“The rationality of index investing and paradoxes of indexation”

AUTHORS	David Eagle Arsen Djatej Robert H.S. Sarikas David Senteney

ARTICLE INFO	David Eagle, Arsen Djatej, Robert H.S. Sarikas and David Senteney (2009). The rationality of index investing and paradoxes of indexation. *Investment Management and Financial Innovations*, *6*(4-1)
RELEASED ON	Wednesday, 10 February 2010
JOURNAL	"Investment Management and Financial Innovations"
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
0	0	0

David Eagle (USA), Arsen Djatej (USA), Robert H.S. Sarikas (USA), David Senteney (USA)

The rationality of index investing and paradoxes of indexation

Abstract

This paper explores the logic of actively investing instead of passive indexation investments. The findings of this investigation incorporate the conventional rationality and tactical investment strategies for active investors. Furthermore, the paper outlines the definition of "investing rationality" from the angle of conventional academic concepts of expected returns and risk. However, given that about 90% of stock is actively invested, the reasons for the continued existence of active investment must go beyond this academic definition of rationality. This paper also elaborates on the possibility of investors being drawn to active investing not for perceived expectational advantages over index investing, but rather simply for the sake of investing.

Keywords: index funds, passive investing, active investing.
JEL Classification: G11, G14.

Introduction

Investors are facing hard choices related to their ability to optimize their portfolio mix by proportionally allocating equity vs. other investments. This scholarly investigation focuses on the investment strategies related to two primary choices of active investing vs. passive or index investing.

Individuals or financial institutions who prefer investment strategies utilizing passive investing vehicles such as index funds, index ETFs, ETNs, etc., expect, by definition, to receive a gross return equal to the market average. On the contrary, investments in actively managed funds or through investment manager can provide a greater or smaller return in comparison to the market. The fundamental question is whether someone expects to do better with index (passive) investing vs. active investing. The conventional wisdom dictates the "rational" choices of "better" or "worse" investment strategy should be directly correlated to higher returns of the two alternatives mentioned above with risk considerations always to be present. Hsu and Campollo (2006) in their scholarly investigation titled "New Frontiers in Index Investing" reinforced the widely accepted notion that index investing is a strong strategy for equity investing since it is inexpensive to implement and absolutely transparent (Hsu & Campollo, 2006, p. 33). Furthermore, according to the authors, the strategy of index investing is highly liquid and naturally well diversified and presents strong capacity. Evidence suggests that a cap-weighted market index is an efficient equity investment. Fundamental indexing eliminates the return drag inherent in cap-weighted indexes (Hsu & Campollo, 2006, p. 58). The methodology preserves the capacity, liquidity, diversification and broad-market participation that are the chief benefits of traditional cap indexes (Hsu & Campollo, 2006, p. 58). According to Mar, Bird, Casaveccia and Yeung, capitalization-weighted indexes provide the basis for passive investment strategies designed to capture market performance (Mar et al., 2009, p. 1). The findings of their research support the overall "proposition that fundamental indexation provides an interesting investment alternative to cap-weighted indexation" however the research didn't provide any answers on whether overall it should be considered "a passive strategy" (Mar et al., 19).

In his article "Fact and Fantasy in Index Investing", Kirzner focuses on the ability of investors to determine suitable investment strategies. According to Kirzner, investors face "two broad choices of methodology – "passive investment" and "active selection" (Kirzner, 2000, p. 7). The author identified two broad dilemmas faced by investors related to investment policies and investment strategies. In his view, investment policy incorporates three crucial steps including: 1) "strategic asset allocation" (which percentage of capital should be devoted to each primary investment class); 2) "dynamic asset allocation" (setting guidelines for when and how to make changes in investment holdings in order to restore overall portfolio mix back to its original target); 3) "tactical asset allocation" or "market timing" (adjustment of the "target portfolio mix in an attempt to increase returns by anticipating changes in economic conditions) (Kirzner, 2000, pp. 7-9). Eventually, Kirzner concentrates on passive investment strategies and presents "four paradigms" for successful index investing including: selection of right strategic asset-allocation mix (Markowitz Portfolio theory and diversification principles), incorporation of large equity component in the portfolio, passive investing supported by extensive studies and focus on a specific implementation strategy (Kirzner, 2000, p. 24). The research attempts to convince that passive investment is the best chance of growing capital since "passive portfolios have closely tracked the target indexes, minus minimal expense adjustments, while active equity investors in Canada and the US have, on average, underper-

formed market indexes – typically by as much as 3% per annum" (Kirzner, 2000, p. 49).

Our primary assumption is based on the notion that all active investors are rational (as we said before, this is a rather strong assumption) and they prefer active investment vehicles instead of passive. Therefore, according with this assumption, the expectations for the return from this type of investing strategies are higher than index investing. Consequently, investors expect to do better utilizing these active investment strategies vs. passive or index investing. However, that means that investors expect to do better than the market since, as it was mentioned above, index investing provides them only with a gross return[1] equal to the market average.

At this point, it is important to integrate the concept of "rational investment" or "rational investor" into the conceptual framework of this investigation. Thus, we assume that in accordance to the concept of "rationality" of investment strategies, investors who would expect "to do worse than the market" prefer passive or index investing rather than active investments. However, since with relatively few exceptions the market consists of only active and passive investors, and the index investors have a gross return equal to the market average, then the rationally expected gross performance of the group of active investors must also be the expected market performance[2] (the definition of "performance" could encompass risk as well as expected return). Therefore, it would be impossible for all active investors to rationally expect to do better than the market. If some active investors expect to do better than the market, then other investors must expect "to do worse than the market", assuming they all had rational expectations.

In a market consists of only investors possessing rational expectations it is impossible to have any active investors at all. As the above paragraph points out, no active rational investor would expect return or performance below market average. Without any investors expecting to do worse than the market, then no investor including active investor will be able to do better than the market. Given that the management fees and transactions costs are greater with active investing than with index investing, this would imply that all investors in such a market would prefer passive investments or indexation. This leads to the "Indexing Paradox:

Assume (i) investors have rational expectations, (ii) investors make rational decisions, (iii) investors have a common risk-averse investment performance measure, and (iv) indexing results in a return equal to the average market return. Under these assumptions, no investor can expect to do better than the market. If the cost of indexing is less than the cost of active investing, then all investors would index, which would result in no mechanism to price the possible investments.

Despite the logic of the Indexing Paradox, we see about 90% of stock under active management. While the existence of the active investment helps preclude a collapse of the stock market, we need to ask why is this stock under active management. By the Indexing Paradox, if there is active management then there must exist irrational investors. If there are irrational investors, then it could make sense for other rational investors to actively manage if they expect to do better than the market. Imagine that someone realized that there are irrational investors out there (irrational in the sense that they are delusional, thinking that they expect to do better than the market when in fact they are expected to do worse than the market). Roughly speaking, up to almost half of the active investors may be rational, rationally expecting to do better than the market, as long as the other half of the active investors would be delusional.

Now if you realized that you yourself were one of these delusional active investors, then you would switch to indexing. Therefore, your continuation at being an active investor indicates that you believe not only that you expect to do better than the market, but you believe that you are one of the rational investors who rationally expect to do better than the market. However, that must be the case with all the active investors, both rational and delusional. In other words, all active investors believe they are not the delusional ones. Thus, even though you believe you can do better than the market, there is a 50% chance that you are one of the delusional active investors rather than one of the rational ones. Therefore, why is it that you believe you are one of the rational active investors and not one of the delusional active investors?

The above paragraph shows why it should be so hard for investors going through rational thought to continue being active. Perhaps, investors do not fully understand the Indexing Paradox and when this and other papers on the Indexing Paradox are

[1] The net return of indexing could be less than the market average because of fund management fees and transactions costs. However, those fees are significantly less for index investing than for active investing.

[2] If this is not obvious, then realize that the expected market average performance will be a weighted average of the expected performance of the active investors and the expected performance of the index investors. Mathematically, where w is the fraction of the value of stock under active management, $\mu = wa + (1-w)\mu$, where μ is the expected market performance and a is the expected performance of the active investing. (Remember that the index investors have an expected performance equal to the expected market performance.) Solving for a gives $a = \mu$.

published, more investors will switch to index investing.

An alternative reason for the existence of active investing can be seen by looking at the analogy with sports. People have the option of engaging in sports or not engaging in sports. If their objectives were to maximize their winning records, with wins being +1s and losses being –1s, then people choosing not to engage in sports would be like index investors getting a win/loss score of 0. On average, those people who do engage in sports will also have a win/loss score of 0. Why do these people choose to engage in sports? Is it because each expects to have a win/loss score greater than zero at the end of the season? If that were the case, half would be irrational because on average the win/loss score would be zero – half of the players would have more wins than losses and half would have more losses than wins.

However, in the real world people will engage in sports even if they are rational, even if they expect they will not have a winning record. People have a natural drive to compete. People engage in sports because they enjoy the game, they enjoy the competition; they do not just play the game with the expectations of having a winning record. They may all "hope" to have a winning record, they may all "strive" to have a winning record, they may all "dream" about having a winning record, but they may realize that they all cannot rationally "expect" to have a winning record, they may realize that the number of winners will equal the number of losers.

Many may approach investing much as a sport. Many may view investing as a grown-up "sport" or game. Just as they are motivated by the excitement of a sports game, they may be so motivated by the investing game, trying to compete with other investors. While they "hope" to do better than the market, while they "dream" about doing better than the market, while they "strive" to do better than the market, they may realize that on average, they will be unable to do so. Nevertheless, they continue to actively invest anyway because of the "sport" of the game of investing. Just as an athelete may identify his/her worth based on his/her performance in his/her sport, so might an active investor's identity being tied to his/her performance in the stock market. For such investors, index investing would be too boring, it would not be a game. To them, actively investing in the stock market is a game.

This is similar to explanations given why people gamble. Some, including Milton Friedman, have rationalized gambling with risk-loving utility functions. However, another explanation of why people gamble when they also possess insurance policies, is because of the game qualities of gambling.

While investors, treating investing as a game, would explain why some investors continue to actively invest despite the evidence that they cannot rationally expect to do better than index investing, this cannot explain why so much of our market actively invests. Much of the stock under active investment is by people who are quite ignorant of the markets and feel uncomfortable making their own investment decisions. These investors are not in the market because they enjoy the competition of the "game" of investing.

Our view of why these investors are actively investing is because of the retail marketing of investment services. Ignorant investors go to full service brokers or financial planners for financial advice about investing their money. Since these brokers or financial planners earn more commissions on active investment options than indexing investment options they provide to their clients, or because the brokers or financial planners enjoy the "game" or "sport" of investing, they steer their clients in the direction of active investing instead of index investing. We in the Finance profession must recognize how much of our financial market place is determined by the marketing of financial services, which often goes against market efficiency. This should be especially clear now following the last year of revelation after revelation about the conflict of interests between analysts and the investment banking services provided by their employers of the companies being analyzed.

This completes the main argument of the paper. However, many readers may be unconvinced about the Indexing Paradox. They may wonder if the Indexing Paradox applies to when stock performance considers risk as well as expected return. They may wonder if it depends on an assumption of market efficiency (It does not). They may wonder if it applies in a world where investors differ in their information or their ability to analyze data. To demonstrate the broad applicability of the Indexing Paradox, the next two sections present a rigorous equilibrium model of expected utility maximizing investors possessing different degrees of comparative informational advantages and disadvantages. Later sections of the paper will use this model to demonstrate the wide applicability of the Indexing Paradox.

1. Basic description of the model

This one-year model consists of a positive number of expected-utility-maximizing investors (m) and a positive number of stocks (n), where the value of the stock one year from now (which is the stock's termination value) depends on a particular probability distribution. For simplicity, this model uses a common distri-

bution for each stock. Investors do, however, have different comparative informational advantages and different information sets and thus generally have differing expectations. So that investors have the same performance measure, we assume that the investors have identical risk-averse utility functions of return. Each investor attempts to maximize the investor's expected utility given the investor's information set by choosing whether to actively invest or to index and, if the investor chooses to actively invest, then choosing what fraction of the investor's initial wealth to invest in each stock.

Fixed quantities of stock exist. A full equilibrium exists when (1) each investor maximizes his/her expected utility given his/her information set, and (2) the resulting demand for each stock equals this fixed supply of each stock. The computation of this full equilibrium is very complex because investors know the equilibrium prices of the stocks, but those equilibrium prices themselves depend on the stock demands of the investors and hence at least partially reflect some information (See Grossman and Stiglitz, 1980). Instead of directly computing the full equilibrium, we present a sequence of quasi equilibria that lead to a full equilibrium. A quasi equilibrium differs from a full equilibrium in that investors do not take into account the informational content of prices when they maximize their expected utility. This sequence of quasi equilibria also tells a story about how the Indexing Paradox would unfold.

The indexing methodology we use is where an indexing investor owns an equal portion of every existing stock. An investor j using this indexing method would invest $\frac{w_j}{\sum_{k=1}^{n} p_k s_k} p_i s_i$ amount of money into stock i, where p_i is the price of stock i, s_i is the supply of stock i, and w_j is the wealth of investor j. This implied index is a weighted average index of all stocks in the stock market[1].

Because of the complexities of the model, we are unable to find a closed-form algebraic solution of the model. Instead, we use a combination of Monte Carlo simulations and computer numerical analysis. Even with the computerization, the task of maximizing expected utility for each investor is too time consuming for our computers. Instead, we maximize a proxy utility function of expected portfolio return and standard deviation that seems to generate results sufficiently consistent with maximization of expected utility.

The Monte Carlo simulation generates values for the random variables of the model. For these random variables the computer iterates through the following process:

1. Using numerical methods, the simulation determines for each investor the fractions of funds that the investor invests in each stock in order to maximize the investor's proxy utility function of expected portfolio return and standard deviation conditional on the information the investor has with the exception that the investor ignores any informational content in prices.
2. The computer determines the excess demand or supply for each stock and then increases or decreases the prices to move toward equilibrium.

Eventually, the computer reaches a quasi equilibrium. The computer then repeats this process by generating a new set of values for the random variables and redetermining the quasi equilibrium for those random variables. For each simulation in this paper, the computer conducted 70 sets of these random variable realizations to create a very good "sample" of the possibilities. We then compare how each investor did relative to the performance of indexers. When the Monte Carol results show an active investor expects to do worse than the indexers, we switch that investor to being an indexer and then repeat the process all over again.

While we do use a proxy utility function to determine the investors' "optimal" choices, we use the actual utility function to compute the average of the utilities across all simulated realizations to get what we call "the after-simulation expected utility" for each investor. Given the theoretical nature of this model and our assumption that investors have the same utility function, we use the after-simulation expected utility as the common performance measure.

The next section discusses the mathematical details of the model. Readers should be able to skip that section if they choose and still be able to get a general understanding of the rest of the paper.

2. Mathematical details of the model

This one-year[2] model assumes there are m investors and n stocks. The investors invest their money at time 0 and spend their money at time 1. Stock i's value at the end of the period is

$$v_i = k_i u_i + (1 - k_i)\eta_i, \quad (1)$$

[1] This indexing methodology and indeed the Indexing Paradox can be extended to any market of risky assets as long as we know the prices and existing quantity of those assets. However, for readability this paper will refer to these assets as stocks.

[2] Many modelers talk about this type of model as being a two-period model. We prefer to think of it as a one-period model with a beginning and an end. Investors invest at the beginning of the period and consume at the end of the period.

where v_i is the value of stock i at time 1, and u, and η_i are independent random variables, each with a standard exponential distribution. Both u_i and η_i represent unsystematic risk. (For simplicity, this model does not include any systematic risk.) However, u_i is somewhat predictable depending on ones comparative informational advantage, while η_i is completely unpredictable for all investors. Equation (1) states that the value of stock i at the end of the period depends on the weighted average of u_i and η_i. For the simulations in this paper, k_i equals one half, where v_i is equally determined by u_i and η_i.

Each investor j has his or her own comparative informational advantage at predicting the value of stock i. Investor j's comparative informational advantage is represented by g_{ij}, which can range between 0 and 1. Each investor j observes a related random variable y_{ij} that gives some information on u_i depending on the value of g_{ij}. The observed random variable is given by:

$$y_{ij} = g_{ij}u_i + (1 - g_{ij})\varepsilon_{ij}, \tag{2}$$

where ε_{ij} is a random variable that has a standard exponential distribution and is independent of u_i and η_i. As stated before, g_{ij} represents investor j's comparative informational advantage at predicting the value of stock i. If g_{ij} equals 0, then y_{ij} provides no predictive information about u_i. If g_{ij} equals 1, then y_{ij} can perfectly predict u_i.

Below are four cases depending on the value of g_{ij} and the conditional expected value of u_i and its conditional variance under those cases:

Case 1: $g_{ij} = 0$. $E_j[u_i \mid y_{ij}] = 1$ and $var_j[u_i \mid y_{ij}] = 1$ as y_{ij} provides no information on u_i. Therefore, $E_j[u_i \mid y_{ij}]$ and $var_j[u_i \mid y_{ij}]$ equal the unconditional expected value and unconditional variance of u_i, both of which equal 1 since u_i has a standard exponential distribution.

Case 2: $g_{ij} = 1$. $E_j[u_i \mid y_{ij}] = y_{ij}$ and $var_j[u_i \mid y_{ij}] = 0$. By equation (2), $y_{ij} = u_i$ which means y_{ij} provides complete information on u_i.

Appendix A derives the results given below for cases 3 and 4:

Case 3: g_{ij}=½. $E_j[u_i \mid y_{ij}] = y_{ij}$ and $var_j[u_i \mid y_{ij}] = \dfrac{y^2}{3}$.

Case 4: $g_{ij} \in (0, ½) \cup (½, 1)$.

Where $\tilde{y}_{ij} \equiv \dfrac{1-2g_{ij}}{g_{ij}(1-g_{ij})} y_{ij}$, the conditional expectations and variances of u_i are:

$$E_j[u_i \mid y_{ij}] = \left(\frac{1-g_{ij}}{1-2g_{ij}}\right)\left(1 + \frac{\tilde{y}_{ij}}{1-e^{\tilde{y}_{ij}}}\right), \tag{3}$$

$$\mathrm{var}_j[u_i \mid y_{ij}] = \left(\frac{1-g_{ij}}{1-2g_{ij}}\right)^2 \left[1 - \left(\frac{\tilde{y}_{ij}}{1-e^{\tilde{y}_{ij}}}\right)^2 e^{\tilde{y}_{ij}}\right]. \tag{4}$$

Next, we need to determine each investor's expected value and variance of each future stock value conditional on their information about y_{ij}. Returning to equation (1), since u_i and η_i are independent, and η_i has a standard exponential distribution,

$$E_j[v_i \mid y_{ij}] = k_i E_j[u_i \mid y_{ij}] + (1 - k_i), \tag{5}$$

$$\mathrm{var}_j[v_i \mid y_{ij}] = k_i^2 \,\mathrm{var}_j[u_i \mid y_{ij}] + (1 - k_i)^2. \tag{6}$$

The return on stock i equals $r_i = \dfrac{v_i - p_i}{p_i} = \dfrac{v_i}{p_i} - 1$ where p_i is the price of stock i at the beginning of the period. Therefore, the expected return on stock i and the variance of that return conditional on y_{ij} are:

$$E_j[r_i \mid y_{ij}] = \frac{E_j[v_i \mid y_{ij}]}{p_i} - 1, \tag{3}$$

$$\mathrm{var}_j[r_i \mid y_{ij}] = \frac{\mathrm{var}_j[v_i \mid y_{ij}]}{p_i^2}. \tag{4}$$

These expectations and variances for investor j are conditional only on y_{ij} for i=1..n and not prices. However, the prices will at least partially reflect the information observed by all investors. Ignoring this informational content of prices could lead to significant expectational errors. However, we will find that the sequence of quasi equilibria that results from low performing active investors switching to indexing does lead to a full equilibrium where investors do not make those expectational errors.

We assume that investors have identical utility functions and that their desire is to maximize their expected utility. Each investor j's utility function is $U(r_j^P) = \ln(1 + r_j^P)$ where ln(.) is the natural logarithm and r_j^P is the return on investor j's portfolio. Given that this is a utility function only of return and not wealth, relative risk aversion should be constant; the logarithmic utility function does have a constant relative risk coefficient of one.

Equilibrium is defined when the following conditions hold:

1. Each investor j maximizes his/her expected utility conditional on his/her information on y_{ij} and p_i for stocks i=1..n by (a) choosing whether to analyze or index, and (b) if an analyst, choosing the fraction of funds to invest in each individual stock.
2. All stock markets clear.

A quasi equilibrium is defined when the investors who engage in active investing and who index are given and the following conditions hold:

1. Each active investor j maximizes his/her expected utility conditional on his/her information on y_{ij} for stocks i=1..n.
2. All stock markets clear.

The differences between a full equilibrium and a quasi equilibrium are two: First, for a quasi equilibrium, whether an investor actively invests or indexes is given; for a full equilibrium, the investor determines whether to engage in active investing or indexing based on expected utility maximization. Second, for a quasi equilibrium, the investor ignores the informational content of the individual stock prices; for a full equilibrium, the investor does take that information into account.

A closed-form solution of the quasi equilibrium of this paper is not possible. Instead we conduct Monte-Carlo simulations, and use computer numerical methods to both solve the investor's maximization problem and to determine the prices where demand equals supply for each stock. To simplify our analysis, we use a proxy for maximizing each individual i's expected utility. This proxy, a utility function of the expected value and standard deviation of the portfolio return, is a straight average of the following two values: $U(1+r_j^P + c*\sigma_j^P)$ and $U(1+r_j^P - c*\sigma_j^P)$ where c is a constant, r_j^P is the return on the portfolio for individual j, σ_j^P is the standard deviation of the portfolio for individual j, and U(.) is the investor's utility function.

n = 10 (# of stocks)
m = 10 (# of investors)
w_j = 1 for each investor j (wealth)
s_i = 1.1 for each stock i (supply of stock)
k_i = 0.5 (portion of stock value related to information variables)

Fig. 1. List of parameter values for basic Monte Carlo simulation

Currently we are using c = 2, which seems to give results sufficiently consistent with true expected utility maximization.

3. Analysis and results

For a simulation of ten investors and ten stocks, Table 1 presents the expected returns, standard deviations, and utilities for each investor depending on how many of the investors are indexers. Investors are ordered from lowest to highest by their comparative informational advantage (Investor j's comparative informational advantage variable, g_{ij}, equals (j-1)/(m-1) for all stocks i and for all investors j, where m is 10, the number of investors). Table 1 depicts a story where investors with lower comparative informational advantages switch to indexing when they realize they are expecting to do worse than the market and, hence, worse than indexing.

When all investors are actively investing, the after-simulation expected portfolio returns for investors 1, 2, 3, and 4 are negative. These investors' before-simulation expected returns were positive. This before-simulation/after-simulation discrepancy in expected returns results from investors, in a quasi equilibrium, making expectational errors because they ignore the information reflected in prices.

Once investors realize that they will make those expectational errors, they take corrective action. One way they can take corrective action is to switch to indexing. To determine if these investors would be better off actively investing or indexing, it is best to look at the after-simulation expected utility of each investor, which accounts for both expected return and risk. When all investors are analysts, investors 1, 2, 3, 4, 5, and 6 have lower expected utilities than the 6.04 centi-utils[1] they would have experienced had they indexed. As a result, those six investors switch to indexing.

When all investors are analysts, investors 7, 8, 9, and 10 expect to do better than the market. However, when the other investors switch to indexing, investors 7 and 8 find their expected utilities being below the market average of 6.05 centi-utils, which is the expected utility of an indexer. The reason is that the active investors as a whole can only do as well as the market average, and, if investors 9 and 10 do better than the market average, then others must expect to do worse than the market average.

Table 1. Simulation results with no margin trading

	Expected returns and standard deviations				Expected utility (centi-utils)			
	Number of indexers				Number of indexers			
Investor	0	6	8	9	0	6	8	9
1	-2.73% (27.8%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	-7.03	6.05	6.06	6.05

[1] A centi-util is defined as 1/100th of util.

Table 1 (cont.). Simulation results with no margin trading

	Expected returns and standard deviations				Expected utility (centi-utils)			
	Number of indexers				Number of indexers			
Investor	0	6	8	9	0	6	8	9
2	-2.76% (28.0%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	-7.15	6.05	6.06	6.05
3	-0.61% (28.1%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	-4.64	6.05	6.06	6.05
4	-0.80% (29.8%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	-5.5	6.05	6.06	6.05
5	2.39% (28.4%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	-1.98	6.05	6.06	6.05
6	11.20% (34.6%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	5.63	6.05	6.06	6.05
7	13.80% (34.4%)	4.97% (32.2%)	9.20% (24.6%)	9.19% (24.6%)	8.49	0.16	6.06	6.05
8	19.40% (32.2%)	6.41% (26.9%)	9.20% (24.6%)	9.19% (24.6%)	13.88	2.75	6.06	6.05
9	25.60% (30.7%)	12.30% (25.7%)	8.66% (25.4%)	9.19% (24.6%)	19.61	8.78	5.35	6.05
10	26.30% (30.6%)	13.10% (25.0%)	9.73% (24.7%)	9.18% (24.6%)	20.5	9.76	6.60	6.05
Indexers	9.18% (24.6%)	9.20% (24.6%)	9.20% (24.6%)	9.19% (24.6%)	6.04	6.05	6.06	6.05

Note: Shaded area represents indexers. A centi-util is one one-hundredth of a util.

Because they expect to do worse than the market if they remain analysts, investors 7 and 8 switch to indexing. While investor 9 expected to do better than the market before investors 7 and 8 became indexers, when all but investors 9 and 10 index, investor 9 has an expected utility less than the market average of 6.06 centi-utils. Therefore, investor 9 also switches to indexing. However, when only investor 10 remains as an active investor, his/her expected portfolio return, standard deviation of return, and expected utility are then the same as the market average.

The quasi equilibrium where only investor 10 is actively investing is also a full equilibrium. Since the model does not assume any cost of active investing (or of indexing), investor 10 is indifferent between active investing and indexing. As a result, investor 10 is maximizing his or her expected utility in this quasi equilibrium. Also, since the only information that can be reflected in prices is the information investor 10 directly observes, investor 10 is already fully using this information. The other investors in the market must also be fully using the information reflected in prices, because obviously they cannot use that information to do better than investor 10 who directly observes that information and they are already doing as well as investor 10 by indexing[1].

This Monte Carlo simulation clearly demonstrates how the Indexing Paradox unfolds in an environment where performance is based on risk as well as expected return and investors have different information sets or different information abilities.

Summary, conclusions, and reflections

The Indexing Paradox states that if all investors are rational having a common performance measure, then in a market consisting of only rational investors, no active investor can rationally expect to do better than the market. Therefore, since there are active investors, we may conclude that some investors are irrational. However, if the active investors realized this, then they would realize that each active investor believes he/she can expect to beat the market. Since half of these active investors must be delusional, the active investors may find it hard to continue actively investing given the 50% probability that they themselves are delusional.

However, the reason that the active investors do not switch to indexing, may be because investors are not only concerned about the risk and return of their

[1] That stock prices could only partially reflect information and not fully reflect information that was shown in a model by Grossman and Stiglitz (1980). The randomness in that model that caused the less-than-full reflection of information stemmed from randomness Grossman and Stiglitz assumed in the supply of the risky asset. However, the supply (the number of shares outstanding) of stock is public information in reality and that public information is the basis for indexing. Nevertheless, using random components on the demand side rather than the supply side can salvage the Grossman and Stiglitz's results. That is the approach taken in this paper. We make no assumptions about the investors' knowledge of other investors' wealth or their utility functions. If individual investors are uncertain of this knowledge, then prices would only partially refect information.

stock portfolio, they may also gain utility from engaging in the "game" or "sport" of investing.

While the "sport of investing" could explain why some remain actively investing in the market, it does not explain why so many investors needing "handholding" end up actively investing. We propose that the reason for this has more to do with the marketing of financial services than rationality. While many have argued about the price efficiency of the financial securities markets, recent events such as the conflict of interests between analysts and their investment banking employers highlight all too well the incongruency between the marketing of financial services and the actual needs of investors. Investors searching out handholding in the financial marketplace, end up being advised into active investing because of the greater amount of commissions to be made by those doing the handholding.

References

1. Bogle, J.C. (2000), "The Death Rattle of Indexing" *Perspectives on Equity Indexing*, 2nd edition (Frank J. Fabozzi Associates – New Hope, PA), pp. 2-19.
2. Bogle, J.C., "The Search for the Holy" March 29, 1998, http://www.vanguard.com/bogle_site/lib/investment.html (accessed May 10, 2001).
3. Buffet, W., quoted from 1996 annual report for Berkshire Hathaway in http://www.financial-planner.com.sg/nov1.html (accessed May 10, 2001).
4. Carhart, M., "On Persistence in Mutual Fund Performance" (1997), *Journal of Finance*, 52(#1): 57-82.
5. Fabozzi, F.J., editor, *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), 2000.
6. Friedman, M., quoted in "Bubblenomics" by Peter Brimelow, *Forbes*, May 3, 1999.
7. Grossman, S.J. and Stiglitz, J.E. "On the Impossibility of Informationally Efficient Markets." (June1980), *The American Economic Review*, Vol. 70, No. 3, pp. 393-408.
8. Hsu, J.C., Campollo, C. (January/February 2006), "New Frontiers in Index Investing. An Examination of Fundamental Indexation", *Journal of Indexes*, pp. 32-58
9. Johns, R.C. (2000), "The Active versus Passive Debate: Perspectives of an Active Quant" *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), pp. 46-69.
10. Kirzner, E., Fact and Fantasy in Index Investing, January 2000, http://www.bylo.org/pdf/fact.fantasy.in.index.investing.pdf (accessed September10, 2009)
11. Malkiel, B.G. (2000), "Indexing: A Continuing Winning Strategy" *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), pp. 20-37.
12. Mar, J., Bird, R., Casavecchia, L., Yeung, D. (June 2009), "Fundamental Indexation: An Australian Investigation", *Australian Journal of Management,* Vol. 34, No 1, pp. 1-20
13. Rubinstein, M. (2001), quoted in "This Crazy Market is Tough to Beat", *Wall Street Journal*, February 1, p. C1.

Appendix

Derivation of Expectations and Variance Terms

This appendix derives the formulas for $E_j[u_{ij}|y_{ij}]$ and $var_j[u_{ij}|y_{ij}]$ for cases 3 and 4.

Case 3: g_{ij}=½:

In all cases, the probability density function of u_i and ε_{ij} is $f(u_i, \varepsilon_{ij}) = e^{-u_i - \varepsilon_{ij}}$ for $u_i \geq 0$ and $\varepsilon_{ij} \geq 0$ as u_i and ε_{ij} are independent standard exponential random variables; this probability density function equals 0 whenever u_i or ε_{ij} is less than zero. In case 3 with g_{ij}=½, equation (2) becomes $y_{ij} = \dfrac{u_i + \varepsilon_{ij}}{2}$. Therefore, the cummulative probability distribution function of u_i and ε_{ij} is

$$F(u_i, y_{ij}) = \int_0^{2y_{ij}} \int_0^{2y_{ij}-u_i} e^{-u_i - \varepsilon_{ij}} du_i d\varepsilon_{ji} = \int_0^{2y_{ij}} \left(e^{-u_i} - e^{-2y_{ij}}\right) du_i .$$

Since the probability density function of u_i and y_{ij} is $\dfrac{\partial}{\partial u_i}\dfrac{\partial}{\partial y_{ij}} F(u_i, y_{ij})$, the probability density function of u_i and y_{ij} is $f(u_i, y_{ij}) = 2y_{ij}e^{-2y_{ij}}$. By integrating out u_i, we get the probability density function of y_{ij} by itself:

$$F_{Y_{ij}}(y_{ij}) = \int_0^{2y_{ij}} 2y_{ij}e^{-2y_{ij}} du_i = 4y_{ij}^2 e^{-2y_{ij}} .$$

The conditional probability density function of u_i given y_{ij} is

$$f(u_i \mid y_{ij}) = \frac{f(u_i, y_{ij})}{F_{Y_{ij}}(y_{ij})} = \frac{2y_{ij}e^{-2y_{ij}}}{4y_{ij}^2 e^{-2y_{ij}}} = \frac{1}{2y_{ij}} .$$

In other words, the conditional probability density function of u_i given y_{ij} is uniformly distributed between 0 and $2y_{ij}$. Therefore, the conditional expectation and conditional variance of u_i are:

$$E\left[u_i \middle| y_{ij}\right] = \int_0^{2y_{ij}} \frac{u_i}{2y_{ij}} du_i = y_{ij},$$

$$\operatorname{var}\left[u_i \middle| y_{ij}\right] = \int_0^{2y_{ij}} \frac{u_i^2}{2y_{ij}} du_i - (E\left[u_i \middle| y_{ij}\right])^2 = \frac{y_{ij}^2}{3}.$$

Case 4: $g_{ij} \in (0, ½) \cup (½,1)$:

The derivation of the conditional expectations and the conditional variance follows the same logic as in case 3, but the resulting equations are much more complex. To simplify the equations some define $\tilde{y}_{ij} \equiv \frac{1-2g_{ij}}{g_{ij}(1-g_{ij})} y_{ij}$. Then equation (2) becomes

$$\tilde{y}_{ij} = \frac{1-2g_{ij}}{1-g_{ij}} u_i + \frac{1-2g_{ij}}{g_{ij}} \varepsilon_{ij}.$$

The cummulative probability distribution of u_i and $\tilde{y}_{ij}$ is

$$F(u_i, \tilde{y}_{ij}) = \int_0^{\frac{1-g_{ij}}{1-2g_{ij}}\tilde{y}_{ij}} \int_0^{\frac{g_{ij}}{1-2g_{ij}}y_{ij} - \frac{g_{ij}}{1-g_{ij}}u_i} e^{-u_i - \varepsilon_{ij}} d\varepsilon_{ij} du_i = \int_0^{\frac{1-g_{ij}}{1-2g_{ij}}\tilde{y}_{ij}} \left(1 - e^{-\frac{g_{ij}}{1-2g_{ij}}\tilde{y}_{ij} + \frac{g_{ij}}{1-g_{ij}}u_i}\right) du_i.$$

The probability density function of u_i and $\tilde{y}_{ij}$ is $\frac{\partial}{\partial u_i}\frac{\partial}{\partial \tilde{y}_{ij}} F(u_i, \tilde{y}_{ij}) =$

$$f(u_i, \tilde{y}_{ij}) = \frac{g_{ij}}{1-2g_{ij}} e^{-\left(\frac{g_{ij}}{1-2g_{ij}}\right)\tilde{y}_{ij} - \left(\frac{1-2g_{ij}}{1-g_{ij}}\right)u_i} \text{ for } u_i \in \left[0, \frac{1-g_{ij}}{1-2g_{ij}} y_{ij}\right].$$

For u_i outside this range, the probability density function equals 0.

By integrating out u_i, we get the probability density function of $\tilde{y}_{ij}$ by itself:

$$F_{\tilde{Y}_{ij}}(\tilde{y}_{ij}) = \int_0^{\frac{1-g_{ij}}{1-2g_{ij}}\tilde{y}_{ij}} f(u_i, \tilde{y}_{ij}) du_i = \frac{g_{ij}(1-g_{ij})}{(1-2g_{ij})^2}\left[1 - e^{-\tilde{y}_{ij}}\right] e^{-\left(\frac{g_{ij}}{1-2g_{ij}}\right)\tilde{y}_{ij}}.$$

The conditional probability density function of u_i given y_{ij} is

$$f(u_i \mid y_{ij}) = \frac{f(u_i, y_{ij})}{F_{Y_{ij}}(y_{ij})} = \frac{1-2g_{ij}}{1-g_{ij}} \left(\frac{e^{-\left(\frac{1-2g_{ij}}{1-g_{ij}}\right)u_i}}{1 - e^{-\tilde{y}_{ij}}}\right) \text{ for } u_i \in \left[0, \frac{1-g_{ij}}{1-2g_{ij}} y_{ij}\right].$$

For values of u_i outside this range, $f(u_i|y_{ij}) = 0$.

Using this conditional probability density function, we can determine the conditional expectation of u_i given $\tilde{y}_{ij}$:

$$E[u_i \mid \tilde{y}_{ij}] = \int_0^{\frac{1-g_{ij}}{1-2g_{ij}}\tilde{y}_{ij}} u_i f(u_i \mid \tilde{y}_{ij}) du_i = \left(\frac{1-g_{ij}}{1-2g_{ij}}\right)\left(1 + \frac{\tilde{y}_{ij}}{1-e^{\tilde{y}_{ij}}}\right).$$

This is equation (3).

Also, using the conditional probability density function, we can determine the conditional expectations of u_i^2 given $\tilde{y}_{ij}$ and then we can determine the conditional variance of u_i given $\tilde{y}_{ij}$:

$$E[u_i^2 \mid \tilde{y}_{ij}] = \int_0^{\frac{1-g_{ij}}{1-2g_{ij}}\tilde{y}_{ij}} u_i^2 f(u_i \mid \tilde{y}_{ij}) du_i = \left(\frac{1-g_{ij}}{1-2g_{ij}}\right)^2 \left(2 + \frac{\tilde{y}_{ij}^2 + 2\tilde{y}_{ij}}{1-e^{\tilde{y}_{ij}}}\right).$$

$$var\left[u_i \middle| \tilde{y}_{ij}\right] = E\left[u_i^2 \mid \tilde{y}_{ij}\right] - \left(E\left[u_i \mid \tilde{y}_{ij}\right]\right)^2 = \left(\frac{1-g_{ij}}{1-2g_{ij}}\right)^2 \left[1 - \left(\frac{\tilde{y}_{ij}}{1-e^{\tilde{y}_{ij}}}\right)^2 e^{\tilde{y}_{ij}}\right].$$

The conditional variance above is equation (4).