Appendix to: "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation"

Appendix A: Attribution Analysis

Is there an upper bound on how much of the index alphas we can hope to explain with factor models based on size- and value-sorted portfolios? To answer this question, we can decompose index alphas into two sources: exposure to passive size- and value-sorted portfolios and stock selection within these portfolios. The decomposition between the two sources of alpha tells us whether the index stocks have different returns relative to other stocks with similar characteristics – for example, whether S&P tends to select higher-alpha stocks for its indices. The stock selection alpha is unlikely to be explained with any size and value factor model, but the rest of the alpha in principle could be explained as it comes from passive and broad-based portfolios of stocks.

Suppose that the assets held by a portfolio *p* are each members of exactly one of *J* benchmark portfolios. The alpha of benchmark portfolio *j* is α_{jt}^{b}. We denote the share of portfolio *p* accounted for by members of benchmark portfolio *j* in month *t* as w_{jt}, and their weighted average alpha as α_{jt}^{a}. Given this setup, the alpha of portfolio *p* can be decomposed into alphas due to the *J* benchmark portfolios, and alphas due to the relative performance $\alpha_{jt}^{a} - \alpha_{jt}^{b}$ of the members of each benchmark portfolio *j* held by *p*. We can call the latter "selection alphas," as they arise from stock selection within a benchmark portfolio, and the former "style alphas" due to asset allocation across benchmark portfolios:[1]

$$\alpha_{pt} = \sum_{j=1}^{J} \underbrace{w_{jt}\alpha_{jt}^{a}}_{\text{active alpha}} = \sum_{j=1}^{J} \underbrace{w_{jt}\left(\alpha_{jt}^{a} - \alpha_{jt}^{b}\right)}_{\text{selection alpha}} + \sum_{j=1}^{J} \underbrace{w_{jt}\alpha_{jt}^{b}}_{\text{style alpha}}. \tag{1}$$

[1] We could also further decompose style alpha to timing alpha and average style alpha, similarly in spirit to Daniel, Grinblatt, Titman, and Wermers (1997). Each month, we compute the return from investing in the 10x12 benchmarks using the actual S&P 500 weights that month, as well as the return from investing in the 10x12 benchmarks using the average S&P 500 weights over all months in the sample period; the difference is our measure of timing alpha that month. In unreported results, we found that timing has a slightly negative contribution to the returns on both the S&P 500 and Russell 2000.

In Panel A of Table A-1, we show for the S&P 500 the total alpha contribution ("active alpha" in equation (1)) coming from S&P 500 stocks within each benchmark portfolio, as well as the selection alpha of the index stocks relative to other stocks in the corresponding benchmark portfolios. As benchmarks for this attribution analysis, we pick the 10x10 Fama-French portfolios which are also the basis for creating the common Fama-French factors. To cover the full universe of the CRSP market index, we again add 10x2 portfolios to include the remaining U.S. stocks as well as the securities with other share codes, as discussed earlier. The attribution analysis is possible since we know the holdings of the S&P 500 index; the holdings of the Fama-French portfolios are not available but we construct them ourselves using the publicly available decile cutoffs. All numbers in the table are time-series averages across the sample period.

The four-factor alpha of the S&P 500 comes overwhelmingly from the top market cap decile, which accounts for 73 bp out of the 81 bp alpha of the index. This is mostly due to the two extreme growth portfolios within the top size decile which have large positive four-factor Carhart alphas of 371 bp and 296 bp per year and which contain about 35% of the value of the S&P 500 index.[2] The second part of Panel A indicates that stock selection by the S&P 500 within the 10x12 benchmark portfolios accounts for only 11 bp of its alpha. Hence, almost 90% of the S&P 500 alpha comes from its exposure to passive benchmark portfolios and not from any well-informed stock selection by the S&P index selection committee.[3]

The Russell 2000 (Panel B) exhibits some negative "stock selection" which amounts to 69 bp per year. However, about 70% of the Russell 2000 negative alpha, 169 bp out of 238 bp per year, still comes simply from its exposure to Fama-French portfolios and could potentially be explained by a factor model. The remaining stock selection alpha comes almost entirely from the upper and lower boundaries of the index (size deciles 2 and 5-6, while deciles 3-4 show very little selection alpha.

[2] This analysis is based on the holdings of Fama-French portfolios and benchmark indices. Because we do not perfectly replicate the 10x10 Fama-French component portfolios, some small discrepancies arise when compared to the 10x10 portfolio returns from Ken French's web site. Nevertheless, the match is economically very close and does not seem to affect our results. The index alphas in this analysis also differ from the official results by a 1-3 bp per year because the attribution analysis requires that we compute index returns from month-end holdings.

[3] The alpha contributions of individual cells do not add up exactly to the marginal portfolio alphas because each cell alpha is estimated separately, and due to time-variation in weights across cells this is not the same as estimating the value-weighted alpha of the marginal portfolio (without time-variation in weights, the numbers would add up exactly). Because portfolio weights across the 100 Fama-French portfolios are more stable across size than across value deciles, the alpha contributions add up better across size deciles.

Appendix B: Robustness of Beta Estimation

When estimating betas to compute the tracking error of a fund in Section 5.1, it is not obvious what the time horizon or the sampling frequency should be. We try four different methods: monthly data over five or three years, and daily data over twelve or six months. Monthly data is convenient to use, but it requires a longer history of returns and it may mismeasure betas if they vary over time. Daily data allows for a large number of data points while keeping the beta estimates current, but it may introduce problems due to stale prices for some stocks. Panels A and C in Table 9 show the average out-of-sample tracking errors across the four estimation methods (Panel A shows the baseline results with 12 months of daily data). The main conclusion from the results is that daily data produces superior beta estimates to monthly data, as evidenced by lower out-of-sample tracking errors.

With monthly data, even a simple benchmark-adjustment performs as well out-of-sample as the Fama-French and Carhart models on excess returns. The four-factor index model performs best, while adding more factors slightly increases out-of-sample tracking error. Whether we use three or five years of data does not matter much for models with only a few factors, but models with at least five factors are clearly better estimated from a longer dataset.

With daily data, it does not matter whether we use six or twelve months of data. In general, the twelve-month estimates perform slightly better, except for the CAPM where we need to estimate only one parameter. Tracking error improves monotonically as we add new factors, at least up to the seven-factor model.

Because daily beta estimates perform so much better out-of-sample than monthly beta estimates, it appears that any staleness in prices does not interfere much with beta estimation. Stale prices would undoubtedly be more important for individual stocks, but mutual funds hold broad portfolios of stocks, so the average staleness in fund return is likely to be close to the average staleness in benchmark index return. Nevertheless, we investigated daily beta estimates further to see whether including leads and lags would lead to even lower out-of-sample tracking errors; we find that it does not.[4]

[4] Results available upon request.

Appendix C: Correlations of Factors

Table A2 presents correlations between the original Fama-French/Carhart factors, our modifications thereof, and our suggested alternative index-based factors.

The market factor with only U.S. stocks has an almost perfect correlation of 99.9% with the CRSP market index, in spite of the 23 bp difference in average return.5 The modified SMB factor with value weights and all U.S. stocks also has a very high correlation of 97.5% with the original SMB factor in spite of the considerable differences in portfolio weights; however, the modifications reduce its correlation with HML from -43.2% to -29.6%, which is desirable when the factors are used together in a model.

When SMB is split into Mid-minus-Big (MMB) and Small-minus-Mid (SMM), the two new factors have a correlation of 56.4%. Splitting HML into BHML (deciles 9-10), MidHML (deciles 6-8), and SHML (deciles 1-5) also produces correlated factors, with correlations ranging from 69.8% between SHML and BHML to as high as 89.4% between SHML and MidHML. To keep them as comparable as possible to the Fama-French factors, we maintain the long-short structure for our index-based factor portfolios. Hence, our index-based version of the Carhart model includes the S&P 500 as the market, Russell 2000 minus S&P 500 as the small-minus-big factor, Russell 3000 Value minus Russell 3000 Growth as the value factor, and the usual momentum factor. Our more comprehensive seven-factor model splits R2-S5 into R2-RM and RM-S5, thus distinguishing between small and midcaps, and divides the value factor into S5V-S5G, RMV-RMG, and R2V-R2G (large, mid, and small-cap value factors), while keeping the same momentum factor. The momentum factor has little impact on our results, but we include it as it has become relatively standard in the academic literature.

The index-based models have generally similar correlations to the modified Fama-French factors. The main exception is our value factor based on Russell 3000 (R3V-R3G), which uses value weights rather than the 50-50 weights between small and large stocks in the original HML factor. The original HML factor is more of a small-cap value factor, as its correlation is 89.7% with R2V-R2G and only 72.5% with S5V-S5G. In contrast, R3V-R3G has a similar correlation with all three value factors.

Table A-1. Attribution analysis of benchmark indices.

Panel A shows how the Carhart alpha of the S&P 500 index arises from the contributions of index stocks in 100 Fama-French portfolios selected by market capitalization and book-to-market ratio, as well as size portfolios for U.S. stocks with insufficient BM data ("None") and for other CRSP securities ("Other"). For each cell, the Carhart betas and monthly alphas of index stocks are computed, then monthly alphas are multiplied by the monthly weight of the index in that cell, and finally the monthly alpha contributions are added up across all months from 1980 to 2005. The alpha contribution of index stocks is also shown relative to all stocks in each cell, using the same weights on the 120 component portfolios as the S&P 500. Panel B repeats the analysis for the Russell 2000. All numbers are in basis points per year.

Panel A: S&P 500

Contribution to alpha

	Other	None	Growth	2	3	4	5	6	7	8	9	Value	All
Large	-5	6	74	45	-11	-12	-13	-13	-6	-11	-2	-3	73
9	0	1	1	4	-2	-2	2	-2	-1	-3	-1	1	-2
8	0	2	4	1	-1	-1	2	-3	0	1	1	-1	7
7	0	0	0	0	-1	0	0	1	-1	-1	1	-1	-3
6	0	0	-1	1	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	-1
4	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0
Small	0	0	0	0	0	0	0	0	0	0	0	0	0
All	-5	7	79	49	-14	-14	-12	-14	-5	-21	-5	-2	81

Alpha relative to Fama-French benchmark

	Other	None	Growth	2	3	4	5	6	7	8	9	Value	All
Large	3	2	-3	2	0	1	1	-2	-8	2	0	0	3
9	0	1	-3	1	0	0	0	0	-1	1	0	1	0
8	0	2	2	1	0	0	2	-1	1	2	0	0	10
7	0	0	0	-1	0	1	0	1	0	0	0	0	-1
6	0	0	-1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0
Small	0	0	0	0	0	0	0	0	0	0	0	0	0
All	3	3	-4	4	0	1	4	-2	-4	1	1	0	11

Panel B: Russell 2000

Contribution to alpha

	Other	None	Growth	2	3	4	5	6	7	8	9	Value	All
Large	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	-1	0	0	0	0	0	0	0	0	0	-3
7	0	-3	6	-1	-1	0	-1	-1	1	0	0	0	2
6	-2	-5	-12	-7	0	-3	-5	-1	-6	-2	-1	-1	-60
5	0	-1	-14	-2	-2	-1	-2	-2	0	-3	4	1	-29
4	-4	-1	-19	0	-7	-1	1	5	1	6	1	-4	-33
3	-3	-8	-16	1	2	-5	4	3	0	2	-2	-1	-29
2	-1	-7	-15	-6	-7	-2	1	-2	3	-2	1	-1	-50
Small	-1	-1	-13	1	-2	3	2	0	0	0	1	-1	-29
All	-14	-35	-93	-15	-20	-14	-5	-3	-3	-1	5	-9	-238

Alpha relative to Fama-French benchmark

	Other	None	Growth	2	3	4	5	6	7	8	9	Value	All
Large	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	-2	0	0	0	0	0	0	0	0	0	-4
7	0	-2	0	-1	-1	1	-1	-1	0	0	0	0	-2
6	0	-2	-5	-9	3	-1	-1	-1	-2	-1	-2	-2	-41
5	2	-6	-6	-1	-2	-3	0	-3	0	-1	4	2	-12
4	2	-3	-1	1	0	2	1	-1	0	1	-2	-1	-2
3	2	-8	1	0	0	-1	1	-1	0	-1	-1	-1	-8
2	0	-6	5	-2	-4	-2	0	-2	0	-2	-1	1	-15
Small	0	6	4	3	-1	1	1	0	0	-1	-1	-1	-4
All	4	-32	-7	-6	-3	-3	-1	-10	-4	-3	-2	-3	-69

Table A-2. Correlations across factors.

Panel A reports the time series correlations of the Fama-French factors with our modified versions of those factors. Panel B reports the correlations of the Fama-French factors with factors based on common benchmark indices: the S&P 500 (S5), Russell 2000 (R2), Russell Midcap (RM), and Russell 3000 (R3). The value and growth components of the indices are represented by V and G. For example, "R2-S5" is long Russell 2000 and short S&P 500, while "R2V-R2G" is long Russell 2000 Value and short Russell 2000 Growth. The time period is 2/1986–12/2005.

Panel A: Original FF factors with modified FF factors

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1)	MktRF as in FF											
(2)	SMB as in FF	18.9										
(3)	HML as in FF	-49.0	-43.2									
(4)	MktRF, share codes 10/11	99.9	17.7	-49.2								
(5)	SMB, value weights	15.6	97.5	-29.6	14.3							
(6)	BHML, Size top 5, H top 3	-37.3	-25.3	90.7	-37.8	-8.2						
(7)	SHML, Size btm 5, H top 3	-51.9	-52.3	93.3	-51.8	-43.5	69.6					
(8)	MMB	19.9	85.9	-22.5	18.4	88.8	-0.2	-38.3				
(9)	SMM	7.7	86.4	-29.7	7.0	88.0	-14.5	-38.4	56.4			
(10)	BHML, Size top 2, H top 7	-35.7	-23.8	86.3	-36.4	-7.7	91.8	69.2	3.1	-17.2		
(11)	SHML, Size btm 5, H top 7	-54.8	-55.4	91.8	-54.6	-46.3	68.5	98.3	-39.7	-41.9	69.8	
(12)	MidHML	-39.9	-57.0	90.3	-39.7	-44.5	77.1	88.5	-39.5	-38.6	76.4	89.4

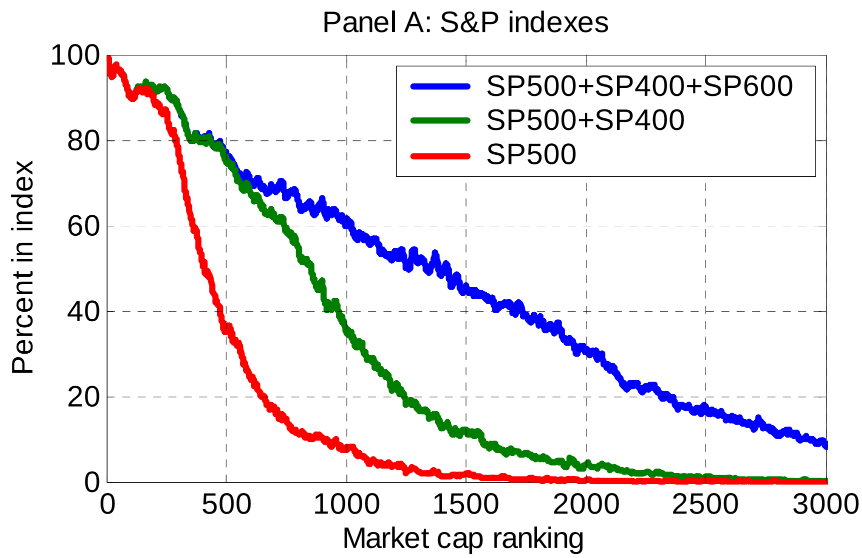
Panel B: Original FF factors with index factors

		(1)	(2)	(3)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1)	MktRF as in FF										
(2)	SMB as in FF	18.9									
(3)	HML as in FF	-49.0	-43.2								
(13)	S5	98.1	2.0	-41.5							
(14)	R2-S5	14.6	93.3	-24.2	-2.4						
(15)	RM-S5	10.7	70.8	-3.9	-4.6	85.7					
(16)	R2-RM	14.7	91.8	-36.0	0.0	90.2	55.1				
(17)	R3V-R3G	-44.2	-36.0	90.5	-37.9	-15.1	5.0	-28.6			
(18)	S5V-S5G	-21.9	-16.1	72.5	-20.0	5.6	23.8	-10.9	84.4		
(19)	RMV-RMG	-48.2	-53.0	89.2	-37.5	-36.4	-20.4	-41.9	89.8	63.3	
(20)	R2V-R2G	-55.6	-53.8	89.7	-44.6	-40.0	-23.7	-45.0	83.2	55.1	92.9

Table A-3. Alphas of benchmark indices until 2010.

This table shows the Carhart four-factor alphas as well as the Fama-French three-factor alphas for common benchmark indices. Alphas are computed from monthly data. The numbers shown are expressed in percent per year, with heteroskedasticity-robust t-statistics in parentheses. The sample period is January 1980 to December 2010, except for the following indices whose return data begin later: S&P 400 (2/1981), Wilshire 4500 (1/1984), S&P 600 (3/1984), and the Growth and Value components of the Russell Midcap (2/1986), S&P 400 (6/1991), and S&P 600 (1/1994). The variance-covariance matrix used in the joint significance test allows for clustering of index returns within time periods.

Main index	Carhart Alpha			Fama-French Alpha		
	Growth	All	Value	Growth	All	Value
Russell 3000	0.64	0.02	-0.54	0.70	-0.03	-0.94
	(1.35)	(0.10)	(0.98)	(1.62)	(0.16)	(1.75)
Russell 1000	1.04	0.27	-0.43	1.08	0.21	-0.84
	(2.00)	(1.35)	(0.80)	(2.28)	(1.11)	(1.55)
Russell Midcap	1.33	0.46	0.09	1.46	0.35	-0.38
	(1.33)	(0.69)	(0.09)	(1.45)	(0.56)	(0.42)
Russell 2000	-2.92	-2.16	-1.23	-2.73	-2.11	-1.65
	(4.05)	(3.44)	(1.37)	(3.73)	(3.51)	(1.93)
S&P 500	1.36	0.65	-0.49	1.81	0.57	-1.33
	(2.43)	(2.21)	(0.86)	(3.14)	(2.05)	(2.24)
S&P Midcap 400	1.70	1.64	1.12	1.29	1.73	1.78
	(0.92)	(1.76)	(0.74)	(0.69)	(1.87)	(1.14)
S&P Smallcap 600	-1.98	-1.98	-0.71	-1.27	-1.91	-1.25
	(1.20)	(2.01)	(0.53)	(0.82)	(2.03)	(0.92)
Wilshire 5000		-0.03			0.01	
		(0.17)			(0.06)	
Wilshire 4500		-0.81			-0.53	
		(1.36)			(0.94)	
P-value of joint test for 26 indices		<.000001			<.000001	



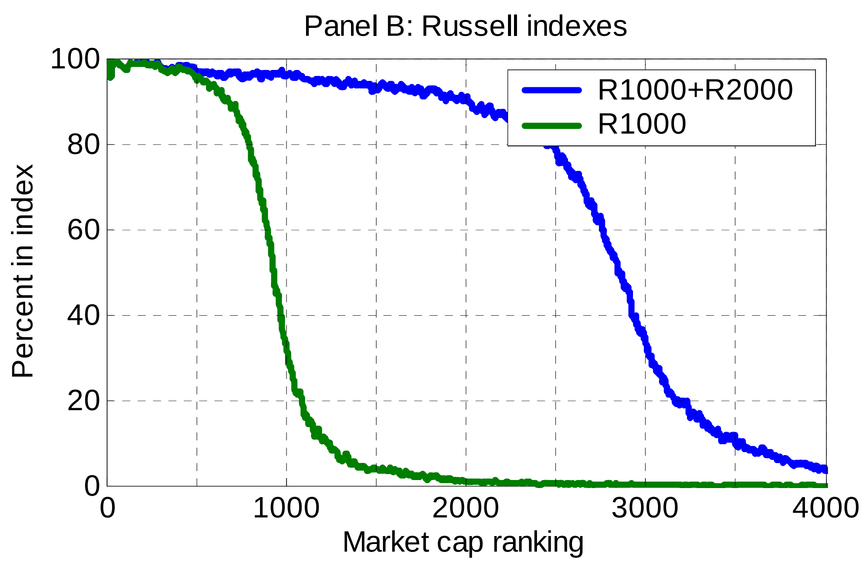
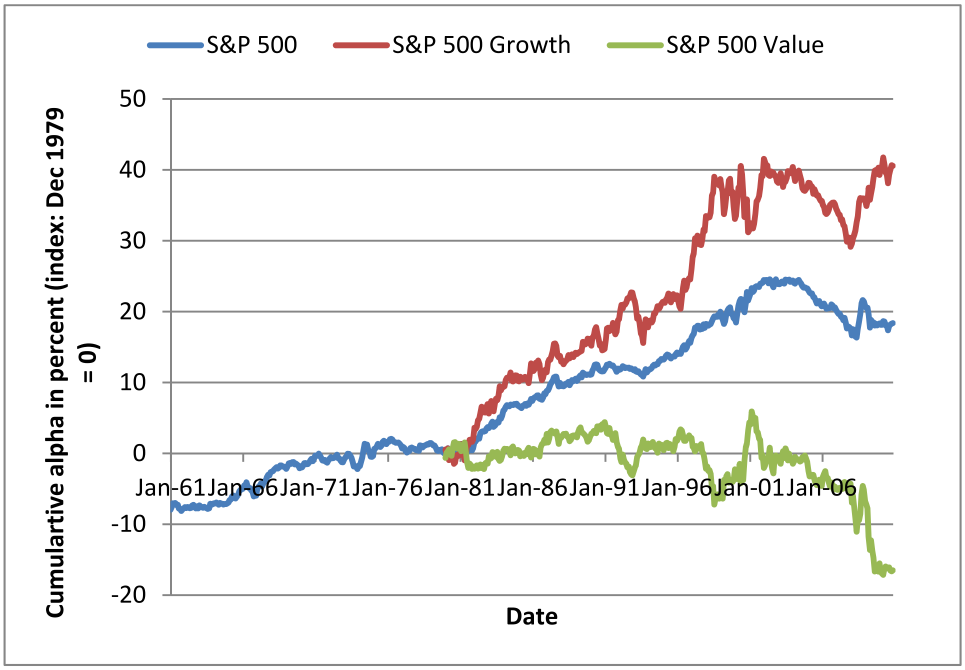


Figure A-1. Index membership as a function of market capitalization.

All U.S. stocks in CRSP are sorted each month based on their market cap. For each market cap rank, we include 10 stocks above and below and then compute the percentage of those 20 stocks that are index constituents that month. The figures show the averages across 120 months from 1996 to 2005.

Panel A. Abnormal returns of S&P 500 and subindices until 2010



Panel B. Abnormal returns of Russell 2000 and subindices until 2010

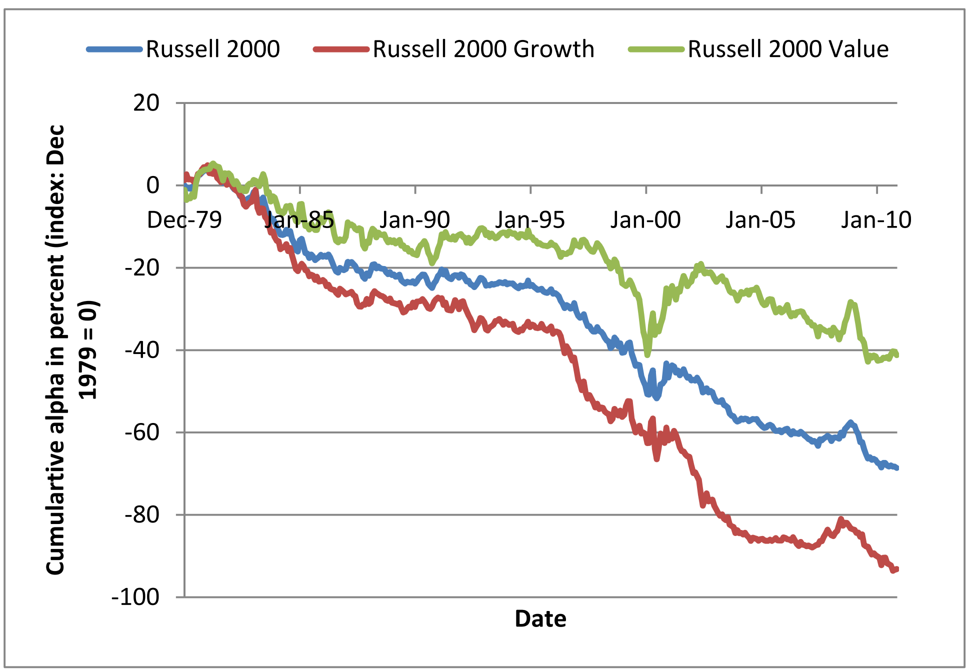
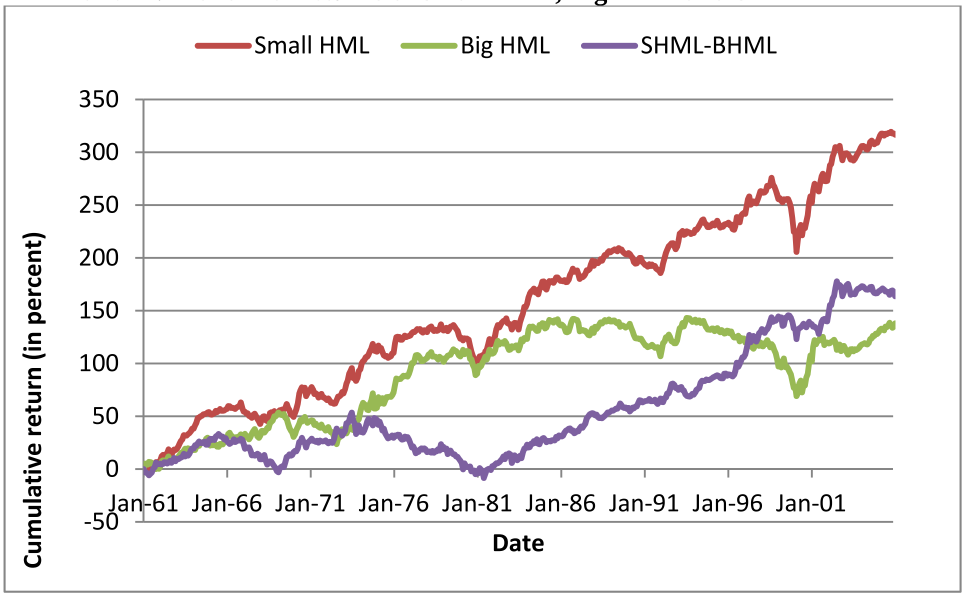


Figure A-2. Abnormal returns of S&P 500 and Russell 2000.

This figure plots the cumulative abnormal alpha from the Carhart model, estimated in-sample from the time period 1980-2010 (except for the S&P 500, which is 1961-2010). Cumulative alphas are expressed in log percentage points, and normalized to zero at the beginning of 1980.

Panel A. Abnormal returns of Small HML, Big HML and SHML-BHML



Panel B. Abnormal returns of the common U.S. stocks in the CRSP VW index

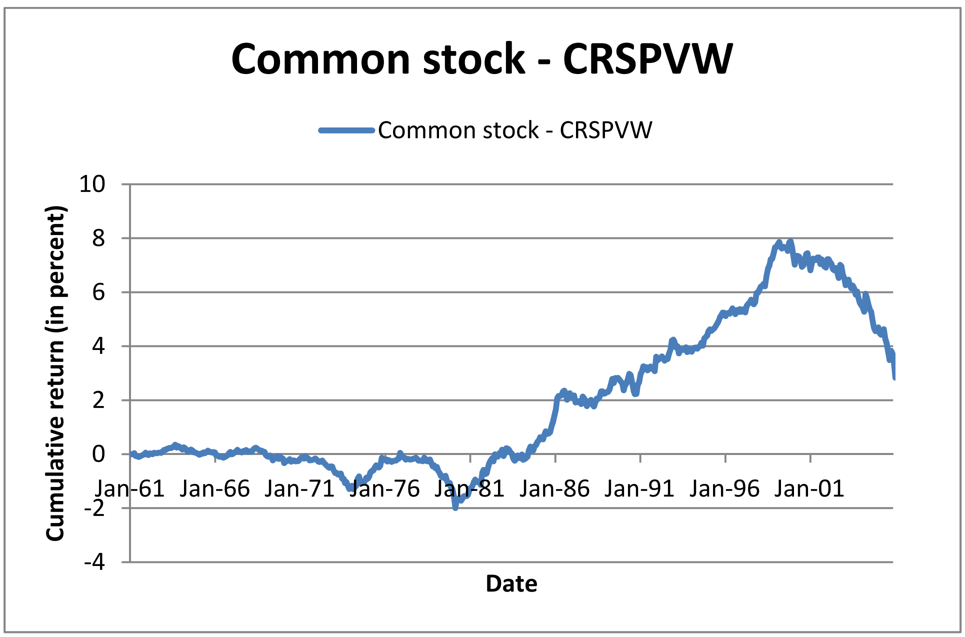


Figure A-3. Abnormal returns of HML components and VW CRSP index.

This figure plots the cumulative abnormal alpha from the Carhart model, estimated in-sample from the time period 1980-2005 in Panel A for the two components of HML: small HML and big HML, and from 1961-2005 in Panel B for the common U.S. stocks in the value-weighted CRSP index. Cumulative alphas are expressed in log percentage points, and normalized to zero at the beginning of the sample.

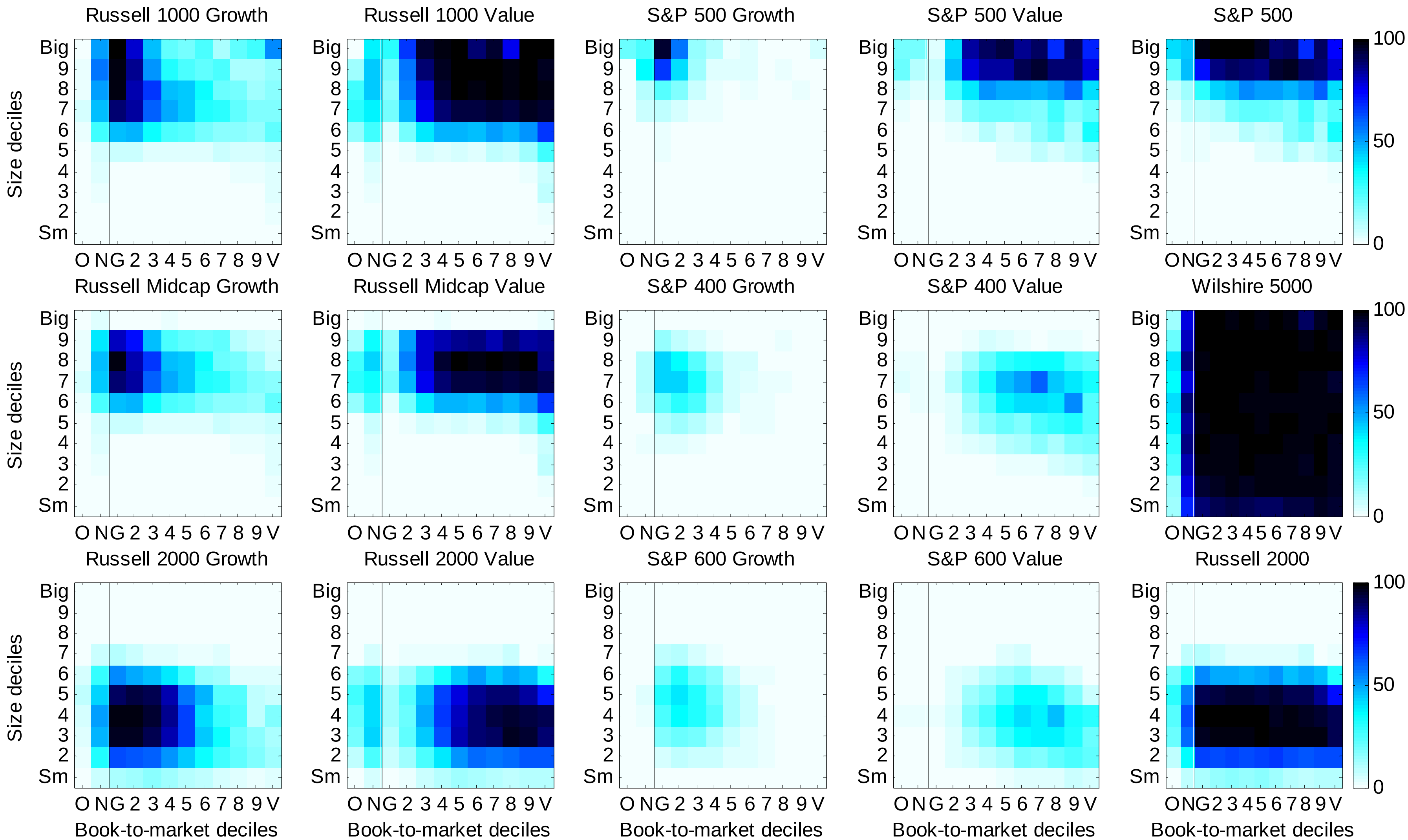


Figure A-4. Index membership across size and value groups.

All securities on CRSP are divided into 10 size groups and one of 12 value groups. For each 10x12 component portfolio, the figures show the fraction of market capitalization that is included in the benchmark index. The component portfolios are determined once a year based on market equity and book-to-market, following the methodology of Fama and French (1993). We also add two new value groups: "N" for those U.S. stocks where the Fama-French inclusion criteria are not satisfied (typically relatively new listings), and "O" for all other stocks. The figures show the mean value from 1997 to 2005, computed across all months. Only ADRs are excluded to mimic the inclusion criteria of the CRSP market index.