilon\,.$$

The results summarized in Exhibit 19 validate the use of the approximation. The regression applied to the pooled sample of 11 sectors and 59 quarters explains 93% of the variation in the returns to perfect skill based on growth and 92% of the variation based on value. The parameter estimates are comparable to what is predicted by theory: The intercept estimates are approximately 0, the coefficient estimates on the mean returns are near 1, and the coefficient on $\sigma_R\;\rho_{R,V}$ is in the neighborhood of 1.4. Exhibit 19 also presents the estimates for individual sectors. The variability of the estimates across sectors is due to the smaller number of observations used in each regression as well as the extent to which sectors deviate from the bivariate normal assumption.

[9] This is a standard result from conditional and truncated normal distributions, as discussed on pages 365-368 of *Limited-Dependent and Qualitative Variables in Econometrics* by G.S. Maddala, Cambridge University Press, Cambridge, 1983.

Exhibit 19: Regression analysis of the returns to perfect skill

	Economic Sector	Adjusted R-Square	Intercept (α)	Coefficient on Mean Return (β)	Coefficient on $\rho^*\sigma$ (γ)	Number of Observations
Growth	All Sectors	0.93	0.00	1.10	1.36	649
	Basic Materials	0.94	0.01	1.07	1.29	59
	Consumer Cyclicals	0.98	0.01	1.08	0.98	59
	Consumer Staples	0.94	0.01	1.13	1.11	59
	Health Care	0.92	0.01	1.17	1.20	59
	Energy	0.93	0.00	1.06	1.50	59
	Financials	0.96	0.00	1.06	1.14	59
	Capital Goods	0.97	0.01	1.10	1.25	59
	Technology	0.96	0.01	1.14	1.35	59
	Communication Services	0.81	0.01	1.14	1.22	59
	Utilities	0.90	0.01	0.92	1.35	59
	Transportation	0.89	0.00	1.11	1.45	59
Value	All Sectors	0.92	0.01	1.06	1.10	649
	Basic Materials	0.96	0.02	1.13	0.89	59
	Consumer Cyclicals	0.96	0.04	1.13	0.58	59
	Consumer Staples	0.95	0.02	1.00	1.14	59
	Health Care	0.89	0.02	1.03	1.21	59
	Energy	0.90	0.01	1.00	1.30	59
	Financials	0.98	0.01	1.11	0.96	59
	Capital Goods	0.96	0.02	1.04	0.95	59
	Technology	0.97	0.00	1.00	1.36	59
	Communication Services	0.79	0.02	1.02	0.65	59
	Utilities	0.96	0.00	1.06	1.02	59
	Transportation	0.88	0.00	1.18	1.39	59

Source: Goldman Sachs Research.

Appendix F: Results for the largest 500 stocks

The broad patterns of the results observed in the primary sample in the main body of the paper also hold in the large-cap stocks. In Exhibits 20 through 24, we repeat the primary results for a sample consisting of only the largest 500 stocks in each quarter.

As we have noted elsewhere, stockpicking is generally somewhat less effective in the very large-cap stocks than in smaller-cap stocks. Thus, we calibrate these simulations to produce average excess return of 3% without sector controls.

Exhibit 20: Effect of sector controls on growth and value—largest 500 stocks

Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Growth			
Without Sector Control	3.0	2.5	1.20
With Sector Control	2.5	2.3	1.07
Difference	**-0.5**	**-0.2**	**-0.13**
Value			
Without Sector Control	3.0	2.6	1.15
With Sector Control	3.3	2.3	1.40
Difference	**0.3**	**-0.3**	**0.25**

Source: Goldman Sachs Research.

Exhibit 21: Sharpe ratio decomposition—largest 500 stocks

Investment Style	Sharpe Ratio Without Sector Control	Sharpe Ratio With Sector Control	Change in Sharpe Ratio	Change Due to Excess Return	Change Due to Tracking Error
Growth	1.20	1.07	-0.13	-0.21	0.08
Value	1.15	1.40	0.25	0.10	0.15

Source: Goldman Sachs Research.

Exhibit 22: Optimal groups—largest 500 stocks

	Growth			Value		
Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Without Sector Controls	3.0	2.5	1.20	3.0	2.6	1.15
11 Sectors, Sector Controls	2.5	2.3	1.07	3.3	2.3	1.40
3 Optimal Groups, Sector Controls	3.2	2.5	1.26	5.7	2.5	2.27
5 Optimal Groups, Sector Controls	3.3	2.5	1.30	6.3	2.6	2.41

Source: Goldman Sachs Research.

Exhibit 23: Optimal groups, focused strategies—largest 500 stocks

	Growth			Value		
Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
With Sector Controls						
3 Optimal Groups	3.2	2.5	1.26	5.7	2.5	2.27
Focused - 2 of 3 Highest Excess Return Groups	3.8	3.1	1.24	6.3	3.4	1.86
5 Optimal Groups	3.3	2.5	1.30	6.3	2.6	2.41
Focused - 3 of 5 Highest Excess Return Groups	3.8	2.9	1.28	7.2	3.4	2.10

Source: Goldman Sachs Research.

Exhibit 24: Out-of-sample group stability—largest 500 stocks

	Growth			Value		
Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Without Sector Control	3.0	2.5	1.20	3.0	2.6	1.15
With Sector Control						
11 Compustat Sectors	2.5	2.3	1.07	3.3	2.3	1.40
3 Optimal Groups	3.2	2.5	1.26	5.7	2.5	2.27
Focused - 2 of 3 Optimal Groups with Highest Excess Returns	3.8	3.1	1.24	6.3	3.4	1.86
5 Optimal Groups	3.3	2.5	1.30	6.3	2.6	2.41
Focused - 3 of 5 Optimal Groups with Highest Excess Returns	3.8	2.9	1.28	7.2	3.4	2.10

create optimal groups from 12/31/1984 - 3/30/1990
run portfolios from 3/30/1990 - 9/30/1999

Source: Goldman Sachs Research.

Appendix G: Results for a more concentrated portfolio

The broad patterns of the results observed with portfolios of 20% of the stocks in the primary sample in the main body of the paper also hold for more concentrated portfolios of 4% of the sample. In Exhibits 25 through 29, we repeat the primary results for the more concentrated portfolios.

The 20% portfolios in the main body of the paper range from roughly 180 to 370 stocks, depending on how many stocks are in the sample on any given quarter. Many portfolio managers hold more concentrated portfolios. Thus, we examine results for 4% portfolios, which range from roughly 33 to 70 stocks. Note that we have renormalized skill so that these more concentrated portfolios produce average excess returns of 5% without sector controls.

Exhibit 25: Effect of sector controls on growth and value—concentrated portfolios

Investment Style	Average Return to Skill (%)	Average Tracking Error (%)	Sharpe Ratio
Growth			
Without Sector Control	5.0	4.5	1.11
With Sector Control	4.5	4.4	1.01
Difference	**-0.5**	**0.0**	**-0.10**
Value			
Without Sector Control	5.0	4.4	1.13
With Sector Control	5.4	4.4	1.23
Difference	**0.4**	**0.0**	**0.10**

Source: Goldman Sachs Research.

Exhibit 26: Sharpe ratio decomposition—concentrated portfolios

Investment Style	Sharpe Ratio Without Sector Control	Sharpe Ratio With Sector Control	Change in Sharpe Ratio	Change Due to Excess Return	Change Due to Tracking Error
Growth	1.11	1.01	-0.10	-0.11	0.01
Value	1.13	1.23	0.10	0.08	0.01

Source: Goldman Sachs Research.

Exhibit 27: Optimal groups—concentrated portfolios

Investment Style	Growth: Average Return to Skill (%)	Growth: Average Tracking Error (%)	Growth: Sharpe Ratio	Value: Average Return to Skill (%)	Value: Average Tracking Error (%)	Value: Sharpe Ratio
Without Sector Controls	5.0	4.5	1.11	5.0	4.4	1.13
11 Sectors, Sector Controls	4.5	4.4	1.01	5.4	4.4	1.23
3 Optimal Groups, Sector Controls	5.3	4.6	1.16	8.1	4.4	1.84
5 Optimal Groups, Sector Controls	5.5	4.6	1.19	8.9	4.5	1.97

Source: Goldman Sachs Research.

Exhibit 28: Optimal groups, focused strategies—concentrated portfolios

Investment Style	Growth: Average Return to Skill (%)	Growth: Average Tracking Error (%)	Growth: Sharpe Ratio	Value: Average Return to Skill (%)	Value: Average Tracking Error (%)	Value: Sharpe Ratio
With Sector Controls						
3 Optimal Groups	5.3	4.6	1.16	8.1	4.4	1.84
Focused - 2 of 3 Highest Excess Return Groups	6.2	5.3	1.16	8.2	4.4	1.86
5 Optimal Groups	5.5	4.6	1.19	8.9	4.5	1.97
Focused - 3 of 5 Highest Excess Return Groups	6.3	5.2	1.22	9.9	4.9	2.01

Source: Goldman Sachs Research.

Exhibit 29: Out-of-sample group stability—concentrated portfolios

Investment Style	Growth: Average Return to Skill (%)	Growth: Average Tracking Error (%)	Growth: Sharpe Ratio	Value: Average Return to Skill (%)	Value: Average Tracking Error (%)	Value: Sharpe Ratio
Without Sector Control	5.0	4.5	1.11	5.0	4.4	1.13
With Sector Control						
11 Compustat Sectors	4.5	4.4	1.01	5.4	4.4	1.23
3 Optimal Groups	5.3	4.6	1.16	8.1	4.4	1.84
Focused - 2 of 3 Optimal Groups with Highest Excess Returns	6.2	5.3	1.16	8.2	4.4	1.86
5 Optimal Groups	5.5	4.6	1.19	8.9	4.5	1.97
Focused - 3 of 5 Optimal Groups with Highest Excess Returns	6.3	5.2	1.22	9.9	4.9	2.01

create optimal groups from 12/31/1984 - 3/30/1990
run portfolios from 3/30/1990 - 9/30/1999

Source: Goldman Sachs Research.

Appendix H: Optimal grouping methodology

This appendix describes a method to construct optimized groups of stocks.

Optimally formed groups of stocks are the basis for a portion of the analysis; this appendix describes the procedure used to construct the optimal groups. By optimal we mean groupings of stocks specifically designed to provide the highest returns to skill. Appendix E developed a relationship between stock selection returns based on the returns of the entire sector, the dispersion of returns within the sector, and the strength of the relationship between returns and fundamentals (e.g., growth, value, and momentum). In constructing optimal clusters, we ignore the first component, which is largely beyond the control of the portfolio manager, and concentrate on the latter two.

It is easiest to first consider forming optimal groups based on a single quarter of data. The objective is to allocate each stock into a group such that we maximize an objective function of the form

$$\text{One quarter objective function} = \frac{\sum_{g=1}^{g=G} N_g \rho_g \sigma_g}{\sum_{g=1}^{g=G} N_g},$$

where G is the total number of groups to be formed,

g subscripts individual groups,

N_g is the number of stocks assigned to group g,

σ_g is the standard deviation of returns within group g, and

ρ_g is the correlation between returns and fundamentals within group g.

The objective function is a count-weighted estimate of the returns to stockpicking. The objective function is easily extended to multiple quarters.

$$\text{Multi-quarter objective function} = \sum_{q=1}^{q=59} \left[\frac{\sum_{g=1}^{g=G} N_{q,g} \rho_{q,g} \sigma_{q,g}}{\sum_{g=1}^{g=G} N_{q,g}} \right],$$

where q indexes quarters of data.

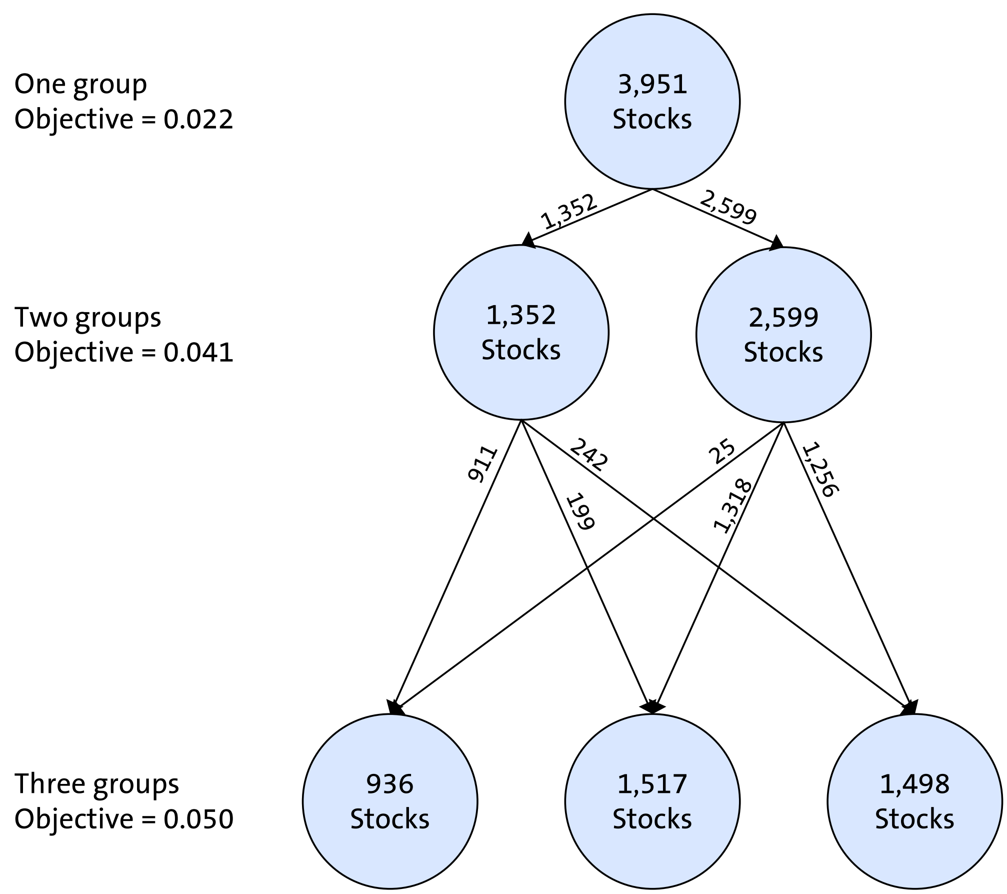
Each stock is assigned to a single group and is a member of that group for all quarters in which it is active.

Algorithm for forming optimal clusters

Stocks are assigned to groups using an iterative process beginning with all stocks in a single group and attempting to form two groups. The algorithm examines each stock sequentially and considers moving it into the other group, which starts out empty. A stock is moved only if the move produces an increase in the multi-quarter objective function. It is possible that a stock is moved from group A to group B and then returned to group A on a subsequent pass through the data. The process is repeated until no further movement of a single stock improves the objective. The one-step-ahead nature of the technique is comparable to that used in CART and CHAID. However, the grouping technique is less parametric as we assign individual observations to the groups without regard for explicit decision rules. The computational burden of the process is reduced by observing that σ_g and ρ_g are straightforwardly characterized as sums of squares and cross products that are easily updated as stocks enter and leave groups.

The number of potential groups is expanded to three once the assignment of stocks into two groups has stabilized. The new group starts out with no members. Each stock is again examined sequentially and placed into the group that maximizes the objective function. The process is repeated until no further movement of a single stock produces an improvement in the objective. Then the number of potential groups expands again. As shown in Exhibit 30, there are no restrictions on the placement of stocks into groups at successive levels.

Exhibit 30: Generating optimal groups for value



Source: Goldman Sachs Research.

Jackknife estimation

The optimal clustering procedure described above is susceptible to overestimating the incremental value of increasing the number of groups. The problem is analogous to a regression model producing a large in-sample R-Squared yet having little out-of-sample predictive ability. A jackknife procedure assesses the degree of over-fitting. To compute the return to skill for quarter q, we construct a set of clusters applying the above procedure to the data omitting that for quarter q. We then compute the return to skill using the data for quarter q and the cluster definitions omitting quarter q. This produces 59 different sector definitions corresponding to the 59 quarters of our sample. By construction, the data used to form the clusters is independent of that used to assess the returns to the skill.

Definitions

RL = Recommended List. Expected to provide price gains of at least 10 percentage points greater than the market over the next 6-18 months.
LL = Latin America Recommended List. Expected to provide price gains at least 10 percentage points greater than the Latin America MSCI Index over the next 6-18 months.
TB = Trading Buy. Expected to provide price gains of at least 20 percentage points sometime in the next 6-9 months.
MO = Market Outperformer. Expected to provide price gains of at least 5-10 percentage points greater than the market over the next 6-18 months.
MP = Market Performer. Expected to provide price gains similar to the market over the next 6-18 months.
MU = Market Underperformer. Expected to provide price gains of at least 5 percentage points less than the market over the next 6-18 months.

Legal/policy ratings and other abbreviations
Definitions

NR = Not Rated. The investment rating and target price, if any, have been suspended temporarily. Such suspension is in compliance with applicable regulation(s) and/or Goldman Sachs policies in circumstances when Goldman Sachs is acting in an advisory capacity in a merger or strategic transaction involving this company and in certain other circumstances.
NC = Not Covered. This company is not covered by Goldman Sachs.
CS = Coverage Suspended. Goldman Sachs has suspended coverage of this company.

NA = Not Available. The information is not available for display.
NM = Not Meaningful. The information is not meaningful and is therefore excluded.

Global Portfolio Analysis: Research highlights

<table>
<tr>
<td>

Beating Benchmarks
A Stockpicker's Reality: Part II,
November 1999

- **Cap-weighted benchmarks contain large risk positions**
- **For a stockpicker's skill to drive the portfolio's outperformance, the PM must either**
  - Reduce the effective risk in the benchmark, which can be done by benchmark weighting the largest stocks in the benchmark, or
  - Take more active risk than the benchmark, which can be done by concentrating the portfolio
- **Portfolio diversification is very fragile with respect to the portfolio weights**
  - Significant deviation from equal-weighing rapidly reduces the diversification value of a position
  - Positions more than twice the average weight actually add back stock-specific risk

</td>
<td>

Style, Size and Skill
A Stockpicker's Reality: Part I,
January 1999

- Different investment styles exploit different types of market inefficiencies and require different types of insights and holding periods
- **Large-cap growth managers** should focus on broad thematic investment insights and have relatively long holding periods
- **Smaller-cap growth managers** should focus on short-term earnings catalysts and have relatively short holding periods
- **Value managers** should focus on long-term earnings but have an aggressive trading attitude toward individual names to maximize results, regardless of capitalization
- On a pure return basis, growth at a reasonable price strategies generally outperform pure style strategies. However, on a risk- adjusted basis, mixes of value and growth and hybrid strategies outperform any single style strategy

</td>
</tr>
<tr>
<td>

Making the Most of Value and Growth Investing,
January 1998

- Value and growth should not be looked at as opposing investing strategies that invest in very different types of stocks but as distinct and potentially complementary ways of evaluating stocks
- **The bulk of the outperformance for both growth and value strategies arises from a common group of firms that are both high-growth and high-value (low P/E)**—without these high-growth/high-value firms, both growth and value strategies tend to underperform
- Avoiding the "bad" stocks can be more important for generating outperformance than finding "good" stocks, especially in the large-cap universe

</td>
<td>

Global Equity Portfolios and the Business Cycle,
April 1997

- **In periods of low global capacity utilization (typically early in the business cycle)**
  - Equity returns are higher than average, and volatility and cross-country correlations are lower than average
  - There is an advantage to active country-selection strategies
- **In periods of high global capacity utilization (typically late in the business cycle)**
  - Equity returns are lower than average, and volatility and cross-country correlations are higher than average
  - There is an advantage to macro-based total market exposure strategies
- Asset allocation strategies based on stable long-run returns and correlations understate the value of global diversification early in the business cycle and overstate the value later

</td>
</tr>
</table>