Portfolio Performance Measurement: A No Arbitrage Bounds Approach

Co-Authors: Dong-Hyun Ahn and H. Henry Cao

Presented by Stéphane Chrétien
Faculty of Business Administration
Laval University

April 18, 2008

Benchmark Model Choice

- Highly controversial issue in performance evaluation;
- Performance evaluation and ranking can change significantly from a benchmark model to another;
- Related literature: Roll (1978), Lehmann and Modest (1987).

Implications of the Controversy

- Need strong assumptions to believe in a particular model;
- Possibility that the model chosen is not adequate;
- Tradeoff:
  - High precision;
  - High possibility of incorrect inference.
- Are there guidelines to help us select a performance model?

Contribution of Chen and Knez (1996)

- Develop conditions for the admissibility of performance measures;
- Most substantial condition: to price correctly some basis assets or passive portfolios;
- Another important condition: not to allow arbitrage;
- Result in an infinite number of admissible performance measures, among which the CAPM and other models might be.

Our Contribution

- We extend Chen and Knez (1996) by directly considering the infinite number of admissible performance measures;
- We do so by developing upper and lower bounds on performance measurements;
- We let the data tell us what is the best and worst possible performance for a fund.

Our Distinctive Characteristics

- We do not impose strong assumptions to specify a model;
- We avoid the 'bad model' problem emphasized in Fama (1998);
- Tradeoff:
  - Possibility of low precision;
  - **No incorrect inference**;
- We provide a conservative view of performance evaluation.

Three Applications of the Bounds

- Performance evaluation of mutual funds: it is often difficult to sign the performance of mutual funds so that a large number of funds are evaluated positively by some measures.

- New methodology to rank mutual funds: ranking is highly arbitrary and subject to important inference error;

- Diagnostic instrument to evaluate the admissibility of candidate measures: inference errors in existing parametric performance measures can be of an important magnitude.

Outline for the Rest of the Talk

- Theory and examples on the performance bounds;
- Methodology to rank mutual funds;
- Diagnosis of candidate performance models;
- Estimation of the performance bounds;
- Data and empirical results;
- Conclusion.

Economic Setup of the Paper

- Economy contains the payoffs x (in gross returns) on N basis assets;
- The market is incomplete (Assumption 1);
- Arbitrage trading strategies are precluded among the basis assets (Assumption 2).

Theoretical Results

- There exists an infinite number of admissible stochastic discount factors M such that $\theta' 1_N = E^P[M \cdot y]$ for all $y = \theta' x$ and $M > 0$ (Proposition 1);

- Let the performance value (or alpha) of a mutual fund be $\alpha = E^P[M \cdot x_{mf}] - 1$. It is possible to find the stochastic discount factors that give the minimum ($\underline{\alpha}$) or maximum ($\overline{\alpha}$) performance measure to mutual fund payoffs not covered by the basis assets (Proposition 2).

Finite State Example

State	Prob P	R_f	x_s	M^{NA}	M^{ALT}
1	0.3	1.05	1.15	0.340	0.765
2	0.5	1.05	1.07	1.020	0.510
3	0.2	1.05	0.99	1.701	2.338

Admissible measures satisfy:

$$E^P[M \cdot R_f] = 0.3 \times M_1 \times 1.05 + 0.5 \times M_2 \times 1.05 + 0.2 \times M_3 \times 1.05 = 1;$$

$$E^P[M \cdot x_s] = 0.3 \times M_1 \times 1.15 + 0.5 \times M_2 \times 1.07 + 0.2 \times M_3 \times 0.99 = 1;$$

$$M_1 \geq 0; \quad M_2 \geq 0; \quad M_3 \geq 0.$$

Finite State Example (Cont)

State	Prob P	R_f	x_s	x_A	x_B
1	0.3	1.05	1.15	1.1	1.07
2	0.5	1.05	1.07	1.05	1.05
3	0.2	1.05	0.99	1.03	1.03

State	M^{NA}	M^{ALT}	$M^{\underline{\alpha}(x_A)}$	$M^{\overline{\alpha}(x_A)}$
1	0.340	0.765	0.000	1.190
2	1.020	0.510	1.429	0.000
3	1.701	2.338	1.190	2.976

Funds	α^{NA}	α^{ALT}	$\underline{\alpha}(x_A)$	$\overline{\alpha}(x_A)$
A	-0.00170	0.00213	**-0.00476**	**0.00595**
B	-0.00476	-0.00476	-0.00476	-0.00476

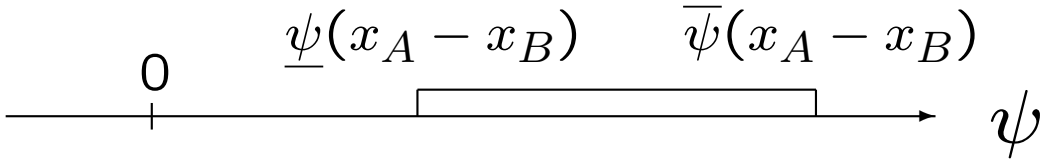
Ranking of Mutual Funds

- How can we use the performance bounds to rank mutual funds?

- We develop three alternative ranking rules (Definition 2):
  - Universal Dominance;
  - Best Case Scenario Dominance;
  - Worst Case Scenario Dominance.

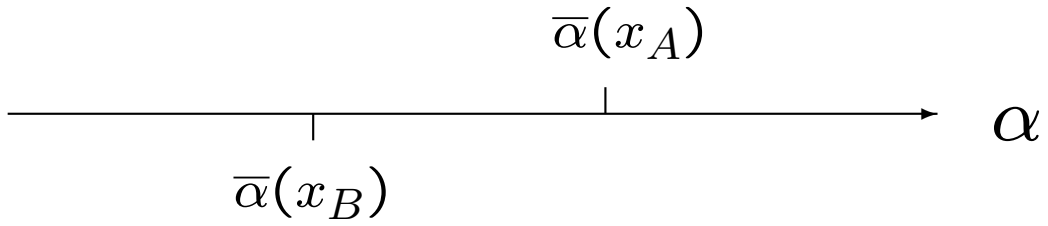
Universal Dominance

Fund A dominates Fund B in the sense of Universal Dominance *if the lower bound on the differential in performance measures of A and B evaluated with the same SDF is positive:*



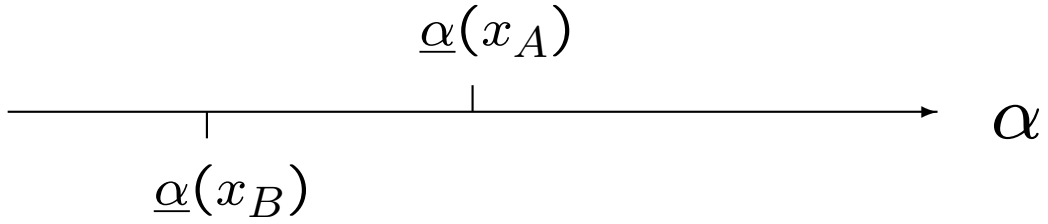
Best Case Scenario Dominance

Fund A dominates Fund B in the sense of Best Case Scenario Dominance *if the upper bound on the performance measure of A is greater than the upper bound on the performance measure of B:*



Worst Case Scenario Dominance

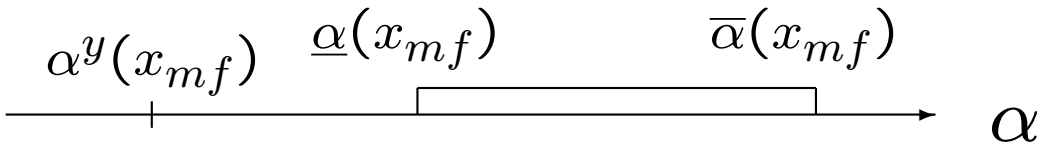
Fund A dominates Fund B in the sense of Worst Case Scenario Dominance *if the lower bound on the performance measure of A is greater than the lower bound on the performance measure of B:*



Diagnosis of Candidate Performance Measures

- How can we use the performance bounds to diagnose candidate performance models?

- A necessary condition for a candidate measure to be admissible is that its α resides within the performance bounds:



Estimation of the Performance Bounds

Assume that

- The economy is replicated over time in a stationary and ergodic manner;
- The N observed basis assets are linearly independent;
- $N < T$, where T is the sample size.

Estimation of the Performance Bounds (Cont)

Then, the performance bounds can be rewritten as:

$$\text{Minimum: } \underline{\alpha}^{\star}(x_{mf}) = \min_{\{M_1,\ldots,M_T\}} (1/T) \sum_{t=1}^{T} (M_t x_{mft}) - 1;$$

$$\text{Maximum: } \overline{\alpha}^{\star}(x_{mf}) = \max_{\{M_1,\ldots,M_T\}} (1/T) \sum_{t=1}^{T} (M_t x_{mft}) - 1;$$

with M_t satisfying the constraints:

$$\mathbf{1} = (1/T) \sum_{t=1}^{T} (M_t \mathbf{x}_t);\ M_t \geq 0\ \ \forall t.$$

Including Conditioning Information

- As argued by Ferson and Schadt (1996), portfolio manager using only public information should not be given a positive performance. Thus, performance evaluation should incorporate conditioning information.

- We do so by using basis assets that are dynamically managed portfolios: $x^c = x \otimes z$, where $z \stackrel{\triangle}{=} \tilde{z}/E[\tilde{z}]$ are predetermined standardized instrumental variables.

Mutual Fund Sample

- 320 U.S. equity mutual funds available in Morningstar from 1984 to 1997;

- Monthly returns include reinvestment of all distributions and are net of management fees, incentive fees, and other fund expenses.

Basis Asset Payoffs

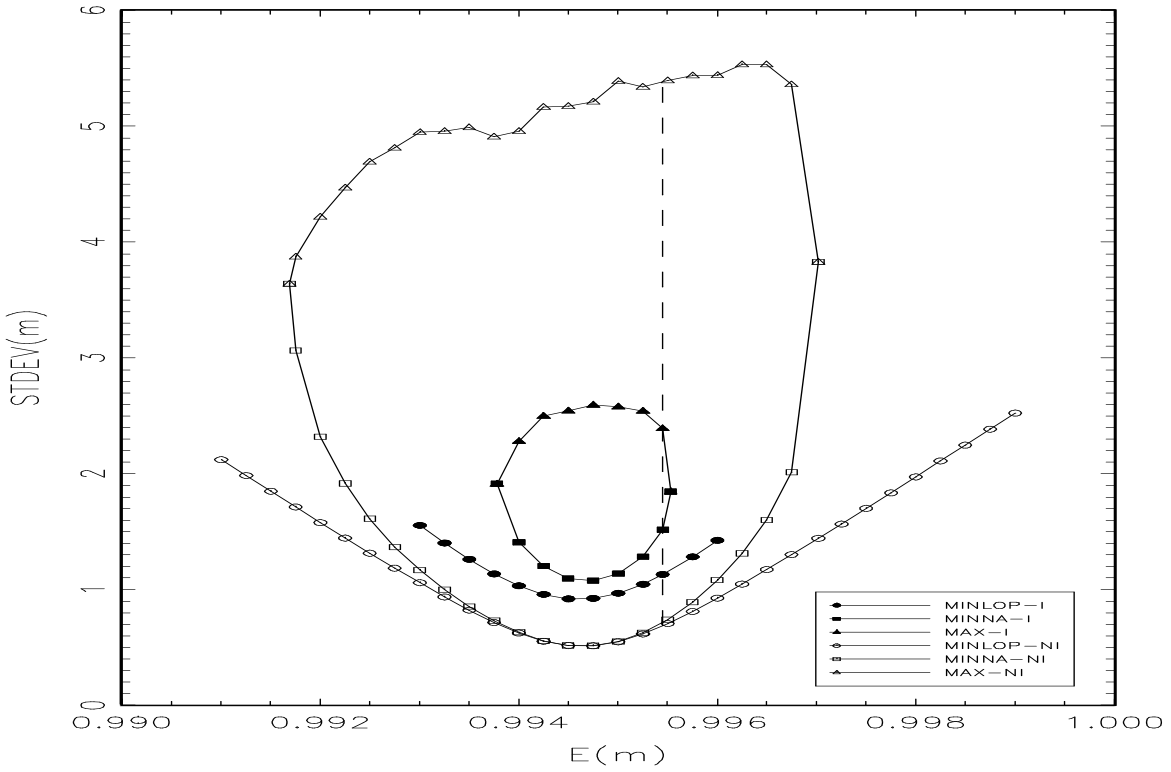
- 20 portfolios formed by industry described in Moskowitz and Grinblatt (1999) (Source: CRSP File);

- 2 bond portfolios (Source: CRSP Fama File):
  1- Short-term (maturity less than one year);
  2- Long-term (maturity greater than 10 years).

- Lagged credit spread (Source: CITIBASE);

- Lagged dividend yield on the CRSP value-weighted index.

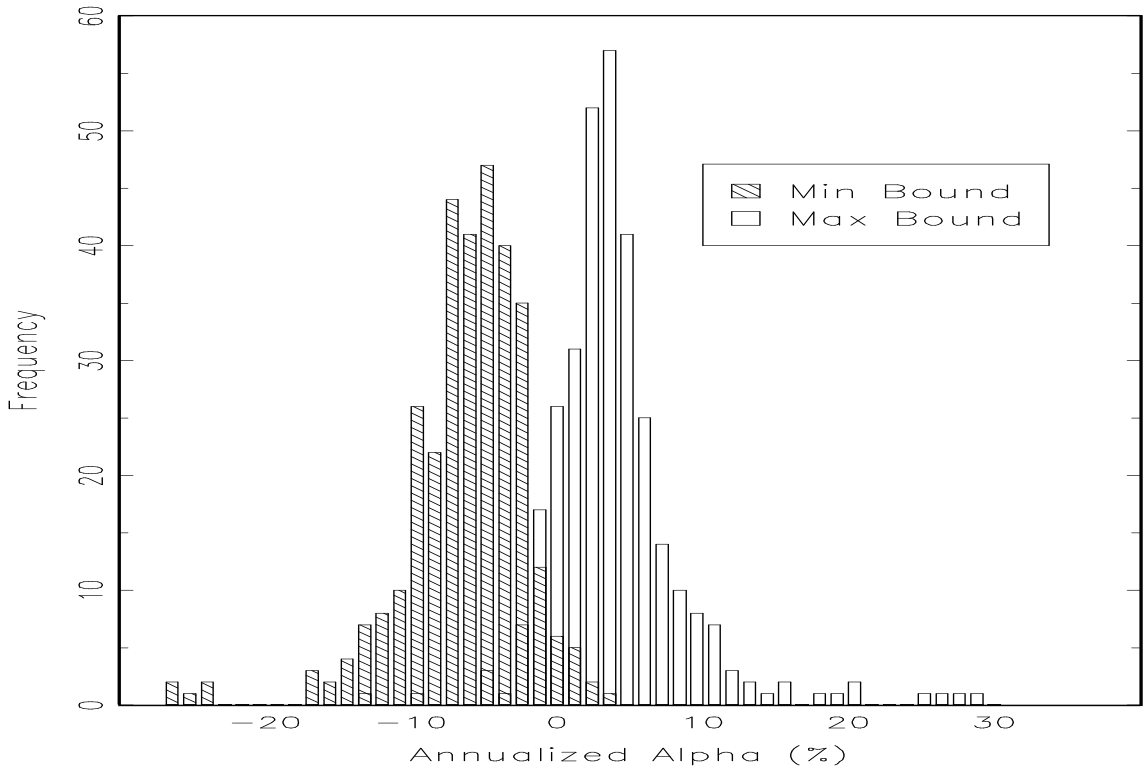
Admissible Performance Measures: Figure 3



Performance Bounds: Table 3, Panel A

Statistics	Lower Bound	Upper Bound	Bound Diff
Mean	-0.00578	0.00316	0.00895
Std Dev	0.00372	0.00543	0.00539
(t-stat)	(-18.0767)	(6.7705)	(19.2747)
Min	-0.02852	-0.01181	0.00311
10%	-0.00953	-0.00088	0.00520
Median	-0.00532	0.00232	0.00754
90%	-0.00224	0.00710	0.01323
Max	0.00243	0.05447	0.05362

Histogram of the Bounds: Figure 4



Performance Evaluation

Performance	Criteria	% of Funds
Positive	Lower Bound > 0	2.5%
Negative	Upper Bound < 0	17.2%

- Simulations show that:
  - A large proportion of the bounds are significantly different from zero;
  - Signing the performance of mutual funds is even more difficult when accounting for sampling error;
  - Precision of the bound estimates is inversely related to the variability of fund returns.

Performance of the Universe of Funds

- Bounds for an equally-weighted portfolio of all funds:
  - Worst performance = -0.337%;
  - Best performance = 0.091%.
- At least a SDF that values positively the universe of funds;
- Does not rule out the 'efficiency with costly information' argument of Grossman and Stiglitz (1980).

Diagnostic of Models: Table 6, Panel B

	CAPM	CAPM-C	FFM	FFM-C
% < Lower Bound	2.19	2.50	1.25	1
% > Upper Bound	6.56	8.44	9.38	7.50

	FSM	FSM-C	POWER	HABIT
% < Lower Bound	0.63	1.25	1.56	1.56
% > Upper Bound	27.19	31.25	50.63	36.56

Conclusions

- We develop performance measure bounds that are free of inference errors;

- In the paper, we show how to use the bounds to
  - Evaluate mutual fund performance;
  - Rank mutual funds;
  - Diagnose candidate performance measurement models.

Implications of our findings 1

- Without making auxiliary assumptions that lead to the 'bad model' problem, it is often not possible to sign the performance of mutual funds:
  - Suggests that inference errors can have a strong effect on performance measurement;
  - Provides a partial explanation for the sensitivity of performance to the benchmark choice documented by the existing literature.

Implications of our findings 2

- In our incomplete market setup, we cannot rule out that a large number of mutual funds are valued positively by some investors:
  - Heterogeneous preferences might explain the disagreement on mutual fund valuation;
  - The performance measure based on a fund's most favorable investor class (the upper bound) give a positive assessment to many mutual fund portfolios, supporting casual observations on the importance of the mutual fund industry.