Should Benchmark Indices Have Alpha?
Revisiting Performance Evaluation

Martijn Cremers (Yale)

Antti Petajisto (Yale)

Eric Zitzewitz (Dartmouth)

How Would You Evaluate These Funds?

- Regress 3 stock portfolios on the four-factor Fama-French-Carhart model

- 26-year period from 1980 to 2005

- Portfolio A:
  - $b = 0.82\%$ ($t = 2.95$) per year

- Portfolio B:
  - $b = -2.41\%$ ($t = -3.35$) per year

- Portfolio C:
  - $b = 5.24\%$ ($t = 3.97$) per year

Even Passive Indices Appear to Have Skill

- All portfolios are entirely passive:
  - *A*: S&P 500
  - *B*: Russell 2000
  - *C*: S&P 500 Growth – Russell 2000 Growth
- Easily replicated by cheap index funds
- Most common US equity benchmarks
  - … significant biases for performance evaluation?
- Together *A* and *B* cover 85% of US equity market
  - … relevant for portfolios with size or size-value tilt

Implications for Active Trading Strategies

- Fama-French 2x3 size-value grid
- Active long-short portfolio
  - Long positions in Large Growth stocks
  - Short positions in Large Value
- 1980 to 2005:
  - Long side underperformed short side by 1.66% p/y
  - Carhart model: positive alpha of 3.90% p/y
  - Fama-French alpha is 4.33%
- Mechanical application may lead to incorrect inferences about performance

Research Questions

1) How large a problem are these index alphas?

- Effect for all common indices

2) Why do we get non-zero index alphas?

- Problems in factor construction

3) What would be a better factor model?

- Modified factors and index-based factors

Main Results

- Benchmark 'alphas' for Fama-French-Carhart model
  - Large impact on performance evaluation
  - Especially across size-value grid
- Benchmark 'alphas' misleading
  - Primarily from FF factor construction
  - Also: CRSP market index construction and (Russell 2000) index reconstitution
- Alternative size and value factors based on benchmark indices
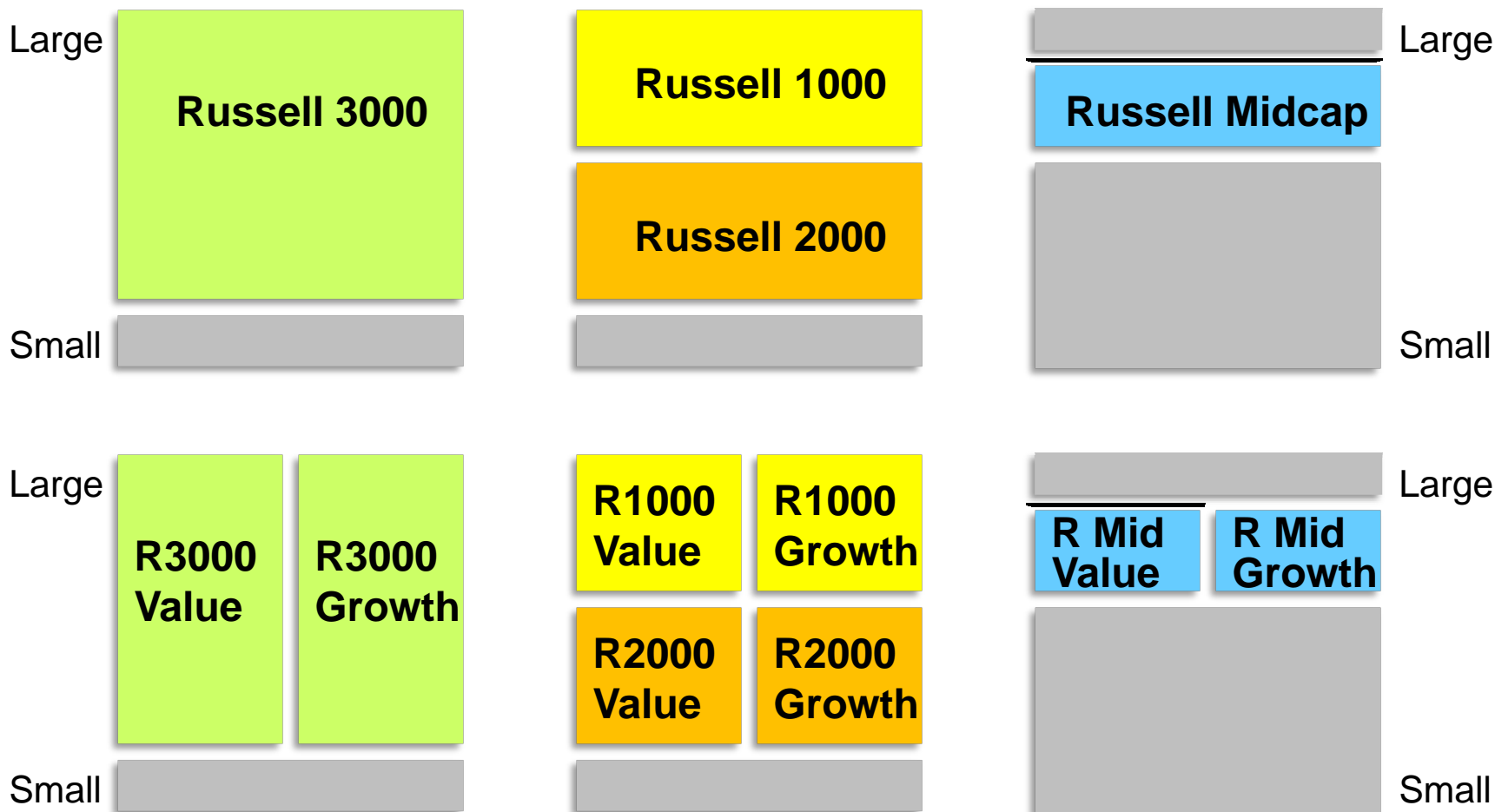 - Improve estimation of fund alphas
  - Improve pricing

Data Sources

- Benchmark index returns
  - Monthly and daily returns directly from index providers: S&P, Russell, DJ Wilshire
- Benchmark index holdings
  - Monthly data directly from index providers
- Mutual fund returns
  - Monthly returns from CRSP
  - Daily returns from CRSP, Yale ICF, and S&P
- Mutual fund holdings
  - Quarterly/semiannually from Thomson Financial
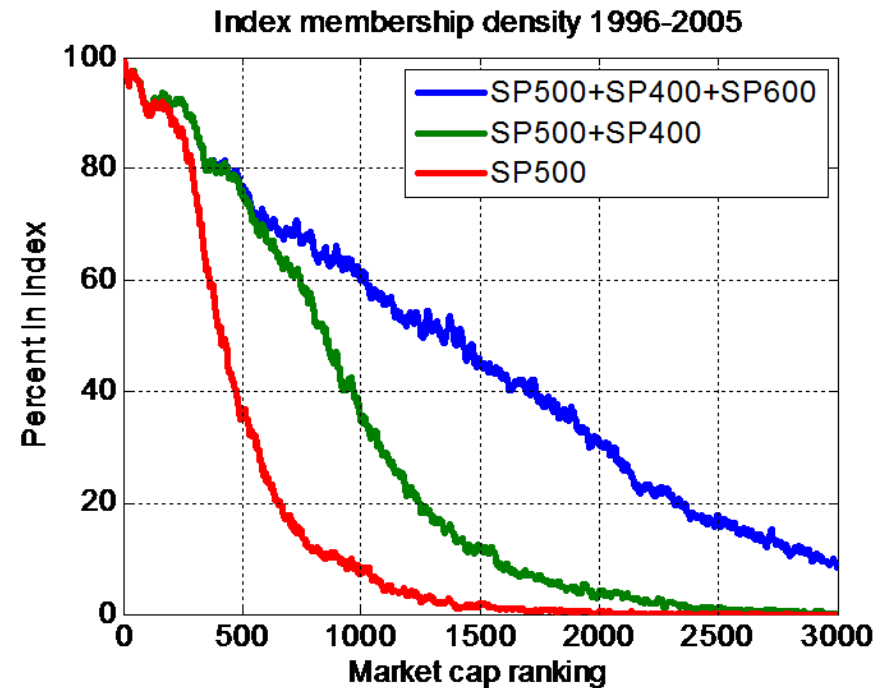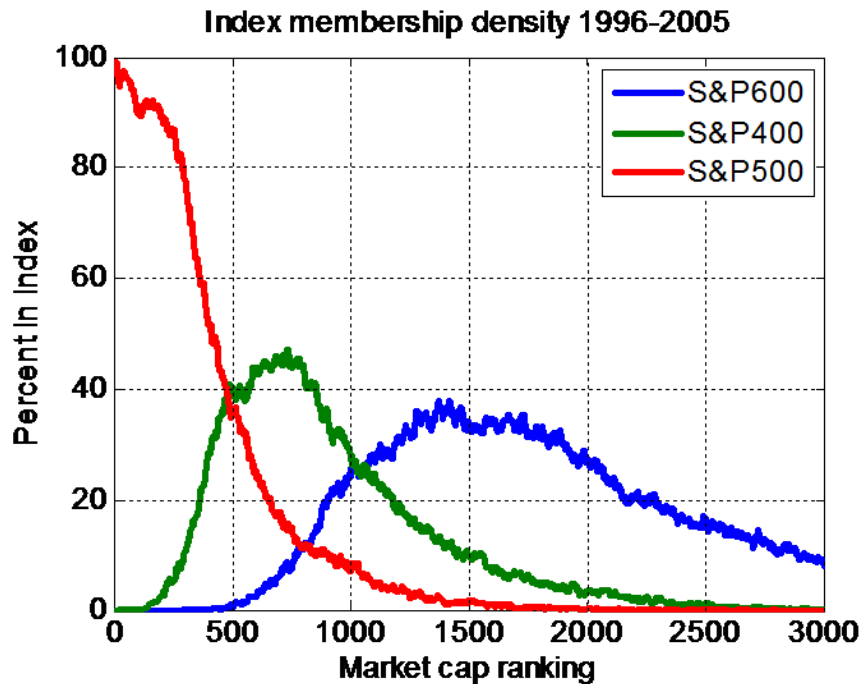 - Stock and firm data from CRSP and Compustat
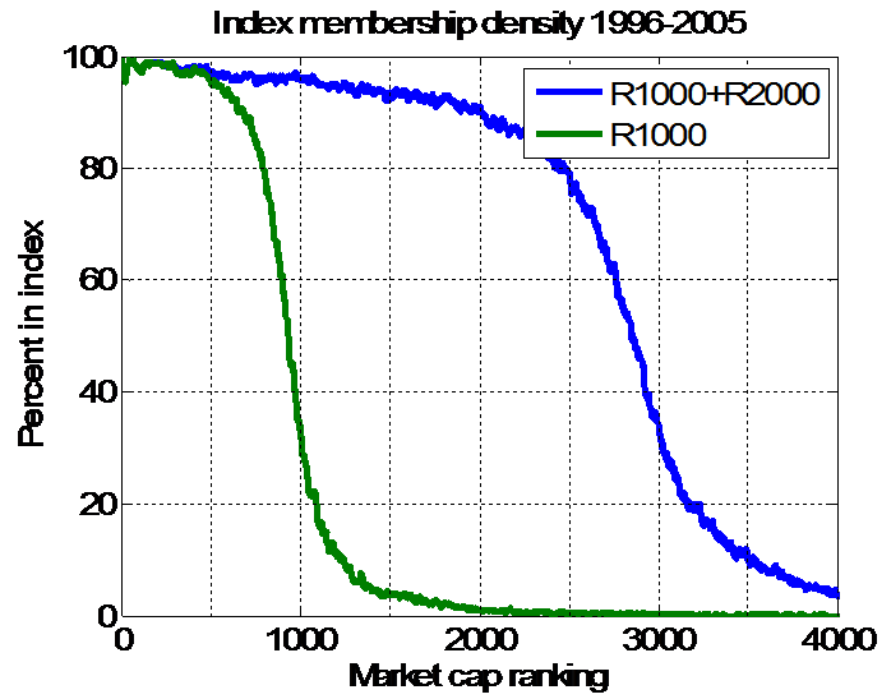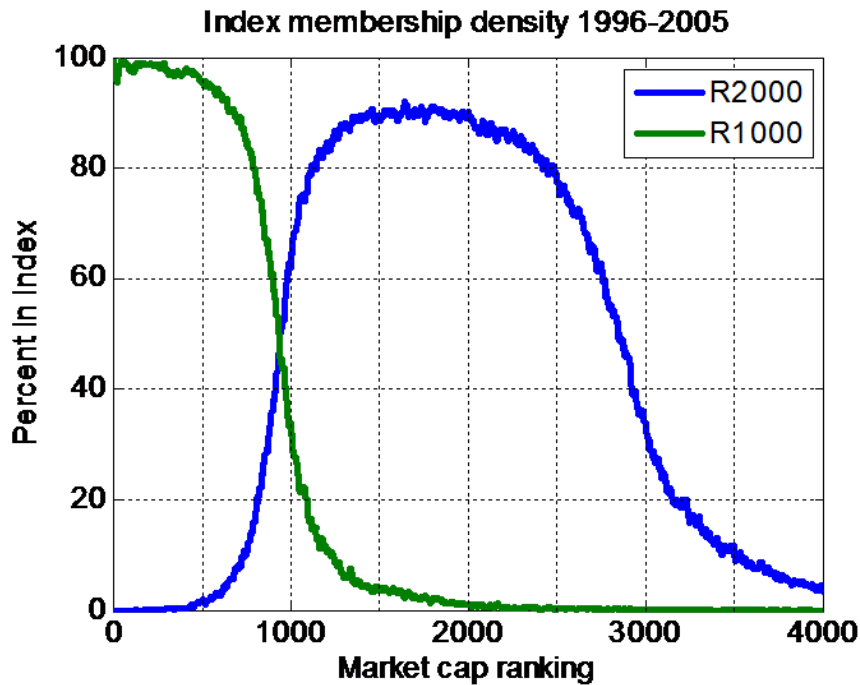
Benchmark Indices: S&P and Wilshire

Benchmark Indices: Russell

Large

Russell 3000

Small

Large

Russell 1000

Russell 2000

Small

Large

Russell Midcap

Small

Large

R3000 Value **R3000 Growth**

Small

Large

R1000 Value **R1000 Growth**

R2000 Value **R2000 Growth**

Small

Large

R Mid Value **R Mid Growth**

Small

S&P Indices Are a Subset of the Market

Index membership density 1996-2005

Russell Indices Include Entire Market

Index membership density 1996-2005

Index membership density 1996-2005

# **Fama-French-Carhart vs. Indices**

Benchmarks implied by FF-C model *versus* indices:

- CRSP-VW includes assets most indices/portfolios do not
- SMB is EW mix of growth and value
  - Indices are VW
  - Small value has outperformed => EW SMB > VW SMB
- SMB excludes stocks with no/negative BtM (e.g., IPOs)
  - Including reduces SMB spread
- FF-C: same value effect for Big and Small stocks
- Boundaries differ
  - FF-C includes Midcaps in Big
 - Indices include both High and Med BtM stocks in Value

Index Alphas 1980-2005

| Main index	Style component		
	Value	All	Growth
Russell 3000	-0.55	0.18	1.02
	(-1.01)	(0.96)	(2.05)
Russell 1000	-0.45	0.47	1.50
	(-0.83)	(2.58)	(2.73)
Russell Midcap	-0.52	0.17	1.61
	(-0.54)	(0.24)	(1.34)
Russell 2000	-1.25	-2.41	-3.41
	(-1.31)	(-3.35)	(-3.87)
S&P 500	-0.35	0.82	1.82
	(-0.69)	(2.95)	(2.76)
S&P Midcap 400	0.84	1.44	0.64
	(0.51)	(1.33)	(0.32)
S&P Smallcap 600	-1.49	-2.59	-3.05
	(-0.89)	(-2.20)	(-1.39)
Wilshire 5000		0.05	
		(0.43)	
Wilshire 4500		-0.56	
		(-0.79)	

Choice of the Market Portfolio:
Not just US Common Stocks in CRSP-VW
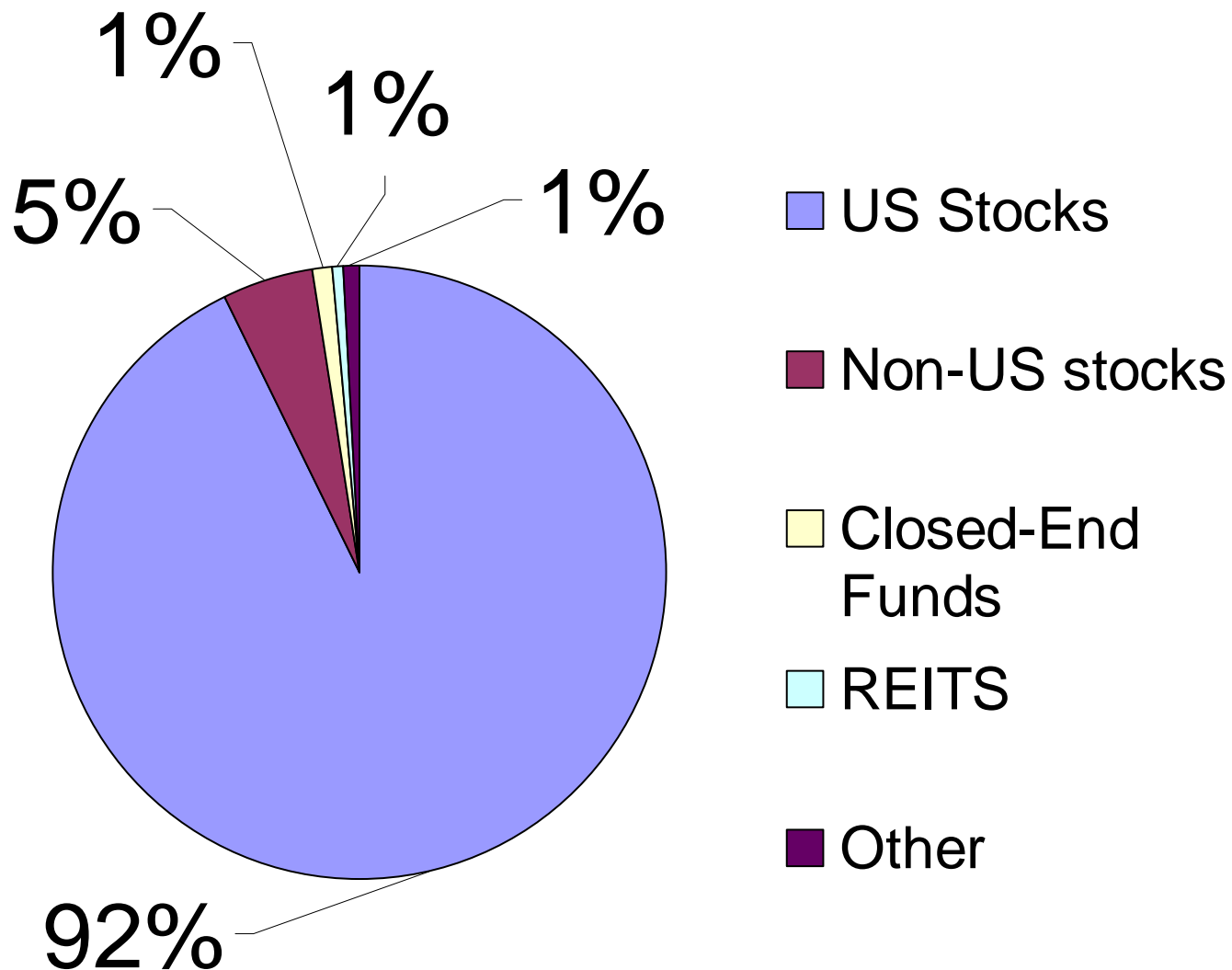
Choice of the Market Portfolio:
FF-C Annualized Alphas for CRSP-VW Categories
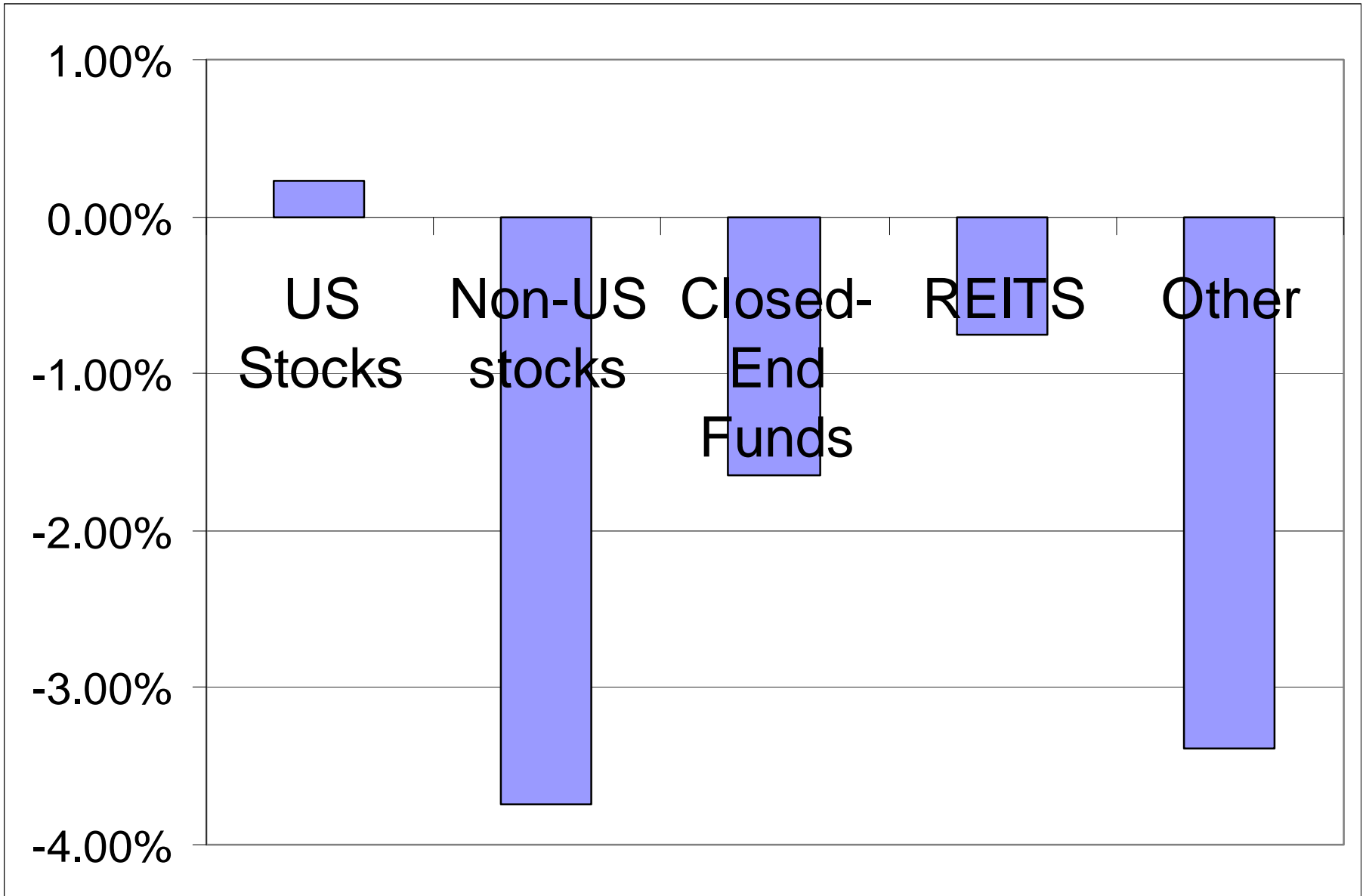
Fama-French Component Portfolios: Market Weights
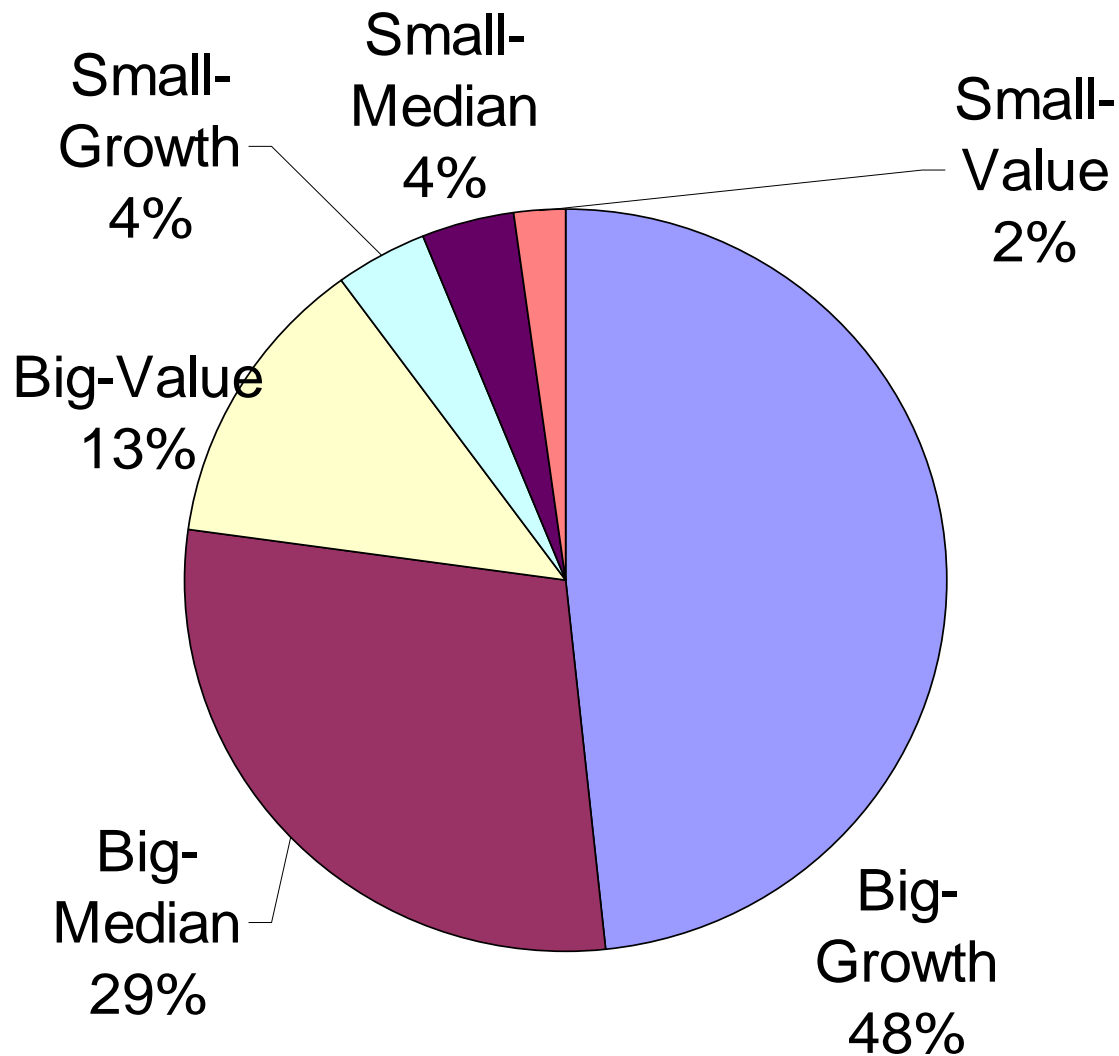
Fama-French Component Portfolios

MktRf weights

	None	Gro	Med	Val	All
Big	7.8	42.6	25.5	11.1	86.9
Small	4.2	3.5	3.4	2.0	13.1
All	12.0	46.1	28.9	13.0	100.0

Average excess return per year

	None	Gro	Med	Val	All
Big	5.92	7.61	8.62	9.20	7.72
Small	6.47	4.85	11.77	13.21	8.29
All	5.87	7.20	8.95	10.02	7.64

Weights of SMB and HML

SMB

	None	Gro	Med	Val	All
Big	0.0	-33.3	-33.3	-33.3	-100.0
Small	0.0	33.3	33.3	33.3	100.0
All	0.0	0.0	0.0	0.0	0.0

HML

	None	Gro	Med	Val	All
Big	0.0	-50.0	0.0	50.0	0.0
Small	0.0	-50.0	0.0	50.0	0.0
All	0.0	-100.0	0.0	100.0	0.0

Size Decile 10 vs. Its 3-Factor Benchmark

Target portfolio:
Size decile 10

	None	Gro	Med	Val	All
Big	0.0	60.0	29.2	10.8	100.0
Small	0.0	0.0	0.0	0.0	0.0
All	0.0	60.0	29.2	10.8	100.0

Benchmark portfolio:
0.967 x MktRf - 0.318 x SMB - 0.086 x HML

	None	Gro	Med	Val	All
Big	7.5	56.1	35.2	17.0	115.8
Small	4.1	-2.9	-7.3	-13.0	-19.1
All	11.6	53.2	27.9	4.0	96.7

Size Decile 4 vs. Its 3-Factor Benchmark

Target portfolio:
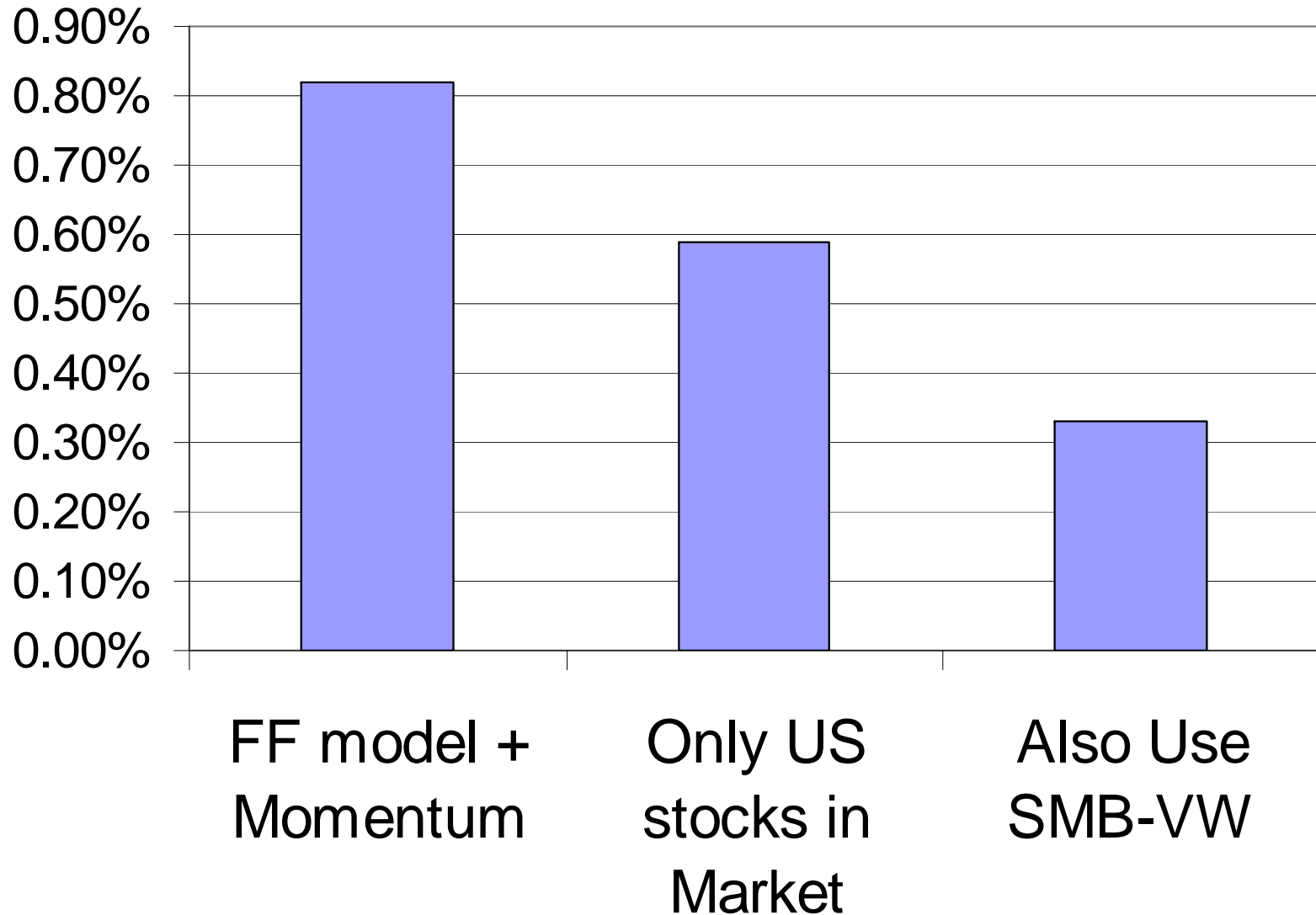Size decile 4

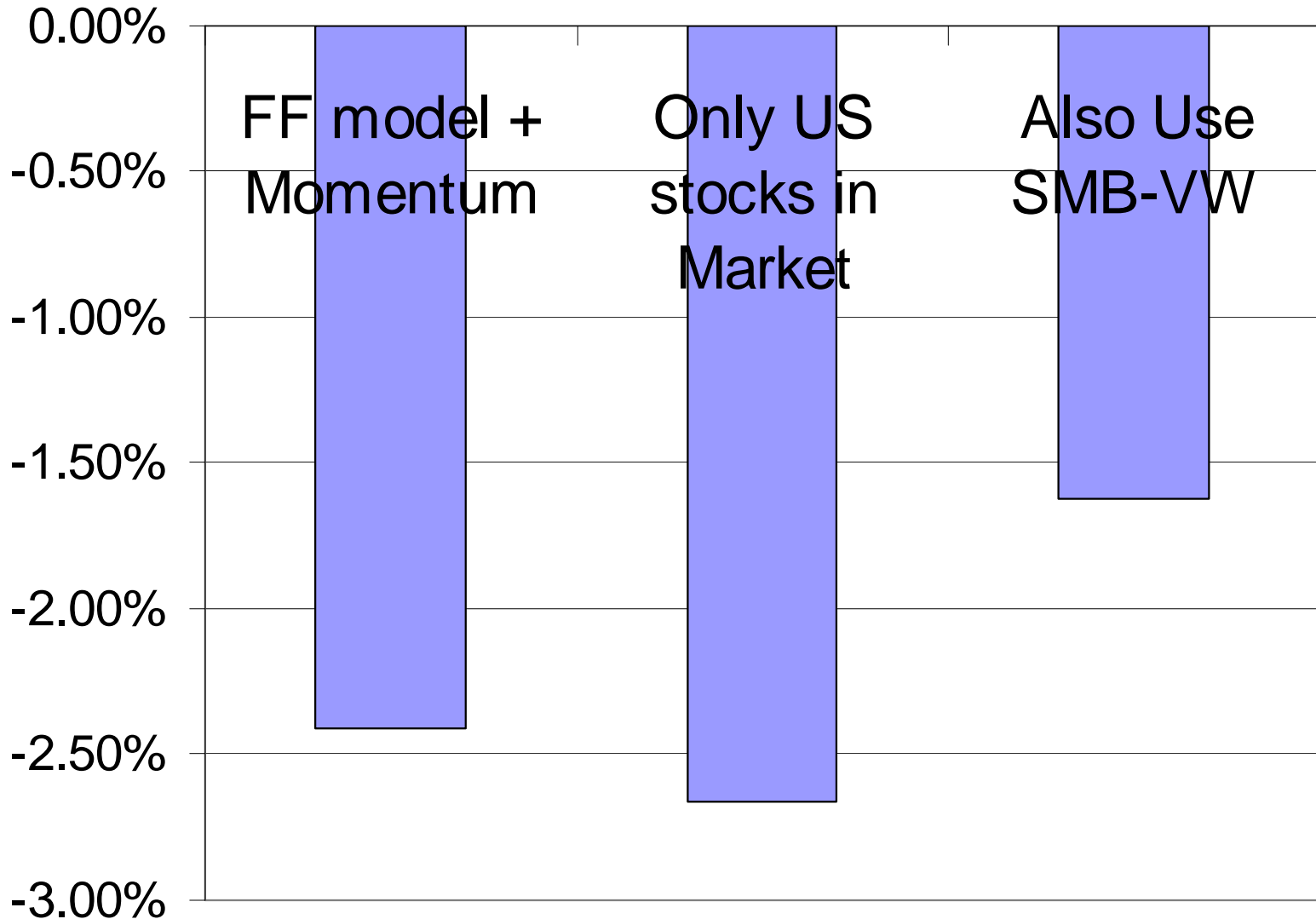|	None	Gro	Med	Val	All
Big	0.0	0.0	0.0	0.0	0.0
Small	0.0	40.7	40.5	18.7	100.0
All	0.0	40.7	40.5	18.7	100.0

Benchmark portfolio:
1.055 x MktRf + 0.799 x SMB + 0.226 x HML

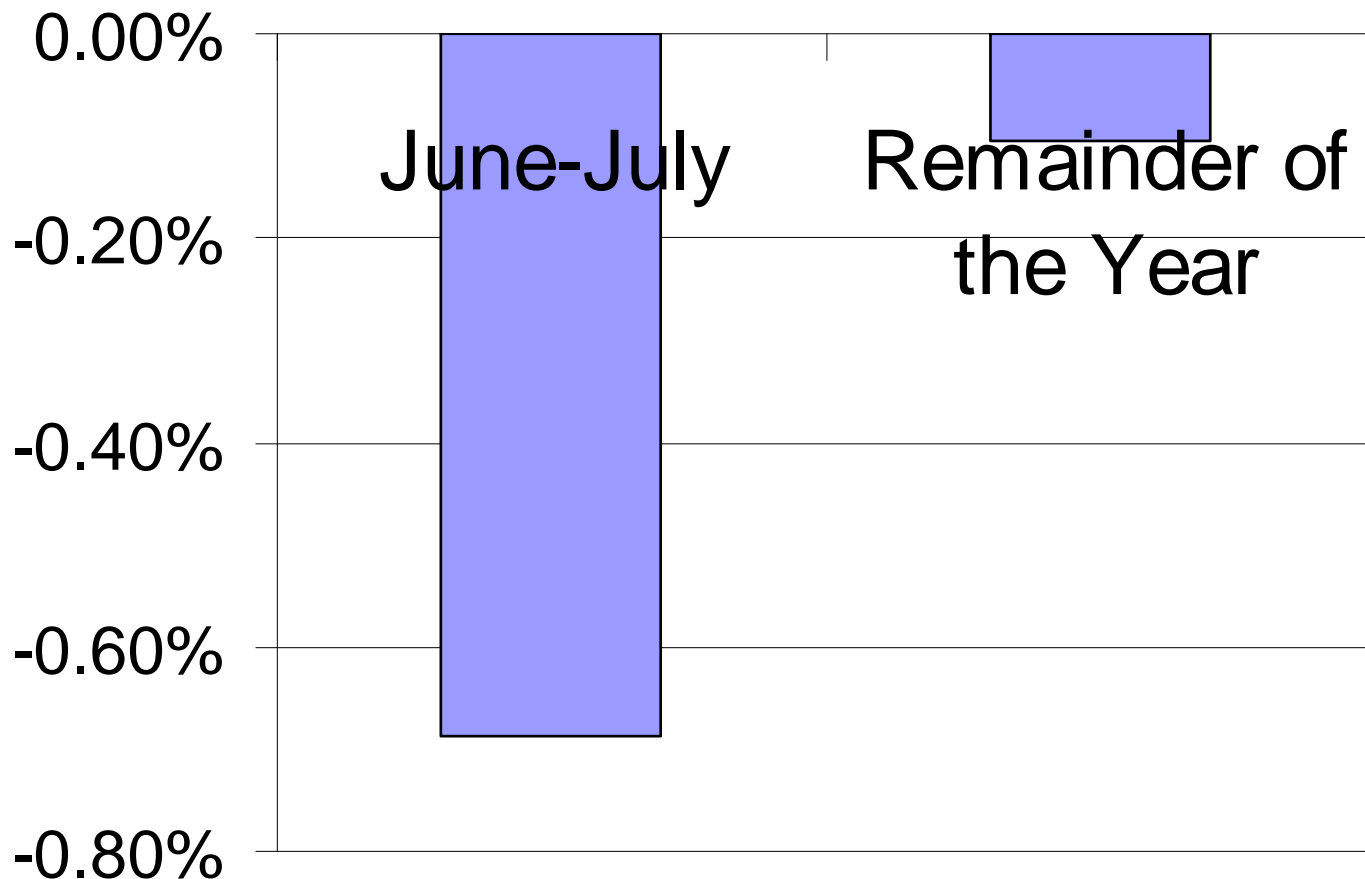|	None	Gro	Med	Val	All
Big	8.2	7.0	0.3	-3.7	11.8
Small	4.5	19.1	30.2	40.0	93.8
All	12.7	26.1	30.5	36.3	105.5

S&P 500 Alpha and Modified FF Factors: Annual Alpha

Russell 2000 Alpha and Modified FF Factors: Annual Alpha

Russell 2000 Reconstitution Effect: Annualized Alpha

- Index reconstitution costs (e.g., Petajisto (2008))

Criteria for a Good Model

- Pricing
  - Simplest possible explaining the cross-section of expected returns
- Benchmarking money managers
  - Most accurate estimate of a manager's value added relative to a passive strategy
- Benchmark model includes pricing model
  - Add non-priced factors
  - Reduce noise in alpha estimates

Selection of Alternative Models

- Carhart: MktRf, SMB, HML, UMD

- Carhart + S&P500 + Russell 2000

- Modified Carhart:
    - Only US stocks, value-weighted SMB and HML
    - Seven-factor model:
        - MktRf, MMB, SMM, BHML, MidHML, SHML, UMD

- Index-based models:
    - S5, R2-S5, R3V-R3G, UMD
    - S5, RM-S5, R2-RM, R3V-R3G
    - S5, RM-S5, R2-RM, S5V-S5G, RMV-RMG, R2V-R2G, UMD

Our Tests for Alternative Models

- Explain common time series variation
  - Tracking error volatility of all-equity mutual funds
- Explain cross-section of average returns
  - Mutual funds
    - 3 x 3 portfolio sort by size and value
  - 100 Fama-French size-value portfolios
    - 10 x 10 portfolio sort by size and value

Index-Based Models Perform the Best

- Out-of-sample Tracking Error of mutual funds
  - Index models: lowest TE (5% - 10% lower)
- Cross-section of average mutual fund returns:
  - Index models
    - close-to-zero alphas across all fund groups
  - Carhart model
    - small-cap funds underperform large-caps;
    - conclusion fully reversed (by 5% per year!) if control for benchmark index
- Cross-section of average return on 100 FF portfolios
  - Index models produce the highest R^2

Mutual Fund Alphas: Fama-French factors
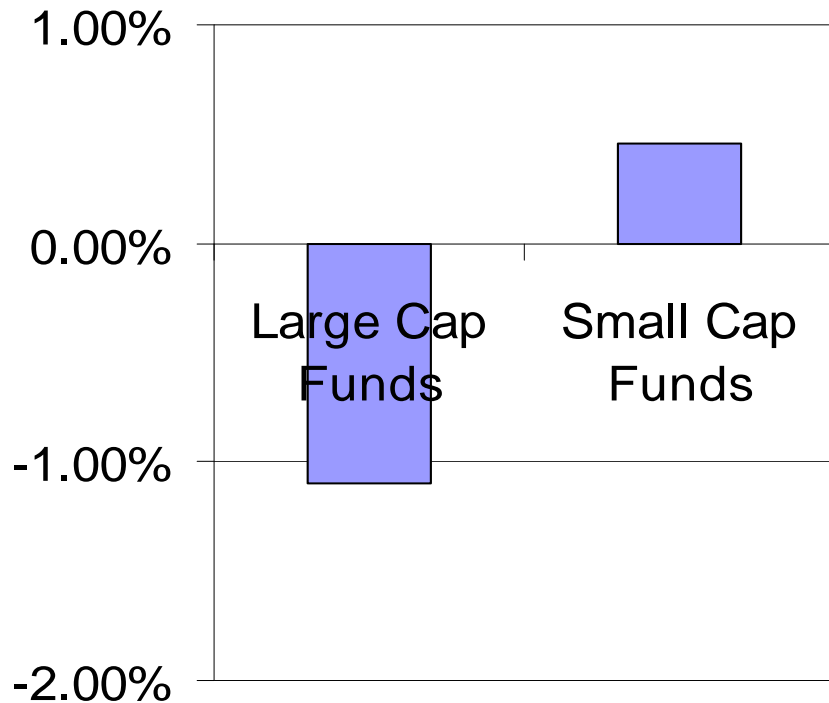
FFC Annualized Alpha of
Fund Return
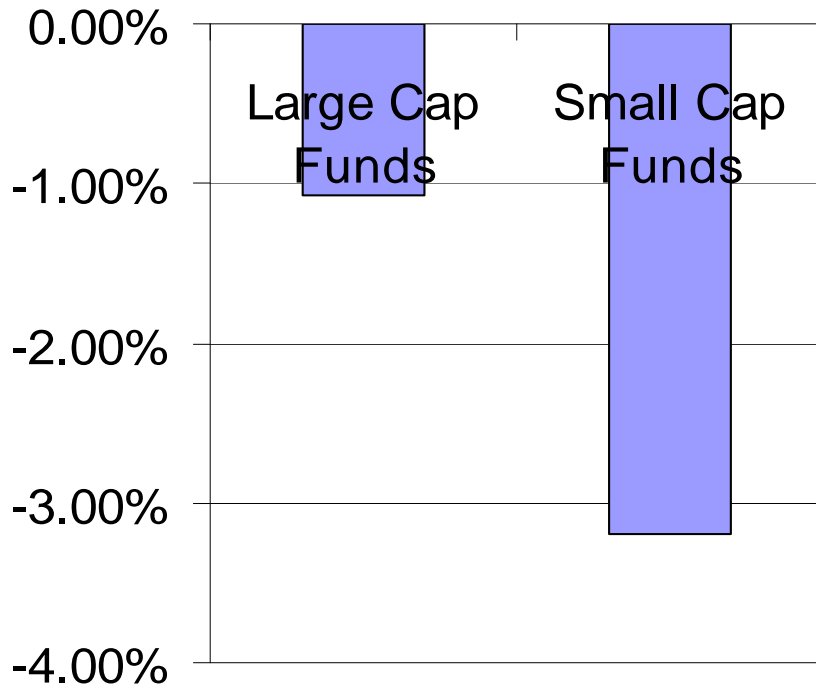
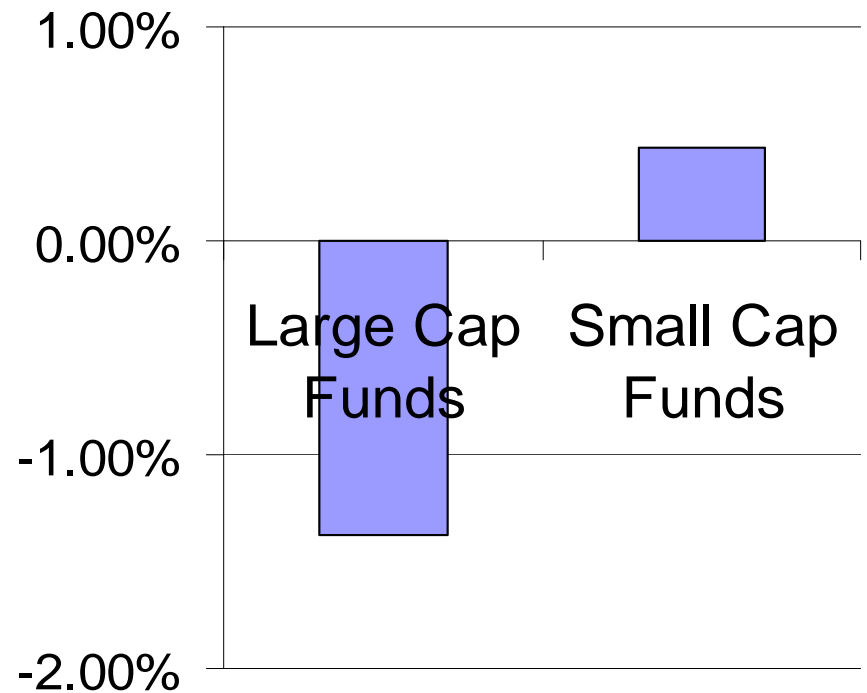FFC Alpha, Benchmark-
adjusted Fund Return
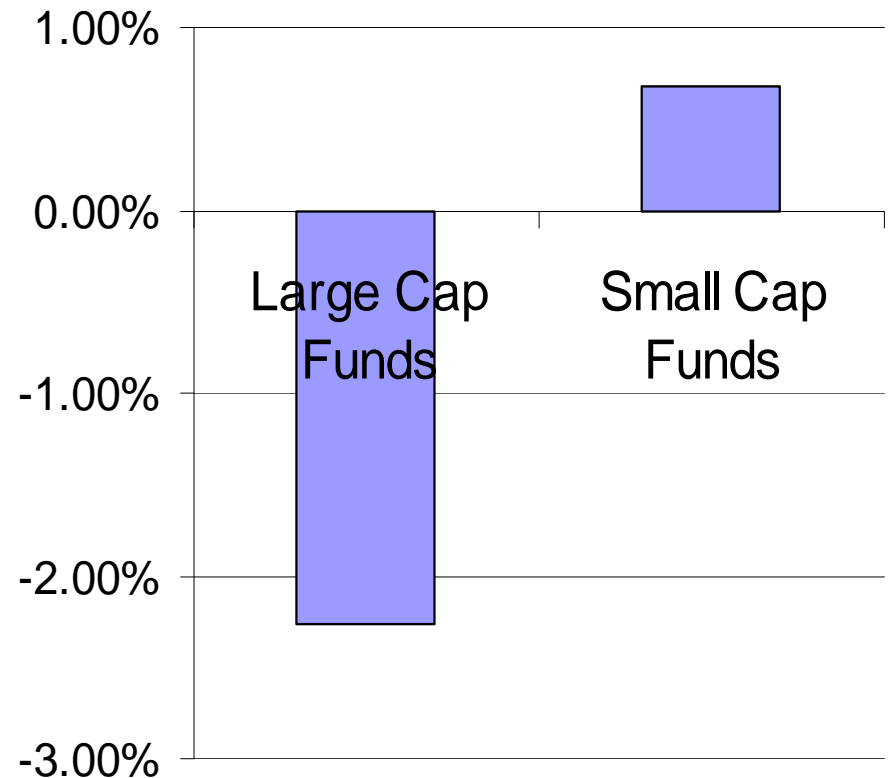
Mutual Fund Alphas: Index factors
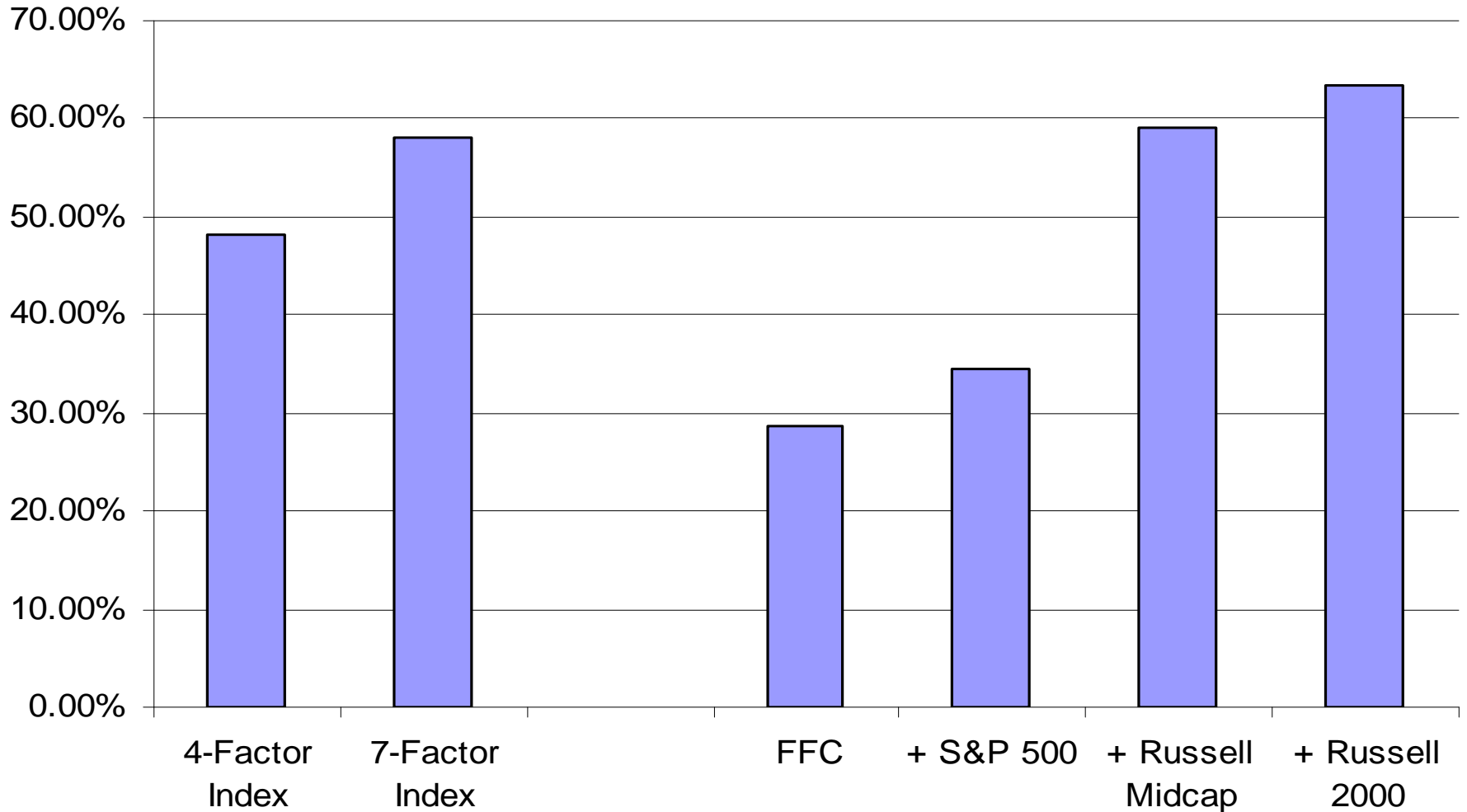
4-factor Index Model
Annualized Alpha,
Fund Return

4-factor Index Model
Annualized Alpha,
Benchmark-adjusted Fund
Return

Cross-Section of 100 Fama-French Portfolios

R2 of Mean Returns on Betas

So What Model Do We Propose?

3 approaches:

1) Modify the Fama-French factors
   - Value-weight SMB, only US stocks, etc.

2) Use benchmark-adjusted returns
   - FF factors relatively harmless

3) Use index-based models
   - 7 factors: S5, RM-S5, R2-RM, S5V-S5G, RMV-RMG, R2V-R2G, UMD
   - 4 factors: S5, R2-S5, R3V-R3G, UMD, or
     S5, RM-S5, R2-RM, R3V-R3G
   - Simplest and most general approach

Further Considerations for
Portfolio Performance Evaluation

Use benchmark factors that

- … include similar assets as portfolio
  - E.g., both only US common stocks
- … are representative of the portfolio asset class(es)
  - E.g., separate index for each asset class
- … use similar weights
  - E.g., value-weighted
  - E.g., similar weights to value and growth, liquid and illiquid, traded and non-traded, high and low transaction costs, etc.

Conclusions

- Common indices
  - Large nonzero Fama-French alphas
  - Affects performance evaluation
- Benchmark FFC alphas are misleading
  - FF factor construction methodology
  - CRSP market and Russell 2000 reconstitution
- Alternative models based on benchmark indices
  - Improve estimation of alphas
  - Improve cross-sectional pricing