14 Week 5b Mutual Funds

14.1 Background

1. It would be natural, and completely sensible, (and good marketing for MBA programs) if funds outperform darts! Pros outperform in any other field.

2. Except for... competition and free entry. The *marginal* fund will be worthless, and the *average* fund may not be very good.

3. The standard empirical approach and view (Jensen 1968):

   (a) There is a big bias problem, and this work is hard! (It’s more severe for hedge funds, venture capital funds)

   i. *Survivor* bias: Funds that close are not in database. Funds are more likely to close if they have bad returns

   ii. *Selection* bias. Funds are more likely to report if they have good returns. (Or to report audited returns)

   iii. *Backfill* bias. Often, when a fund is included in a database or index, all past history of that fund is then included. This biases the results towards winners. There were other funds around during the backfill period that lost money and did not get included in the database.

   iv. *Incubator* bias. Fund families start “incubator funds,” and then only open the ones that do well. They then report the entire history. It’s amazing that the SEC lets them do this. This bias remains in the CRSP database.

   v. (My addition) *Academic interest bias*. We’re only interested in hedge funds since they did so well! All t statistics in the Journal of Finance are above 2.1.

(b) You have to correct for beta. A fund which simply leverages or holds higher beta stocks will seem to outperform with no skill. Thus, run CAPM regressions on funds and examine alphas.

(c) $\sigma/\sqrt{T}$ again. Returns are so volatile, it’s very hard to measure mean returns.

(d) After addressing all these problems, the standard result: The average fund underperforms S&P by about 1%; typically even before fees.
Notes: Average returns of mutual funds over the Treasury bill rate versus their

(e) “Why the average fund?” you might object. Of course the average fund is bad — competition and easy entry into the fund business mean anyone will join. Why aren’t we looking at “the good funds?” What about Warren Buffet?

(f) Answer: There is no way to tell after the fact if a given fund was lucky or skillful. We must look at average performance in a group; we have to look at some strategy we might plausibly have used for picking funds ahead of time, and then see what happens to all funds we might have picked.

i. Just as with stocks, we sort all funds at $t$ based on some indicator of “good,” then we watch those funds through the following year $t + 1$. We have to make sure we include all funds we would have picked at the time, and we have to make sure you keep track of everyone including the losers and the dead in the following year.

ii. Early answers: They looked (as in practice) at funds that had done well in previous years. They found that there is a random walk in funds too: Past winners are no more likely to keep winning. Here “Good group” means past good performance.

(g) What we can’t do: “why is Warren Buffet so good?” (Was he lucky? Where is Elmo Buffet?) “Sure average alpha is bad, but our track record is great.” (Why does everyone who walks in the door have above average alpha?)

(h) What we can do: Find everyone who “looked like” Warren Buffet at $t$, invest in them through $t + 1$, track returns of winners and losers.

4. More background

(a) Most variation in ex-post fund returns is due to different strategies (small, large, value, growth, sector) etc. not to differences in particular stocks that funds pick (“style” not
“selection”). Quantitatively, suppose you run
\[ R_{t} = \alpha_{i} + \beta_{im} R_{t}^{em} + \beta_{11} f_{1}^{i} + \beta_{22} f_{2}^{i} + \varepsilon_{t}^{i}; \ t = 1, 2, \ldots, T \] for each \( i \)
(Here, we’re doing a style analysis or variance analysis, so the factors don’t have to be “priced” factors. Even before smb and hml, people understood that stocks moved when their industries moved together. The CAPM just says that these betas don’t raise expected returns.) Mutual funds hold quite diversified portfolios, so even individual funds (not just portfolios of funds as in Carhart) have very high \( R^{2} \). More importantly, when you look at average returns in a sample, say 5 years, (\( E \) here means a 5 year average)
\[ E \left( R_{t}^{i} \right) = \alpha_{i} + \beta_{im} E \left( R_{t}^{em} \right) + \beta_{11} E \left( f_{1}^{i} \right) + \beta_{22} E \left( f_{2}^{i} \right), \]
then think about how \( E(R^{i}) \) varies across mutual funds in that sample, you find almost all variation across funds in \( E(R^{i}) \) comes from variation across funds in their \( \beta \) choices and luck-of-the-draw in how the factors \( f \) that they load on happen to do, not from variation across funds in the alphas they achieve. (In 5 year samples, many “factors” such as industry that give no long-term average return and thus are not “priced” will nonetheless do well or poorly for 5 years.) Again, “most variation in results comes from style not selection.”

(b) Puzzles, to me:

i. Why are there so many styles, and these are so unrelated to any betas or sensible risks I’ve ever heard of? What does “growth and income” mean? Is this just marketing, so we can always have some type of fund that did well last year to tout?

ii. If value/size is important, why are we arguing about performance relative to the Market/CAPM? Why do value funds not blow away the graph? FF answer: Most funds were not really following value. (Davis, Fama, French, below; h breakpoints are -0.08 and 0.3 in my graph above). OK, but why not? (New facts, p.51) If funds didn’t know about the value premium, how can it be an equilibrium risk premium? (Their answer: “growth is where I can find alpha.”)

iii. For 40 years we have been railing about the wasted money on active management. From Jensen 1968 to Ken French 2009. Yet it persists. Everywhere else, free market economists say “if something persists, it must be serving a function,” not “Active management persists because investors are dumb.” We don’t allow the “investors are dumb” story for price movements, how do we allow it for active management?

iv. Related, Carhart’s Table III shows that fees and turnover are bad for investors. Why? This makes no sense from a market-at-equilibrium perspective. Management wants you to think more fees pay for more research, which raises alpha, and they split that alpha with you. OK, a perfectly cynical competitive market perspective would say that returns to investors must always be the same, so fees raise alpha just enough to pay the fees. How in the world can fees hurt investors in a long-run (40 year) equilibrium? It’s fun to be cynical but it doesn’t make any sense.

(c) What are we asking here, really? Once upon a time, mutual fund performance was a test of “semi-strong form” efficiency – can you do better with public, but hard to get information? In this analysis, the question is, implicitly, do funds have “stock picking ability” to find “undervalued stocks,” i.e. “inefficiencies,” “information not reflected in market prices,” and the alternative is “can I replicate fund returns by passive or mechanical portfolio formation strategies (without paying high fees)?”
Note, I think this question is really passé. The real issue for most active management is that the managers may understand multifactor betas you don’t understand, a view you will see emerge over the next two weeks. But nobody knew multifactor betas existed in 1970, and old habits die hard.

5. Building up to Carhart: Reexamining the evidence, several authors did find “hot hands” – Last year’s funds do better this year. (p. 57). We see this like momentum – by forming portfolios of last year’s winners and noticing they do well. Why was it missed earlier? Well, it wasn’t really. This is like my analysis of momentum; we’re using a new telescope to make an old puzzle seem more important than it used to. There is still only a very small autocorrelation of individual fund returns.

Suppose a winner is only 52% likely to win again, which is the old version of the facts. In the old days, you’d say “there’s next to no skill; funds that went up last year are almost exactly as likely to go up as to go down next year.” Like momentum, though the same fact has a much more dramatic-looking implication for a portfolio of funds. The portfolio of winning funds went up 100% last year, so only a 0.01 $R^2$ means a 10% mean return for this portfolio next year. The chance that the portfolio does better is much more than 52/48 because you diversify across fund risk. This is just like momentum in stocks – you don’t have to buy just one stock! Recall how momentum is a way of magnifying a small return autocorrelation. Thus, by looking at portfolios, we confront the fact that Maybe that 52% is meaningful, and has important economic implications.
14.2 Carhart on funds – questions

The introduction summarizes his conclusions:

1. Momentum in stocks accounts for momentum in funds. Funds that did well last year have stocks that went up and those stocks will keep rising a bit. It is not
   (a) Persistent skill, or 
   (b) Good returns for momentum funds.

2. Momentum funds do poorly after transactions costs.

3. There is some persistent under performance

4. Survivor bias free data – includes funds that die. (Lots of hard work by Carhart, and another great CRSP dataset.)

We need to look for the facts! Find the facts behind these assertions in the paper.

Now Questions for Class:

1. On p.61 Carhart defends the four-factor model as a performance attribution model.
   (a) Why is it OK to use a “momentum factor” even if that is not a “state variable for investment opportunities?”
   (b) What question are we using the multifactor model to answer, and how is that question different from Fama and French’s question?
   (c) Suppose you find, looking at very long samples, that $E(smb) = 0$. Might you still use sml for performance attribution in a shorter sample? If so, why?

2. Does Carhart’s “momentum factor” solve the return puzzles in Fama French’s last table – does it account for returns in momentum-sorted stock portfolios? (There is no table here, but he does report some results verbally. What does he say?)

3. (Hint: table III is the most important. Spend most of your time to understand it.) How does Carhart form portfolios of mutual funds - -what are Portfolio 1A 1B...10C in column 1 of Table III?

4. Do funds that did well last year continue to do well next year? Point to numbers – is this serious or a tiny effect?.Is the phenomenon stronger among winners or among losers?

5. Do the funds that went up last year always continue to go up? How much risk is there in this investment strategy? Can you guesstimate, of 100 funds in 1A last year, how many of them will make money this year?

6. Wait a minute. Knowing there are “hot hands” but that individual funds are risky, we could form a portfolio of the funds that went up last year (portfolio I). As we increase the number of funds in the portfolio, the portfolio variance should keep going down and our Sharpe ratios should keep going up. Why is the anticipated risk-reduction of diversification not happening?
7. Perhaps the funds that do better in the test year continue to do well on average because they have high CAPM betas. What does Carhart say about that?

8. What lessons do you learn from the CAPM $R^2$ values in Table 1?

9. Are all the alphas zero after the 4-factor model is done, or is there a puzzle? Who seems still to be outperforming and who is underperforming?

10. Which factor’s betas seem to be accounting for the “hot hands” (spread in average returns)?

11. What puzzle does Table IV address? What is its conclusion?

12. Fund managers claim that fees and turnover do not reduce returns to investors. How could charging more money not reduce returns to investors? (Try to be a good salesperson for a high-turnover high-fee fund. Why should I give you my money? Then try to be a good supply-demand economist. What should the equilibrium relationship be between fees, expenses and returns to investors?)

13. (Table V. Make sure you understand how this table was created. How are Table IV and Table V different?) What does Carhart find about fees and turnover? How much does a 1% change in fees change returns to investors? How much does turnover – selling one stock and buying another – change returns to investors?

14. What are the two hypotheses that Carhart spends the rest of the paper distinguishing (see bottom of p. 70)

15. What is the point of Figure 2? (Hint: what would it look like if the sort on one year performance indicated skill?)

16. What does Carhart say about momentum funds – funds that seem to follow a momentum strategy, as revealed by high loadings on the momentum factor? (Hint: no table, but text on p. 73)

17. One year lagged returns are probably mostly luck, not skill. What if you sort funds by the more common 5-year performance averages? (Hint: Figure 3)

18. Does Carhart suggest any trading strategies? (Hint. Look on p. 80)

### 14.3 Fama and French Mutual Fund Performance Questions

So far, we have been looking for “skill” by guessing some characteristic associated with skill – past returns, MBA by manager, etc. – and looking at the return of a sorted portfolio going forward. This paper tells us whether there is any skill at all, without us taking a stand on what characteristic can be used to find good funds. It answers the question “sure the average fund is mediocre, but there are some good funds.”

1. What do Fama and French mean by “Equilibrium Accounting? (1915)

2. Table 2 tells us about the performance of mutual funds as a whole – it studies the portfolio of all mutual funds. What does it tell us? Do mutual funds as a whole outperform benchmarks? What kinds of indices are most similar to the performance of all mutual funds? Do mutual funds as a whole generate alpha, before or after fees?
3. Fama and French focus on the alpha t statistic. Why not look at alphas or information ratios?
   Why don’t FF just use the t distribution, to judge how many funds should have alpha t statistics above a certain cutoff? Explain the bootstrap procedure. (Hint, p. 1924)

4. Explain the numbers in Table 3-4.
   (a) What would they look like if all funds had zero true alpha, but the pattern of luck fully conformed to the assumptions of the t distribution (normal, independent, etc.)?
   (b) What would they look like if there were some funds with +5% alpha and other funds with -5% alpha, so that the average fund was not skilled but some were good and some were bad?
   (c) Why is the probability of a t greater than 2 or less than -2 not the usual 5% value that we expect for a t statistic?

14.4 Berk Questions

1. What happens to future returns and flows, according to Berk, if a manager does have some skill?

2. Berk says, unlike FF, that managers do have some skill even though alphas are all zero. How can that be?

3. Berk says that when investors chase past returns, investing in funds that have done well in the past, they are not being irrational, even though future returns are no better than average. How can this be?

4. Berk says that even though skill is permanent, returns will not be persistent. Why not?
14.5 Carhart on funds – questions and answers

The introduction summarizes his conclusions:

1. Momentum in stocks accounts for momentum in funds. Funds that did well last year have stocks that went up and those stocks will keep rising a bit. It is not
   (a) Persistent skill, or
   (b) Good returns for momentum funds.

2. Momentum funds do poorly after transactions costs.

3. There is some persistent under performance

4. Survivor bias free data – includes funds that die. (Lots of hard work by Carhart, and another great CRSP dataset.) (p. 58)

We need to look for the facts! *Find the facts behind these assertions in the paper.*

Now Questions for Class:

1. On p.61 Carhart defends the four-factor model as a performance attribution model.
   (a) Why is it OK to use a “momentum factor” even if that is not a “state variable for investment opportunities?”
   (b) What question are we using the multifactor model to answer, and how is that question different from Fama and French’s question?
   (c) Suppose you find, looking at very long samples, that $E(smb) = 0$. Might you still use smb for performance attribution in a shorter sample? If so, why?

   A: To measure stock picking ability, performance relative to mechanical portfolios is enough, whether or not it is a “true” multifactor model. What you want to know is whether you can replicate the fund’s performance with (cheap) index or mechanical strategies, not whether the returns from such a strategy or “style” are justified from fundamental risk. Thus, Carhart’s model is market, smb, hml, momentum, and we don’t argue about whether momentum is a “real” factor or not. Anything that you could (and might have!) realistically programmed a computer to do on the right hand side goes here.

   SMB, for example. Yes for performance evaluation, as you would for an industry fund. Then in a period (5 years) where small stocks happened to do well, you could tell if a fund’s good performance was due to skill, or due to the fact that they invest in small caps and small caps happened to go up. This is an instance of “style” vs. “selection.”

   Even better see Fama and French p.1918 pp2 prose.

2. Does Carhart’s “momentum factor” solve the return puzzles in Fama French’s last table – does it account for returns in momentum-sorted stock portfolios? (No table here, but he does report some results)

   A: Yes, p.62.
3. (Hint: table III is the most important. Spend most of your time to understand it.) How does Carhart form portfolios of mutual funds - -what are Portfolio 1A 1B...10C in column 1 of Table III?  
A: Re-formed once per year based on the previous year’s performance. Then look at monthly returns in the portfolios. Again, we have to form portfolios, then watch. Keep in mind, all of Carhart’s evidence is about portfolios of funds (or average behavior of funds of a given type), not individual funds.

4. Do funds that did well last year continue to do well next year? Point to numbers – is this serious or a tiny effect? Is the phenomenon stronger among winners or among losers?  
A: Yes. See column 2 Table III Variation across portfolios of funds is 1% per month or 12% per year – a lot! Most of the action is in the losers. Most amazingly, 10C lose money in the following year!

5. Do the funds that went up last year always continue to go up? How much risk is there in this investment strategy? Can you guesstimate, of 100 funds in 1A last year, how many of them will make money this year?  
A: For risk, std dev column. 1A is 0.75% with $\sigma = 5.45\%$. Thus, the chance of going up at all is $\Phi(0.75/5.45) = 55.47\%$. So this is still a 55 up / 45 down phenomenon, and that is among a group of funds in the highest 1/10 of performance for the previous year!

And that’s for the portfolio of funds – individual funds in the portfolio are even more volatile, so the chance of an individual winner fund doing well next year is even lower. Just a reminder that returns have a lot of risk with them! Don’t confuse alpha with arbitrage opportunity. Alpha means an average return, but not a good return every year.

Also, as with momentum, this is not a “new phenomoneon” it’s a “new way of looking at something we knew all along.” There is no conflict here with the conventional wisdom that funds who won last year are almost 50% likely to fall this year. If you can wrap your mind around that, you will have really understood the first column!

6. Wait a minute. Knowing there are “hot hands” but that individual funds are risky, we could form a portfolio of the funds that went up last year (portfolio I). As we increase the number of funds in the portfolio, the portfolio variance should keep going down and our Sharpe ratios should keep going up. Why is the anticipated risk-reduction of diversification not happening?  
(More explanation of the question: If there is a residual standard deviation of 10% (big), since there are about 6000 funds, a portfolio of the 600 top funds should have a standard deviation of $10/\sqrt{600} = 0.41\%$. Yet, this risk-reduction does not seem to be happening. The portfolio 1 standard deviation is 5.04% per month implying individual-fund standard deviations of $\sqrt{600} \times 5.04 = 123\%$. Per Month. Worse, if you divide it into three (1A 1B and 1C), the sub-portfolio standard deviations are almost unchanged. Dividing one portfolio in to three should raise the standard deviations by a factor of $\sqrt{3} = 1.73$.)

A: You already know there is a common factor in the winner funds returns, because you’re not seeing diversification kick in. The individual funds in this portfolio must all be moving together. We learn from the regressions what that factor is, but we know it’s there the minute we see these standard deviations. in There is a common component. Last year’s winners are all likely to win/lose together. On Wall street they say “momentum didn’t work this year,” and this is what they mean.
7. Perhaps the funds that do better in the test year continue to do well on average because they have high CAPM betas. What does Carhart say about that?

A: Good point – $E(R)$ is not a puzzle, only alpha is a puzzle. CAPM columns of TIII show that betas are all about 1, so ER variation is all alpha. These are single regression betas, so we are really using the CAPM. Anyway, what does it mean to use information to produce negative alpha?? (p.63)

8. What lessons do you learn from the CAPM $R^2$ values in table 1?

A: Especially in the middle, they’re huge. Individual companies get about 40%; FF stock portfolios got about 65%, here we’re seeing 98%. Portfolios of active funds are almost exactly replicating the market index. Note the $R^2$ tail off as we go up and down the table. To be extreme, you have to stray far from the benchmark. The 1 and 10 portfolios actually have very low $R^2$ meaning huge tracking errors, for funds.

9. Are all the alphas zero after the 4 factor model is done, or is there a puzzle? Who seems still to be outperforming and who is underperforming?

A: Alphas from -0.1% (1% / year, as before) to -0.2% (9). Then a big increase in the bottom end to an amazing $\alpha = -0.64\%$! Puzzle – how can you lose money in a diversified portfolio?? Efficient markets mean you can’t do this either (or you and I short what they are long). Expenses? Coming up.

Note: negative alphas is not that surprising. The indices do not include transactions costs. Real-world performance is usually less than the index.

Note: FF4F have no transactions costs or short costs (hml is perpetually short small growth stocks, without any cost), and assume you buy at midpoint of b/a. Some negative alpha is natural.

10. Which factor’s betas seems to account for the “hot hands” (spread in average returns)?

A: A bit of SMB, mostly PR1YR. Point: get in the habit, as with FF, of looking at the beta pattern. (SMB: to get in the top 1/10 you have to take a big bet!)

11. What puzzle does Table IV address? What is its conclusion?

A: Table IV wants to know how the bad funds can be so bad. In an efficient market, you can’t get a negative alpha either! All you can do is hold too much $\sigma$ (poor diversified). You can waste money in fees and trading, and that’s what Carhart is trying to measure here. Is the remaining spread in alphas due to too much expenses and turnover (paying transactions costs)? Note we’re done with overperformance. Turnover and expense ratios are a bit worse for 9,10, but not enough. p. 65 bottom.

12. Fund managers claim that fees and turnover do not reduce returns to investors. How could charging more money not reduce returns to investors? (Try to be a good salesperson for a high-turnover high-fee fund. Why should I give you my money? Then try to be a good supply-demand economist. What should the equilibrium relationship be between fees, expenses and returns to investors?)
A: the claim is that fees pay for superior ability, and that turnover is losing dogs and buying
good stocks so helps. Standard S=D economics says we should see zero effect. It is a surprise
to see any effect! .

13. (Table V. Make sure you understand how this table was created. How are Table IV and Table
V different?) What does Carhart find about fees and turnover? How much does a 1% change
in fees change returns to investors? How much does turnover – selling one stock and buying
another – change returns to investors?

A: Table V. and p. 67. Table V is based on individual funds, while Table IV looks at the
portfolios of Table 1. Expenses and turnover are all bad. Expenses are more than 1-1 bad. Turnover corresponds to a 0.95bp/transaction cost. (Seems large; more than 1-1 as with expenses). I think this is a big puzzle, we should see zeros here in a market in equilibrium!
You can’t say “investors are dumb” in mutual funds but “investors are smart” in trading!

14. What are the two hypotheses that Carhart spends the rest of the paper distinguishing (see
bottom of p. 70)
A: Read that carefully. Is the performance of funds due to the funds’ stable strategy, or do funds just happen to own stocks that go up in a year?

15. What is the point of Figure 2? (Hint: what would it look like if the sort on one year perfor-
mance indicated skill?)

A: Skill shouldn’t disappear, so the fact that returns seem to revert (2 year return is worse
than 1 year return) suggests momentum in stocks (which does dissipate over time). If it were
skill, the lines would be flat. This is the crucial evidence that it’s momentum in underlying
stocks, not skill. Also p. 71, 80% turnover in top funds each year.

16. What does Carhart say about momentum funds – funds that seem to follow a momentum
strategy, as revealed by high loadings on the momentum factor? (Hint: no table, but text on
p. 73)

A: p.73. Momentum funds – sorted by PR1YR loadings – do not earn higher returns, let
alone four-factor alphas. They lose it all in transactions costs. Hence, chance holders of the
winners get a little boost. This is his evidence that it’s momentum in stocks, not momentum
funds. This also raises a strong suspicion about how easy it is to earn momentum returns in
practice, once you account for trading costs.

17. One year lagged returns are probably mostly luck, not skill. What if you sort funds by the
more common 5 year performance averages? (Hint: Figure 3)

A: Figure 3. The initial expected return spread declines! 5 years is less revealing than one
year! Other components stay the same. There is less effect in the 5 year sort.

18. Does Carhart suggest any trading strategies? (Hint. Look on p. 80 bottom)

241
A: yes: p. 80-81. Buy winning funds for one year to get momentum strategy without trading costs, if you can get in at NAV. But you have to buy a fund of funds to do this, and many of Carhart’s funds have loads too. Also, alas, you can’t short funds.

19. Conclusion

(a) 8% / year spread between 1 year winners and losers
(b) 4.6% from factors, 0.7% from expenses, 1% from transactions costs. Most unexplained comes from 9-10 (worst two). “Cold hands.”
(c) Expenses, turnover, fees lower performance 1-1. Pre-load returns are 80bp lower in load funds!
(d) Buy last year’s winning fund gets you momentum for free (since you buy at NAV, not bid/ask). Funds are starting to charge fees to discourage this.

20. Comments/big take-away points:

(a) Selection bias.
(b) Evaluating the average of a group, identified on ex-ante information, not “why is Warren Buffet rich?”
(c) Evaluate funds by regression of fund returns – how they actually behave not by word “style.” The “portable alpha” problem – If you’re a good bond fund you can become a good stock fund with a very small investment in index futures.
(d) A Performance attribution model – the question is only “can I replicate this performance with passive indices or a computer program, or do I need to pay for stock-picking (or style-picking) skill?”
(e) google shows lots of momentum funds, http://www.gobcafunds.com/portfolios/momentum.asp. Whatever it is, it’s certainly not “overlooked.”
14.6 Fama and French Mutual Fund Performance Q&A.

So far, we have been looking for “skill” by guessing some characteristic associated with skill – past returns, MBA by manager, etc. – and looking at the return of a sorted portfolio going forward. This paper tells us whether there is any skill at all, without us taking a stand on what characteristic can be used to find good funds. It answers the question “sure the average fund is mediocre, but there are some good funds.”

1. What do Fama and French mean by “Equilibrium Accounting? (p. 1915 top)
   A: The average investor must hold the market. Anything else is a zero sum game. (p. 1915)

2. Table 2 tells us about the performance of mutual funds as a whole – it studies the portfolio of all mutual funds. What does it tell us? Do mutual funds as a whole outperform benchmarks? What kinds of indices are most similar to the performance of all mutual funds? Do mutual funds as a whole generate alpha, before or after fees?
   A: Mutual funds as a whole are delivering the market index, with no style deviation towards small or value. The gross returns might have a slight positive alpha – but mostly in small funds – though that is insignificant. Net returns have negative alpha. As a whole they’re just swallowing fees.

3. Fama and French focus on the alpha t statistic. Why not look at alphas or information ratios?
   A: Short lived funds are more likely to deliver big alphas. (1924)

4. Why don’t FF just use the t distribution, to judge how many funds should have alpha t statistics above a certain cutoff? Explain the bootstrap procedure. (Hint, p. 1924)
   A: Well, read the two paragraphs. They’re doing a very careful derivation of the actual distribution of the t statistic in this dataset. When statisticians say “here is the distribution of the t statistic” they are making a lot of assumptions.

5. Explain the numbers in Table 3-4.
   (a) What would they look like if all funds had zero true alpha, but the pattern of luck fully conformed to the assumptions of the t distribution (normal, independent, etc.)?
   (b) What would they look like if there were some funds with +5% alpha and other funds with -5% alpha, so that the average fund was not skilled but some were good and some were bad?
   (c) Why is the probability of a t greater than 2 or less than -2 not the usual 5% value that we expect for a t statistic?
   A: Tables 3-4 shows you how many funds would have (say) 2.0 alpha t statistic if there really were no alpha but some got lucky. If the t distribution were right, 5% would have alphas greater than 2.0. The “sim” tell us the actual distribution which is a little different than the actual t, but that’s why numbers around 2 are at the 9x% tails. Overall, that’s very much what we see. If there were “good” and “bad” funds, the distribution of alphas would be much more spread out than the simulated distribution. For example, if half the funds were truly +5% and half truly -5%, then we would see a distribution with two humps, and a lot more funds with t>2 or t <-2 than the roughly 5% in these tables. See also the figures which make the point visually.
14.6.1 Fama and French Basics – see detailed notes below

• The average fund earns the market, but are there good ones and bad ones? So far, we had to guess a characteristic (past returns) to select them. This paper avoids that step

• You observe funds, some with good returns, some with bad returns. If nobody had any skill at all, how many funds would we see with positive alphas? If we see more than this number, there is some alpha out there. (And some “negative alpha” whatever that is)

• A simple version: Each fund generates returns

\[ R_t^i = \alpha_i + \varepsilon_t^i \]

The sample alphas for a given fund will be

\[ \hat{\alpha}_i = E(R_t^i) = \alpha_i + \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^i \]

Thus, the estimated alpha has a distribution

\[ \hat{\alpha}_i \sim N(0, \sigma_{\varepsilon}^2 / T), \]

and the alpha t statistic has a t distribution

\[ \frac{\hat{\alpha}_i}{\sigma(\varepsilon)/\sqrt{T}} \sim t \]

The t distribution is just a bit wider than normal and corrects for the fact that \( \sigma(\varepsilon) \) is also estimated. This is the good old t distribution.

• Assume the \( \varepsilon^i \) have the same variance for each fund and are uncorrelated across funds. Then the distribution of alpha t statistics across funds should follow a t distribution. Looking across funds is then like “looking across universes,” and we do get to see more than one draw of \( \hat{\alpha} \).

• If true alphas are all zero, we should see no more than 5% of the funds with alpha t statistics greater than two.

• Fama and French: What if the \( \varepsilon \) are correlated across funds, funds follow styles too, and returns are not normal? Answer: do a big simulation. But it’s just a fancy version of the distribution of a t statistic and doesn’t really change things a lot. Notice the 5% cutoffs are pretty near 2.

• Fama and French result: the actual distribution is shifted to the left and just a tiny bit wider than the t distribution centered around zero. That means the average alpha is a bit negative (?) But there are some funds with more or less skill.
Assume true alpha is zero

\[ \text{Luck} = \frac{\sigma}{\sqrt{T}} \]

Simulated distribution of $t(\alpha)$ across funds

What if this were the distribution in the data?

Then we would know it had to come from this distribution of true skill
Plus the same amount of luck

- JC: we can back out the distribution of “true” skill from how much wider the alpha $t$ distribution is than the simulation.

Fama French in graphical form
• It looks grim - the actual looks just like simulation shifted to the left. But look hard, the actual is just a bit wider – there is some skill out there!

• But unwinding t distributions to actual alphas looks better: I have to assume fund age and tracking error
14.7 Berk Q&A

1. What happens to future returns and flows, according to Berk, if a manager does have some skill?
   A: Berk’s central idea: If someone has skill, he can earn better returns. Investors flow in, but all trading strategies have limited scale. The first investors make some alpha, but soon the alpha is driven down to zero. All that happens is that the fund has gotten bigger.

2. Berk says, unlike FF, that managers do have some skill even though alphas are all zero. How can that be?
   A: In the equilibrium above, skill implies larger funds, but no alpha.

3. Berk says that when investors chase past returns, investing in funds that have done well in the past, they are not being irrational, even though future returns are no better than average. How can this be?
   A: Past returns indicate skill. Then investors flow in as above. This drives the future returns back to normal. But it’s not irrational.

4. Berk says that even though skill is permanent, returns will not be persistent. Why not?
   A: Same story. The “true” skill is persistent, but actual returns get eaten up.

14.7.1 Chevalier and Ellison, “Risk Taking by Mutual Funds”

- Background for Berk

- Figure 1 1177 for young funds, and Figure 2 for old funds. Doing well brings money in. Interestingly, you lose fewer customers by doing poorly. As in Carhart, there seem to be some awful funds that keep their customers.

- They see the convexity as an incentive to gamble.

- JC: Is this real? If so, funds would grow forever as luck bounces them from one to the other category.

14.7.2 Berk Comments

- Facts:
  1. Performance at best persists a year (Carhart)
  2. Investors not only choose active management, but they chase past returns. (Chevalier and Ellison)
  3. Managers are paid a lot, in a competitive market.

- Instead of saying “folly” let’s try to explain this as a normal competitive market

- 5 Hypotheses that we have taken for granted
  1. Returns (alpha) to investors measure skill
2. Average returns (alpha) do not beat the market so the average manager is not skilled.
3. If a manager has skill, returns should persist.
4. Since returns don’t persist, investors who follow past returns are irrational.
5. Compensation based on assets under management doesn’t reward performance. (JC: This is pretty obvious)

- Suppose you have skill, you have 6% alpha and can run $10m. Above that, you have to index because your strategy can only work at a $10m scale. You charge 1%
  1. Year 1: You have $10m AUM, you earn 6%, return 5% to investors.
  2. Investors see the good return and flock in. Year 2 Week 1. You’ve got $20m AUM. You earn 6% α on first $10m, 0% α on the second $10m. Thus overall you earn 3% α, charge 1%, return 2% to investors. More join.
  3. Where does this stop? Year 2 week 2. You have $60m AUM, you still earn 6% α on the first $10m, indexing the remaining $50m. You charge 1% of $60m. Your investors now get 0% (The index). Still, there is no reason for them to leave, and if any do, returns rise again.

- What do we see in this example
  1. Skill lasts forever
  2. Alpha to investors dies out after 2 weeks.
  3. In equilibrium, alpha is uncorrelated with skill.
  4. Skillful managers get paid a lot.
  5. Investors are perfectly rational to chase past performance.
  6. You are able to “raise your fee” to 5%, even though the stated fee is always 1%.

14.8 Extras from optional papers

14.8.1 Davis, Fama and French.

1. Most funds tilt towards small and growth. Even the 1/10 deep value funds aren’t that value.
2. Post ranking betas are a lot less dispersed than pre-ranking betas. Betas are harder to measure than you think.
3. A positive alpha at last, and almost significant! 13 bp per month, in the (pretty small) growth portfolio! (Since small growth underperforms the FF3F model, this is even better than a “characteristic adjustment” would suggest.)
14.8.2 Chevalier and Ellison, “Are some managers better than others?”

1. We’ve sorted on past return. How about sorting on who the managers are – MBA, SAT?

2. Table II p. 882. Yahoo! MBA = 63bp (p. 881) But is it style or skill? And of course the $R^2$ is pretty low.

3. Table III. Being smart and having an MBA means you have higher beta, lower turnover, lower fees!
4. Table IV, first two columns. Having an MBA raises your returns... Oh well, not your alpha – you get higher returns just because you took higher beta. SAT still has a small effect. Maybe

14.8.3 Cohen, Frazzini and Malloy “Small world”

1. Table III. Divide stocks held by a mutual fund into those where the fund manager and company CEO went to school together vs the others. Surprise, they do better in the stocks of their friends!

14.8.4 Carhart, “Leaning for the tape”

1. Table 1. Fund returns on last day of the year/first day of the next year.
2. The effect is strongest in small cap growth.
3. The next day return being lower suggests it’s not just luck, year end effects.
4. Figure 2 p. 667 the effect is strongest for funds near the top
5. Carhart interpretation: funds mark aggressively, and also buy their own holdings at the very end of the day to push up NAV of the stuff they already have

![Diagram of financial market with labels: Spuriously high price, Large return at t, Low return at t+1]
14.9 Fama and French Mutual Funds notes

Why the Fama-French simulation works to detect skill, even without knowing the characteristics of skill.

The genius of the Fama-French simulation is that it lets us see if there are some “skilled” managers without trying to identify which funds are skilled – without trying to identify a characteristic (past returns, past alphas, Morningstar rating, MBA of manager, etc.) that we think is associated with skill, form portfolios, and watch returns going forward. Of course that also means we don’t learn how to identify those skilled managers either.

How does it work? Consider a very simple example – a world of Bernie Medoffs; well, of the guy Bernie was pretending to be. Suppose there are two managers, one with positive alpha who always generates 5% returns, and one with negative alpha, who always generates -5% returns. Now, look at a sample. Subtract from each fund its sample mean — 5% in the first case, -5% in the second case – and simulate what’s left over. There’s nothing left over, so in this “no skill” (true alpha = 0) simulation, the distribution of estimated alphas should always stick at 0. The fact that we see “fat tails” in the distribution of estimated alphas tells us that there is some skill and some “negative skill.”

More generally, the amount of return variability over time tells you how much return could possibly be generated by luck. This observation is really just the standard error of the mean. If returns vary by \( \sigma(R) \) over time, then the mean return should vary by only \( \sigma(R)/\sqrt{T} \) across funds. This time we get to see lots of draws, unlike the usual application of standard error of the mean where you only see one draw. If mean returns vary across funds by more than luck can account for, then we know there is some skill. By looking at how much fatter the actual distribution of mean returns is than would be generated by luck, we can find the distribution of true underlying skill.

Fama and French don’t look at mean returns. First, they adjust for factor exposures; for each fund they run the usual regression \( R_i^t = \alpha_i + b_i r_m f_t + b_i hml_t + s_i smb_t + \varepsilon_i^t \), and they look at the estimated \( \hat{\alpha}_i \), and save the residuals \( \hat{\varepsilon}_i^t \) (hats mean estimates). Then they do a bootstrap: They keep the \( \hat{\varepsilon}_i^t \) of all the funds at each moment in time together and re-sample months. (We’ll do a bootstrap later if this is confusing.) This controls for the fact that returns might not be normally distributed, and by resampling the entire month of returns for all funds together they capture the fact that returns might be correlated across funds, so each fund is not really a separate experiment. Finally, they evaluate each fund by its t statistic \( \hat{\alpha}_i / \left[ \sigma(\varepsilon)/\sqrt{T} \right] \). This is a good idea. If we just looked at the distribution of alphas, some funds that are alive for shorter times or have larger tracking error \( \sigma(\varepsilon) \) are more likely than others to achieve big alphas by chance, and to figure out “what is the chance of seeing so many big alphas” we’d have to keep track of all that. But these are refinements on the basic idea.

Here’s a simple example. Ignore factors and suppose returns are normal. Fund returns are generated by fund-specific skill \( \alpha_i \) (which is constant over time – this is the key to measuring it) and luck,

\[
R_i^t = \alpha_i + \varepsilon_i^t
\]

Assume the \( \varepsilon_i^t \) have the same variance for each fund and are uncorrelated across funds, to make the example simple.

Now, let’s think about the distribution of average returns = estimated alpha. The estimated
alphas for a given fund will be

\[ \hat{\alpha}_i = E(R_i) = \alpha_i + \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^i \]

Thus, the estimated alpha has a distribution

\[ \hat{\alpha}_i \sim \alpha_i + N(0, \sigma^2_{\varepsilon}/T) \]

These estimated alphas capture some skill as well as some luck. This is our standard error of the mean formula. It means “if you ran this fund over and over again, the mean would fill out this distribution.” But since the funds are the same in this example, it also means “this is how much variation we expect across funds.”

If there is no skill, then, we expect to see a distribution of estimated alphas across funds that has variance \( \sigma^2_{\varepsilon}/T \).

Now, what if there is some skill? Skill is also distributed across funds. To keep it simple, I’ll write that the distribution of alphas across funds is also normal,

\[ \alpha_i \sim N(\mu_\alpha, \sigma^2_{\alpha}). \]

This is the distribution of true skill across funds; I assume that skill is picked once and for all at the start of the sample and stays constant over time once picked. Since the sum of normals is normal, and the subsequent luck is independent of the true alphas, the distribution of estimated alphas across funds will now be

\[ \hat{\alpha}_i \sim \left[ N(\mu_\alpha, \sigma^2_{\alpha}) + N(0, \sigma^2_{\varepsilon}/T) \right] = N(\mu_\alpha, \sigma^2_{\alpha} + \sigma^2_{\varepsilon}/T). \]

We have a normal distribution of estimated alphas. Some of the variation across funds comes from luck (\( \sigma_{\varepsilon} \)) and some from skill (\( \sigma_{\alpha} \)). You can see how adding some distribution of true skill fattens up the distribution of estimated alphas across funds. If we see too many good funds and too many bad funds, we can infer that there is some underlying skill, though we still cannot tell whether a particular fund was skilled or lucky.

Let’s see how the FF procedure separates these two components. They take out the sample mean, so they form

\[ e_t^i = R_t^i - \frac{1}{T} \sum_{t=1}^{T} R_t^i = \alpha_i + \varepsilon_t^i - \frac{1}{T} \sum_{t=1}^{T} (\alpha_i + \varepsilon_t^i) = \varepsilon_t^i - \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^i \]

Notice that this always removes the true alpha, and leaves us a series that is just demeaned luck. Reshuffling (simulating) the \( e_t^i \), their simulation delivers the predicted cross-sectional distribution of estimated alphas, under the assumption that true alpha is zero for everyone,

\[ \hat{\alpha}^{sim} = \frac{1}{T} \sum_{t=1}^{T} e_t^i \]

Using the standard error of the mean logic, the variance variance across funds of this quantity is

\[ \sigma^2(\hat{\alpha}^{sim}) = \sigma^2 \left( \frac{1}{T} \sum_{t=1}^{T} e_t^i \right) = \sigma^2 \left( \varepsilon_t^i - \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^i \right) = \frac{(T-1) \sigma^2_{\varepsilon}}{T} \]

252
Compare this to the actual distribution of estimated alphas, from above, with variance

$$\sigma^2(\hat{\alpha}_i) = \sigma^2_\alpha + \frac{\sigma^2_\varepsilon}{T}. $$

If there were no real alpha, then the simulation would give the variance (and hence distribution) of the actual estimated alphas. The simulation would be a little narrower, because of the $(T - 1)/T$ term. This should be small, (and FF could correct for it.) But if there is substantial true alpha, then the distribution of the simulated alphas $\sigma^2(\hat{\alpha}_{sim})$ will be much narrower than the distribution of the true alphas, with variance $\sigma^2(\hat{\alpha}_i) = \sigma^2_\alpha + \sigma^2_\varepsilon/T$. We would see a lot more large alphas then there should be by chance.

Here is a picture that will make it all clear. The top graph shows the simulation. Assume all funds have zero alpha, and by simulation (or standard error of the mean), we see what the distribution of estimated alphas should be. Now suppose we saw a wider distribution as in the second graph. There is only so much luck one can have (that was the Bernie Medoff tipoff – if returns are incredibly stable over time, it’s impossible to have that much good luck.) Thus, we know that the distribution of observed fund alpha t statistics had to come from a distribution of skill (green) plus the luck.

For example, look at Table 4, top left block. The 5 Pct and 95 Pct numbers mean this on the graph, and as I eyeball the results it looks something like the following
Rather depressing overall.

**Beyond Fama and French**

This discussion says we can go beyond what FF did, and directly estimate the distribution of alphas across funds. If the distributions are normal, the variance of the actual distribution of estimated alphas minus the variance of the simulated distribution of estimated alphas is a direct estimate of the variance across funds in true alpha. Of course, the mean of the distribution of estimated alphas is the mean of the distribution of true alphas which we knew all along – “the average fund underperforms the market.”

It’s also possible to directly estimate the distribution of true alphas. The distribution of alpha t statistics comes from luck and skill as follows.

\[ f(t_\alpha) = \int_{-\infty}^{\infty} f(t_\alpha | \alpha) \times f(\alpha) \, d\alpha \]

Finding the unknown \( f(\alpha) \) that “fattens up” \( f(t_\alpha | \alpha) \) just enough is easy if you do the probabilities at discrete points

\[ f(t_{\alpha i}) \Delta t_{\alpha i} = \sum_j f(t_{\alpha i} | \alpha_j) \Delta t_{\omega i} \times f(\alpha_j) \Delta \alpha_j \]

This is a matrix multiplication so we can find \( f(\alpha_j) \) as a simple matrix inversion. (Well, keeping it numerically stable is hard, but conceptually, this is just a matrix inversion.)

The simulation tells us \( f(t_\alpha | \alpha = 0) \). In general, we would have to do a lot of simulations to find \( f(t_\alpha | \alpha) \)– assume a new value of \( \alpha \), then simulate again, and so forth. However since luck is independent of skill, \( f(t_\alpha | \alpha) \) is the same as \( f(t_\alpha | \alpha = 0) \) but just shifted over to the right. (This procedure uses the same assumption as in the simulation, that tracking error and fund life are independent of alpha.)

Alas, from the numbers in Fama and French’s paper we can only recover the distribution of the “true \( t_\alpha \)” meaning the true alpha divided by the standard error of estimated alphas. To go back from true \( t_\alpha \) to “true alpha” we need to know \( \sigma(\varepsilon) \) and \( T \). Fama and French can easily calculate it, but we don’t have those numbers. I make some calculations below.

In sum, if we write the sample as \( \hat{t_\alpha} = \hat{\alpha} / (\sigma(\varepsilon) / \sqrt{T}) \) and the true alpha divided by standard error as \( t_\alpha = \alpha / (\sigma / \sqrt{T}) \) then we have

\[ f(t_{\hat{\alpha i}}) \Delta t_{\hat{\alpha i}} = \sum_j f(t_{\hat{\alpha i}} | t_{\alpha j}) \Delta t_{\alpha i} \times f(t_{\alpha j}) \Delta t_{\alpha j} \]
We can figure out the \( f(t_{\hat{\alpha}}|t_{\alpha}) \) from the simulation

\[
f(t_{\hat{\alpha}}|t_{\alpha}) = f(t_{\hat{\alpha}} - t_{\alpha}|t_{\alpha} = 0)\]

\[
f(t_{\hat{\alpha}}|t_{\alpha}) \Delta t_{\hat{\alpha}} = [f(t_{\hat{\alpha}} - t_{\alpha}|t_{\alpha} = 0) \Delta t_{\alpha}] \frac{\Delta t_{\hat{\alpha}}}{\Delta t_{\alpha}}
\]

and now we can invert to find \( f(t_{\alpha}) \).

To illustrate, I will make graphs corresponding to FF’s Table 4, bottom left panel. These are four-factor alphas.

Fama and French tabulate the cumulative distribution. To plot it as a density, I made histograms using the midpoints. For example, if the 95% value happened at \( t_{\alpha} = 1.5 \) and the 96% value happened at \( t_{\alpha} = 1.6 \), I drew a histogram with 1% probability at \( t_{\alpha} = 1.55 \). Of course, FF could easily calculate histograms as well. The first figure presents that result.

You can see that the actual distribution of \( t_{\alpha} \) is shifted to the left and somewhat spread out compared to the distribution we expect if all true alphas are zero.

The table below gives the moments. Under the simulation, the mean alpha t-stats are zero (up to simulation error) and the standard deviation across funds is 1.18. This is what we would observe if there were no skill at all and all alphas were zero. The actual distribution is shifted to the left, with a mean -0.53, and a wider distribution with standard deviation of 1.32. Assuming normality, then, we can immediately infer that the distribution of the “true alpha” t-statistic (true alpha/(\( \sigma(\varepsilon)/\sqrt{T} \))) has a mean of -0.53 and a standard deviation across funds of 0.60.

<table>
<thead>
<tr>
<th>sim. ( \alpha = 0 )</th>
<th>actual</th>
<th>true ( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(t_{\alpha}) )</td>
<td>0.03</td>
<td>-0.53</td>
</tr>
<tr>
<td>( \sigma(t_{\alpha}) )</td>
<td>1.18</td>
<td>1.32</td>
</tr>
</tbody>
</table>

For the other calculations, it’s easier for me to infer the histogram on evenly spaced values on the x axis. (This isn’t really necessary, but it really helped the programming for this quick example. Also, FF can easily calculate these if they want to.) I did this by smoothing across the histogram.
The next figure shows you the FF “histogram” and the histograms I use with a finer and even set of breakpoints. (This is a “kernel density estimate.”) As you can see, it’s a decent approximation.

![FF distributions and interpolated finer histograms](image)

Before doing the “nonparametric” (i.e. histogram) of the true alpha distribution, I plot here the normal value. The “true $t_\alpha$” here has normal distribution as calculated above with $\mu = -0.53, \sigma = 0.60$. The “simulated with varying $t_\alpha$” line then is formed from the simulation multiplied by this “true $\tau_\alpha$” $\int_{-\infty}^{\infty} f(t_\alpha | t_{\alpha}) \times f(t_\alpha) \, dt_\alpha$. This calculation comes quite close to the “actual” which is the actual distribution of $t_\alpha$ across funds. The fact that blue and red lines are so close means that the normal distribution of true alphas is an excellent approximation. (I used these as starting points for a more accurate calculation that finds $f(t_\alpha)$ that will exactly match the actual distribution, but that calculation isn’t working yet.)
OK, now we have the implied distribution of true alpha t-statistics. What does this mean about true alphas? I don’t know average fund ages and tracking errors, but we can make some simple assumptions. Here are the implied distributions of true alphas for $T = 10$ years and $T = 5$ years, and various tracking errors.
These distributions wider than you might have thought, given that the distribution of fund alpha t statistics didn’t look much wider than the simulated distribution. However, we add squares of standard deviations. It took a good deal of alpha t statistic to widen up the fund distribution; $0.60^2 + 1.18^2 = 1.32^2$. 