

基礎研 レポート

パフォーマンス評価 評価基準は常識なのか迷信なのか

金融研究部門 上席主任研究員 暹澤 秀一
(03)3512-1848 chizawa@nli-research.co.jp

1—はじめに

AIJ投資顧問問題が発覚した際、公表されたパフォーマンス・データは「良過ぎてありえない」とかねてより思っていたとの声が資産運用業界の中で続出した。だが、運用評価機関がAIJ投資顧問を顧客に推奨しなかったのは情報開示姿勢に問題ありとしたからであって、過去のパフォーマンスが理由ではなかった。統計的にほとんど起こりえないということと、実際に起こったか起こらないかは別の話である。パフォーマンスが「良すぎる」ことが問題だということと、思考停止してしまえば、ありうる程度の数値に捏造されたデータを見抜けないことになる。また、本当にスキルのある運用会社を見落とすことにもなる。

米国においてもバーナード・マドフ事件という、一説によれば被害総額 650 億ドルに上る詐欺事件が起こったことがある。このような事例は例外的ではないかと思われるかもしれないが、2003 年には米国のミューチュアル・ファンド業界を舞台に不正問題が発生した。これは詐欺事件とは異なりフロント・ランニング等のスキャンダルであったが、資産運用会社が顧客重視の姿勢を貫いているのか、という運用会社としての原点そのものが問われるものであった。それを受けて 2004 年に投信評価会社である Morningstar 社は Stewardship grades というファンドの定性評価を開始した。それまでは過去のパフォーマンスによってファンドのレーティングを付与していたのに対し、経営陣の質や企業文化などを含む5項目から成っている、新たなレーティングを創設したのである。2004 年の公表以降、データの蓄積が進んだこともあり、米国では Stewardship grades によるファンドの定性評価とパフォーマンスとの関係の実証分析が行われている。

運用のパフォーマンス評価に関しては、運用評価会社やコンサルティング会社などから多くの情報が顧客に提供されているが、評価および評価方法の問題点や将来のパフォーマンスとの関連性については明確でないことも多い。そこで本稿では、海外の実証分析結果の概略を紹介しながら、パフォーマンス評価(定量評価・定性評価)にまつわる論点を整理することを試み、パフォーマンス評価に対してどのような姿勢で臨めばよいのかを考えたい。

1 | 定量評価の問題点

過去のリターンに基づく定量評価はパフォーマンス評価の基本データである。しかし、そこには限界がある。その限界が発生する理由として下記のことが考えられる。第一にリスク調整後のリターンを計算するのに、どのような指標を使うかという問題がある。株式投資の場合、リスクの捉え方の背後にあるモデルとして、CAPM、Fama-French の3ファクター・モデルやそれを拡張した4ファクター・モデル等があるが、実務的にはシャープ・レシオ(Sharpe Ratio)が使われることが多いようだ。一方でアカデミックな検証では、複数の指標を比較することが多い。第二にデータ数(トラック・レコードの長さ)の問題がある。パフォーマンスを統計的に検定するのは、運とスキルを区別することが目的だ。つまりパフォーマンスが良かったのは、たまたまなのか運用能力が高いのかを判別したいからである。ところが、統計的に有意な結果が得られるだけのデータ数を確保できないことが実務的には多い。そうするとパフォーマンスを定量評価しても、大半のファンドでは最初から統計的に有意な結果が得られないのが普通になる。逆に有意な結果が得られるファンドはたまたま投資戦略が経済環境にマッチしただけで高いリターンを得たということも起こりかねない。第三は再現性の問題である。投資戦略の成果は経済環境に依存するので、結局のところ、投資成果は一定の条件下における結果でしかない。経済や金融の状態は絶えず変化し、構造変化も起こりうる以上、良好なパフォーマンスを収めたファンドがその後も高いリターンを得るという保証はどこにもないのである。

特にシャープ・レシオの推計誤差について論じたものに Lo(2002)がある。シャープ・レシオとは、超過収益率の平均値を標準偏差で割ったものである。分母・分子とも直接観察できないので推計誤差を持ち、その商であるシャープ・レシオも推計誤差を持つことになる。要するにシャープ・レシオの正確さは収益率の特性に依存することになり、投資戦略や投資環境の影響を受けることになる。一般的に言えば、リターンのぶれが大きい投資戦略を採用するポートフォリオは、ぶれが小さいポートフォリオよりもパフォーマンスを正確に測定するのが難しくなる。また、投資戦略の特性(平均回帰、モメンタム、系列相関)もシャープ・レシオの推計に影響を与える。特にヘッジファンドは系列自己相関が高いので、通常、パフォーマンスが過大評価されがちになる。月次で収益率を計測して得られたシャープ・レシオに $\sqrt{12}$ を乗じて年率換算することは良く行われるが、これは特別なケースを除けば厳密には成立しない。投資戦略と投資環境がマッチしている時期と mismatches の時期とでは、リターンの水準が異なるだけでなくリターン特性も異なるので、過大あるいは過小評価が起きる。

2 | 意思決定が時を刻む

思考実験のため極端な例を考えよう。あるポートフォリオ・マネジャーが、新規資金でポートフォリオを構築・買い付けした直後に交通事故にあい、意識不明のまま生死の淵をさまようことになったとする。運用組織は彼の死亡が確定するまで、当該ファンドの担当者を彼のままにしておくことを決定した。30年後、奇跡的に意識を取り戻したポートフォリオ・マネジャーが、ファンドのパフォーマンスを確認してみると、統計的に有意にベンチマークを上回っていたとする。この場合、彼は運用ス

キルがあると言えるのだろうか。常識的には単に運がよかっただけではなかったか、と思う人が大半であろう。投資においては現状維持や当面静観という行動方針を決めることも意思決定の一つと考えられるが、このケースではそれすらしていないからだ。しかし何の付随する情報を与えずに、運用評価機関に30年間のトラック・レコードを与えると、定量評価では運用スキルが優れているという評価になってしまうことも起こりうるだろう。この場合、30年という期間は単に物理的に時間が流れただけであり（ニュートン時間）、投資においては意思決定の度に時を刻む（内的時間の存在と物理的時間との不一致）と解釈して、十分なデータがないと見なす方が合理的だろう。だがそうなると、一定期間、一定間隔でパフォーマンスを評価しても、リバランスやトレーディングの頻度やタイミングがファンドそれぞれで異なるため、その背後にある意思決定の頻度やタイミングも異なることになり、厳密な意味でのパフォーマンス評価は不可能だということになりかねない。理論はさておき実務的には、パフォーマンス評価基準に即した運用を行っているファンドを評価・比較していると解釈すべきだろう。

また、意思決定は単一階層だけで行われるのではなく、重層的・連鎖的になされる。たとえば投資プロセスがトップ・ダウン型の場合、まずマクロ経済環境や市場環境の判断があり、それからセクターや個別銘柄の投資判断へと流れていく。そのとき、上位レベルの判断と下位レベルの判断とは無関係ではありえない。ここで問題となるのは、下位レベルの意思決定よりも上位レベルの判断の方が回数は少なくなることである。ある特定期間について下位レベルの投資判断が的確であったとしても、それは上位レベルの判断能力や洞察力があることを証明するものではないし、判断回数が少ない上位レベルを統計的に検定することは難しい。López de Prado(2008)はヘッジファンドのリターンシリーズの問題を論じ、十分なトラック・レコードが得られないときはサンプリング頻度を上げて必要期間を短くすることを提唱している。しかし前述の通り、上位レベルの意思決定回数は変わらないので、下位レベルでの統計値はともかく上位レベル（経済的・ファンダメンタルズ）の統計的信頼度を高めることにはならないと考えられる。

3 | 定量評価における外れ値

安定的に優れたパフォーマンスを出し続けるのはむしろかしいという主張に対して、よく反例として挙げられるのが、米国の富豪ウォーレン・バフェット氏である。彼の投資会社バークシャー・ハサウェイの運用実績は米国の代表的な株価指数である S&P500 を長期間に亘って凌駕していることで有名だ。だが、その事実の解釈として、大勢でじゃんけんトーナメントを行えば、誰かが勝ち残ることになるので、多くの投資家の中には長期間高いリターンを挙げた人間がいたとしても不思議ではないと言われている。また、バフェット氏は株を大量保有しているので、一般投資家よりも経営者と直接接する機会も多く情報面で優位に立っている点を指摘する人もいる。

だが、ここでは前節を踏まえて別の解釈を挙げよう。1991年にソビエト連邦が解体され、旧社会主義陣営の国々が資本主義化された結果、西側・東側に分断されていた世界市場が統合され、拡大されることになった。たとえば、コカコーラといえばバフェット銘柄として有名だが、市場拡大の恩恵を最も受けた銘柄であることは間違いない。米国にとってゴールデン'60sと言われ繁栄を極めた1960

年代から、石油危機やベトナム戦争によって国力を消耗した 1970 年代を経て、1980 年代になると大きな家、大きな車、コーラ、ハンバーガー、ハリウッド映画という米国のライフスタイルのあらが西側諸国にも見え出した。そこに旧共産圏という手付かずの巨大市場が現れたのである。1990 年代以降は米国のブランド企業はその恩恵を十分に受けることができた。だが、潤沢なキャッシュフローとブランド価値を重視するバフェット流のボトム・アップ銘柄選択と、旧ソ連崩壊による平和の配当を受益して安定成長できることを予測できたこととは直接結びつかない。つまり、洞察力による意思決定であったとは限らないし、また、最上位レベルの意思決定の回数は一度と言えるのかもしれない。その場合、この 20 年間のパフォーマンス評価はどうあるべきなのかという問題は残っている。

4 | 小括

パフォーマンス評価の基本は過去のリターンに基づく定量評価である。しかし十分なトラック・レコードが得られないことが多いこと、技術的問題（リスク調整、リターン・データの特性補正）もあること、経済や市場の環境に依存し再現性があるわけではないことなどの限界がある。そのため定量評価を補完・補強するため定性評価も合わせて判断することが重要になる。

3——定性評価について考える

1 | なぜ定性評価が重要なのか

前章で述べた通り、定量評価には限界がある。このことは広く知られているため、定量評価を補完するため定性評価も重視して資産運用会社やファンドを評価することが半ば常識化していると言えよう。だが、定性評価がパフォーマンスと果たしてどこまで関連付けられるのかという点に関しては、確固たる根拠があるわけではない。また、定性評価の方法も運用評価機関やコンサルティング会社によって差異があると考えられる。

冒頭に述べたように、米国の投信不正問題を契機に、2004 年に投信評価会社である Morningstar 社は定性評価基準として Stewardship grades の公表を始めた。このデータの蓄積を利用して、米国では実証分析の土俵に乗りにくかった定性評価の検証が行われるようになった。これは運用会社の「顧客重視」の姿勢が、パフォーマンスに反映しているかを調べるものである。

またそれとは別に、米国では資産運用会社の人材の質とパフォーマンスとの関係を調べる研究も盛んに行われている。運用評価機関の定性評価では、ファンド・マネジャーやアナリストの経験年数、定着率、学歴、資格等を評価項目としていることが多い。しかしこれらがファンド・マネジャーやアナリストとしての優秀さを果たして担保するものなのかどうかは、いまだ実証されていないのが実情である。

本章では海外での実証分析結果を紹介しながら、人材の質、経営陣の質、企業文化等がパフォーマンスに直結しているのか、資産運用会社の質を担保するものなのかについて考えてみたい。

2 | 人材の質

アクティブ運用においてファンド・マネジャーの質がきわめて重要であることは言うまでもない。運用会社から顧客に提供される資料の中には、担当ファンド・マネジャーの略歴が提示されることも少なくない。また、運用評価機関が運用会社に求める情報の中でも、ファンド・マネジャーやアナリストの経歴は重要な位置を占める。運用組織のファンド・マネジャー評価項目として一般に注目されるのは、経験年数と定着率だろう。それは経験豊富なファンド・マネジャーの方が優れている、少なくとも業界内で生き残ってきただけの実績と能力を持っているという評価が下されるからである。また従業員の高い定着率は、待遇や企業文化の面を含めて組織上の問題点が少ないとの評価と結び付く。それ以外にも学歴や資格なども評価対象に含まれるが、ファンド・マネジャーとして望ましい属性は何かは判然としていないし、実証された根拠に乏しいのが実情である。たとえば、ベテランの方が若手よりも本当に優秀なのだろうか、教育や資格はファンド・マネジャーとしての優秀さを意味するのだろうか。

人材の質あるいは属性と、投資パフォーマンスとの関連性については、大別してファンド・マネジャーを対象にした実証分析と、広く一般投資家を対象にした研究とがある。いわばプロに絞った検証と、アマを中心にした検証である。まず、ミューチュアル・ファンドやヘッジ・ファンドのマネジャーに関する米国の実証研究結果を紹介しよう。

米国ではミューチュアル・ファンドのファンド・マネジャーの属性とパフォーマンスとの関係に関する実証分析が行われているが、結果に関して必ずしもコンセンサスが得られているわけではない。検証対象となる属性としては、学歴、資格（証券アナリスト資格）、性別、経験年数等が挙げられるが、時期、データの取り扱い方法（途中で検証対象から外れるファンドの影響、リスク調整方法等）や投資スタイルによって逆の結果が得られることも珍しくない。

まず、Chevalier and Ellison(1996)の結果を紹介しよう。この論文では、ミューチュアル・ファンドのパフォーマンスをファンド・マネジャーの属性（年齢・学歴・在任期間・リスクの取り方・投資スタイル等）によりクロスセクションで比較している。結論をかいつまんで学歴と年齢に分けて述べよう。なお彼らのいう学歴には、大学・学部のランキングも含んでいる。

SAT (Scholastic Assessment Test、大学進学適性試験、Critical Reading、Writing、Mathから成る)の高いスコアが必要な大学の出身の方がそうでないファンド・マネジャーよりもパフォーマンスが良く、手数料が低く、規模が大きいファンドの運用をしているという傾向が見られた。MBA取得者のパフォーマンスもよいが、彼らは高ベータ・ファンドを運用している。パフォーマンスを比較する場合、リスクの取り方にも注意を払う必要がある。高いリスクを取った結果として高いリターンを得たのであれば、運用スキルがあるとは言えないからである。MBA取得者の場合、リスクを調整すると超過収益率は消えるという。

また年齢に関しては、若いファンド・マネジャーの方がベテランよりもパフォーマンスが良いという結果を得ている。ベテランは手数料が高く、規模が小さいファンドの運用をしている傾向があるという。投資信託のパフォーマンスを比較する場合、運用手数料とファンドのリスクを考慮する必要がある。これらの結果は、手数料とリスクを考慮しても変わらないとの分析だ。

高SATスコアを要する大学の出身者のパフォーマンスが良い理由として、資質に恵まれ、良い教

育を受けていることが考えられる。しかし、そもそもファンド・マネジャーの採用に当たっては、優秀な卒業生を選考しているはずなので、むしろ縁故採用されたファンド・マネジャーの影響ではないかと、彼らは指摘している。

この結果が日本にも当てはまるかどうかについては、資産運用業界の状況の相違と大学入試制度の相違の2点を考慮すべきだろう。前者に関しては、米国の場合、運用会社の数が日本よりも桁外れに多く、規模、経営体制、質が千差万別である。悪く言えば玉石混交であり、ファミリー・ビジネスに近い運用会社も見られるため、ファンド・マネジャーの質も日本よりも差が大きいと考えられる。また、後者に関しては、米国では入学者選考における学力の比重が日本より低いこと、SATも高校での履修科目の習得度というよりも大学での学業に対する基礎的資質を見ることに主眼が置かれていることが異なる点として挙げられる。SATスコアが高い大学というのは日本で言えば偏差値の高い大学に相当するというイメージを持たれるが方もおられるかもしれないが、これらの点を勘案すると、この実証分析結果がそのまま日本に当てはまるとは限らないことに留意すべきである。

一方で、属性の異なるファンド・マネジャーは、異なる属性のファンドを運用しているという。たとえば、高SATスコアを要する大学の出身者は、手数料率が低く、サポート・スタッフも充実し、意欲を高めるインセンティブが用意されているファンドや運用会社で働いている傾向にあり、それが高いパフォーマンスの一因だと考えられるとしている。働きやすく、モチベーションを高く保てる環境で働く方が高いパフォーマンスを実現できるというのは、日本にも当てはまると思われる。その意味で、運用会社の環境、企業文化等を定性評価することは、一般的には有意義だと考えられる。

また、ファンド・マネジャーの年齢に関しては、若い方がベテランよりも好成績であったというのは興味深い結果であろう。著者らは、ベテランは雇用面でより強固な立場にあるため働く動機付けが薄れている可能性や、教育が時代遅れになり能力面で若手に凌駕される可能性を指摘している。それとは別に、優秀なファンド・マネジャーはミューチュアル・ファンド業界から若いうちに退出してしまいがちだということも理由の一つに挙げている。たしかに米国では、実績を挙げたミューチュアル・ファンド・マネジャーがヘッジ・ファンド業界に転出することも多いので、このような傾向があっても不思議ではないかもしれない。

つぎに、Gottesman and Morey(2006)を紹介しよう。この論文ではMBA取得者を入学選考に用いるGMAT (Graduate Management Admission Testの略、Analytical Writing Assessment、Quantitative、Verbalから成る)のスコアによって分けて(Business Weekのランキングに基づく)いる。ランキング最上位のMBAプログラム取得者はそれに準じる取得者よりもパフォーマンスが良いが、その差はわずかであった。しかしランキング下位のMBA取得者はMBA非取得者よりもパフォーマンスが悪かったという。また、Chevalier and Ellison(1996)の結果とは異なり、SATスコアによる相違は見られなかった。しかしパフォーマンスの悪いファンドが途中で検証対象から外れることによるバイアスを修正すると、SATスコアとパフォーマンスとの間には相関が観察された。他の教育プログラム、たとえば、CFA (米国証券アナリスト)、リベラル・アーツ教育、MBA以外の修士号、博士号は好パフォーマンスにつながらなかったという。

ただしファンド・マネジャーの属性とパフォーマンスとの関連性が明確に出た結果ばかりではない。

Switzer and Huang(2007)では、経験年数、性別、MBA取得、CFA保有等を説明変数にして調べたが、条件によって必ずしも一貫性のある結果が得られていない。

また、Dincer, Gregory-Allen and Shawky(2010)は経験年数、CFA保有とMBA取得に的を絞った実証分析を行った。その結果、これらはファンドのリターンに明確な影響を与えないが、リスクで相違があるという。経験年数が長くなるとポートフォリオのベータを低めに抑えるようになり、CFA保有者はMBA取得者よりもベータでもトラッキング・エラーでも低リスク運用になった。

教育プログラムとパフォーマンスとの関係に関する実証分析で、ミューチュアル・ファンドではなくヘッジ・ファンドのマネジャーを対象にした研究結果を紹介しよう。Li, Zhang and Zhao(2009)は、ヘッジ・ファンド・マネジャーの属性とパフォーマンスとの関係を分析した。高いSATスコアを要求する大学出身者のパフォーマンスはリスク調整後リターンで比較すると優れており、リスクも小さい傾向があり、ファンドへの新規資金も多いという結論を得ている。

ミューチュアル・ファンドとヘッジ・ファンドとでは、ファンド・マネジャーのインセンティブ構造が異なることを明らかにしたのが、Boyson(2003)である。ミューチュアル・ファンド業界ではファンド・マネジャーは経験を積むほどリスクを取るようになるのに対して、ヘッジ・ファンド業界では逆に若手がリスクを取り、ベテランほどリスクを低下させるという。ミューチュアル・ファンド業界では、若手は経験不足で自信がないだけでなく、雇用面でのリスクを取りたくないため他のマネジャーと類似の投資行動を取りがちで(Herding)、リスクを抑制しがちであるという。一方、ヘッジ・ファンド業界では若手はファンド規模が小さいため、定率手数料だけでは運営できないため成功報酬を求めて高いリスクを取る傾向にある。首尾よく高いリターンが得られれば、成功報酬が増加するだけでなく、ファンドへの資金流入が期待できるので、高リスク運用へのインセンティブが高くなる。ところが、ベテランになるとファンド規模が大きくなり、定率報酬部分の金額も大きくなる。また、自分の資産をファンドに投入して運用していることもあり、リスクを低減させるインセンティブが働く。つまり個人資産だけでなく業界内での評判も含めて、失うものが大きくなるため守りに入りがちになるという。またベテランになると再起して新ファンドを立ち上げるのも難しくなるというのも理由の一つとして挙げられる。そのため、ヘッジ・ファンドでは若手の方がベテランよりもパフォーマンスが良く、その差の半分はリスクの取り方で説明できる。若手の方が高いリスクを取ることの代償としてファンドの生存率は、最初の6年間で35%だという。その次の6年間の生存率は9割近いことを考えると、ファンド創設期がいかにハイリスク・ハイリターンであるかがわかる。この実証分析結果から示唆されるのは、報酬体系とファンド・マネジャーの属性によって、ファンド・マネジャーのインセンティブが大きく異なり、また、リスクに対する姿勢も変わるので、顧客サイドとしてもこれらに無関心ではいけないということである。

ここまで紹介してきた研究は、ミューチュアル・ファンドであれヘッジ・ファンドであれ、言わば「プロ」を対象にしていた。つぎに一般投資家を対象にIQと投資との関係を調べた研究を紹介しよう。フィンランドの徴兵制では新兵時代にIQテスト(数学的能力、言語的能力、論理的推論)を行っている。Grinblatt, Keloharju and Linnainmaa(2012)は、このIQと株式売買データを突き合わせ

て、IQと投資行動との関係を調べた。その結果、IQが高い分位に属する投資家は気質効果（値上がりするとすぐ売りがたり、値下がりすると購入価格に戻るまで売りがたらない傾向）の影響を受けにくく、節税のための実現損を出す取引にも積極的で、パフォーマンスも特にマーケット・タイミング効果を修正すれば低い分位に属する投資家よりも良いという。また、IQが高い分位に属する投資家は過去1年間のリターンが相対的に高く、1ヶ月のリターンが相対的に低い銘柄を買い付けている。つまり12ヶ月モメンタム、1ヶ月リバーサルに注目している。一方で、IQが低い分位に属する投資家は銘柄を買い付ける際に他の投資家の動向の影響を受けやすいという。

この結果には二つのインプリケーションがある。第一は従来の実証研究の結果の信頼度を補強する意味があるという点である。従来の研究はデータの制約からSATスコア等は大学・学部に対して与えていたので、個々のファンド・マネジャーのスコアを直接扱ったわけではなかった。ところが、本研究では個人のデータを直接処理しているので、より信頼できる結果と言える。第二は一般投資家を対象にして得られた結果であるため、たとえば確定拠出型年金等の制度設計に示唆するところが大きい点である。もちろん投資家教育の実情や国民性などの相違があるため、本研究の結果が日本にそのまま当てはまるわけではない。しかし、投資家教育だけではカバーできない投資家の属性の影響が残る可能性は、投資家保護の観点からも無視できないだろう。

ここまで紹介してきたファンド・マネジャーの属性とパフォーマンスとの関係に関する実証分析結果を整理しておく（図表1）。パフォーマンスと有効な関係ありとの結果を得た項目には○、逆の場合には×、検証項目でない場合には－を記している。

図表1 ファンド・マネジャーの属性とパフォーマンスとの関係

	SAT	GMAT	MBA	CFA	若さ
C & E ¹	○	－	△	－	○
G & M ²	○	○	×	×	×
S & H ³	－	－	×	×	×
D, G & S ⁴	－	－	×	×	－
L, Z & Z ⁵	○	－	－	－	－
B ⁶	－	－	×	－	○

注1: Chevalier and Ellison(1996)

注2: Gottesman and Morey(2006)

注3: Switzer and Huang(2007)

注4: Dincer, Gregory-Allen and Shawky(2010)

注5: Li, Zhang and Zhao(2009)

注6: Boyson(2003)

ここで、資産運用業界における人材の質、あるいは、もっと広く投資家の属性に関する実証研究のインプリケーションを総括しておこう。実証結果を整理すると、基本的な知的能力の高さは投資に有用だが、教育や資格はパフォーマンスと特段の関係は見られなかったということになる。また、経験年数に関しても、報酬体系や雇用の安定度とあいまってリスクの取り方に影響を与えるが、経験が長ければパフォーマンスがよいというわけではなかった。

高いSATやGMATのスコアを要求する教育機関の出身者のパフォーマンスが良好なことについて

では、つぎのような解釈がありうる。

第一にSATやGMATで測定する計数能力、言語能力、論理的推論などに代表される基本的な知的水準が高い方が投資には有利だという考え方である。投資とは情報を収集、分析した上で意思決定を行うことであるから、知的水準とは無関係ではありえないということである。特に一般投資家を対象とした検証では、IQとパフォーマンスとの間の関係が確認されていることも、この解釈を補強する材料になるだろう。

第二は高いスコアを要求する学校では良い教育が行われており、これが投資に役立つという考え方である。たとえばMBAプログラムで習得する経済、金融、証券等に関する知識がファンド・マネジャーとしての準備や訓練になっているというだけでなく、教育による知的訓練全般が投資にも有効だということである。しかしながら、MBAプログラム全般がパフォーマンスに必ずしも明確な関係を示さず、SATスコアやGMATスコアが影響するというのであれば、この解釈は説得力が乏しいと言えるだろう。

第三は高いレベルの学校では有効な人脈形成がはかられ、これがその後の活動を有利に運ぶ助けになるという考え方である。Cohen(2007)によれば、ファンド・マネジャーは投資において学生時代のネットワークによるつながりがある銘柄に重きを置く傾向にあり、関係のある銘柄はそうでない銘柄に対して年8.4%アウトパフォームしたという。人的コネクションの影響は有形・無形のさまざまな形態が考えられるので実証するのは困難が伴うが、この有効性を否定する材料は現時点ではないようである。

3 | Morningstar 社の Stewardship grades — 資産運用会社の定性評価

ファンドや資産運用会社の定性評価を考えるに当たり、指標として使われている Morningstar 社の Stewardship grades について説明しておく。2003年に米国で起きた投信不祥事に直面した顧客はファンドのガバナンスに対する関心を高めた。そこで、Morningstar 社は 2004年8月に Stewardship grades を創設した。従来から公表されていた Star rating が過去のパフォーマンスに焦点を当てているのに対して、Stewardship grades はファンド運用会社そのものを評価対象にしており、運用会社の文化(Corporate Culture)、経営陣の質(Board Quality)、ファンド・マネジャーのインセンティブ(Manager Incentive)、手数料(Fees)、規制(Regulatory History)の5項目からなっている。

本稿では、特に経営陣の質と企業文化に注目する。なお、Morningstar 社は 2007年に企業文化のウエイトを20%から40%に引き上げて最重視する体系に変更した。それまでは5項目を等ウエイトで扱っていたのである。

Morningstar 社は経営陣の質に関しては下記を注目している。

- ・ファンドが投資家に報いていない場合、経営陣は何らかの策を講じたか
- ・経営陣は自社ファンドに投資しているか
- ・担当するファンド数が多過ぎて、経営陣が個々のファンドを十分フォローできない状況に陥っていないか

- ・社外取締役の比率がSECの必要条件の最大値に達しているか

Morningstar社が重きを置いているのは、経営陣がファンドのパフォーマンスを十分に監視した上で、必要に応じてコミットする姿勢を持っているかである。

また、企業文化に関して Morningstar 社のアナリストが注目するのは下記の点である。

- ・資金を集めるため、「売れ筋」商品をラインナップに加えていないか
- ・ファンドは運用に適正な規模で売り止めをしているか
- ・ファンドの短期売買を回避するような償還手数料を設定しているか
- ・キー・パーソンを引き止めるような対策を講じているか
- ・ソフト・ダラー（売買執行以外のサービスのコストを売買手数料に含める）を使用しているか

これらは Morningstar 社のアナリストがインタビューで重視している質問である。彼らが企業文化として注目しているのが、顧客のパフォーマンスを重視する姿勢であることが見て取れる。

では、Jay and Xhou(2007)によるStewardship gradesの実証分析の結果を紹介しよう。彼らはStewardship grades全般や内訳に関しては下記の結果を得た。

- ① Stewardship gradesが高いファンドはリスク調整後のリターンが高い。
- ② 経営陣の質が高いほど、また、手数料が低いほどパフォーマンスがよい。
- ③ ガバナンスの質が高いファンドは、ガバナンスが良好な企業を保有する傾向にある。
- ④ 高gradesファンドには資金が流入していることから見て、投資家はこのgradesを活用している。

図表2 Stewardship gradesとパフォーマンスとの関係

Stewardship Grade	○
Regulatory Issues	—
Board Quality	○
Manager Incentives	—
Fees	○
Corporate Culture	—

注：2004年9月～2006年12月までの実証分析結果。
○は統計的に有意な結果だったことを意味する。

つぎにStewardship gradesの中の経営陣の質と企業文化に焦点を当てた実証分析結果を紹介しよう。

(1) 経営陣の質(Board Quality)

Lai, Tiwari and Zhang(2009)が、Morningstar社のStewardship gradesの中のboard quality ratingsに焦点を当てた実証結果を紹介しよう。Jay and Xhou(2007)でboard qualityとパフォーマンスとが相関を持つことが示されたが、彼らはパフォーマンスの継続性と資金流入に注目した。

- ① 過去のパフォーマンスが悪い場合、経営陣の質の評価が低いと、資金流入が少ない。

- ② 経営陣の質とパフォーマンスの継続性とは関係がある。経営陣の質が低いファンドでは、パフォーマンスが悪い時期が続く傾向にある。経営陣の質の高いファンドでも継続性は観察されるが、質の低いファンドよりも悪い時期の継続性は弱い。パフォーマンスが良い時期は、質の高い経営陣のファンドでは継続性があるが、質の悪いファンドではパフォーマンスの良い時期の継続性はずっと弱くなる。
- ③ パフォーマンスが悪かった場合のファンドの運用戦略変更は、経営陣の質が高い場合の方が低い場合よりも起こりやすい。

ということが見て取れるという。

(2) 企業文化(Corporate Culture)

企業文化は Morningstar 社の Stewardship grades の中でも全体を 4 割のウエイトを占め、最重視されている。Gottesman and Morey(2011)は下記の理由で良好な企業文化はパフォーマンスに好影響を与えるという仮説を立てた。

- ① 良好な企業文化は従業員の勤勉さ、生産性の高さにつながる。
- ② 営業志向ではなく顧客志向・パフォーマンス志向である、すなわち、
 - (a) ファンドを適正規模に保つ
 - (b) 売れ筋商品に乗らない
 - (c) 手数料が公正である
 - (d) 頻繁な売買を抑制する手数料体系
 - (e) ソフト・ダラーを使わないという運営は、結局のところ、パフォーマンス改善につながる。
- ③ 顧客への良好な情報開示は、長期保有につながる。

彼らは Morningstar 社の corporate culture grades とリスク調整済みリターンとの関係を 2005 年から 2010 年までのデータについて調べた。結論は、パフォーマンスの悪いファンドが途中で計測対象外になるサバイバーシップ・バイアス (survivorship bias) を考慮したところ、corporate culture grades とパフォーマンスとはあまり関係が見られなかったというものであった。企業文化はたしかにスキャンダル予防効果はあるが、マーケットに勝てるわけではなかったと結論付けている。実証分析による他の知見として、

- ① 企業文化の評価が高いファンドは、手数料が安く、回転率も低い傾向にある。
- ② 企業文化の評価が高いと、従業員の勤続年数が長くなる。
- ③ 評価が低いファンドは、途中償還やファンド合併が起こりやすい。

を挙げている。corporate culture grades が高い評価の場合、ファンドの手数料が安く回転率も低い傾向にあるにもかかわらず、パフォーマンスとの間に明確な関連性が見出されなかったのは、意外な結果である。回転率が低い方が運用面で有利であるし、手数料は手数料控除後のパフォーマンスに直接結び付くからである。また、従業員の勤続年数の長さもパフォーマンスに必ずしも結び付いていない。

彼らもこの結果の解釈には苦慮したようで、企業文化とパフォーマンスとの間に明確な関係が見られなかった理由として、つぎの二つを挙げている。第一に市場はかなり効率的なので、市場に一貫して勝つのはそもそも難しい。企業文化がよいというだけでは市場に勝つには不十分である。第二は企業文化が良好なことは、実は組織が硬直的になっており、環境変化に対して速やかな転換がはかれないことを意味している。第二の論点については、Sørensen(2002)の一般企業に対する研究結果を援用して説明している。彼は良好な企業文化は一般企業（資産運用会社に限らない）のパフォーマンスに下記のような影響を与えているとしている。

- ① 協力と管理を強める。規範違反は速やかに発見され是正される。
- ② 企業と従業員のゴールを合致させられる。
- ③ 従業員の意欲、勤勉さを高める。

このようなメリットがある反面、危機的状況で変化が激しい場合、企業文化が強く浸透していると迅速に対処できない傾向にあるという。つまり、企業文化の強靭さは環境変化に対する適応力とトレード・オフの関係にある。強い企業文化は現在のコンピテンシーを活用するのに長けるが、変化する環境の中で新たなコンピテンシーを発見するのが難しくなる。強靭さと柔軟性、効率性と変化への適応力は両立し難いということだ。Sørensenによればこの傾向は企業全般に当てはまる。資産運用会社は変化が速い「マーケット」を相手にしているのだから、環境変化に対する適応力が劣るのは運用会社としては弱点になりうる。だが、Morningstar社のcorporate culture gradesの中身を見ると、顧客志向の徹底や従業員の待遇などを重視しており、前述の企業文化のトレード・オフ仮説はむしろboard qualityの経営陣のコミットメントに重なる部分が多いため、企業文化に関する理由として挙げるのは説得力に乏しいと考えられる。

むしろ理由付けとしては、corporate culture grades が個々のファンドではなく資産運用会社に付与されること、そして運用会社として望ましい特性を評価しているものの、それらのパフォーマンスとの関連性は間接的であることに求めるべきだろう。つまり高いcorporate culture grades が良好なパフォーマンスを意味するのであれば、その運用会社のファンド全般のパフォーマンスが良いことを意味することになる。だが実際には、個々のファンド間のばらつきを考慮すると、そこまで頑健な結果が得られない。また、会社の方針として顧客志向が徹底されていたとしても、実際のパフォーマンスに反映されるのは具体的な方策であって、基本姿勢だけでは関連性が薄く積極的意味合いに欠けるということも考えられる。

5 | 小括

大別すれば、人材の質（ファンド・マネジャーの属性とパフォーマンスとの関係）と、Morningstar社のStewardship gradesに関する海外の実証分析結果を紹介した。

前者に関しては、従来、ファンド・マネジャーの経歴、資格、経験年数などが評価基準として重視されてきたが、実証分析でパフォーマンスと関連付けられたのは、SATのスコアのような基本的な

知的能力だけである。MBA取得、CFA保有、その他の教育プログラムなどはリスク調整後のリターンで有意な結果は得られていない。また、年齢に関してはベテランよりも若手の方が良好である。ただし、ミューチュアル・ファンド、ヘッジ・ファンドの業態によって、若手・ベテランのリスクの取り方が異なる。そのため、ファンド・マネジャーのキャリア・パス、待遇を含めたインセンティブ体系に投資家は無関心であってはならない。パフォーマンスにも影響を及ぼす可能性があるからである。

後者の **Stewardship grades** に関する実証結果では、**Stewardship grades** そのものはパフォーマンスと相関を持つことが確認された。しかし5項目のうち、リターンと有意に関連付けられるのは、**Board Quality** と **Fees** だけであった。特に全体の40%を占める **Corporate Culture** はパフォーマンスとあまり関係がなかった。

4—終わりに

本稿では、パフォーマンス評価に関して定量面と定性面の双方について、米国の実証分析結果の紹介を交えて再検討を行った。

ファンドのパフォーマンス評価は過去のリターンを定量評価することから始まる。だが定量評価するにしても、リスク調整後のリターンを算出するには、どのパフォーマンス指標を採用するのかが問題になる。得られた数値を統計的に検定しようとしても、実際には統計的に有意な結果を得るだけのトラック・レコードの蓄積がないことが多い。それに加えてパフォーマンス指標は推計誤差を含むので、使用するリターン系列の特性、すなわち、投資戦略や投資環境の影響を受けることになる。特に系列自己相関の高いヘッジ・ファンドはパフォーマンスが過大評価されがちなことを指摘した。また、過去のパフォーマンスの定量評価は、ある投資環境下における条件付きの評価に過ぎない。経済環境が変化した場合の適応力の評価は別の問題であるので、再現性は保証されない。

だからこそ定量評価のほかに、判断の補完・補強材料として定性評価が重視されることになる。本稿では、人材の質と Morningstar 社の **Stewardship grades** による定性評価に焦点を当てた。実際の運用評価機関の定性評価でも、ファンド・マネジャーの経歴、資格、経験年数や従業員の定着度などの形で人材の質は評価されるし、「運用哲学」「運用方針」の一貫性も強く求められている。だが実証分析の結果は、研究者の間でコンセンサスが形成されるほど一貫性のある結果が必ずしも得られていなかったり、パフォーマンスと関連性があるとする仮説を裏付けなかったりすることも多かった。

現時点では、情報を処理する上で基本となる知的能力は投資に有用でパフォーマンスと関連付けられるが、教育プログラムや資格とは明確な関係が見られなかったということを確認したにとどまる。また、ファンド・マネジャーの経験年数も長い方がパフォーマンスが良いとも言えなかった。報酬体系やキャリア・パスも含むインセンティブが意味を持っており、投資家・委託者としてもこれらには注意を払う必要があるとの示唆を得た。

Morningstar 社の Stewardship grades については、全体の評価、経営陣の質 (Board Quality)、手数料 (Fee) がパフォーマンスに影響を及ぼすことが確認された。一方で企業文化 (Corporate Culture) はパフォーマンスと明確な相関を示していない。経営陣の質と企業文化の相反するパフォーマンスへの影響をどう解釈するかが問題となる。従来、運用評価機関の定性評価では「運用哲学」「運用方針」の一貫性が強く求められてきたし、このことに対して疑義がはさまれることはなかった。また、従業員の定着率や勤続年数に関する評価も然りである。

Morningstar 社は、ファンドが不振の場合にアドバイザーを解雇することも board quality grades でプラス評価するとしている。ただし、この 15 年間にその回数が数えるほどであるという条件付きだ。つまり、「変えない」ことが常に評価されるわけではなく、ファンド運用に関して根本的な問題がある場合は対策を講じることこそが board member として求められるということだ。もちろん、経営陣が投資環境の短期的変動に対しても介入せよという意味ではない。抜本的な対策が必要な場合や、構造変化への対応が求められる場合には、「一貫性」よりも「動くべきときに動く」ことが重要だと解釈できる。結局のところ、それが顧客のパフォーマンスにつながると言える。要するに、一貫性の維持を大義名分に何もしないこと自体が目的となり経営陣の不作為の罪が黙認されたり、逆に見識なき無節操な経営陣の介入が頻発したりするようでは、顧客の利益につながる board quality を維持できていないということなのだろう。

また従業員に関しても、従来は従業員を大切にする企業文化や待遇が評価されるので、定着率の高さや勤続年数の長さが評価基準として重視されてきた。しかし実証分析結果はそれほど単純ではなく、報酬体系やキャリア・パスも含めて、インセンティブを付与していくことが求められている。単に勤続年数や定着率だけで評価するのでは、従業員の待遇が「ぬるま湯」指数になりかねないからである。

さて本稿では、海外の実証研究のサーベイも含めてパフォーマンス評価について考えてきた。海外の実証分析結果がそのまま日本に当てはまるとは限らないが、日本の資産運用業界においても常識とされてきたことがパフォーマンスと必ずしも結び付かない可能性があることは指摘しておきたい。そのことを踏まえると、現在の運用評価機関やコンサルティング会社の評価基準は形式的・硬直的過ぎる可能性がある。人材の質、経営陣の質、企業文化に関しても、一般に思われているようなパフォーマンスとの関連性は実証分析によって裏付けられているわけではない。特に定性評価に関してはパフォーマンスとの関連性に対する実証分析が不足していることもあり、今後、さらなるデータの蓄積が必要であるという現状を考慮すれば、定性情報を機械的に適用するのではなく総合的に判断して評価していくことが求められる。また、資産運用会社は運用評価機関の基準を妄信しそれに迎合するのではなく、健全な懐疑主義に基づき、最終的には顧客のために合理性を貫徹するという姿勢を堅持することが大切だろう。

参考文献

Boyson, N., 2003, "Why do experienced hedge fund managers have lower returns?", Working paper

Cesarini,D., Johannesson,M., Lichtenstein,P., Sandewall,O., and Wallace,B., 2010, "Generic variation in financial decision-making", *Journal of Finance*, 65, 1725-1754

Chen, C. and Huang,Y., 2011, "Mutual Fund Governance and Performance: A Quantile Regression Analysis of Morningstar's Stewardship Grade", *Corporate Governance: An International Review*, 2011, 19(4), 311-333

Chevalier,J. and Ellison,G., 1996, "Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance", *Journal of Finance*, 54(3), 875-899

Chevalier,J. and Ellison,G., 1997, "Career concerns of mutual fund managers", *Quarterly Journal of Economics*, 114, 889-432

Chou,W., Ng,L. and Wang,Q., 2007, "Do Governance Mechanisms Matter for Mutual Funds?", Working paper

Cohen,L., Frazzini,A., and Malloy,C., 2008, "The small world of investing:board connections and mutual fund returns", *Journal of Political Economy*, 116, 951-979

Dincer,O., Gregory-Allen,R., and Shawky,H., 2010, "Are You Smarter than a CFA'er? Manager Qualifications and Portfolio Performance", Working paper

Grinblatt,M., Keloharju,M., and Linnainmaa,J., 2011, "IQ and stock market participation", *Journal of Finance*, 66, 2121-2164

Grinblatt,M., Keloharju,M., and Linnainmaa,J., 2012, "IQ, trading behavior, and performance", *Journal of Financial Economics*, 104, 339-362

Gottesman, A. and Morey, M., 2006, "Manager education and mutual fund performance", *Journal of Empirical Finance*, 13, 145-182

Gottesman, A. and Morey, M., 2011, "Mutual Fund Corporate Culture and Performance", Working paper

Lai, S., Tiwari, A., and Zhang, Z., 2009, "Mutual Fund Flows, Performance Persistence, and Board Quality", Working paper

Li,H., Zhang,X., and Zhao,R., 2009, "Investing in Talents: Manager Characteristics and Hedge

Fund Performances”, Working paper

Lo, A., 2002, "The Statistics of Sharpe Ratios", Financial Analyst Journal, Vol.58, No.4(july/August), 36-52

López de Prado M., 2008, “How long should a track record be?”, Working paper

Sørensen, Jesper B., 2001, "The Strength of Corporate Culture and the Reliability of Firm Performance", Administrative Science Quarterly, 47,1, 70-91

Switzer,L. and Huang,Y., 2007, “Management characteristics and the performance of small & mid cap mutual funds”, Working paper

Wellman,J. and Zhou,J., 2007, "Corporate governance and mutual fund performance: A first look at the Morningstar Stewardship Grades", Working paper

渡辺慧、2012、「時」、河出書房新社