The Indexing Paradox

Be Thankful for Irrational Analysts

by David Eagle, Ph.D.

Revised December 29, 2005

Abstract

This paper introduces the indexing paradox, which states that it if all investors are rational with rational expectations and have a common risk-averse investment performance measure, then 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 with no mechanism to price the possible investments. This paradox relies merely on understanding averages. It does not rely on markets being "informationally efficient," as demonstrated in a model where different investors have differing degrees of informational advantages and disadvantages.

David Eagle
Associate Professor of Finance
Department of Management
College of Business and Public Administration
Eastern Washington University
668 N. Riverpoint Blvd., Suite A
Spokane, WA 99202-1660

phone: 509-358-2245
fax: 509-358-2267
email: deagle@ewu.edu

The Indexing Paradox
Be Thankful for Irrational Analysts

by David Eagle, Ph.D.

One of the most important decisions facing stock investors is whether to use active investing or index investing (also called passive investing). According to Buffet (1996), "Most investors, both institutional and individual, will find that the best way to own common stocks is through an index fund that charges minimal fees. Those following this path are sure to beat the net results delivered by the great majority of investment professionals."

Bogle (1998) explains how the much higher costs associated with the average active mutual fund causes that fund's performance to be much less than a low-cost index fund's performance:

> " … the costs the fund incurs — advisory fees, operating expenses, marketing …, plus the cost of buying and selling portfolio securities … can be conservatively estimated at upwards of 2.0 percent per year. … [In] the past 15 years … an all-market index fund, operated at a cost of 0.2 percent … would have provided an annual return of 16.5 percent — or nearly 99 percent of the market return … [compared to]… 86 percent for the [average] managed fund…. In fact, the terminal value of an initial investment of $10,000 in the index fund would have been worth $98,800 (97 percent of the market result), while the terminal value of the same investment in the traditionally managed active fund would have been $74,200, only 73 percent of the market."[brackets added]

An additional advantage of index funds is that they pass less taxable capital gains onto their taxable investors because their turnover rate is so much less than active funds.[1]

---

[1] To be fair, one should list the disadvantages of indexing as well as the advantages. However, while there are some critics of indexing, their criticisms primarily are criticisms of the S&P 500 index funds in that this index is for large-cap stocks and thus does not represent the whole market. I agree with this criticism. However, the indexing underlying the indexing paradox must be a whole market index.

Given the arguments in favor of index funds, some may wonder why anyone would choose to invest in an actively managed fund when they could index. Responding to that thought, Friedman (1999) argued, "… if all owners of equities used index funds, there would be nothing to decide the prices of anything. It's the people who don't use index funds who are essentially setting the relative prices of different stocks." This paper presents a different perspective about the possibility of everyone indexing. While agreeing with Friedman's conclusion that the market would collapse if everyone were to index, this perspective also states it makes no economic sense that anyone would actively invest in a market consisting of all rational investors. If this sounds contradictory, it is, and this is the reason we call this perspective, "The Indexing Paradox."

While this paper is the first to formally label and identify "The Indexing Paradox," pervious writers have to some extent realized the paradox existed. Jones (2000) recognized this paradox but mistakenly assumed it depended on the existence of efficient markets. The present paper shows that the Indexing Paradox holds even in the absence of efficient markets.

The next section presents and proves the Indexing Paradox, which consists of four assumptions and a conclusion. To better understand this paradox, we then develop an equilibrium model of stock allocation involving utility-maximizing investors having comparative informational advantages. We use the model first as an example to show how the Indexing Paradox unfolds. Then we look more closely at two of the assumptions of the paradox to see whether those assumptions could be violated in reality, and if so, then how changes in those assumptions would affect the Indexing Paradox. At the end of the paper, we summarize the conclusions reached in the paper and reflect upon the real world implications of the Indexing Paradox.

A Succinct Statement of 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 with no mechanism to price the possible investments.

The Indexing Paradox stems from the often-ignored reality that in order for some to do better than the average, others must do worse than the average. This paradox assumes rational expectations in the sense that investors have unbiased expectations, which implies that the average of all the investors' expected investment performances equals the expected performance on the overall market.

Under rational expectations, some investors may expect to perform better than market, but then other investors must expect to perform worse than the market. However, it would be irrational for investors to engage in active investing with submarket expected performances when they could index and always get the market return. Therefore, in equilibrium, no investor would expect to do worse than the market, which implies that no investor would expect to do better than the market. In other words, no investor would expect to do better than indexing.

The above paragraph does not take into account the cost differentials between active and passive investing. If the cost of indexing is less than the cost of active investing as very clearly seems to be the case in reality, then no investor in such a world of rational investors would choose to actively invest; in other words, all investors would index. As the Indexing Paradox states, if all investors index, then there is no mechanism to price the possible investments.

We acknowledge that universal indexing has yet to materialize. Less than 10% of the equity market is under index management. Nevertheless, the Indexing Paradox still has implications to reality. In particular, if not everyone indexes, we must ask ourselves why. Is it because investors are irrational? Is it because investors do not have rational expectations, that instead they are, on average, delusional? Is it because investors are different, possessing different utilities and different levels of risk aversion, thus needing different performance measures? Or is it that indexing (before costs) does not actually do as well as the market?

Also, the Indexing Paradox may have relevance to the direction the market might head in the future. If investor irrationality is why not all investors index, education of the benefits of indexing, education about decision-making errors and bias, and more computerization to supplement or supplant human decisions may eventually lead to an increasing share of the market switching to indexing, which could bring us closer to the market demise that the Indexing Paradox predicts.

Discovering a new paradox is one thing; understanding it is another. The rest of this paper explores the Indexing Paradox in the context of an equilibrium model of expected utility maximizing investors possessing different degrees of comparative informational advantages and disadvantages.

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 distribution 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 instead 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.[2]

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] 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. 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.

Mathematical Details of Model

This one-year[3] 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

$$v_i = k_i u_i + (1 - k_i)\eta_i \qquad \textbf{(1)}$$

where v_i is the value of stock i at time 1, and u_i, 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:

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

where ε_{ij} is a random variable that has a standard exponential distribution and is independent from u_i and η_i,. As stated before, g_{ij} represents investor j's comparative

[3] 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.

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 $\text{var}_j[u_i \mid y_{ij}] = 1$ as y_{ij} provides no information on u_i. Therefore, $E_j[u_i \mid y_{ij}]$ and $\text{var}_j[u_i \mid y_{ij}]$ 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 $\text{var}_j[u_i \mid y_{ij}] = 0$. 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 $\text{var}_j[u_i \mid y_{ij}] = \frac{y^2}{3}$.

Case 4: $g_{ij} \in (0, ½) \cup (½,1)$. Where $\tilde{y}_{ij} \equiv \frac{1-2g_{ij}}{g_{ij}(1-g_{ij})} y_{ij}$, the conditional expectations and variances of u_i are:

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

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

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,

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

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

The return on stock i equals $r_i = \frac{v_i - p_i}{p_i} = \frac{v_i}{p_i} - 1$ 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:

$$E_j[r_i \mid y_{ij}] = \frac{E_j[v_i \mid y_{ij}]}{p_i} - 1 \quad \textbf{(7)}$$

$$\mathrm{var}_j[r_i \mid y_{ij}] = \frac{\mathrm{var}_j[v_i \mid y_{ij}]}{p_i^2} \quad \textbf{(8)}$$

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* σ_j^P) and U(1+ r_j^P -c* σ_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. Currently we are using c=2, which seems to give results sufficiently consistent with true expected utility maximization.

Figure 1: List of Parameter Values for Basic Monte Carlo Simulation

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)

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[4] they would have experienced had they indexed. As a result, those six investors switch to indexing.

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
	(27.8%)	(24.6%)	(24.6%)	(24.6%)				
2	-2.76%	9.20%	9.20%	9.19%	-7.15	6.05	6.06	6.05
	(28.0%)	(24.6%)	(24.6%)	(24.6%)				
3	-0.61%	9.20%	9.20%	9.19%	-4.64	6.05	6.06	6.05
	(28.1%)	(24.6%)	(24.6%)	(24.6%)				
4	-0.80%	9.20%	9.20%	9.19%	-5.5	6.05	6.06	6.05
	(29.8%)	(24.6%)	(24.6%)	(24.6%)				
5	2.39%	9.20%	9.20%	9.19%	-1.98	6.05	6.06	6.05
	(28.4%)	(24.6%)	(24.6%)	(24.6%)				
6	11.20%	9.20%	9.20%	9.19%	5.63	6.05	6.06	6.05
	(34.6%)	(24.6%)	(24.6%)	(24.6%)				
7	13.80%	4.97%	9.20%	9.19%	8.49	0.16	6.06	6.05
	(34.4%)	(32.2%)	(24.6%)	(24.6%)				
8	19.40%	6.41%	9.20%	9.19%	13.88	2.75	6.06	6.05
	(32.2%)	(26.9%)	(24.6%)	(24.6%)				
9	25.60%	12.30%	8.66%	9.19%	19.61	8.78	5.35	6.05
	(30.7%)	(25.7%)	(25.4%)	(24.6%)				
10	26.30%	13.10%	9.73%	9.18%	20.5	9.76	6.60	6.05
	(30.6%)	(25.0%)	(24.7%)	(24.6%)				
Indexers	9.18%	9.20%	9.20%	9.19%	6.04	6.05	6.06	6.05
	(24.6%)	(24.6%)	(24.6%)	(24.6%)				

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

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

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

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.

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.[5]

---

[5] That stock prices could only partially reflect information and not fully reflect that information 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'

Table 2: Simulation Results With 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
	(40.9%)	(24.6%)	(24.6%)	(24.6%)				
2	-20.80%	9.18%	9.19%	9.20%	-INF	6.04	6.05	6.06
	(38.7%)	(24.6%)	(24.6%)	(24.6%)				
3	-13.40%	9.18%	9.19%	9.20%	-INF	6.04	6.05	6.06
	(39.5%)	(24.6%)	(24.6%)	(24.6%)				
4	-11.80%	9.18%	9.19%	9.20%	-33.05	6.04	6.05	6.06
	(39.4%)	(24.6%)	(24.6%)	(24.6%)				
5	-4.60%	9.18%	9.19%	9.20%	-12.64	6.04	6.05	6.06
	(33.5%)	(24.6%)	(24.6%)	(24.6%)				
6	10.00%	9.18%	9.19%	9.20%	-INF	6.04	6.05	6.06
	(42.1%)	(24.6%)	(24.6%)	(24.6%)				
7	16.60%	1.10%	9.19%	9.20%	6.76	-5.82	6.05	6.06
	(50.6%)	(37.1%)	(24.6%)	(24.6%)				
8	35.40%	6.90%	9.19%	9.20%	23.48	2.84	6.05	6.06
	(49.1%)	(27.8%)	(24.6%)	(24.6%)				
9	50.40%	14.20%	8.63%	9.20%	34.73	10.4	5.32	6.06
	(52.2%)	(26.4%)	(25.4%)	(24.6%)				
10	51.70%	14.60%	9.74%	9.19%	35.85	10.99	6.60	6.05
	(52.0%)	(25.6%)	(24.7%)	(24.6%)				
Indexers	9.18%	9.20%	9.20%	9.19%	6.04	6.05	6.06	6.05
	(24.6%)	(24.6%)	(24.6%)	(24.6%)				

Note: Shaded area represents indexers.

In the simulations in table 1, we did not allow any margin trading because we restricted the fraction of funds invested in one stock to range between 0 and 1. To show that selling short and buying on the margin does not affect the indexing paradox, we ran new simulations where investors' fraction of funds in one stock could range between -1 and 2. This very generous margin trading allowance allowed investors to sell short the value of a particular stock equal to the amount of the investor's wealth. It also allowed

wealth or their utility functions. If individual investors are uncertain of this knowledge, then prices would only partially refect information.

investors to buy on the margin equal to twice that of the investor's wealth. No other restrictions were made other than the sum of the fractions of funds in each stock still needed to add up to one for each investor.

The results of this simulation with margin trading are shown in Table 2. The results are similar to the results without margin trading, except that often investor expectational errors due to ignoring the informational content of prices caused the investors to have expected utilities of negative infinity. Because the complexity of margin trading does not seem to affect the validity of the indexing paradox, the rest of the paper assumes no margin trading.

Now that we have seen the unfolding of the Indexing Paradox with this paper's model, we can investigate how changes in the assumptions of the Indexing Paradox can affect whether the Indexing Paradox still holds. The first two assumptions of the Indexing Paradox involve rationality, which the summary and conclusions section of this paper discusses.

The third assumption of the Indexing Paradox is that investors all have the same performance measure. Our model met this assumption by investors having identical utility functions; hence we used expected utility as the uniform measure of performance. We now revise the model to allow for two different utility functions depending on whether the investor's identifying number is odd or even. Even investors continue to have the logarithmic utility function, which has a coefficient of relative risk aversion of one. Odd investors, on the other hand, now have the utility function, U(r) = $2\left(\sqrt{1+r}-1\right)$, which has a coefficient of relative risk aversion of one-half.

Table 3: Simulation Results With Different Utility Functions

	Expected Returns and (Standard Deviations)				Expected Utility (in centi-utils)				
	number of indexers				number of indexers				
investor	0	6	8	9	0	6	8	9	
	(31.08%)	(24.60%)	(24.58%)	(24.59%)					
2	-2.62%	9.17%	9.19%	9.19%	-6.78	6.03	6.05	6.05	
	(27.37%)	(24.60%)	(24.58%)	(24.59%)					
3	-5.31%	9.17%	9.19%	9.19%	-8.02	7.57	7.60	7.59	
	(31.43%)	(24.60%)	(24.58%)	(24.59%)					
4	-1.03%	9.17%	9.19%	9.19%	-5.62	6.03	6.05	6.05	
	(29.21%)	(24.60%)	(24.58%)	(24.59%)					
5	3.10%	9.17%	9.19%	9.19%	0.27	7.57	7.60	7.59	
	(33.38%)	(24.60%)	(24.58%)	(24.59%)					
6	10.65%	9.17%	9.19%	9.19%	4.88	6.03	6.05	6.05	
	(35.59%)	(24.60%)	(24.58%)	(24.59%)					
7	18.89%	4.39%	9.19%	9.19%	15.58	1.80	7.60	7.59	
	(36.78%)	(33.68%)	(24.58%)	(24.59%)					
8	18.39%	5.30%	9.19%	9.19%	12.76	1.32	6.05	6.05	
	(32.85%)	(27.89%)	(24.58%)	(24.59%)					
9	27.36%	14.18%	9.40%	9.17%	23.97	12.29	7.73	7.58	
	(31.66%)	(26.03%)	(25.23%)	(24.59%)					
10	26.41%	12.79%	8.98%	9.19%	20.43	9.51	5.97	6.05	
	(30.27%)	(24.65%)	(24.09%)	(24.59%)					
Indexers	9.10%	9.17%	9.19%	9.19%	7.51	7.57	7.60	7.59	odd
	(24.55%)	(24.60%)	(24.58%)	(24.59%)	5.97	6.03	6.05	6.05	even

Note: Shaded area represents indexers.

For this next simulation, we assign each investor's comparative informational advantage parameter, g_{ij}, to be (j-1) div 2 divided by 4. ("div" represents the integer divide operation.) This results with the comparative informational advantage parameters being 0 for investors 1 and 2, 0.25 for investors 3 and 4, 0.50 for investors 5 and 6, 0.75 for investors 7 and 8, and 1.0 for investors 9 and 10.

The results of the simulation with these different utility functions are shown in Table 3. When all investors actively invest, investors 1, 3, and 5 have lower expected utilities than the 7.62 centi-utils they would have experienced had they indexed, and investors 2, 4, and 6 have lower expected utilities than the 6.05 centi-utils they would have experienced had they indexed. As a result, investors 1, 2, 3, 4, 5, and 6 switch to indexing.

When only four investors are actively investing, investors 7 and 8 have lower expected utility than the 7.59 and 6.02 centi-utils that they would have respectively experienced had they indexed. Therefore investors 7 and 8 switch to indexing. When only investors 9 and 10 index, investor 10 has a lower expected utility than if he/she had indexed. Therefore, investor 10, who is more risk-averse relative to investor 9, switches to indexing. Therefore, the conclusion of the Indexing Paradox holds in this example even though investors have different utility functions, which would cause them to have different measures of portfolio performance.

The fourth and final assumption of the Indexing Paradox is that indexing results with the market average return. We have found no reason to doubt this assumption for the indexing method used in this paper. Some may think that the rebalancing needs of an index funds could provide active investors with an opportunity to profit at the expense of the indexers. They may think that when active investors learn new information, these active investors can react to that information before indexing changes the weights the index fund uses for each stock." However, while index funds do need to adjust the fraction of funds they invest in each stock when prices change, they do not need to rebalance. With the indexing method used in this model, $f_{ij}^{I} = \frac{p_i s_i}{\sum_{k=1}^{n} p_k s_k}$ where f_{ij}^{I} is the fraction of funds invested in stock i by the jth index investor (assuming that investor j is an indexer). Define λ_{ij} to be the fraction of stock i that index investor j owns, which should equal the ratio of d_{ij}^I, the demand for stock i by index investor j, divided by the supply of stock i:

$$\lambda_{ij} = \frac{d_{ij}^I}{s_i} = \frac{f_{ij}^I w_j / p_j}{s_i}$$

Substituting in for f_{ij}^I gives:

$$\lambda_{ij} = \frac{p_i s_i}{\sum_{k=1}^{n} p_k s_k} \frac{w_j}{p_i s_i} = \frac{w_j}{\sum_{k=1}^{n} p_k s_k}$$

This shows that λ_{ij} is the same for all stocks i. In other words, index investor j owns an equal fraction of each stock. As a result we will drop the i subscript. The resulting λ_j represents investor j's "slice" of the market for each stock.

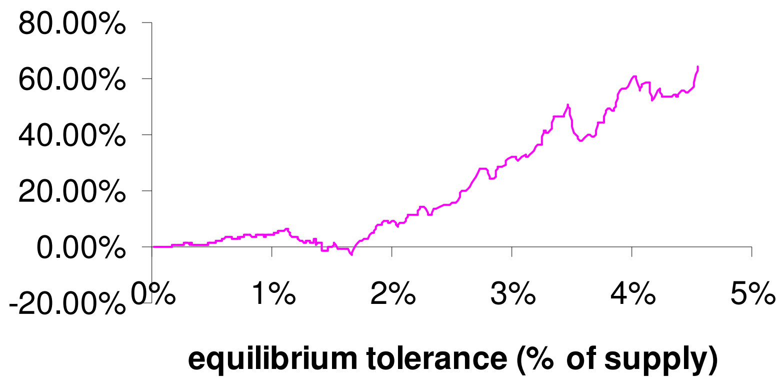
Now suppose the prices change after λ_j is set. Index investor j's new wealth will be $w_j^{new} = \sum_{i=1}^{n} p_i^{new} \lambda_j^{old} s_i = \lambda_j^{old} \sum_{i=1}^{n} p_i^{new} s_i$. Therefore, the new "slice" of each stock for the indexer will be:

$$\lambda_j^{new} = \frac{w_j^{new}}{\sum_{i=1}^{n} p_i^{new} s_i} = \frac{\lambda_j^{old} \sum_{i=1}^{n} p_i^{new} s_i}{\sum_{i=1}^{n} p_i^{new} s_i} = \lambda_j^{old} .$$

We conclude that index investor j's "slice" of each stock does not vary with prices. Therefore, index investor j need not rebalance as a result of price changes.

A second argument why indexing could do worse than the market average is that the market in reality uses market makers instead of a Walrasian auctioneer. Therefore, when they learn new information, some investors can react to that information before the market can change to a new equilibrium. This is part of the reason for the bid-ask spread of market makers – to buffer the market maker from investors responding to new information before the market maker learns this information.

Figure 2: Comparison of Utilities between Last Active Investor and Indexers (% Difference)



This paper's model can be used to investigate this argument by changing the tolerance in the numerical technique of determining equilibrium. With the numerical technique used, when the absolute difference between demand and supply for a specific stock is less than the allowed tolerance, the computer calls the result a quasi equilibrium. The resulting difference between demand and supply could be interpreted as being filled from the inventories of the market makers. The graph in Figure 2 shows how the difference in utility between the last active investor and the indexers varies with the equilibrium tolerance. For small equilibrium tolerances, the performance of the active investors usually but not always exceeds the performance of the indexers. However, for larger equilibrium tolerances, the active investors do outperform the indexers and the greater is the equilibrium tolerance, the greater is the degree to which active investors outperform the indexers.

However, in this paper's model, we have no one absorbing the cost of the equilibrium tolerance. When active investors react to information before the market makers respond, and when that information allows active investors to benefit relative to the market, the market markets are hurt. The market makers will pass on those costs through higher bid-ask spreads. All investors who buy or sell shares of stock will pay these costs. Since active investors are buying and selling much more frequently than do the indexers, the vast majority of these costs will be absorbed by the active investors rather than by the indexers. Therefore, this second argument is unlikely to cause the return for indexers to differ from the market return.

Summary, Conclusions, and Reflections

The Indexing Paradox states that if four assumptions hold, no active investor can do better than indexers. We explored the Indexing Paradox with a model where we saw investors, who were expecting to do worse than indexing, switching to indexing until only one active investor was left, which was our full equilibrium. However, this "full" equilibrium was a precarious equilibrium. The model had no cost of active investing and no cost of indexing. If we were to add those costs, with the cost of active investing exceeding the cost of indexing, then even that last active investor will switch to indexing, resulting with a collapse of the market as no one will be left to set prices of the individual stocks.[6]

Since the conclusion of the Indexing Paradox rests on its assumptions, we then investigated whether those assumptions necessarily hold and if they may not hold, then how would the conclusion of the index paradox change when the assumption changed.

[6] In our model, we assume investors are pricetakers even with only one investor actively investing. Clearly as the number of investors shrink, the assumption of competitive behavior becomes less justifiable. However, while our model does not consider noncompetitive behavior, the Indexing Paradox itself does not assume competitive markets.

We contemplated possible violations to the fourth assumption, that the return on indexing must equal the market return, and found no grounds to doubt this assumption.

We looked at the third assumption of one performance measure and considered two performance measures (utility functions) for two different groups having different levels of relative risk aversion. In our simulations, the Indexing Paradox survived our relaxing this assumption.

We, however, know that active investors are currently very plentiful in the real world. (See Rubinstein, 2000). Therefore, at least one of the four assumptions of the Indexing paradox must be violated. If the return on indexing must equal the market return (assumption 4) and if violating assumption 3 on one performance measure does not affect the conclusion of the indexing paradox, then one of the first two assumptions on rationality must be violated. Therefore**, because some investors do actively invest, some investors must be irrational**. Either some investors have irrational expectations or some make irrational decisions.

Most likely the assumption of irrational expectations is violated in reality. In order for investors to remain active investors, they must all expect to do better than the market. However, by basic principles of averages, only half of them will do better than the market average. Therefore, half of these active investors must be delusional, expecting to do better than the market when in fact they probably will do worse. If all investors realize that half of the investors are delusional, then how could an active investor, who expects to do better than the market, be sure that they are part of the rational half of active investors and not part of the delusional half?

From the behavioral side of finance and from recent psychological research, evidence has accumulated that people (including investors) make irrational decisions and have biased expectations. Perhaps, some active investors or money managers may expect that they have a comparative informational advantage because of being trained not to make these errors. On the other hand, maybe they are just using this behaviorial side of finance as a rationalization for their remaining an active investor when in fact they may be of the half who are delusional in thinking that they expect to do better than the market.

In one sense this paper can be looked on favorably by the finance behaviorialists in that it does point out that the very existence of our stock market depends on the existence of some investors being irrational or having irrational expectations. However, even behaviorialists themselves cannot escape the Indexing Paradox. Imagine that the behaviorialists are right that investors do make these irrational decisions and have irrational expectations. Imagine also, that, thanks to the behaviorialists, investors sometime in the future are educated not to make those errors or biases or that "rational" computers are used to augment or supplant human decision making to correct for those errors or biases. Also suppose investors are educated about indexing. Once rationality is brought to the stock market, the Indexing Paradox may appear with the market demise that it predicts.

While the Indexing Paradox predicts a market collapse when all investors index, as long as some investors are active (albeit for irrational reasons) the rest of the market is not significantly affected. Tables 1, 2, and 3 show that, as long as at least one investor remains active, the expected utility of indexers changes only slightly and in no clear direction when the active investors switch to indexing.

Both rational active investors and indexers should be thankful for the existence of irrational active investors in the market. The indexers benefit because without those irrational active investors, there would be no market. The rational active investors benefit because the irrational active investors allow the rational active investors to beat the market. However, the indexers have the comfort of knowing that they are not among the irrational active investors. Most active investors will never know with certainty that they are among the rational half of the active investors rather than among the irrational half.

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 $y_{ij} = \frac{u_i + \varepsilon_{ij}}{2}$. Therefore, the cumulative probability distribution function of u_i and ε_{ij} is

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

Since the probability density function of u_i and y_{ij} is $\frac{\partial}{\partial u_i}\frac{\partial}{\partial y_{ij}} F(u_i, y_{ij})$, the probability density function of u_i and y_{ij} is $f(u_i, y_{ij}) = 4y_{ij}e^{-2y_{ij}}$. By integrating out u_i, we get the probability density function of y_{ij} by itself:

$$F_{Y_{ij}}(y_{ij}) = \int_0^{2y_{ij}} 4y_{ij}e^{-2y_{ij}} du_i = 8y_{ij}^2 e^{-2y_{ij}}$$

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

$$f(u_i \mid y_{ij}) = \frac{f(u_i, y_{ij})}{F_{Y_{ij}}(y_{ij})} = \frac{4y_{ij}e^{-2y_{ij}}}{8y_{ij}^2 e^{-2y_{ij}}} = \frac{1}{2y_{ij}}$$

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:

$$E[u_i \mid y_{ij}] = \int_0^{2y_{ij}} \frac{u_i}{2y_{ij}} du_i = y_{ij}$$

$$\text{var}[u_i \mid y_{ij}] = \int_0^{2y_{ij}} \frac{u_i^2}{2y_{ij}} du_i - (E[u_i \mid y_{ij}])^2 = \frac{y_{ij}^2}{3}$$

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 cumulative probability distribution of u_i amd $\tilde{y}_{ij}$ 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

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

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 $\tilde{y}_{ij}$:

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

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

The conditional variance above is equation (4).

References

Bogle, J. C., “The Death Rattle of Indexing,” *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), 2000, pp. 2-19.

Bogle, J. C., "The Search for the Holy," March 29, 1998, http://www.vanguard.com/bogle_site/lib/investment.html (accessed May 10, 2001)

Buffet, W., quoted from 1996 annual report for Berkshire Hathaway in http://www.financial-planner.com.sg/nov1.html (accessed May 10, 2001).

Carhart, M., "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52(#1):57-82, 1997.

Fabozzi, F. J., editor, *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), 2000.

Friedman, M., quoted in "Bubblenomics," by Peter Brimelow, *Forbes*, May 3, 1999.

Miller, R. “Measuring the True Cost of Active Management by Mutual Funds,” Finance 0506010, Economics Working Paper Archive EconWPA, revised 08 Jul 2005.

Reinker, Kenneth S. and Edward Tower, “Index Fundamentalism Revisited”, *Journal of Portfolio Management*, January 2005.

Grossman, S. J. and Stiglitz, J. E. "On the Impossibility of Informationally Efficient Markets." *The American Economic Review*, Vol. 70, No. 3, (June1980), pp. 393-408.

Johns, R. C., “The Active versus Passive Debate: Perspectives of an Active Quant,” *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), 2000, pp. 46-69.

Malkiel, B. G., “Indexing: A Continuing Winning Strategy,” *Perspectives on Equity Indexing*, 2nd edition, (Frank J. Fabozzi Associates – New Hope, PA), 2000, pp. 20-37.

Rubinstein, M., quoted in "This Crazy Market is Tough to Beat," *Wall Street Journal*, February 1, 2001, p. C1.