Yale ICF Working Paper No. 00-50
August 2001

A THEORY OF MUTUAL FUNDS: OPTIMAL FUND OBJECTIVES AND INDUSTRY ORGANIZATION

Harry Mamaysky
Yale School of Management

Matthew Spiegel
Yale School of Management

August 20, 2001

Yale SCHOOL of MANAGEMENT

A Theory of Mutual Funds: Optimal Fund Objectives and Industry Organization

by

Harry Mamaysky†

and

Matthew Spiegel‡

August 20, 2001

Comments Welcome

†Yale School of Management, P.O. Box 208200, New Haven CT 06520-8200. Phone: 203-436-0649, Fax: 203-436-0630, email: **harry.mamaysky@yale.edu**, web page: **http://som.yale.edu/~hm68**.

‡Yale School of Management, P.O. Box 208200, New Haven CT 06520-8200. Phone: 203-432-6017, Fax: 203-432-8931, email: **matthew.spiegel@yale.edu**, web page: **http://som.yale.edu/~spiegel**.

For helpful discussions and comments we thank Utpal Bhattacharya, Simon Gervais, A. Subrahmanyam and seminar participants at Rutgers University, Tuck School of Management, University of Connecticut, the 2001 CEPR/JFI Symposium at INSEAD, and the 2001 Cowles Foundation Conference on Missing Financial Markets at Yale University.

A Theory of Mutual Funds:
Optimal Fund Objectives and Industry Organization

Abstract

This paper presents a model in which investors cannot remain in the market to trade at all times. As a result they have an incentive to set up trading firms or financial market intermediaries (FMI's) to take over their portfolio while they engage in other activities. Previous research has assumed that such firms act like individuals endowed with a utility function. Here, as in reality, they are firms that simply take orders from their investors. From this setting emerges a theory of mutual funds and other FMI's (such as investment houses, banks, and insurance companies) with implications for their trading styles, as well as for their effects on asset prices. The model provides theoretical support for past empirical findings, and provides new empirical predictions which are tested in this paper.

JEL Classification: G20, G12

Banks, investment houses, and mutual funds have in recent years created a wide array of vehicles that trade on behalf of investors. In 1999, for example, U.S. equity funds managed roughly 6 trillion dollars; in 1990 this number was only 300 billion. Presumably, such financial market intermediaries (FMI's) meet some particular investor demand. A number of empirical papers have noted that the plethora of existing financial institutions exhibit a wide range of trading behaviors, many of which are difficult to reconcile with the existing theory. Gruber (1996) notes that while most models predict that investors will create only a small number of passively managed funds, in reality thousands of funds exist and most of these are actively managed. Moreover, as Sharpe (1992) and Brown and Goetzmann (1997) demonstrate, mutual funds encompass a fairly wide array of dynamic trading strategies. This paper attempts to bridge existing data and theory. At the same time it produces a number of new hypotheses, several of which are tested and found to hold within the U.S. mutual fund industry.

This paper takes a first principles approach to the issue of financial market intermediation. It does so by using as primitives the population of investors, their preferences, and trading technologies. FMI's are defined as corporations that can trade securities on behalf of their investors. Thus, the model can be applied to both mutual funds and the trading arms of an investment bank. Since FMI's are corporations, society can create them as it sees fit and can give them whatever instructions it likes. From this setting emerges a distinct theory of such institutions and their behavior.[1]

The multitude of managed funds noted by Gruber (1996) arise naturally within our model. Essentially, funds and other financial firms cater to a population of individuals with different desired trading strategies. While each investor might like to see his or her optimal trading strategy carried out, this level of customization is economically infeasible, unless FMI's can be produced and staffed at no cost. A second best solution relies on a small number of intermediaries, each of which trades along what can be thought of

---

[1]This paper's analysis differs from the traditional principal-agent paradigm often used to examine investor-fund manager interactions (Allen and Gorton (1993), Dow and Gorton (1997), Ou-Yang (1997), Das and Sundaram (2000), Nanda Narayanan, and Warther (2000)). These papers generally focus (at least in part) on how to design a contract that will induce a fund manager to take unobservable actions on behalf of his investors. In contrast, this paper assumes that the agent's (FMI's) actions are contractible. Hence the agent may be induced to carry out exactly the principal's orders. The paper therefore asks what those orders should be given certain technological restrictions. Obviously, principal-agent problems are an important issue for FMIs and this paper does not suggest otherwise. Rather it offers a complimentary theory that helps explain some issues that seem unrelated to moral hazard or adverse selection issues.

as a unique strategy basis. Investors then carry out their preferred trading strategies by selectively buying shares in these different fund types to match as closely as possible their preferred trading strategy.[2]

Note the distinction here between trading strategies and securities: FMI's offer agents particular trading strategies, rather than simply committing to hold some fixed portfolio of securities. Since the number of desired trading strategies may be quite large (even in a two security world), the number of funds may bear very little resemblance to the number of traded securities.

In most models of market microstructure, FMI's are assumed to behave as utility maximizing individuals. These firms (or people) are then assumed to be either risk neutral (as in Kyle, 1985), or risk averse (as in Ho and Stoll, 1983, and Campbell and Kyle, 1993).[3] While this paradigm has proven itself extremely useful for many applications, it ignores the fact that financial firms are corporations that simply carry out their investors' instructions. This is how they are modeled here. As a result, the equilibrium comparative statics and thus this paper's empirical predictions differ from those produced by prior models. For example, if trading firms are risk averse individuals then additional firms reduce price risk and the equilibrium discount rate. In our setting, the creation of additional trading firms increases price risk, and has no direct effect on the discount rate.

Extending the model to allow for multiple risky assets produces a theory of fund families. Each family may be thought of as having a research department which collects information about the state of the economy. Funds in a family then share this information, and trade portfolios of securities based on the restrictions in their prospectuses. In equilibrium, fund families offer to trade portfolios which are useful to the subset of the population which is concerned with the economic information to which that fund family has

[2]While people generally think of mutual funds as the stand ins used by investors to trade on their behalf, other institutions perform the same function. Someone purchasing a share of Merrill Lynch allows Merrill to trade on their behalf. Thus, the paper's use of the word "fund" is meant to encompass other FMI's as well, unless otherwise noted.

[3]The number of papers on the impact risk averse utility maximizing FMI's have on prices has been expanding quite rapidly as of late. Articles include Biais, Foucault, and Salanié (1998), Pirrong (1999), Massa (2000), and Viswanathan and Wang (2001). In contrast to the current paper, these models take the FMI's objective function as an exogenous part of the environment.

access. The result of this analysis is a theory that provides an explanation for why fund families exist, and why people frequently invest in several funds both within and across fund families.

An empirical prediction of the model is that newly created fund families should provide trading *strategies* which are maximally different from those of existing funds. At the fund family level, new funds should be created which allow investors to take advantage of the firm's strategy in new ways. To understand the difference consider two funds, both of which own only one stock, say IBM. Further suppose that on *average* each hold 80% of their portfolio in this stock and 20% in bonds. While both hold IBM each fund's strategy will differ if they buy and sell IBM at different times. On the other hand if each family were to introduce a new fund the model predicts that this new fund would trade in a stock other than IBM in order to allow investors to take advantage of the fund's information set in a new way. The data confirm these predictions. For example, when a fund family with only a few funds introduces a new fund it will typically use a strategy that places the new fund in a different Morningstar category than its older siblings. Another test of the model comes from an examination of the asset allocation decisions made by funds. According to the model investors should value funds that help them to time their entry into and out of particular parts of the market. In fact, over 1,700 funds exist which move of at least 20% of the portfolio in and out of stocks during their lifetime.

The model predicts that investors value a range of dynamic trading strategies. One place to look for this effect is in fund loadings (or betas) on economic risk factors (such as the market portfolio). The model suggests that these betas should be relatively highly correlated for funds within a family, as intra-family funds share a common signal from the family's research department. Across families the correlations should be lower. The data support both predictions.

A more subtle prediction of the model is that when a fund family starts up is should follow a relatively unique strategy. This means the time varying betas (or equivalently factor loadings) of such funds should exhibit particularly low correlations with the funds offered by other families. For later funds in a given family the correlations should increase as families fill in the strategy space, thus forcing the introduction of products closer to those already offered elsewhere. Our analysis finds exactly this pattern. Funds introduced earlier in a family's life exhibit lower time varying beta correlations with other funds in our sample than do those funds which are introduced later. This pattern is monotonic across pairings. That is pairing the two oldest funds across families produces (on average) a lower correlation in their time varying betas than pairing the two second oldest funds, or the oldest from one family with the second oldest from another.

The paper is organized as follows. Section 1 provides an example that lays out the general problem. Section 2 presents the model and our equilibrium concept. Section 3 studies a special case of the model where there is a single risky security. Section 4 provides examples of sample economies. Section 5 extends the analysis to the case of an economy with multiple risky securities. Section 6 presents empirical evidence supporting the model. Section 7 relates the current paper to the existing literature, and Section 8 contains the paper's conclusions.

1 Outline of the Problem

Consider a market at one particular point in time labeled period two. In the real world, some investors will appear in the market to trade on their own behalf. Other investors, the vast majority in fact, will find that their other commitments prevent them from participating directly in the market. Instead they will rely on mutual funds, specialist firms, and investment banks to take their place. These institutions trade on behalf of their investors with those who are currently present in the market. In order to model this phenomenon, a mechanism is needed which generates trade amongst a group of investors.

Consider an example economy. There exists a riskless bond with a normalized price of one and a return of zero. In addition, imagine the market contains K risky assets with normally distributed terminal dividends D, a $K\times 1$ vector with means of zero and variance-covariance matrix Σ_D, that will be paid out in period three. In period two agents receive a mean zero normally distributed endowment shock vector $N(i)$ of length K with variance-covariance matrix $\Sigma_{N(i)}$. This vector represents shares of risky securities whose payoffs occur in period three. For simplicity, these payoffs are perfectly correlated with the dividends of the tradeable securities just described. Hence, trading in the appropriate set of available risky securities provides a hedge against the period three payoffs from the endowment shock vector.

These endowment shocks are a modeling device which generates a marketable (or tradable) form of heterogeneity in the economy. One interpretation is that they arise from non-tradable assets whose cash flows are correlated with financial securities such as human capital, or real estate. Along these lines Davis and Willen (2000a,b) find that innovations in people's labor incomes are correlated with returns on certain financial securities and that these correlations vary across socioeconomic groups. Another interpretation uses consumption differences to produce this type of heterogeneity. For example, people in the Northeast region of the U.S. are relatively large consumers of oil. Thus, they may be happy to hold stocks whose returns are positively correlated with the price of oil, even if these stocks have relatively low returns. Another (very loose) interpretation is that endowment shocks proxy for heterogeneity in agents' beliefs about future performance of the risky securities (although this paper does not consider any issues of asymmetric information). Whatever the source, the model only requires that period two asset demands vary across people in ways that are not totally predictable by them ahead of time.

This still leaves open the question of the degree to which endowment or consumption shocks individuals might yield meaningful differences among investors. Ultimately, of course, this is an empirical question. But, there is anecdotal evidence which suggests that these differences may be quite important. First, Morningstar rates mutual funds within categories. A five star rating implies that a fund is a top performer relative to others using the same strategy. Note, this means that Morningstar does not rate funds by whether or not they "beat" the market, but by whether or not they outperformed other funds inside their objective category. The popularity of these ratings suggests that investors seek heterogenous fund strategies for idiosyncratic reasons. Second, articles in the popular press indicate that investors seek funds that will "fit" with their risk needs. For example, an article by Hechinger (2001) in *The Wall Street Journal* describes some recent changes at Fidelity Investments that caught some investors unawares:

> Many investors placed Mr. Vanderheiden in the value camp. Gordon Jackson . . . says he started buying [Destiny I because of its] . . . cautious approach. Mr. Jackson, who now works at a technology firm, figures he lost about $120,000 because of the strategy shift [into high tech] at Destiny I and may have to put off his retirement for several years.

Third, investment banks advertise that they can help their wealthy clients with hedging and diversification. For example, Merrill Lynch recently held a series of seminars on, "Hedging, Diversifying, and Monetizing Your Wealth." They also ran an advertisement in the *New York Times Magazine* (2001) stating "Your risk profile drives investment decisions." Clearly, Merrill Lynch must believe that its clients differ in the risks they face, and thus differ in what investments they will prefer.

Returning to the example, if investors have exponential (CARA) utility functions over period three consumption then standard arguments show that individual *i* will demand

$$X_2(i) = -X_1(i) - N(i) - c(i)P_2 \qquad \textbf{(1)}$$

shares of the $K\times 1$ security vector. Equation **(1)** states that *i*'s demand (represented by $X_2(i)$) is a linear function of three variables. The first vector, $X_1(i)$ represents the number of shares of the risky securities held by the investor coming into period two. The $N(i)$ term equals the trader's endowment shock of the risky assets. Following standard practice, assume $N(i)$ and P_2 are independent random vectors. The term, c_k is a constant matrix dependent on Σ_D, and the trader's risk aversion coefficient. Finally, P_2 equals the security vector's market price.

Obviously, every investor would like the FMI's that he has invested in to mimic this trade. But an FMI cannot do so without knowing $X_1(i)$, $N(i)$, and $c(i)$. Since these variables are investor specific, and since the economy may have many investors, this level of customization is infeasible. FMI's are companies and investors are able to buy shares of these companies. Hence investors divide the FMI's trading profits in proportion to their investment in the FMI. For example if Fidelity's Contra Fund purchases 1,000,000 shares of IBM, then Contra Fund investors split the trade in proportion to the number of fund shares they hold. The Contra Fund does not provide a "personalized" service by performing specific trades on behalf of individuals. Instead it provides a service by following a particular, and pre-announced, rule for buying shares. Investors then customize their own exposure to this trading rule by investing varying amounts of money into the Contra Fund.

While an FMI cannot offer personalized trades, it can offer to trade in some particular set of securities on the basis of some signal. That is an FMI can offer to follow a strategy that may be useful to individuals with particular values of $X_1(i)$, $N(i)$, and $c(i)$. Clearly this offers investors a potentially useful service. While they may not wind up with the exact position they would have had by trading on their own behalf, investors can at least get part way there without spending all of their time watching the market. This naturally raises two questions, which are the focus of this paper: What instructions will a diverse population of investors give to the set of FMI's? And what FMI's should exist in the economy?

2 The Formal Model

There are three dates ($t = 1,2,3$). Investors receive utility from a single consumption good (cash) and it serves as the numeraire. There exist two measure one continua of investors labeled date one and date two respectively. All investors have identical information sets. Date one investors have the ability to create financial market intermediaries that can implement trades in date two.

2.1 Financial Assets

Each of the securities in the economy may be traded at either date one or two. A position in a $K\times1$ column vector X of the risky securities and in B shares of the bond at the end of t=2 will result in a t=3 payment equal to $X'D+B$ units of the consumption good.

2.2 Financial Market Intermediaries

Each FMI $j\in\{1,2,...,J^*\}$ observes a t=2 random variable, e_j. The signal e_j provides the FMI with information about the period one investors' t=2 endowment shocks. Informative nonpublic signals are costly to acquire, which ensures that only a finite number of such FMI's will be created. Public signals are free.

FMI's are "robotic entities" that obey whatever instructions they have been given by their date one investors. Each FMI j has a technology which allows it to purchase a $K\times1$ dimensional vector $f_j^i(e_j,I)$ of shares of the risky assets on behalf of date one investor i. Here I refers to all publically available time two information. For example, $P_2 \in I$ but $e_j \notin I$ for all j. For notational convenience, the paper suppresses the dependence on I. Such purchases occur at the prevailing t=2 market prices, P_2, and hence provide investor i with $f_j^i(e_j)'(D-P_2)$ units of the consumption good at t=3. Each date one investor can choose the function $f_j^i(\bullet)$ which he submits to FMI j.

These functions are clearly more general than the trades that real FMI's are able to implement. In particular, the dependence of $f_j^i(\bullet)$ on i is troubling: after all, Fidelity cannot tailor its funds to make specific trades on behalf of individual investors. Based upon this observation, call the instructions submitted to a fund *feasible* if they satisfy the following criteria.

Definition: *A trade request* $f_j^i(\bullet)$ *to an FMI is* feasible *if it is of the following form*

$$f_j^i(e_j) = x_j^i f_j(e_j) \tag{2}$$

for some constant x_j^i *and some function* $f_j(e_j)$*, which is not investor specific.*

This definition insures that one FMI cannot use information acquired by another FMI. Furthermore, it insures that FMI's act as firms, and only allocate their trades in proportion to an investor's position in the FMI. Thus, somebody owning twenty shares of an FMI receives twice the allotment of the fund's trading revenues as somebody owning ten shares. The paper will show that, within the model, the optimal trade requests made by a period one investor are indeed feasible.

By convention, for FMI's with $j \leq J$ the signals e_j are random variables that are measurable with respect to t=2 information. For some set of FMI's with $J<j\leq J^*$ the signal is e_j=1 (the *constant* funds), and for another set of FMI's also with $J<j\leq J^*$ the signals are given by $e_j=P_2(k)$ (the *price* funds). The model assumes that FMI's which trade only on public signals can be set up at no cost, and thus exist to whatever degree investors demand. Thus, if investors wish, funds J+1 and above always exist to the extent necessary. Hence the choice variable with respect to the creation of costly FMI's becomes J which can assume any value 0, 1, 2, 3, . . . etcetera .

2.3 Date One Investors

Date one investors, referenced by $i\in[0,1]$, enter the model at the beginning of t=1 endowed with $X_0(i)$ shares of the risky securities (a $K\times 1$ dimensional vector), and $B_0(i)$ units of the consumption good. At t=2, investor i receives a random endowment shock of $N(i)$ shares of the risky securities, and $N_0(i)$ units of the consumption good (cash). The endowment shock results in no immediate payments, but at t=3, it produces $N(i)'D+N_0(i)$ units of the consumption good for the investor.

Due to other full time commitments (such as employment) date one agents cannot participate in the financial markets at t=2. Instead, at t=1, they submit a set of J^* demands of the form $f_j^i(\bullet)$ to the funds in the economy. Given these instructions, an agent receives $\left(\sum_{k=1}^{J+K+1} f_j^i(e_j)\right)'(D-P_2)$ units of the consumption good at t=3. Additionally, an agent can trade in financial markets at the end of date one, after having learned $X_0(i)$, but before having seen $N(i)$ or $N_0(i)$. Refer to an agent's holdings of the risky asset after trade at t=1 as $X_1(i)$. The agent and FMI budget constraints imply that

$$C_3 = [X_0(i) - X_1(i)]' P_1 - \sum_{j=1}^{J+K+1} f_j^i(e_j)' P_2 + N_0(i) + \left[N(i) + X_1(i) + \sum_{j=1}^{J+K+1} f_j^i(e_j) \right]' D. \quad \textbf{(3)}$$

where C_3 equals date three consumption. Then given his date one information set, each date one agent maximizes his expected utility from date three consumption with respect to X_1, and the set of fund instructions $\{f_1^i, f_2^i, \ldots, f_{J^*}^i\}$. Hence each date one investor solves

$$\sup_{X_1, \{f_1^i, f_2^i, \ldots, f_{J+K+1}^i\}} E_1[U(C_3)], \quad \textbf{(4)}$$

where the expectation is taken with respect to time one information.

2.4 Date 2 Investors

Date two investors enter the market at t=2 with an endowment of $\widetilde{X}_0(i)$ shares of the risky securities. Assume $\widetilde{X}_2(i)$ is normally distributed, independently of the model's other random variables, with a possibly nonzero mean. Date two investors can trade only in period two, and care about t=3 consumption. Hence each date two investor chooses $\widetilde{X}_2(i)$ to maximize his expected utility over date three consumption subject to the budget constraint $\widetilde{C}_3 = (\widetilde{X}_0 - \widetilde{X}_2)P_2 + \widetilde{X}_2 D$.[4] Hence the date 2 investors solve

$$\sup_{\widetilde{X}_2} E_2[U(\widetilde{C}_3)], \quad \textbf{(5)}$$

where the expectation is taken with respect to time two information.

2.5 Equilibrium

The equilibrium concept is a standard rational expectations, Walrasian type equilibrium. It specifies a collection of J^* mutual funds, agents' trading policies $\{X_1(i), f_j^i(\bullet) \forall j, \widetilde{X}_2(i)\}$, a time one price P_1, and a price process $P_2\left(N(i), e_1, \ldots, e_J, \widetilde{X}_0(i)\right)$ which must satisfy the following conditions:

1. Date one agents' demands $\{X_1(i), f_j^i(\bullet) \forall j\}$ are optimal.
2. Date two agents' demands $\{\widetilde{X}_2(i)\}$ are optimal.
3. The financial markets clear at date one,

[4] For simplicity, period two agents receive no endowment shocks beyond $\tilde{X}_0(i)$.

$$\int_0^1 [X_1(i) - X_0(i)]di = 0. \quad \textbf{(6)}$$

4. The financial markets clear at date two,

$$\int_0^1 \left[\widetilde{X}_2(i) + \sum_{j=1}^{J^*} f_j^i(e_j) \right] di = \int_0^1 \left[X_0(i) + \widetilde{X}_0(i) \right] di. \quad \textbf{(7)}$$

Agents take as given the funds and the fund signals, as well as the price process at t=2. They then submit their optimal demands in the financial markets and instructions to the mutual funds.

Since FMI's are set up at t=1, it is natural to assume that they are created in a way that is maximally beneficial for the investors who are in the economy at that time (and hence not the t=2 investors). While there exit different ways to model this process, this paper takes the view that in the long run one expects the equilibrium configuration to maximize the social welfare of the period one traders. For those that prefer to view the problem though a decentralized mechanism, one can equivalently assume that the following process occurs. In period one any traders can seek funding to start FMI's, by obtaining the maximum amount any investor would pay via first degree price discrimination if necessary. They can also offer to purchase existing FMI's and potentially alter the type of signal that they may acquire (for example, Citibank recently purchased Salomon Brothers and then altered Salomon Bothers' operations). FMI's are then added until no individual can raise enough money to cover the welfare costs of entry κ. In the event that there exists a unique social welfare optimum configuration of FMI's one can easily show the two approaches are equivalent.

Based upon the above arguments the set of FMI's solves the following problem

$$\sup_{\{J, e_1, \ldots, e_J\}} \int_0^1 E[U(C_3(i))] di - J\kappa. \quad \textbf{(8)}$$

This determines the number of FMI's, as well as the signal chosen by each. Since the benefit of each additional FMI must decrease once enough funds are in the economy, a cost for each additional fund insures that an interior solution for J exists.

3 Equilibrium with a Single Risky Security

To help gain some insight into the full model assume for now that there exists only one risky asset. In conformance with standard notation, represent the now scalar dividend variance by σ_D^2, and the scalar endowment shock $N(i)$ by $\sigma^2_{N(i)}$. To help simplify the algebra, assume $N_0(i)$ is given by $-N(i)P_2$. This implies that investors must pay for their endowment shocks out of their holdings of the consumption good.[5]

The following table provides a comparison of this paper's one security model to a typical model of FMI's found in the microstructure literature.

Summary 1: A Comparison with the Assumptions In a Standard Model of Financial Market Intermediation		
Assumption	Standard Model	First Principle's Model
Security Payoffs	Risky asset with normally distributed payoff. Risk free asset.	Same
Investor Entry and Exit Dates	Agents enter for one period, trade, and leave.	Same
Investor Utility Functions	Negative exponential.	Same
Financial Market Intermediary (FMI) Entry and Exit Dates	FMI's can trade in both periods one and two.	FMI's are formed in period one and trade in period two.
FMI Objective Function	Assumed risk neutral or negative exponential utility function.	Determined by the firm's investors.

3.1 Solution of the One Security Economy

Standard arguments yields the following equilibrium demand schedule for a date two investor: $\widetilde{X}_2(i) = -\widetilde{X}_0(i) - P_2/(\gamma\sigma^2_D)$. For a date one investor, given the signals of the FMI's his second period endowment shock can be written as $N(i) = \alpha_1(i)e_1 + ... + \alpha_J(i)e_J + \eta(i)$, where the α's represent weights on the signals e, and $\eta(i)$ an uncorrelated residual. If a date one trader could trade at t=2, but only observed the funds' signals $\{e_j\}$ and not his own endowment shock $N(i)$, his optimal trade δ is given by the following lemma.

Lemma 1*: Investor i sets* δ *to solve*

$$\delta = -X_1(i) - \sum_{j=1,...,J} \alpha_j(i)e_j - \frac{1}{\gamma}\sigma_D^{-2}P_2. \qquad \textbf{(9)}$$

Proof: See the Appendix for this and all other proofs.

[5]Charging the trader P_2 to obtain the untraded asset simply reduces the model's algebraic complexity. It has no qualitative impact upon the paper's results. Similar assumptions have been made previously in the literature (for example, see Vayanos (2001)).

The trade given in Lemma 1 represents an upper bound on what any investor can hope to accomplish without personally taking part in the market. It is an upper bound since an investor can only carry out the trade δ by using the information of every single FMI, an information set possessed by neither the trader nor any single FMI. Thus, if one can somehow give the set of FMI's the instructions needed to carry out δ no trader could hope to do better. Lemma 2 shows how to accomplish this task.

Lemma 2*: The period one investors can implement δ by sending the following instructions $f_j^i(\bullet)$ to the FMI's:*

- *FMI's with $j \in [1,J]$ receive instructions given by $f_j^i(e_j) = -\alpha_j(i) e_j$.*
- *FMI $j=J+1$ has $f_{J+1}^i(1) = -X_1(i)$.*
- *FMI $j=J+2$ has $f_{J+2}^i(P_2) = -P_2/(\gamma\sigma_D^2)$.*

If an FMI offers a trading strategy *f*, then buying *x* shares in that FMI results in a period three payout of $xf(D-P_2)$. It is clear from the above lemma that an FMI's trading technology may be restricted to be linear and independent of individual investors (i.e. $f_j^i(e_j) \propto e_j$, and independent of *i*) without changing investors' allocation decisions. This implies that it is feasible for corporate FMI's to carry out the program given by Lemma 2. The result is summarized as:

Theorem 1. *Given a set of J+2 funds with signals $\{e_1,...,e_J,1,P_2\}$, funds 1 through J offer to buy e_j number of shares, fund J+1 offers to buy 1 share, and fund J+2 offers to buy P_2 shares in the market at t=2. Furthermore, the optimal demands of period 1 investors $f_j^i(e_j)$ are feasible.*

There are three important points to note. First, trading via the *J*+2 funds is a second best solution for the period one traders. Each trader would rather trade based upon his own endowment shock. However, given that the trader has to engage in other activities a second best solution for *all* traders can be found via the formation of *J*+2 funds. Second, the funds do not behave like individuals: risk neutral or otherwise. The first *J*+1 funds trade without regard to equilibrium prices. No individual trader, trading on his own account would act like this. However, funds are not individuals. They are simply robotic entities that carry out their prescribed instructions. Investors allow for this by using fund *J*+2 (the price fund) to adjust their final positions in response to the price. Third, even though there is only one underlying stock there naturally arise a large number of funds.

Some evidence regarding the above results can be found in Table 1. It shows that since 1996 the number of funds capable of investing in equities has exceeded the number of available stocks in the U.S., until today when there are approximately eleven funds for every eight stocks. If one adds in FMI's beyond mutual funds the ratio of FMI's to stocks becomes even more impressive. While puzzling within a standard CAPM framework, this empirical fact is fully compatible with conclusions found in Theorem 1. Simply put, investors want a wide range of funds because they want a wide range of trading strategies available for their own personal use while absent from the markets.

Another facet of the model born out in Table 1 is Theorem 1's statement that investors wish to have a price sensitive fund manage part of their portfolio. Within the traditional nomenclature, Theorem 1's price sensitive fund looks like an "asset allocation fund."[6] After all, both fund types move an investor's portfolio into and out of stocks and bonds depending on whether equity prices are relatively low or high. In 1999 while the number of funds that invested only in equities (funds 1 through J in the model) equaled 1,996, there were also 1,775 asset allocation funds. However, more tellingly, while investors placed 934,607 million dollars in equity only funds they put 1,802,193 million into asset allocation funds, about twice as much. This is in line with the model's prediction. While investors like a variety of equity funds, they want to invest in an asset allocation fund too.

Another issue to note is that Theorem 1 provides a recipe for the optimal construction of FMI's. People want funds to provide a *strategy* basis. Investors then produce their desired strategy vector by appropriately purchasing shares in each fund. In this way a relatively small number of funds (relative to the size of the population) serve to produce the varied investment strategies a diverse population demands.

Having solved for the demands of both the period two traders and the investment funds one can now solve for the period two price. Summing over the demands of the date one and date two agents yields the following condition for market clearing at date two

$$-\widetilde{X}_0 - P_2/(\gamma\sigma_D^2) - \int_0^1 \left(X_1(i) + \sum_j \alpha_j(i)e_j + P_2/(\gamma\sigma_D^2) \right) di = 0 \qquad \textbf{(10)}$$

[6]Defined here as a fund that has changed it percentage holdings in stocks by at least 20 points sometime during it life. For example, from 60% to 80% equities.

which produces an equilibrium price of

$$P_2 = -\frac{\gamma\sigma_D^2}{2}\left\{\int_0^1\left(X_1(i) + \sum_j \alpha_j(i)e_j\right)di + \widetilde{X}_0\right\}. \tag{11}$$

Notice the demands of the first J+1 funds are not price sensitive. Yet, the price itself looks like one that would arise if the period one traders had to trade on the set of signals e_j rather than their own personal endowment shocks $N(i)$. Another feature is that the expected price ($\overline{P}_2$) does not depend upon the number of trading firms established by the period one traders. By comparison, in a standard model with risk averse intermediaries additional firms increase the expected price as they add to the pool of traders willing to absorb risk. While this particular contrast depends strongly on the assumptions employed here it does point to a general difference between the models. In a standard model FMI's are assumed to be risk averse individuals, while here they are firms and as such do not absorb risk themselves.[7] The only way for FMI's to affect the risk premium is for them to change the effective risk aversion of the investors in the population. Hence introducing more FMI's does not, holding constant the implied risk aversion of the investors in the economy, have any impact on the price level.

3.2 Period Two Stock Price Volatility with Financial Market Intermediation

Equation **(11)** fully describes the equilibrium price in terms of the model's primitives one of which is the number of FMI's. Note, that increasing their number increases the summation and thus can be expected to impact the price variance. This in fact occurs, and while the derivation is algebraically complex the result is straightforward.

Theorem 2. *The volatility of the t=2 price,* σ_P^2 *,increases with the number of FMI's.*

Theorem 2 shows that increasing the trading options available to people via third parties increases price volatility. Rather than smooth out prices, the FMI's increase volatility because they are trading on behalf

---

[7]The traditional FMI demand is of the form $-P/(\gamma\cdot\sigma^2)$, whereas in this model it is, for example, e_j. Clearly having many FMI's with the former demands will decrease the risk discount in the price, whereas having many FMI's with the latter demands does nothing to the risk discount but (potentially) makes the price more volatile.

of their owners. These owners do not create FMI's to provide services to the period two traders, but rather to themselves. This "selfishness" leads the owners to produce corporations whose trading pushes prices, on average, away from the mean. Compare this with the results from a model in which the intermediaries are modeled as entities with utility functions. In that case, price volatility declines as the pool of individuals for risk sharing increases.

3.3 The Period One Risk Discount and Financial Market Intermediation

The period one equilibrium price naturally depends upon what FMI's have been created since traders take into account how investing in these firms impacts their future portfolios. This can be seen in the period one demands and the equilibrium price for the risky asset.

Lemma 3*: The optimal holdings $X_1(i)$ of the date one investors is given by*

$$X_1(i) = \frac{\overline{P}_2}{\gamma\sigma_P^2} - \frac{\sigma_D^2 + \sigma_P^2}{\gamma\sigma_D^2\sigma_P^2} P_1. \qquad \textbf{(12)}$$

Market clearing $\int_0^1 (X_0(i) - X_1(i))di = 0$ *implies that the t=1 price P_1 is given by*

$$P_1 = \frac{\sigma_D^2}{\sigma_D^2 + \sigma_P^2}\left(\overline{P}_2 - \gamma\sigma_P^2 \int_0^1 X_0(i)\, di\right). \qquad \textbf{(13)}$$

The period one price contains what may be called the risk-weighted mean of the time two price, as well as a risk discount for the time one supply.

Note that the number of FMI's trading in period two affects the period one equilibrium price. This interaction occurs via the time two price volatility σ_P^2. When new FMI's are introduced the volatility of the time two price increases relative to the volatility of the time three dividend. This can either increase or decrease the period one price. The following lemma states the relevant result.

Lemma 4. *The period one price decreases (or equivalently the risk discount increases) as the number of period two FMI's increases if and only if*

$$\overline{P}_2 > -\gamma\sigma_D^2 \int X_0(i)\, di. \qquad \textbf{(14)}$$

To understand this lemma, consider what happens in the limit as σ_P^2 goes to either zero or infinity. As $\sigma_P^2 \to 0$ then $P_1 \to \overline{P}_2$ since without any time two price volatility, buying the stock at time one is identical to buying the stock at time two. On the other hand, when $\sigma_P^2 \to \infty$ one finds that $P_1 \to -\gamma\sigma_D^2 \int X_0(i)di$, since an infinitely high period two price variance induces those holding shares today to hold them (in expectation) until time three when the dividend payout occurs. Hence the time one price is equivalent to the price in a one period economy. Moving from a low to a high price volatility simply moves the price from one extreme ($\overline{P}_2$) to the other ($-\gamma\sigma_D^2 \int X_0(i)di$).

Note that the presence of an FMI at time two may affect asset prices at time one. The introduction of a fund, by increasing period two price volatility, actually serves as a wedge between the time two and time one supply shocks. More funds at time two imply that the anticipated time two supply shock will have less of an affect on period one prices. This is somewhat counterintuitive. For example, if traders know that a future supply shock will decrease tomorrow's price, then they ought to push today's price towards that future price. However, if there is sufficient price volatility tomorrow, then even though they anticipate a lower (in expectation) price, their risk aversion will prevent them from trading on it today. Thus, additional FMI's (which makes markets more complete, but also tends to increase price volatility) actually inhibit people from pushing today's price closer to tomorrow's expected price.

3.4 Comparison to Traditional Framework

Again, compare the result in Lemma 4 with what one obtains from a framework in which the intermediaries are modeled as humans with utility functions. In such a model the addition of intermediaries increases the period one price, since there are additional people to share the market wide risk. Here, additional FMI's cannot add to the risk sharing capacity of the economy. As a result, the impact of additional firms has an ambiguous influence on the period one risk premium since it depends on the current risk-return tradeoffs already available.

Summary 2: A Comparison with the Equilibrium Properties In a Standard Model of Financial Market Intermediation		
Result	Standard Model	First Principle's Model
Demands of the Period Two Traders	Typical mean variance demands.	Same.
Equilibrium Clearing Condition	Market supply equals the period one holdings of the FMI's and the period two trader endowment.	Market supply equals the period two trader endowment.
Demands of the FMI's in Period Two	Typical mean variance demands.	Three different types. Type 1 trades a constant amount. Type 2 trades based upon the price. Type 3 trades based upon a private signal, and ignores prices.
Equilibrium Prices	Equal to that derived from a representative agent holding the sum of the FMI's period one holdings, plus the endowment of the period two traders.	Equal to that derived from a representative agent holding the sum of the estimated endowment of the period one traders based upon the signals obtained by the FMI's, plus the endowment of the period two traders.
Price Volatility as the Number of FMI's Increase	Declines	Increases
Period One Price as the Number of FMI's Increase	Increases	Depends upon the model's parameters.

3.5 Welfare Impact of a New Fund

Consider the central planner's problem in the one security economy of the previous section. Adding FMI's increases the ability of the period one agents to hedge their t=2 endowment shocks. It may therefore seem that social welfare always increases if an additional period two FMI can be created at no cost. However this intuition is incorrect. While introducing a new FMI expands the hedging opportunities available to the agents in the economy, it also increases the period two price volatility. The former effect is beneficial. The latter effect may not be. Which dominates determines the welfare implications of new FMI entry.

Consider an economy with $J+2$ FMI's. From **(8)**, a social planner, considering whether to introduce an additional FMI with signal e_{J+1}, solves the following problem

$$G_{J+1} = \max_{e_{J+1}} \int_0^1 \left[V(i, J+3) - V(i, J+2) \right] di - \kappa, \tag{15}$$

where $V(i,J+2)$ is the value function of date one investor i. The sign of G_{J+1} determines whether entry of the new FMI is beneficial. Going forward, assume κ=0.

To gain some insight into how G_{J+1} behaves consider the impact of an additional FMI on the time one welfare of an individual investor. Although not strictly true, assume for the moment that the number of FMI's J can be chosen from the positive real numbers (hence allowing for fractional firms). Differentiating $V(i,J+2)$ with respect to J yields

$$\frac{dV(i,J+2)}{dJ} = \frac{\partial V}{\partial \sigma_\eta^2}\frac{\partial \sigma_\eta^2}{\partial J} + \frac{\partial V}{\partial \sigma_P^2}\frac{\partial \sigma_P^2}{\partial J}. \qquad \textbf{(16)}$$

Clearly, $\sigma_{\eta(i)}^2$ must be decreasing in J since additional signals can only reduce residual endowment uncertainty. Furthermore, from Theorem 2, the period two price volatility is increasing in J. This leaves only the signs of $\partial V / \partial \sigma_{\eta(i)}^2$ and $\partial V / \partial \sigma_P^2$ unknown in equation **(16)**. Both of these are established in the following lemma.

Lemma 5. *The expected utility of a trader V(i,J+2) is concave and decreasing in* σ_η^2 *, i.e.* $\partial V / \partial \sigma_\eta^2 < 0$ *and* $\partial^2 V / \left[\partial \sigma_\eta^2\right]^2 < 0$ *. Also trader i's expected utility increases in the price variance, that is* $\partial V / \partial \sigma_P^2 > 0$ *, if*

$$\sigma_P^2 + \sigma_D^2 \geq (\overline{P}_2 + \gamma \sigma_D^2 X_1(i))^2 . \qquad \textbf{(17)}$$

If this condition holds for some J, then for any I > J social welfare will be higher with I FMI's than with J FMI's (i.e. $G_I > 0$ *for I > J).*

The last result in the lemma has a particularly nice economic interpretation. Consider the simplest case when $\overline{P}_2 = 0$ and $X_1(i) = 0$. Here an increase in time two price volatility aids the period one investors. The reason for this is that in a full information setting price volatility (due to non-dividend information) changes the expected return of the stock. With the ability to go long or short at will, investors always profit from anticipated variability in an asset's expected return. For example, suppose there is a large positive shock to the time two price. Since the price is now high relative to fundamentals, and the expected return is low, investors will short shares of the stock. This selling must be beneficial or they would not do it. Similarly, large negative shocks must also be beneficial. The investors essentially have a free option to go long or short in period two, and the value of this option increases in the volatility of the time two price.

In the case where $\overline{P}_2 \neq 0$ or $X_1(i) \neq 0$ an increase in price volatility may hurt or help investors. It helps for the same reasons given previously. However, consider the agents' time three wealth. Recall that period one investors wish to trade away $X_1(i)$ in period two. Thus, if $X_1(i) \neq 0$, then time two price volatility adds noise their time three wealth, decreasing their utility. If $\overline{P}_2 \neq 0$, then from the wealth equation the uninsurable part of the time two endowment shock equals $\eta(i)\left(-\overline{P}_2 + \varepsilon_P\right)$ where ε_P is the price uncertainty. As the fixed part of the time two price increases, the negative utility impact of ε_P will become more pronounced (this is because after taking an expectation with respect to $\eta(i)$ an agent's welfare will be proportional to $\exp\left(\sigma^2_{\eta(i)}(\overline{P}_2 - \varepsilon_P)^2\right)$). Overall then, the effect of a new FMI on agents occurs through two channels. First the improved hedging opportunities against time two endowment shocks make investors better off at time one. Second, the increased price volatility from that very use makes them worse off.

An immediate consequence of Lemma 5 is that once price variability becomes sufficiently large, the entry of any new FMI has to make society strictly better off. The reason is easily seen from equation **(17)**. The right hand side does not depend on the number of time two FMI's. But the left hand side, namely the price variability, does. Hence once enough FMI's have entered to allow **(17)** to hold, it will continue to hold for all new FMI's. Note, however, that it is possible for society to be better off with no FMI's at all. The reason for this is that **(17)** may never hold, and the negative effects of increased price variability may dominate the positive effects of increased ability to hedge endowment shocks.

4 Examples

This section presents some examples of the single security economy.

4.1 Example 1

Consider an economy with two date one investor classes. Both consist of a continuum of investors of measure 0.5. Let e_1, e_2, e_3, η_1, and η_2 be iid standard Normal random variables. The endowment shocks of the two date one investor classes are given by

$$\begin{aligned} N(1) &= e_1 + 2e_2 + e_3 + \eta_1, \quad \text{and} \\ N(2) &= 2e_1 + e_2 + 3e_3 + \eta_2. \end{aligned} \qquad \textbf{(18)}$$

The initial endowments X_0 of both sectors are zero, and all agents' risk aversion parameter γ=1. Assume that four FMI's exist in the economy, and that their t=2 demands are e_1, e_2, 1, and P_2 respectively. Given this structure, the optimal demands of both investors, from **(9)**, are

$$\begin{aligned} \delta_1 &= -e_1 - 2e_2 - P_2/\sigma_D^2, \quad \text{and} \\ \delta_2 &= -2e_1 - e_2 - P_2/\sigma_D^2. \end{aligned} \tag{19}$$

For example, the type one agents can buy -1, -2, and $-1/\sigma_D^2$ shares of FMI's e_1, e_2, and P_2 respectively. If there are no date two investors, then the stock price is given by $P_2 = -1.5\sigma_D^2(e_1 + e_2)$. The variance of the price is $4.5\sigma_D^4$. The residual variances (i.e. the variance of the unhedged part of the endowment shock) of the two agents equal Var($e_3 + \eta_1$)=2 and Var($3e_3 + \eta_1$)=10. Since the mean of the date two price and the initial endowment of the date one agents are both zero, condition **(17)** insures that agents would be better off by having an additional fund.

Since η_1 and η_2 are independent, social welfare is increased by introducing a new FMI which trades based on e_3. Agents' optimal demands become

$$\begin{aligned} \delta_1 &= -e_1 - 2e_2 - e_3 - P_2/\sigma_D^2, \quad \text{and} \\ \delta_2 &= -2e_1 - e_2 - 3e_3 - P_2/\sigma_D^2. \end{aligned} \tag{20}$$

In this case, the type one agents can buy -1, -2, -1, and $-1/\sigma_D^2$ shares of FMI's e_1, e_2, e_3, and P_2 respectively. Notice that the price is now given by $P_2 = -1.5\sigma_D^2\left(e_1 + e_2 + 4e_3/3\right)$. The variance of the price is now $8.5\sigma_D^4$ which, as shown in a theorem, is higher than the price in the case where FMI e_3 did not exist. Intuitively, increasing the number of FMI's allows agents to better hedge their date two endowment shocks, and hence trade more actively. Since all of the random variables have zero means, the expected price has not changed, even though an extra FMI has entered the economy. The residual variances of both agents (Var(η_1) and Var(η_2)) have fallen to one. The introduction of the additional FMI, by allowing better risk sharing to occur in the economy, has made both agents better off.

4.2 Example 2

Assume that investors are evenly spread in characteristic space on an *n*-dimensional sphere. One interpretation is that of physical location. In this case, investors in Chicago receive endowment shocks that have a higher correlation with each other than with investors in California. The primary reason for assuming that investors are on a sphere is simply to eliminate end point problems when funds are attempting to determine what signal to acquire.

Imagine there *I* groups of investors, indexed via 1, ..., *I*. Assume that all endowment shocks have the same variance. The correlation between endowments *i* and *j* is given by exp(-δ(*i*,*j*)), where δ(*i*,*j*) measures the distance between the two groups, and is given by

$$\delta(i,j) = \left\| [\cos(i2\pi/I)\ \sin(i2\pi/I)]' - [\cos(j2\pi/I)\ \sin(j2\pi/I)]' \right\|, \quad \textbf{(21)}$$

where $\|\bullet\|$ is the Euclidean norm. For example the correlation between groups one and *I* is the same as the correlation between groups two and three. The FMI's owners choose to have it acquire the signal of some group *i*, and then trade *N*(*i*) shares of the risky asset on their behalf.

Figure 1 shows that when the total share endowment at times one and two equals zero investors will create FMI's that are as "spread out" as possible. The circles in the figure represent the characteristic space. Circles with larger radii correspond to economies with more time two FMI's (for example, the second innermost circle corresponds to a two firm economy – which are the dots on the circle). By choosing minimally correlated signals for the FMI's, their creators maximize the fraction of society's time two endowment shocks which is ex-ante insurable.

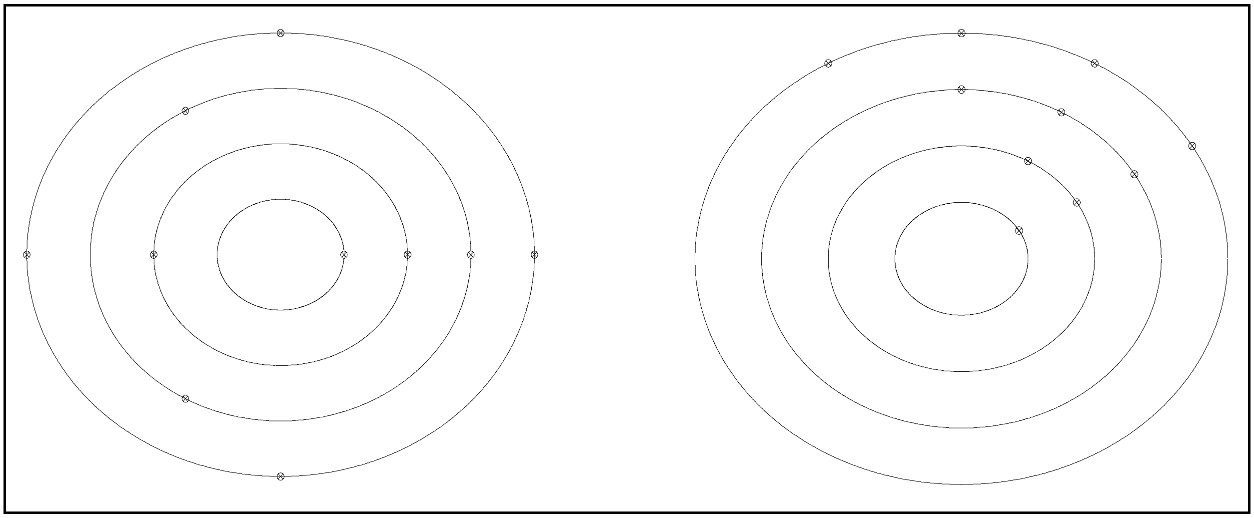

Figure 1. This figure shows the way that investors would choose to position the signals received by the optimal FMI's in the agents' spherical characteristics space. The innermost circle corresponds to the one firm economy, and the outermost circle corresponds to the four firm economy. The parameter values for the right graph are γ = 1, $\sigma_D^2 = 0.1$, $\tilde{X}_0 = 10$, $X_0(i) = 0$, and the number of investor groups is 12. The left graph has $\tilde{X}_0 = 0$. In both cases the variance of $\tilde{X}_0$ equals zero.

Consider, however, when time two supply is non-zero. As has already been pointed out, now there are two factors at play when funds decide on whether or not to enter the market. First investors wish to create FMI's that will provide them with the maximal amount of time two hedging. This leads them to spread out the FMI's as much as possible on the characteristic sphere. On the other hand, the entry of a new fund increases the price variability and this can lead the period one traders to cluster the FMI's. To see when and why this will occur, note that when there is a positive supply which will enter the market at time two the time two risk discount in the price will be higher. Recall that, assuming $X_0(i)$=0, the time two unhedged endowment shock is proportional to

$$\eta(i)(\overline{P}_2 + \varepsilon_P) \propto \eta(i)\left(\widetilde{X}_0 + \int_i \alpha(i)' e\, di\right). \qquad \textbf{(22)}$$

This implies that increasing the time two supply $\widetilde{X}_0$ decreases the willingness of agents to bear time two price risk. That is, variability coming from $\int_i \alpha(i)' e\, di$ hurts investors more. The founder of an FMI therefore has an incentive to choose signals e_j which are minimally informative (and hence which minimize the magnitude of the $\alpha(i)$ vectors). This is exactly the opposite incentive that existed when there was no time two supply.

For a given agent, the introduction of a new FMI, while decreasing the amount of non-hedgeable risk, can still make that agent worse off because so many of the other agents in the economy will use that firm to hedge their endowment shocks, thereby increasing price variability. Indeed, from **Figure 1**, when time two supply is non-zero, FMI's are optimally clustered by their owners, even though investors are still evenly distributed on a spherical characteristic space.

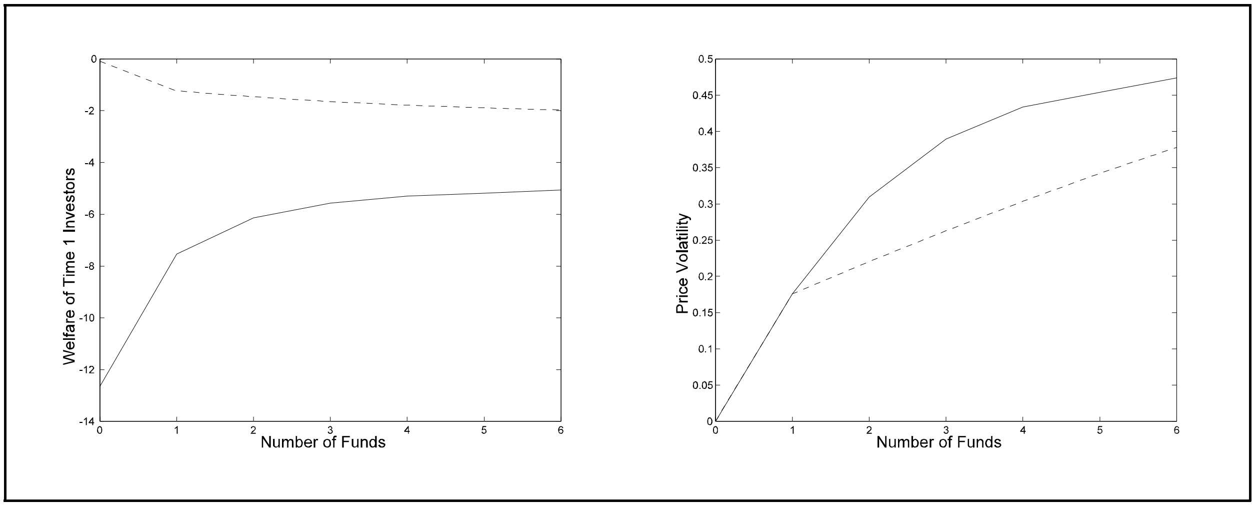

Figure 2. The left graph shows the behavior of social welfare as a function of the number of FMI's in the economy. The right graph shows the volatility of the period 2 price as more firms enter the economy. The dashed line refers to the economy with non-zero time 2 supply, and the solid line refers to the economy with zero time 2 supply. The parameter values for the dashed lines are $\gamma = 1$, $\sigma_D^2 = 0.1$, $\tilde{X}_0 = 10$, $X_0 = 0$, and the number of investor groups is 12. For the solid line $\widetilde{X}_0 = 0$.

This clustering, by minimizing the amount of new hedging service that agents will have access to, also minimizes the amount of price variability that each new FMI creates. **Figure 2** shows an example of this effect. In the economy with no time two supply, the price variability increases rapidly with new firm entry, as does social welfare. However, in the economy with a sufficiently large time two supply, new firm entry decreases social welfare. This can be seen from the downward sloping dashed line in the left graph of **Figure 2**. To minimize the welfare cost of new firm entry, additional firms are positioned in a way which minimizes the price volatility (and which leads to the clustering evident in **Figure 1**). Indeed, time two price volatility (the dashed line in the right graph of **Figure 2**) in the non-zero supply case increases much more slowly than the price volatility in the zero supply economy.

4.3 Comparison to the Standard Model

The following table compares some of the results of our model to those in the standard paradigm.

Summary 3: A Comparison with the Social Welfare Properties In a Standard Model of Financial Market Intermediation		
Result	Standard Model	First Principle's Model
Utility of Period One Traders with an Increase in FMI's	Increases due to improved risk sharing.	Ambiguous, depending upon whether the improved hedging opportunities outweigh the cost arising from additional price volatility.
Number of FMI's with Costless Entry	Infinite	Infinite or zero.
Signal Choice in the Absence of Period Two Supply	Select divergent signals to minimize competition.	Select divergent signals to maximize informativeness.
Signal Choice with Period Two Supply	Select divergent signals to minimize competition.	Cluster on single choice to reduce period two price volatility.

5 Equilibrium with Multiple Risky Securities

Even with a single risky security the model produces a rich theory of FMI's. However, an important role played by FMI's and mutual funds in particular is the selection of securities to be held from a universe of many risky assets. Consider once again the model from Section 2. Since it now becomes useful to talk about FMI's as families that provide a range of trading services that clients can select among, it is easier to use the nomenclature from the mutual fund industry. Although, again, the theory applies more broadly to any institution that offers to trade on an investor's behalf.

Define a fund family as a research entity j that acquires some signal e_j. A fund i within family j is a corporate entity that offers to trade the vector of securities based upon some function $f_j^i(e_j, I)$ of e_j. Thus, funds within a family are assumed to share the results of the family's research department. Since the goal of this paper is to analyze how fund families should structure their offerings, rather than how extensive their research departments should be, assume that each family's research department can only collect one signal.

However fund families can add however many additional funds their investors demand at a cost of $\hat{\kappa}$ per fund.

Recall that acquisition of a signal e_j carried a per capita disutility of κ. In a one security world, a fund can simply offer to trade the stock in proportion to its signal e_j. However, in the case of K securities matters become more complicated. For example, consider a world with two stocks. It is possible that one investor will submit a 2×1 demand vector of the form $f_j^i(e_j) = [1 \;\; 2.2]' e_j$, whereas another investor's demand may be $f_j^{i+1}(e_j) = [4 \;\; 2]' e_j$. Since these two vectors are not co-linear, the previous results in Theorem 1 no longer hold, and it is no longer clear wether or not investors will agree on what policies the set of FMI's should follow. However, say that two funds exist, both of which traded based on e_j, but where one always performed a trade given by $[1 \; 0]'e_j$, and where the other performed a trade given by $[0 \; 1]'e_j$. Then, as long as they were able to buy different numbers of shares of each fund, each investor would be able to implement his optimal trade $f_j^i(\bullet)$ by buying an appropriate number of shares in both of these e_j funds. Then Theorem 1 would once again hold because each fund could offer a linear trading strategy in its signal, which was independent of any investor i's preferences. However, as a practical matter, simple observation indicates that having each fund family set up one fund per stock is economically infeasible. Thus, the derivation of the instructions traders will give to the funds must be found via some other route.

Assume that fund family j offers K_j securities vectors, all of which trade on a signal e_j. Then buying x shares of a K×1 vector ω results in a t=3 cash flow given by $e_j x \omega'(D\text{-}P_2)$. Refer to the $K \times K_j$ matrix of securities vectors offered by family j as W_j. As in Section 2, assume that funds J+1,...,J^* exist and are able to execute arbitrary trades of the form $f_j^i(\bullet)$ on behalf of investors.

The economic structure above conforms reasonably well with the way in which fund families operate. From **(48)** in the Appendix, an agent's optimal demand conditional on knowing the families' signals e_j is of the form $\delta = Ae + BP_2 + c$, for some $K \times J$ matrix A, $K \times K$ matrix B, and scalar c. Hence, as long as agents are able to buy different numbers of shares of the funds inside a given family, they are able to exactly implement their optimal trades $f_j^i(\bullet)$. Each fund family offers funds that trade based upon a linear function of the family's signal, and which does not depend on any individual i. Thus, the trades can be feasibly carried out in a corporate environment.

The benefit of this approach is that (1) it solves for the optimal demands of investors faced with a particular fund structure, (2) it solves for the socially optimal industry organization, and (3) it characterizes equilibrium prices in a special case of the multiple security model. In particular, the social planner in this case solves the following problem

$$\max_{\{J, e_1, ..., e_J, W_1, ..., W_j\}} \int_0^1 E[U(C_3(i))]di - J\kappa - \left(\sum_{j=1}^{J} |W_j|\right)\kappa'. \qquad \textbf{(23)}$$

Here $|W_j|$ is the column rank of matrix *Wj*. Thus one can characterize the optimal structure of the fund industry, in terms of the optimal number of families *J*, the optimal number of funds inside each family $|W_j|$, the optimal research strategies of the families e_j, and the optimal fund structure within each family W_j. Furthermore, the Appendix establishes that under certain regularity conditions, one can relate the optimal number of funds inside a family to the factor structure of endowment shocks in the economy.

6 Empirical Evidence on Fund Families

The model presented here produces a theory of fund families that spread out their offerings in strategy space. There is, of course, an alternative hypothesis. One can imagine a world in which economies of scale encourage fund families to specialize. That is one family in this world might advertise itself as being expert in growth stocks, while another would claim expertise in value stocks. Among other regularities, one would expect these fund families to produce numerous funds within the same Morningstar categories.

Alternative Hypothesis: Scale economies induce fund families to specialize in funds with related strategies and security holdings.

6.1 How Quickly Do Fund Families Spread Out Their Offerings in Strategy Space?

Morningstar maintains a list of 48 categories for mutual funds. The categories are an attempt to classify funds according to their investment styles and objectives. Table 2 provides evidence regarding the number of mutual fund families and funds in each Morningstar category. In general, it appears that fund families try to spread their funds across a variety of categories. While some categories are more popular than others, there are approximately two and a half to three funds per fund family in each. Consider now the investor's optimization problem. The primary restriction comes from the W_j matrix. To the degree that a fund family only offers a few funds, or funds that form a poor strategy basis, W_j will restrict an investor's ability to use that fund family's funds to enhance his welfare. Thus, fund families have an incentive to set up multiple funds, all of which select among different security sets. The finding in Table 2 that fund families appear evenly spread out, is therefore equivalent (within the model) to saying that they are trying to provide investors with funds that cause W_j to impose as few restrictions as possible.

Further evidence that fund families try to provide a wide strategy basis for their investors can be found in Table 3. The first column sorts fund families by the number of funds they offer, while the second column sorts them by the number of Morningstar categories their funds fall into. The column labeled "# of Fam." displays the number of fund families meeting the criteria given by the first two columns. Cells in which fund families have 75% or more of their funds in separate Moringstar categories have been shaded to show just how spread out funds within a family tend to be. While the very big families have a great many funds within the same set of categories, this is largely due to the fact that there are only 48 categories available. Thus, fund families with say 100 or more funds will necessarily have on average over two funds per category. However, fund families with more modest offerings clearly attempt to spread out the strategies they offer to investors. This can also be seen in the following graph.

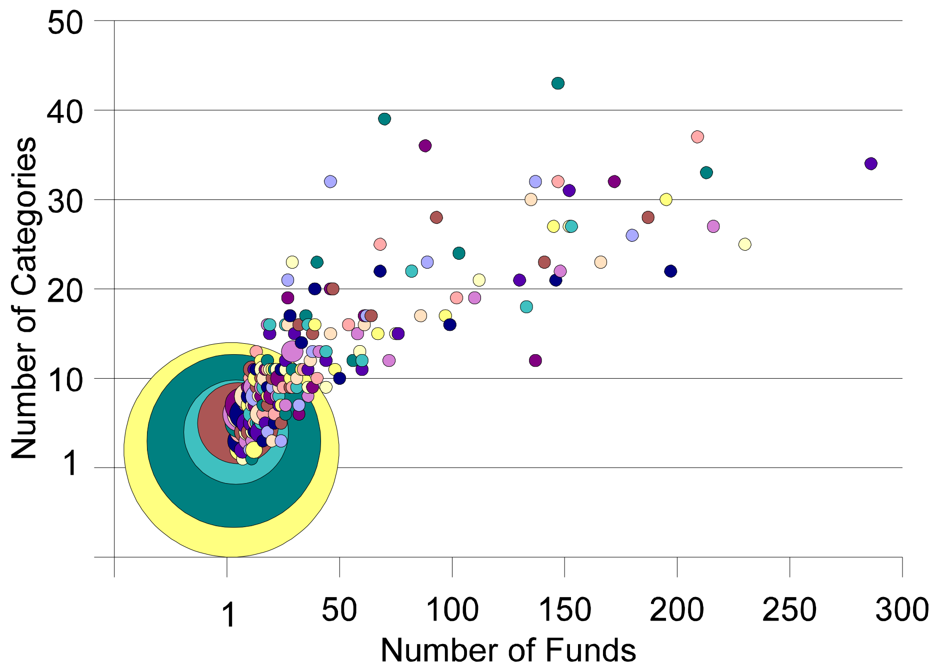

Figure 3: Each circle represents a fund family with coordinates given by (number of funds, number of Morningstar categories into which the family's funds fall). The size of each circle indicates how many fund families lie at a particular coordinate.

Figure 3 displays the data from Table 3 in graphical form.[8] Each circle represents a set of fund families with a particular number of funds, and funds in separate Morningstar categories. A circle's size indicates the number of funds at a particular point. As one can see fund families quickly seek to fill in about thirty or so Morningstar categories with their offerings. Translating this to the model's parameters, a fund family that offers funds within about thirty Morningstar categories provides a sufficiently rich W_j matrix that it becomes unprofitable to fill in additional parts of the strategy space. Indeed, only 13.3% of all fund families offer more than thirty funds altogether.

Within the model, one can also explain the existence of small fund families. In reality it seems unlikely that investors are arrayed uniformly within characteristic space. Thus, fund families that cater to

[8]The graph includes data on fund families with more than 32 funds, data which due to space constraints would not fit in Table 3.

investors in the less populated parts of the characteristic space are going to have insufficient demand for their product to profitably offer more than a few individual funds. These (numerous) fund families then show up in the graph as part of the large circles along the 45 degree line near the axis. Still, even though such fund families may not be able to offer many funds, the model indicates that those they do offer should span as much of the strategy space as possible. This corresponds well with the fact that such families tend to have most of their funds located in separate Morningstar categories.

6.2 Hypotheses and Tests Regarding Fund Returns and Time Varying Fund Betas

While ours is a static model, it nevertheless points towards some interesting dynamic effects. Consider a repeated game each round of which consists of our three period model. In the first round, fund families and funds are created as per our discussion. These fund families and funds are carried over into subsequent trading rounds, in which a technological innovation occurs which lowers the costs of signal acquisition and fund creation, while the investor population is always drawn from an identical distribution. Hence new fund families may start, and existing fund families may introduce new funds. Our empirical hypotheses revolve around the dynamics of fund creation in such a repeated game.

6.2.1 Hypotheses

Within the model investors seek fund families whose signals allow people to engage, via their fund proxies, in market timing strategies within particular stocks. This represents two different testable hypotheses. The latter hypothesis has to do with the span of the security space provided by fund families. Section 5 of the paper shows that investors care about the rank of the matrix W_j for each fund family j. Now consider how such fund families should be set up over time. When a family has but one fund it should seek out a vector of securities that most of the population will find useful. Intuitively, this is likely to be rather similar across fund families. These funds, since they are holding similar securities, should therefore have relatively high return correlations. Now consider the introduction of subsequent funds. As additional funds are introduced a family begins to better, and better span the security space. At this point they can introduce funds with more specialized holdings which better cater to those investors that favor the type of research done by the family. This in turn should lead to lower and lower return correlations as one looks across funds that are introduced later in a family's life.

Let a fund's "organization number" equal one if the fund is the first introduced by a family, two if it the second fund, three if the third, etcetera. Then the above discussion leads to the following empirical hypothesis.

Hypothesis 1: The return correlation between funds in different families should decrease in their organization number .

A second aspect of the model, is that fund families also provide a research and timing function for their investors. "Good" families are those whose research departments signal their managers to get in and out of stocks at times their investors wish this to happen. This should show up in several different places. Consider where these signals come from. Every fund in a family has access to the same research and thus one expects that fund returns within a family should be more strongly correlated than returns for funds in different families. However, a more subtle test can be found in fund time varying betas. Because funds within a family have access to similar research (which in our model is proxied by their trading on identical signals) they should be more likely to enter or exit the market at similar times relative to funds from other families. We thus expect the time variation in betas for funds within a family to be higher than for funds which are in different families.

Hypothesis 2: Time varying betas should show a higher level of correlation for funds within a family, than for funds across families.

Finally let us consider what the model tells us about how the correlation of time varying betas for funds across families should depend on those funds' organization numbers. When a new family is started, the family chooses a signal which maximally differentiates it from existing families. Hence the dynamic strategy followed by a family's early funds is relatively unique at the time of those funds' creation. As the family ages, and in particular, as new families are introduced into the economy, more funds will exist which trade on a similar signal. To the degree that later funds are introduced during trading rounds when more families already exist, the beta correlations of later funds with funds from outside families should be higher than the beta correlations of the earlier funds. Note as well, that this argument suggests that if we were to take a given fund within a family and compute its beta correlations with outside (the family) funds over the first half and the second half of its existence, the average correlation over the first half should be lower than that over the second half. We focus on the first of these implications in this paper.

Hypothesis 3: The beta correlation among funds in different families should increase in their organization number.

Note that the difference between Hypothesis 1 and 3 is that between averages, and changes over time. Hypothesis 1 comes from the fact that if two funds trade in similar securities then they will likely have highly correlated returns. In contrast, Hypothesis 3 points out that even if two funds trade in similar securities they can shift their fund in and out of the market at very different times. As an example consider two funds with holdings governed by the following matrix. Further assume each fund can only invest in either the market portfolio or the risk free asset.

	Fraction of each fund's portfolio in the market portfolio. All remaining funds are invested in the risk free asset.	
	State A	State B
Fund X	.6	.8
Fund Y	.8	.6

If the two states are equally likely then their returns have a correlation of .96, even though their time varying betas have a correlation of -1. Hypothesis 1 is about the .96 value, and Hypothesis 2 the -1.

6.2.2 Methodology

Data from the Morningstar January 2000 CD was combined with the CRSP mutual fund database. The Morningstar database includes active funds as of December 31, 1999. For each fund Morningstar lists its fund family, its origination date, and its ticker symbol, if any. These origination dates were then used to produce an organization number for each fund. This does mean that the organization numbers suffer from survivorship bias. However, expected returns are not at issue here, so at most this simply adds noise to one of the independent variables thereby biasing the results away from finding any particular patterns.

Using the ticker symbols the Morningstar data was then matched with the CRSP data to produce a return series for each fund. Across fund return correlations were only calculated if there existed twelve months of common data. Using the same data month-by-month betas were calculated for each fund using a 10 month weighted rolling regression. Only the five months before and after a particular month were used.

No estimate includes the month in question. Thus, for example, a June beta was estimated via a weighted regression using data from January to May, and July to November. Weights were selected using a tri-cube kernel (see Appendix).

Once a beta series was calculated for each fund, funds were paired with each other. The database includes a fund pair only if they overlap in time for a sufficient number of years to produce at least 13 months of comparable data. Using the overlapping months during which beta estimates exist a correlation coefficient for the pair's time varying beta is calculated.

According to the model fund managers use information from their shared research department to determine what to buy or sell given the particular set of stocks they have been assigned to trade. This type of behavior should show up if one compares returns, market betas, or relevant factor betas. Since many funds advertise themselves as either small cap, mid cap, or large cap funds regressions were run using these factors as well as that of the overall market. For the purposes of this paper large cap was defined as the return on the largest CRSP decile, mid cap the fifth largest decile, and small cap the ninth largest decile. The ninth decile was used to avoid some of the microstructure problems associated with measuring the returns on the very smallest stocks.

6.2.3 Results

Tables 4, 5, 6, 7, and 8 contain the results. Tables 4, and 5 report the fund correlations, while the others look at the time varying betas. The top entry provides the average value for funds meeting the cell's criteria. Below that is the standard deviation of the correlations and the number of observations per cell. Note that the tables support both Hypothesis 1 and Hypothesis 2. Table 5 shows that in terms of Hypothesis 1 the return correlations decline with a fund's organization number. In terms of Hypothesis 2 the return correlations in Table 4 are uniformly higher than those in Table 5. Similarly, the beta correlations in Table 6 are uniformly higher than those in Tables 7 and 8. In fact, the magnitude of those in Table 6 are nearly double those in Tables 7 and 8. With respect to Hypothesis 3, as predicted, the average time varying beta correlations across funds increases as one moves to the right or down in the tables.

Panels B, and C of Tables 5 and 7 break out the fund pairs that do and do not include index funds. As noted in Section 3 while the model predicts such funds will exist, there is no reason to believe they provide particularly unique strategies for the investing public as they do not trade based upon informative signals. Thus, one expects their time varying beta correlations to be much higher than those of managed funds. Panel C of each table shows that the methodology picks this up.

Clearly neither Table 6, 7, nor Table 8's panels provide firm estimates of the correlation between any fund pair's time varying betas. A ten month weighted regression undoubtably contains a great deal of noise. However, Table 7's Panel C does show just how high even this noisy measure can get. It thus provides a benchmark against which one can measure the results in Table 6, panels A and B of Table 7, and all of Table 8. The fact that the index funds have estimated correlation coefficients two to three times higher than those observed by the managed funds indicates support for the hypotheses put forward in the paper. While the estimators are no doubt noisy they do come out in the predicted order.

Table 8 provides some evidence that fund families that initially trade small company stocks are the most successful at staking out a unique part of the strategy space. However, as a family adds funds it appears to be somewhat easier to add funds with relatively unique large firm strategies. This can be seen by examining the (1,1) cells and then comparing across or diagonally. Initially the correlation across funds with respect to time variation in the small capitalization stock beta is lowest among those estimated. But by the time families are up to their third fund (the (3,3) cell) it is the large capitalization stock beta that shows the lowest time varying correlation. Nevertheless, as indicated by the model all of these correlations indicate as one moves from a families first fund to its third.

7 Relationship to the Existing Literature

For the most part papers that explore the topic of intertemporal trading take the existence of trading firms as given, along with some utility function for those firms. Instead research has looked at the design of the institutions through which such trade takes place. For example, Glosten (1994) asks whether trade via a limit order book dominates all other forms. A recent paper by Pirrong (1999) looks at the design of the exchange itself. In his paper individuals with risk averse utility functions form exchanges on which trade takes place.

Previous research has generally looked at how compensation schemes influence investment managers, and in turn how those managers then impact observed prices and trading patterns. Papers in this area include Allen and Gorton (1993), Dow and Gorton (1997), Ou-Yang (1997), Das and Sundaram (2000), and Nanda, Narayanan, and Warther (2000). Given the constant rating of funds by various publications, principal agent issues are important as investors clearly wish to have the best managers handle their money. However, it is worth noting that the current paper compliments this literature. Where they look at the agency problem to be solved, we take the existence of a solution as given and instead look at the strategy the investors wish management to pursue.

The closest paper to the current article is probably that of Massa (2000). His paper assumes that investors have preferences over mutual fund types and then asks how the industry will form to satisfy those preferences. This paper differs from that one in that here preferences are homogenous (in that everybody is identical). Nevertheless, the current setting still produces a wide range of funds in equilibrium.

The security design literature is also related to the present paper. Duffie and Jackson (1990) study the optimal design of futures contracts by exchanges which are trying to maximize their trading volume. They find that exchanges have an incentive to offer contracts which are maximally correlated with a linear combination of the unspanned portion of investors' endowment shocks. Allen and Gale (1989,1991) consider the structure of financial markets when firms optimally issue costly new securities. They find that firms have an incentive to split up their income streams to allocate payoffs in a given state to those investors which most value them. Willen (1999) and Davis and Willen (2000a,b) analyze the social benefits of adding a security to a population of investors who have heterogeneous endowment shocks. Some of the single security results presented here do parallel those in the security design literature. In both cases those providing financial services seek to attract investors though the proper division of a security's payoff. However, once one moves to the multiple security case the models diverge. With fund families, investors must accommodate their portfolio choices to the fact that families differ in what signals they acquire, the number of funds they offer, and the securities their funds trade. For obvious reasons, there are no parallel restrictions in the security design literature.

Many empirical papers have focused on the ability of mutual fund managers to outperform some benchmark, and on the effects that such performance has on future fund flows (see for example Carhart (1997), Chevalier and Ellison (1999), Lynch and Musto (2000), Khorana and Servaes (2000), Ackerman, McEnally, and Ravenscraft (1999), among others). Three general conclusions have emerged from this literature. First, very few fund managers can outperform consistently once risk loadings have been accounted for. Second, funds which have done well in the past tend to attract investors in the future. Third, fund families which charge lower fees and offer a wide range of products tend to have a higher market share. This paper sheds light on the proper basis for analysis of fund performance: risk adjusted alphas may not properly measure a manager's performance when his objective may dictate the pursuit of a very different strategy than the one which maximizes alpha. Furthermore, the model suggests that fund families have an incentive to differentiate themselves by their product offerings exactly so as to attract investors. Similarly researchers have documented a wide range of fund styles and many funds (see for example Gruber (1996), Brown and Goetzmann (1997), and Fung and Hsieh (1997)). These findings are roughly consistent with the current model in that it predicts that many funds ought to exist in order to provide investors with a maximally spanning set of trading strategies.

8 Conclusion

This paper examines the organization of financial market intermediaries as corporations. Here FMI's do not come endowed with utility functions, but instead carry out instructions provided to them by investors. The paper provides either a new or the first explanation for a number of observed phenomena as well as new empirical support for several of its own conclusions:

- Mutual funds outnumber traded securities.
- Mutual funds have a number of objective functions which lead to observably different trading styles. These styles include:
    - A wide variety of firms that trade on news regarding the endowment shocks of individuals. In the real world this might correspond to fundamental research about how various parts of the economy are doing.
    - A price fund that trades only on the equilibrium price – essentially a technical trading fund.
    - A fund that simply trades a fixed amount of stock – basically an index fund.

- Adding mutual funds to the economy increases stock price volatility. This result contrasts sharply with models where funds act like people with utility functions. In those papers additional funds reduce volatility.
- Since many funds have demands which are completely price inelastic, the introduction of an additional such fund does not directly affect the risk discount in the prices of risky securities. There may be such an effect, but it operates through a change in investors' implied risk aversions.
- Generally, new funds should be endowed with trading strategies which are maximally different from those of existing funds.
- Multiple risky securities lead to the existence of mutual fund "families." Fund families allow several funds to use the same research information, and thus afford investors a more varied set of potential trading strategies based upon that information.
- The correlations between time varying betas are lowest for pairs of funds which include the first fund started inside a given family, and these correlations increase for pairs of funds which have been started later in a given fund family's life.

References

Allen, F. and D. Gale, 1989, "Optimal Security Design," *Review of Financial Studies,* 1 (3), 229–263.

Allen, F. and D. Gale, 1991, "Arbitrage, Short Sales, and Financial Innovation," *Econometrica,* 59 (4), 1041–1068.

Allen, F. and G. Gorton, 1993, "Churning Bubbles," *Review of Economic Studies*, 60 (4), 813-836.

Ackerman, C., R. McEnally, and D. Ravenscraft, 1999, "The Performance of Hedge Funds: Risk, Return, and Incentives," *Journal of Finance,* 54 (3), 833–874.

Biais, B., T. Foucault, and F. Salanié, 1998, "Floors, Dealer Markets and Limit Order Markets," *Journal of Financial Markets,* 1, 253-284.

Brown, S. and W. Goetzmann, 1997, "Mutual Fund Styles," *Journal of Financial Economics,* 43, 373-399.

Campbell, John, and Albert Kyle, 1993, "Smart Money, Noise Trading and Stock Price Behaviour," *Review of Economic Studies*, 60, 1-34.

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

Chevalier, J. and G. Ellison, 1999, "Are Some Mutual Fund Managers Better Than Others? Cross-Sectional Patterns in Behavior and Performance," *Journal of Finance*, 54 (3), 875–899.

Das, S.R. and R.K. Sundaram, 2000, "Fee Speech: Signalling, Risk-Sharing, and the Impact of Fee Structures on Investor Welfare," working paper.

Davis, S.J. and P. Willen, 2000a, "Occupation-Level Income Shocks and Asset Returns: Their Covariance and Implications for Portfolio Choice," *NBER working paper 7905.*

Davis, S.J. and P. Willen, 2000b, "Using Financial Assets to Hedge Labor Income Risk: Estimating the Benefits," working paper.

Dermine, J., D. Neven, and J.F. Thisse (1991), "Towards an Equilibrium Model of the Mutual Funds Industry," *Journal of Banking and Finance,* 15, 485 – 499.

Dow, J., and G. Gorton (1997), "Noise Trading, Delegated Portfolio Management, and Economic Welfare," *Journal of Political Economy*, 105, 1024-1050.

Duffie, D. and M.O. Jackson, 1990, "Optimal Innovation of Futures Contracts," *Review of Financial Studies,* 2 (3), 275–296.

Ferson, W.E. and R.W. Schadt, 1996, "Measuring Fund Strategy and Performance in Changing Economic Conditions," *Journal of Finance,* 51, 425-461.

Fung, W. and D.A. Hsieh, 1997, "Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds," *Review of Financial Studies*, 10 (2), 275–302.

Glosten, Larry, 1994, "Is the Electronic Open Limit Order Book Inevitable?," *Journal of Finance*, 49, 1127-61.

Gruber. M., 1996, "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance*, 51, 783–810.

Hardle, W., 1990, *Applied Nonparametric Regression*, Econometric Society Monographs No. 19. Cambridge University Press, Cambridge.

Hechinger, J., May 14, 2001, "All in the Timing: 2 Fidelity Funds SufferAfter Manager's Exit, Move to Tech Stocks," *Wall Street Journal*, Dow Jones and Company Inc., C1.

Ho, Thomas and Hans Stoll, 1983, "The Dynamics of Dealer Markets Under Competition," *Journal of Finance*, 38,, 1053-1074.

Khorana, A. and H. Servaes, 2000, "What Drives Market Share in the Mutual Fund Industry," working paper.

Kyle, Albert, 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

Lynch, A.W. and D.K. Musto, 2000, "How Investors Interpret Past Fund Returns," working paper.

Massa, Massimo, 2000, "Why So Many Mutual Funds? Mutual Fund Families, Market Segmentation and Financial Performance," INSEAD working paper.

Merrill Lynch advertisement, August 5, 2001, *New York Times Magazine*, 2-3.

Nanda, V., M.P. Narayanan, and V.A. Warther, 2000, "Liquidity, Investment Ability, and Mutual Fund Structure," *Journal of Financial Economics*, 57, 417–443.

Ou-Yang, Hui, 1997, "Optimal Contracts in Continuous Time Principal-Agent Problems," University of North Carolina at Chapel Hill working paper.

Pirrong, Craig, 1999, "The Organization of financial Exchange Markets: Theory and Evidence," *Journal of Financial Markets*, 2, 329-358.

Sharpe, W., 1966, "Mutual Fund Performance," *Journal of Business*, 39, 119–138.

Sharpe, W., 1992, "Asset Allocation: Management Style and Performance Measurement," *The Journal of Portfolio Management*, 18, 7-19.

Theil, Henri, 1971, *Principles of Econometrics*, John Wiley & Sons, New York.

Vayanos, D. 2001, "Strategic Trading in Dynamic Noisy Markets," *Journal of Finance*.

Viswanathan, S., and J. Wang, "Market Architecture: Limit Order Books Versus Dealership Markets," 2001, 5, forthcoming *Journal of Financial Markets*.

Whittle, Peter, 1990, *Risk-Sensitive Optimal Control*, John Wiley & Sons, New York.

Willen, P., 1999, "Welfare, Financial Innovation and Self Insurance in Dynamic Incomplete Market Models," working paper.

Appendix

A 1 Proofs from Text

Lemma 1*: Investor i sets δ to solve*

$$\delta = -X_1(i) - \sum_{j=1}^{J^*} \alpha_j(i) e_j - \frac{1}{\gamma}\sigma_D^{-2} P_2. \qquad \textbf{(9)}$$

Proof: Based upon the notation and discussion in Section 2.3 trader *i*'s final wealth equals

$$[X_1(i) + N(i)]D + \sum_j \left[f_j^i(e_j)D + \hat{B}_j(i) \right] + B_1(i) + N_0(i) \qquad \textbf{(25)}$$

where $\hat{B}_j$ represents fund *j*'s holdings of the risk free asset. There are now two budget constraints that must also be satisfied. The first is the standard one for the trader

$$[X_1(i) - X_0(i)]P_1 + B_1(i) + B_0 = 0 \qquad \textbf{(26)}$$

and the second a budget constraint for each fund

$$f_j P_2 + \hat{B}_j = 0. \qquad \textbf{(27)}$$

Plugging **(26)** and **(27)** into equation **(25)** produces the scalar version of **(3)** for the trader's terminal wealth. Since the e_j terms represent normally distributed signals about the endowment shocks of traders at particular locations in characteristic space, one can write the endowment shock for trader *i* as

$$n(i) = \sum_j \alpha_j(i) e_j + \eta(i). \qquad \textbf{(28)}$$

The $\alpha_j(i)$ terms equal the loading on signal *j* used to calculate the expected value of $N(i)$ given the correlation across signals and $N(i)$. Due to the assumption of joint normality this leaves a residual endowment shock $\eta(i)$ that is itself normally distributed with mean zero, variance σ_η^2, and independent of the signals.

By using **(28)** in **(3)** and then employing the risk sensitive certainty equivalence principle from Whittle (1990) the solution to the investor's problem for the f_j can be found via the following program

$$\underset{f_j, D, \eta}{ext} \quad \begin{aligned} & X_1(i)[D-P_1] + \left\{\eta(i) + \sum_j \alpha_j e_j + f_j\right\}[D-P_2] + X_0(i)P_1 + B_0(i) \\ & + \frac{1}{2\gamma}\left[\sigma_D^{-2}D^2 + \sigma_\eta^{-2}\eta^2\right]. \end{aligned} \qquad \textbf{(29)}$$

after substituting in the assumption that $N_0(i)$ equals $-N(i)P_2$. Borrowing Whittle's terminology the expression *ext* stands for "extremization" and implies that the objective is maximized with respect to the controls (f_j) and minimized with respect to the unknowns (D, and η).

The resulting first order conditions with respect to the f_j,

$$D - P_2 = 0, \qquad \textbf{(30)}$$

the dividend D,

$$X_1(i) + \eta(i) + \sum_j \alpha_j e_j + f_j + \frac{1}{\gamma}\sigma_D^{-2}D = 0, \qquad \textbf{(31)}$$

and finally the residual endowment shock η,

$$D - P_2 + \frac{1}{\gamma}\sigma_\eta^{-2}\eta = 0 \qquad \textbf{(32)}$$

are found via differentiation of **(29)**. Using the above three equations to eliminate D and η leads to the solution for f_j given in the lemma. Q.E.D.

Theorem 2. *The volatility of the t=2 price, σ_P^2, increases with the number of FMI's.*

Proof: FMI's only impact price volatility through the $\int_{i_1}\sum_j \alpha_j(i)e_j di_1$ term in **(11)**. Therefore if one knows how the volatility of $\int_{i_1}\sum_j \alpha_j(i)e_j di_1$ changes with the total number of FMI's then one knows how price volatility changes with the total number of FMI's.

From **(11)** the part of the price variance (σ_{PJ}^2) influenced by the number of FMI's equals

$$\sigma_{PJ}^2 = E\left\{\left[\int_x \sum_j \alpha_j(x)e_j dx\right]\left[\int_y \sum_j \alpha_j(y)e_j dy\right]\right\} \qquad \textbf{(33)}$$

where x and y are dummy variables of integration. The vector α can be found via the standard arguments and equals $\Sigma_{N(i),e}{}'\Sigma_e^{-1}$, where $\Sigma_{N(i),e}$ equals the 1×J covariance vector between the endowment shock to period one trader i and the set of J informative signals e. The Σ_e term equals the J×J variance-covariance matrix for the signals e. Substituting out for the α terms and using the fact that $E\left[e_j e_j{}'\right]$ = Σ_e produces

$$\sigma_{PJ}^2 = \int_x \int_y \Sigma_{N(x),e}{}'\Sigma_e^{-1}\Sigma_{N(y),e}dxdy. \tag{34}$$

Next, consider the change in price volatility when going from J-1 funds to J FMI's. Let ${}_{J-1}\Sigma_{N(i),e}$ represent the J-1×1 covariance vector between the endowment shock $n(i)$ and the first J-1 signals. Let ${}_{J-1}\Sigma_e$ represent the J-1×J-1 variance-covariance matrix of the first J-1 signals, and ${}_{J-1,J}\Sigma_e$ the J-1×1 covariance vector between the first J-1 signals and the J^{th} signal. Then one can write

$$\sigma_{PJ}^2 = \int_x \int_y \begin{bmatrix} {}_{J-1}\Sigma_{N(x),e} \\ \sigma_{N(x),e_J} \end{bmatrix}' \begin{bmatrix} {}_{J-1}\Sigma_e & {}_{J-1,J}\Sigma_e \\ {}_{J-1,J}\Sigma_e{}' & \sigma_{e_J}^2 \end{bmatrix}^{-1} \begin{bmatrix} {}_{J-1}\Sigma_{N(x),e} \\ \sigma_{N(y),e_J} \end{bmatrix} dx \tag{35}$$

Using the formula for the partitioned inverse this reduces to

$$\begin{aligned} \sigma_{PJ}^2 = & \int_x \int_y {}_{J-1}\Sigma_{N(x),e}{}' \; {}_{J-1}\Sigma_e \; {}_{J-1}\Sigma_{N(y),e} + \\ & \left({}_{J-1}\Sigma_{N(x),e}{}' \; {}_{J-1}\Sigma_e \; {}_{J-1,J}\Sigma_e - \sigma_{N(x),e_J}\right)\left(\sigma_{e_J}^2 - {}_{J-1,J}\Sigma_e{}' \; {}_{J-1}\Sigma_e \; {}_{J-1,J}\Sigma_e\right)^{-1}\left({}_{J-1}\Sigma_{N(y),e}{}' \; {}_{J-1}\Sigma_e \; {}_{J-1,J}\Sigma_e - \sigma_{N(y),e_J}\right)dxdy. \end{aligned} \tag{36}$$

Recall that $\Sigma_{N(i),e}{}'\Sigma_e^{-1}$ equals α(i). Let ${}_{J-1}\alpha(i)$ represent the vector $\alpha(i)$ when only the first J-1 signals are available. Then **(36)** further simplifies to

$$\begin{aligned} \sigma_{PJ}^2 = & \int_x \int_y {}_{J-1}\Sigma_{N(x),e}{}' \; {}_{J-1}\Sigma_e \; {}_{J-1}\Sigma_{N(y),e} + \\ & \left({}_{J-1}\alpha(x)' \; {}_{J-1,J}\Sigma_e - \sigma_{N(x),e_J}\right)\left(\sigma_{e_J}^2 - {}_{J-1,J}\Sigma_e{}' \; {}_{J-1}\Sigma_e \; {}_{J-1,J}\Sigma_e\right)^{-1}\left({}_{J-1}\alpha(y)' \; {}_{J-1,J}\Sigma_e - \sigma_{N(y),e_J}\right)dxdy. \end{aligned} \tag{37}$$

The first term in the integral of **(37)** equals the price variance when there exist J-1 FMI's. Thus, one only needs to sign the second term in the integral to determine if the price variance increases or decreases in the number of FMI's.

The center term in the product of **(37)**, $\left(\sigma_{e_J}^2 - {}_{J-1,J}\Sigma_e{}'_{J-1}\Sigma_{e\ J-1,J}\Sigma_e\right)^{-1}$, is the J^{th} diagonal element of the inverse of Σ_e. Since Σ_e is a positive definite matrix, the diagonal terms of its inverse must be positive and thus this term is positive. Also, note that $\left(\sigma_{e_J}^2 - {}_{J-1,J}\Sigma_e{}'_{J-1}\Sigma_{e\ J-1,J}\Sigma_e\right)^{-1}$ is a positive scalar that does not vary by investor. Thus, one can change the order of integration of the product and rewrite it as

$$\begin{aligned}&\int_x \left({}_{J-1}\alpha(x)'{}_{J-1,J}\Sigma_e - \sigma_{N(x),e_J}\right)dx \int_y \left({}_{J-1}\alpha(y)'{}_{J-1,J}\Sigma_e - \sigma_{N(y),e_J}\right)dy = \\ &\left[\int_x \left({}_{J-1}\alpha(x)'{}_{J-1,J}\Sigma_e - \sigma_{N(x),e_J}\right)dx\right]^2 > 0.\end{aligned} \quad \textbf{(38)}$$

Q.E.D.

Lemma 3*: The optimal holdings $X_1(i)$ of the date one investors is given by*

$$X_1 = \frac{\overline{P}_2}{\gamma\sigma_P^2} - \frac{\sigma_D^2 + \sigma_P^2}{\gamma\sigma_D^2\sigma_P^2}P_1 \quad \textbf{(39)}$$

Market clearing $\int_0^1 [X_0(i) - X_1(i)]di = 0$ *implies that the t=1 price P_1 is given by*

$$P_1 = \frac{\sigma_D^2}{\sigma_D^2 + \sigma_P^2}(\overline{P}_2 - \gamma\sigma_P^2 X_0). \quad \textbf{(40)}$$

Proof: Let $V(W_1,J)$ represent the expected utility of a period one trader given his initial wealth and given that there exist J FMI's that trade based upon informative signals. After plugging **(9)** into **(3)** and recalling the assumption that $N_0(i)$ equals $-N(i)P_2$ this trader's terminal wealth equals

$$W_3(i) = X_1(i)(P_2 - P_1) + \left(\eta(i) - P_2/\gamma\sigma_D^2\right)(D - P_2) + X_0(i)P_1 + B_0(i) \quad \textbf{(41)}$$

where D, P_2, and $\eta(i)$ are unknown at time one. Taking the expectation of $\exp(-rC_3(i))$ and denoting period one wealth as $W_1(i) = X_1(i) + B_1(i)$ produces

$$\begin{aligned}&V(W_1,J) = E\left[U(W_3(i))\right] = \\ &-\left[\left(1+\sigma_P^2\sigma_D^{-2}\right)\left(1-\gamma^2\sigma_D^2\sigma_\eta^2(i)\right)\right]^{-1/2}\exp\left\{-\gamma W_1 - \frac{\overline{P}_2^2 + 2\gamma\overline{P}_2\sigma_D^2X_1 - \gamma^2\sigma_D^2\sigma_P^2X_1^2}{2(\sigma_D^2+\sigma_P^2)}\right\}\end{aligned} \quad \textbf{(42)}$$

where the expected value of P_2 equals $\bar{P}_2$, and its variance is given by σ_P^2. Note that in equation **(42)**, the two terms which depend on the number of funds J are $\sigma_\eta^2(i)$ and σ_P^2. In particular, the introduction of a new fund will decrease $\sigma_\eta^2(i)$ and will increase σ_P^2 (see Theorem 2).

Instead of taking X_1 as given, consider how traders optimally choose their asset holdings in period one. The time one budget constraint is given by

$$W_1 = W_0 - X_1 P_1. \tag{43}$$

Taking the first order condition of $V(W_1;J)$ and solving for X_1 yields **(12)**. Notice that this demand does not depend on a given agent's residual endowment risk from period two (this is due to the assumption that agents pay the market price for their time two endowment shocks of shares). Hence the expression gives yields the aggregate demand in period one as a function of the period one price. Setting this demand equal to X_1^0 (the aggregate time one supply) the market clearing price equals **(13)**. Q.E.D.

A 2 Multiple Security Case

A 2.1 The Control Problem

The $J\times1$ vector e includes the signals e_j for families 1,...,J. The $K\times J$ covariance matrix of the signals e and the endowment vector $N(i)$ equals $\Sigma_{eN(i)}$. From Bayes' rule, given the set of signals e, i's endowment vector can be written as

$$N(i) = A(i)e + \eta(i) \tag{44}$$

where $A(i)$ is a $K\times J$ matrix and $\eta(i)$ a $K\times1$ residual vector. Thus, one can now write the trader's final wealth as

$$X_1(i)'[D - P_1] + [\eta(i) + A(i)'e + f]'[D - P_2] + X_0(i)'P_1 + B_0(i), \tag{45}$$

where $X_1(i)$, P_1, and f are now $K\times1$ vectors, and Σ_η the $K\times K$ variance covariance matrix of the residuals η. Note, that f now represents the vector of aggregate holdings across all funds that the investor entrusts with his period two trades (i.e. $f = \sum_j f_j$, where f_j is a $K\times1$ vector).

Assume that fund family *j* offers $K_j \le K$ funds. Let W_j represent the $K \times K_j$ weighting matrix used by fund family *j*. Thus, w_{kmj} represents the relative number of shares of stock *k* fund *m* in fund family *j* will purchase given a signal e_j. For technical simplicity assume the rank of the weighting matrix W_j equals K_j. (If not, then simply eliminate the redundant funds and reduce K_j accordingly.) Thus, an investor in fund family *j* can independently set his holdings in K_j stocks given any signal e_j or any other variable, such as the price of the risky securities in the market. Having done so the remaining K-K_j stocks are then set according to the weighting matrix used by that fund family.

Let $W_{K_+,j}$ represent the $K_j \times K_j$ matrix encompassing the first K_j rows of W_j. Note, that this square matrix is of full rank, and thus invertible. Let $W_{K_-,j}$ equal the $[K\text{-}K_j] \times K_j$ matrix encompassing the last K-K_j rows of W_j. Thus, if investor *i* holds $\hat{a}_j(i)$ shares in the funds offered by fund family *j* he will then own $f_{j+} = W_{K_+,j}\hat{a}e_j$ shares of stocks one through K_j and $f_{j-} = W_{K_-,j}\hat{a}e_j$ shares in the remaining stocks via fund family *j*. (Implying $\left[f_{j+}{}' \;\; f_{j-}{}' \right] = f_j'$). Thus, one can write

$$f_{j-} = W_{K_-,j} W_{K_+,j}^{-1} \; f_{j+}, \qquad \textbf{(46)}$$

since $\hat{a} e_j = W_{k+,j}^{-1} \; f_{j+}$.

In addition to the restriction given by **(46)** the analysis has to deal with the fact that the orders assigned to fund family *j* can only depend upon signal e_j and the publicly observed prices. This problem can be solved via a somewhat indirect approach. Imagine that an investor has given trading orders to every fund family other than *j* and now wishes to maximize his utility via *j*'s orders. In this case one can take as given the orders assigned to every other fund family, and treat the resulting endowment shocks as exogenous events. That is the ($f_{i1}, f_{i2}, \ldots, f_{iK}$) terms for fund family $i \ne j$ are taken as given when searching for the optimal values for ($f_{j1}, f_{j2}, \ldots, f_{jK}$). Also define $A_{\bullet j}(i)$ as the column vector used to update the investor *i*'s endowment shock given the signal e_j, and

$$\Omega_j = \begin{bmatrix} I \\ W_{K_-,j} W_{K_+,j}{}^{-1} \end{bmatrix} \qquad \textbf{(47)}$$

as the matrix yielding f_j given holdings f_{j+}, that is $f_j = \Omega_j f_{j+}$. The resulting solution to the investor's optimization problem is characterized in the next lemma.

Lemma 6: *In a multiple security environment in which some or all fund families have fewer funds than stocks the solution to the investor's optimization problem for funds available in family j must satisfy*

$$f_{j+} = -\left[\Omega_j' M^{-1}\Omega_j\right]^{-1}\Omega_j'\left(M^{-1}A_{.j}(i)e_j + \left(M^{-1}\gamma\Sigma_\eta + I\right)P_2\right) \quad \textbf{(48)}$$

where $M = \frac{1}{\gamma}\Sigma_D^{-1} - \gamma\Sigma_\eta$ *, and where we have assumed for simplicity that* $X_1(i) = 0$.

Proof: If there exist as many funds in each fund family as stocks then the objective function **(29)** can be expressed in matrix notation as

$$\underset{f,D,\eta}{ext} \quad \begin{array}{l} X_1(i)'[D-P_1] + [\eta(i) + A(i)e + f]'[D-P_2] + X_0(i)'P_1 + B_0(i) \\ + \frac{1}{2\gamma}\left[D'\Sigma_D^{-1}D + \eta'\Sigma_\eta^{-1}\eta\right]. \end{array} \quad \textbf{(49)}$$

However, when fund family j has fewer than K funds problem **(29)** further transforms to

$$\underset{f_{j+},D,\eta}{ext} \quad \begin{array}{l} X_1(i)'[D-P_1] + \left[\eta(i) + A_{.j}(i)e_j + \hat{A}_{.j}(i)P_2\right]'[D-P_2] + f_{j+}'[D_{j+} - P_{2,j+}] \\ + f_{j+}'W_{K_+,j}^{-1\prime}W_{K_-,j}'[D_{j-} - P_{2,j-}] + X_0(i)'P_1 + B_0(i) + \frac{1}{2\gamma}\left[D'\Sigma_D^{-1}D + \eta'\Sigma_\eta^{-1}\eta\right] \end{array} \quad \textbf{(50)}$$

for fund j, where the problem now allows for the fact that only K_j stock demands can be selected as free variables. Here, D_{j+}, and D_{j-} represent the dividends paid by the first K_j stocks and the remaining K-K_j stocks respectively. The variables $P_{2,j+}$, and $P_{2,j-}$ represent the analogous period two prices for the stocks. The $A_{.j}(i)$ and $\hat{A}_{.j}(i)$ terms represent column vectors that convert the signals e_j and P_2 into estimates of the total shares owned by investor i through both his endowment shock $N(i)$ and orders to the other investment fund families. The residual η is then the error term from these two estimates. Thus, one implicitly has

$$N(i) + \sum_{h \neq j} f_h = A_{.j}(i)e + \hat{A}_{.j}(i)P_2 + \eta(i). \quad \textbf{(51)}$$

Differentiating **(50)** with respect to f_{j+}, D, and η yields first order conditions of,

$$D_{j+} - P_{2,j+} + W_{K_+,j}{}^{-1\prime} W_{K_1,j}{}' (D_{j-} - P_{2,j-}) = 0, \tag{52}$$

$$X_1(i) + \eta(i) + e_j A_j(i)' + P_2{}' \hat{A}_j(i)' + \begin{bmatrix} f_{j+} \\ W_{K_-,j} W_{K_+,j}{}^{-1} f_{j+} \end{bmatrix} + \frac{1}{\gamma}\Sigma_D^{-1} D = 0, \tag{53}$$

and

$$D - P_2 + \frac{1}{\gamma}\Sigma_\eta^{-1}\eta = 0 \tag{54}$$

respectively.

Use equations **(52)** and **(54)** to eliminate D and η from **(54)**. Next use the definition for Ω_j given by **(47)** and write the solution to f_{j+} as **(48)**, where the assumed existence of a complete set of price funds and constant funds have been used to eliminate the X_1 and P_2 vectors. Q.E.D.

To obtain some insight regarding the solution to j_{j+} first consider the case where the fund family contains as many funds as there are stocks. In this case, Ω_j is simply a $K \times K$ identity matrix, and the solution to f reduces to $-A_j(i)e_j - (\gamma\Sigma_D)^{-1}P_2$. Thus, in the "complete markets" case the weights assigned to each stock returns to the previously discussed condition $(a_{1j}, a_{2j}, \ldots, a_{Kj})' e_j$.

Consider, however, what happens when markets are incomplete. Recall that the endowment vector of investor i may be written as $N(i) = A(i)e + \eta(i)$, where $A(i)$ is a K by J vector. Let us write $A_c(i)$ for the c^{th} column of $A(i)$. Ideally, the investor wants to ask each fund family to trade the vector $A_c(i)e$. However, due to an incomplete set of funds, this may not be possible. The investor therefore solves the following problem

$$\min_f \text{ Distance Between } [\Omega\, f, A_c(i)] \tag{55}$$

where the appropriate measure of distance has yet to be determined. One solution is to simply project, in a least squares sense, $A_c(i)$ into the space spanned by Ω to get

$$f = (\Omega'\Omega)^{-1}\Omega' A_c(i), \tag{56}$$

the standard regression solution. Equation **(48)** seems similar to equation **(56)**.

However, this is not quite optimal. Since the projection cannot set the distance to zero, investors' risk preferences may cause them to deviate away from the least squares solution. In particular, consider a security which pays a very variable dividend at time three. Clearly it is more costly to leave oneself unhedged with respect to endowment shocks in this security than it is to leave oneself unhedged to endowment shocks in a security whose time three dividend is known for sure. Optimally an investor therefore sacrifices a precise hedge of the latter security's endowment shock, if this allows him to better hedge the former security's endowment shock. A second effect is also in play. Consider two securities whose residual endowment variances (i.e. the part which is not spanned by the signals collected by the fund families) are small and large, respectively. It would be better to hedge one's endowment shock with respect to the less variable security than the shock with respect to the more variable security. In a sense hedging the more variable security is pointless because the residual variance of the endowment shock will still induce the investor to be very exposed to shocks in that security. However, by hedging out one's exposure to the security with the less variable endowment shock, the investor is able to almost fully eliminate any risk coming from it.

Hence, risk averse investors find it optimal to deviate away from the least squares solution. In fact, they optimally choose a weighting on the fund family's offered portfolios given by

$$f = (\Omega' M^{-1}\Omega)^{-1}\Omega' M^{-1} A_c(i) \tag{57}$$

where the matrix $M = (\gamma\Sigma_D)^{-1} - \gamma\Sigma_\eta$ reflects investors' responses to the two issues just discussed: hedging distances from stocks with more variable dividends are penalized more, and hedging distance from stocks with more variable residual endowment shocks are penalized less. Furthermore, we see that as investors approach risk neutrality (i.e. as $\gamma \to 0$), their optimal portfolio weights reflect more the dividend hedging motive. Conversely, as investors become more risk averse (i.e. as $\gamma \to \infty$), their optimal weights reflect more their desire to only hedge those securities which have lower residual endowment variances.

Though the price vector contains information otherwise unavailable to each fund manager the investors do not want the fund to trade on the basis of that information. Instead, investors only ask the fund manager to trade in response to that manger's signal. The investors then create the portfolio strategy they prefer by combining funds that gather information with those that do not. Instead of asking the managers gathering signals to trade on the basis of the revealed price vector, investors prefer to adjust their holding weights (the *A* matrix) to account for whatever information the market may reveal in period two.

Here, as in previous sections of the paper, one sees that modeling financial market intermediaries as corporations has a substantial impact on the conclusions one draws about the economy. In a model where such intermediaries maximize a risk averse utility function, each firm seeks to trade into the market portfolio. Thus, effectively all intermediaries pursue the same strategy, which is not surprising since they are assumed to act like identical human beings. By contrast, when such firms are modeled as corporations their investors demand that they pursue a wide range of strategies. That way the investing public can craft individual trading strategies via the appropriate purchase of positions in the individual investment funds.

A 2.2 Comparison to the Standard Model

The following table contrasts the results of our intermediation model to those which apply in the traditional setting.

Summary 4: A Comparison of the Equilibrium FMI Strategies In a Standard Model of Financial Market Intermediation with Multiple Risky Securities

Result	Standard Model	First Principle's Model
Do Fund Families Exist?	No. Firms do not have an incentive to create additional competitors.	Yes. Funds families offer investors a richer strategy space for the utilization of any particular signal.
Strategies	Identical, all FMI's hold the market portfolio.	Divergent. Funds obtaining an informative signal trade only on the basis of their signal, and ignore prices. Other fund types that trade only on the basis of prices also exist, as well as funds that trade constant amounts independently of any signal.
Prices	Satisfy unrestricted mean-variance equilibrium conditions.	Prices depend both upon the dividend and supply variance via a risk weighted projection into the space of trading strategies available via existing FMI's.

A 2.3 The Equilibrium with an Incomplete Set of Funds

To gain further insight into the model's equilibrium characteristics this section adds the simplifying assumption that each fund family provides the same set of L funds (i.e. $\Omega_j = \Omega$ for all fund families). Funds thus differentiate themselves by trading on the basis of their specific signal e_j.

Given this simplification, and recalling the fact that one can write the endowment vector of each individual as $N(i) = A(i)e + \eta(i)$, the following result describes each individual's optimal trade.

Lemma 7: *In a K security environment in which all fund families offer the same set of L traded funds (with L < K), the optimal fund demand of investor i is given by*

$$f(i) = -\left[\Omega' M^{-1}\Omega\right]^{-1}\Omega'\left(M^{-1}\left(A_1(i)e_1 + \cdots + A_J(i)e_J\right) + \left(M^{-1}\gamma\Sigma_\eta + I\right)P_2\right) \quad \textbf{(58)}$$

where $M = \left(\gamma\Sigma_D\right)^{-1} - \gamma\Sigma_\eta$.

Notice that this is an L by 1 vector which specifies which funds in each fund family investor i will choose. The trade can be implemented by giving the L by 1 vector

$$-\left[\Omega' M^{-1}\Omega\right]^{-1}\Omega' M^{-1}A_j(i) \quad \textbf{(59)}$$

to each family j, and the L by K vector (recall there are K securities)

$$-\left[\Omega' M^{-1}\Omega\right]^{-1}\Omega'(M^{-1}\gamma\Sigma_\eta + I) \quad \textbf{(60)}$$

to the price fund family (which allows trades in each of its L available funds to be some function of the K by 1 price vector P_2.)

The time 2 traders will optimally demand X_2 shares, where

$$X_2 = -\frac{1}{\gamma_2}\Sigma_D^{-1}P_2 \quad \textbf{(61)}$$

Hence the market clearing condition sets

$$\int_i \Omega\, f(i)\, di + X_2 = \widetilde{X}_0 \quad \textbf{(62)}$$

If one makes the further assumption that $\Sigma_{\eta(i)}$ is the same for all investors, then the time two price vector satisfies

$$\left(\Omega(\Omega' M^{-1}\Omega)^{-1}\Omega'(M^{-1}\gamma\Sigma_\eta + I) + \frac{1}{\gamma_2}\Sigma_D^{-1}\right)P_2 = -\widetilde{X}_0 - \Omega(\Omega' M^{-1}\Omega)^{-1}\Omega' M^{-1}\overline{A}e \quad \textbf{(63)}$$

where $\overline{A} = \int_i A(i)di$ is a J by 1 vector.

This equation reveals the following insight: since Σ_D is of full rank, a K dimensional price vector exists which will support the equilibrium for any supply vector $\widetilde{X}_0$. This would not necessarily be true if period two traders were not present in the market since demands of the period one traders only have L (with $L<K$) degrees of freedom. In this case an equilibrium price exists only when the period two supply is restricted to a linear combination of the offered funds (i.e. one needs $\widetilde{X}_0 = \Omega\widehat{X}_0$ for some L dimensional vector $\widehat{X}_0$).

Some straightforward manipulations express the time two price equation as follows

$$\Omega'\left(I - \gamma^2\Sigma_D\Sigma_\eta\right)^{-1}\left(P_2(1+\gamma/\gamma_2) + \gamma\Sigma_D(\widetilde{X}_0 + \overline{A}e)\right) = 0. \quad \textbf{(64)}$$

Notice that if Ω were the K dimensional identity matrix, the price is given by

$$P_2 = -\frac{\gamma\Sigma_D}{1+\gamma/\gamma_2}(\widetilde{X}_0 + \overline{A}e) \quad \textbf{(65)}$$

which is exactly the multidimensional analog of the price in the one security world. However, given the fact that fund families offer only a subset of the possible funds, this standard K dimensional equilibrium condition does not have to hold. An L dimensional restriction of this condition must hold instead.

A 2.4 Transforming an Incomplete Economy into a Complete One

Define a K-L economy to be an economy with K securities and L (with $L<K$) fund vectors. Given the discussion in Section 5, this economy is incomplete in the sense that fewer funds exist than securities. While it is possible to characterize the equilibrium of this economy, the incompleteness of the fund space makes analysis somewhat difficult. In some circumstances, however, one can transform an incomplete K-L economy into a complete L-L economy. In such cases, equation **(65)** gives the equilibrium price in the L-L economy.

Rewrite equation **(64)** as follows

$$\widehat{P}_2 \equiv \Omega' M^{-1} \Sigma_D^{-1} P_2 = -\frac{\gamma}{1+\gamma/\gamma_2} \Omega' M^{-1} (\widetilde{X}_0 + \overline{A}e). \quad \textbf{(66)}$$

This is an L dimensional restriction on the K dimensional price vector. Consider the L by K matrix

$$R \equiv \Omega' M^{-1} \Sigma_D^{-1}. \quad \textbf{(67)}$$

This matrix gives us the L dimensional combination of the K securities in the incomplete economy whose price $\widehat{P}_2$ is given by equation **(66)**. Call two economies *identical* if, after trade at time one, the distribution of agents' residual endowment shocks at time two are everywhere equal (note that this is a stronger condition than saying that the endowment shocks should be equal with probability one).

Given certain conditions on the endowment shocks of the agents and on the time two supply vector, $\widetilde{X}_0$, it is possible to create an L-L economy which is identical to the original K-L economy. Assume that for some $L<K$ there exists an K by L dimensional funds matrix Ω such that the following conditions hold

$$\begin{aligned} \widetilde{X}_0 &= \Sigma_D^{-1} M^{-1} \Omega \widehat{X}_0, \\ A(i) &= \Sigma_D^{-1} M^{-1} \Omega \widehat{A}(i) \qquad \forall i, \\ \eta(i) &= \Sigma_D^{-1} M^{-1} \Omega \hat{\eta}(i) \qquad \forall i. \end{aligned} \quad \textbf{(68)}$$

where $\widehat{X}_0, \widehat{A}(i), \hat{\eta}(i)$ have dimensions $L \times 1$, $L \times J$, and $L \times 1$ respectively. These conditions say that for a sufficiently large L the endowment shocks which agents receive at time two, as well as the time two supply, are really L dimensional. The following theorem states the transformation result.

Theorem 3: *Assume that condition **(68)** holds. Then a* K-L *economy can be transformed into an equivalent* L-L *economy where the payoffs of the* L *securities is given by*

$$\widehat{D} = R\,D \quad \textbf{(69)}$$

(for the matrix R *from equation **(67)**), where agents' endowments are given by*

$$\widehat{N}(i) = \widehat{A}(i)\,e + \hat{\eta}(i), \quad \textbf{(70)}$$

where the funds matrix is given by

$$\widehat{\Omega} = (\Omega' M^{-1}\Sigma_D^{-1} M^{-1}\Omega)^{-1}(\Omega' M \Omega), \tag{71}$$

and where the L *dimensional price vector is given by*

$$\widehat{P}_2 = R P_2 = \Omega' M^{-1}\Sigma_D^{-1} P_2 \tag{72}$$

Proof: First the proofs shows that the endowment shocks in the two economies are indeed the same. Given an L dimensional vector x of securities in the L-L economy, their payoff will be equal to $R'x$ (a K dimensional vector) in the K-L economy. Hence one needs to show that $N(i) = R'\widehat{N}(i)$, which is true by construction.

The next goal is to show that the trade vector *f(i)* will be the same in the two economies. Define

$$\hat{\Sigma}_D \equiv R\Sigma_D R' = \Omega' M^{-1}\Sigma_D^{-1} M^{-1}\Omega. \tag{73}$$

Using **(58)** and the conditions in **(68)**, the demand in the K-L economy is given by

$$f(i) = -(\Omega' M^{-1}\Omega)^{-1}\hat{\Sigma}_D \widehat{A}(i) e - (\Omega' M^{-1}\Omega)^{-1}\frac{1}{\gamma}\widehat{P}_2. \tag{74}$$

Similarly, since the economy is complete in the L-L case agent's demands are given by

$$f(i) = -\widehat{\Omega}^{-1}\widehat{A}(i) e - \widehat{\Omega}^{-1}\hat{\Sigma}_D^{-1}\frac{1}{\gamma}\widehat{P}_2. \tag{75}$$

Given the definition of $\hat{\Omega}$ from equation **(71)** these are indeed identical.

That the price vector is given by **(72)** can be seen by observing that the complete markets price equation **(65)** in implies that

$$\widehat{P}_2 = -\frac{\gamma\hat{\Sigma}_D}{1+\gamma/\gamma_2}(\widehat{X}_0 + \widehat{\widehat{A}}e). \tag{76}$$

The results follows from the conditions in **(68)**, as well as the definition of $\hat{\Sigma}_D$ given above. Q.E.D.

The significance of this theorem lies in the interpretation of the number of funds in an economy relative to the number of securities. First, given this theorem think of **(66)** not as an L dimensional restriction on the K prices in the K-L economy, but rather as the prices of the appropriately defined L securities in a complete *L-L* economy. More importantly, though, the true dimensionality of an economy is not in the number of securities which exist (even if the payoffs of all of these are independent), but rather in the endowment structure of the economy. The number of fund vectors which are needed for full spanning is not equal to the number of independent security payoffs, but rather is equal to the number of factors which drive endowment shocks.

A 3 Time Varying Beta Regressions

The estimated vector β_{t_0} , comes from a weighted OLS regression of the form $y_t = \alpha_{t_0} + \beta_{t_0} x_t + \varepsilon_t$, where t belongs to the interval t_0-t_s and t_0+t_s, exclusive of t_o. Thus, t_s equals the half window length. The weight assigned to the period *s* observation in the OLS estimate of the regression, is given by a tri-kernel of the form

$$w_s = \frac{\left[1 - \left(\frac{|s - t_0|}{2t_s}\right)^3\right]^3}{\sum_{\substack{s=t_0-t_s \\ s \neq t_0}}^{t_0+t_s} \left[1 - \left(\frac{|s - t_0|}{2t_s}\right)^3\right]^3}. \qquad \textbf{(77)}$$

In the above equation t_0 equals the month for which parameter estimates are desired. The variable t_s equals the half window length, five for the estimates in this paper. Finally, *s* equals the period from which data has been drawn. For example, if the goal is to estimate the model parameters for June, and data comes from the month of April, then $|s\text{-}t_0|$ equals two. Hence more weight is assigned to times *s* which are close to t_0. The period t_0 has been excluded to avoid any possibility that the estimated correlations are due simply to correlated returns in the most highly weighted month of data.

The tri-cubed kernel was selected because it is known to perform well under a wide range of conditions. However, the choice is not that critical since most kernels produce similar results, see Hardle (1990).

Table 1: Division of Equity Funds by Type.

Year	Number of Equities[1]	Equity Funds[2]		Funds Primarily Invested in Equities[3]		Equity Only Funds[4]		Equity Index Funds[5]		Asset Allocation Funds[6]	
		#	TNA[7]	#	TNA	#	TNA	#	TNA	#	TNA
1990	6,635	785	296,361	182	73,469	37	15,344	5	2,631	465	188,566
1991	6,529	738	395,528	218	118,617	39	26,232	6	5,318	430	262,492
1992	6,630	3,998	1,454,405	729	166,378	281	58,108	36	11,837	1,035	429,002
1993	6,790	5,392	1,910,069	1,049	282,405	467	95,953	52	20,347	1,297	624,560
1994	7,554	7,167	1,970,881	1,656	437,746	696	120,200	57	23,206	1,521	658,459
1995	7,982	8,262	2,562,696	2,071	560,718	991	190,866	68	40,303	1,714	871,821
1996	8,221	9,027	3,103,536	2,654	904,693	1,257	295,774	87	75,592	1,832	1,073,854
1997	8,770	10,741	3,945,730	3,604	1,444,976	2,018	476,416	134	128,992	1,907	1,339,362
1998	8,822	11,322	4,805,749	4,410	2,067,695	2,059	678,513	146	198,588	1,877	1,551,096
1999	8,435	11,882	5,809,238	4,936	2,769,340	1,996	934,607	178	296,102	1,775	1,802,193

1. Total companies with available prices on the NYSE, AMEX and NASDAQ combined as of the first trading day in January.
2. Any fund that held 5% or more of its net asset value in equities at some time during its life.
3. Any fund with 90% or more of its portfolio invested in equities in that year.
4. Funds that held 90% or more of their net asset value in equities over their entire life with a minimum of three years in operation.
5. Any fund with “Index” or “Idx” as part of its name and with 90% or more of its portfolio invested in equities in that year.
6. Any fund in which equity holdings as a fraction of all assets vary by 20 or more percentage points during the life of the fund.
7. Total net assets in millions of dollars.

Table 2: Total Number of Fund Families and Funds in Each Morningstar Category as of December 31, 1999.

Morningstar Categories	# of Fund Families[1]	# of Funds[2]	Morningstar Categories	# of Fund Families	# of Funds
Convertibles	24	57	Muni CA Long	38	111
Diversified Emerging Mkts	72	168	Muni NY Interm	18	38
Diversified Pacific/Asia Stock	21	51	Muni NY Long	39	87
Domestic Hybrid	216	747	Muni National Interm	88	181
Emerg Mkts Bond	21	46	Muni National Long	98	307
Europe Stock	55	152	Muni Short	44	103
Foreign Stock	214	668	Muni Single State Interm	58	250
High Yield Bond	105	321	Muni Single State Long	69	717
Intermediate Government	102	308	Pacific/Asia ex-Japan Stock	35	117
Intermediate-term Bond	209	554	Short Government	73	174
International Bond	63	196	Short-term Bond	94	216
International Hybrid	31	82	Small Blend	110	238
Japan Stock	24	52	Small Growth	150	375
Large Blend	242	934	Small Value	105	235
Large Growth	191	645	Specialty-Communication	12	21
Large Value	182	614	Specialty-Financial	25	65
Latin America Stock	25	53	Specialty-Health	25	59
Long Government	28	50	Specialty-Natural Res	24	65
Long-term Bond	56	120	Specialty-Precious Metals	21	42
Mid-Cap Blend	115	262	Specialty-Real Estate	53	128
Mid-Cap Growth	138	393	Specialty-Technology	48	104
Mid-Cap Value	105	275	Specialty-Utilities	29	97
Multisector Bond	43	129	Ultrashort Bond	35	54
Muni CA Interm	25	39	World Stock	83	279

Notes: Total of 48 Categories, 630 Fund Families, and 10,979 Funds.
1. Number of fund families with at least one fund in each category.
2. Number of individual funds in each category.

Table 3: Distribution of Funds within Morningstar Categories by Fund Family.

# of Funds[1]	# of Cat.[2]	# of Fam.[3]	# of Funds	# of Cat.	# of Fam.	# of Funds	# of Cat.	# of Fam.	# of Funds	# of Cat.	# of Fam.
1	1	168	10	2	1	15	4	3	22	8	1
2	1	23	10	4	2	15	5	1	22	11	1
2	2	45	10	7	2	15	6	3	23	7	1
3	1	8	10	8	2	15	8	1	23	8	1
3	2	9	10	10	1	15	9	1	23	9	1
3	3	36	11	1	1	15	10	1	23	10	2
4	1	2	11	3	1	15	12	1	24	3	1
4	2	6	11	4	1	16	3	1	24	5	1
4	3	13	11	5	2	16	6	1	24	7	1
4	4	21	11	6	2	16	7	1	24	11	1
5	2	1	11	7	1	16	9	2	25	9	1
5	3	6	11	8	1	16	11	2	26	6	1
5	4	7	11	9	2	17	5	1	26	7	1
5	5	16	11	11	2	17	9	2	26	11	1
6	2	3	12	2	2	17	10	2	26	12	1
6	3	4	12	4	1	17	11	1	26	16	1
6	4	3	12	5	1	18	4	1	27	16	1
6	5	5	12	8	1	18	7	1	27	19	1
6	6	6	12	9	3	18	8	1	27	21	1
7	1	1	12	10	1	18	9	1	28	9	1
7	2	2	12	11	1	18	11	1	28	10	1
7	4	2	13	5	1	18	12	1	28	17	1
7	6	5	13	6	1	18	16	1	29	9	1
7	7	6	13	7	2	19	11	1	29	11	1
8	3	1	13	8	3	19	15	1	29	13	3
8	4	3	13	9	1	19	16	1	29	23	1
8	5	2	13	10	2	20	3	1	30	15	1
8	6	5	13	11	1	20	8	2	31	9	1
8	7	1	13	13	1	20	9	1	31	10	1
8	8	3	14	5	1	20	10	1	32	6	1
9	3	2	14	6	1	20	11	1	32	7	1
9	4	2	14	8	1	21	5	1	32	16	1
9	5	3				21	6	1			
9	7	1				21	11	1			
9	8	4									
9	9	1									

Notes: Fund families with more than 32 funds not displayed.
1. Number of funds offered by the fund family.
2. Number of Morningstar distinct categories in which the funds appear.
3. Number of fund families within the grouping.
Light shading 100% > # of Categories/# of Funds > 75%,
Dark shading # of Categories = # of Funds.

Table 4: Return Correlations Within Families by Fund Introduction Number				
	Average	1 With 2	1 With 3	2 With 3
All Funds	0.4209	0.5977	0.5313	0.5426
	(0.4002)	(0.3598)	(0.3636)	(0.3525)
	306,932	408	354	351
Correlations ≠ ±1	0.4208	0.5977	0.5313	0.5426
	(0.4002)	(0.3598)	(0.3636)	(0.3525)
	306,901	408	354	351
Correlations ≠ ±1 & No Index Funds	0.4206	0.5976	0.5376	0.5408
	(0.3992)	(0.3607)	(0.3665)	(0.3539)
	296,515	403	344	342

Notes: Correlations are only estimated when there exist twelve months or more of overlapping return data for two funds. Fund introduction number refers to the order in which funds (in existence as of December 1999, and tracked by Morningstar) were introduced within a family. The oldest operating fund within a family is labeled 1, the second oldest 2, etcetera.

Cell entries: Top – average return correlation between funds. Middle (in parentheses) – standard *deviation*. Bottom – number of observations.

Table 5: Return Correlations Across Fund Family Funds by Introduction Number

	Panel A: Fund Introduction Number Within a Family			
	All Pairs			
	All	1	2	3
All	0.3820 (0.3879) 34,712,391	0.4091 (0.3724) 4,470,118	0.3961 (0.3832) 6,496,828	0.3825 (0.3775) 2,122,379
1		0.4607 (0.3618) 152,157	0.4413 (0.3693) 602,592	0.4373 (0.3630) 294,149
2			0.4193 (0.3827) 331,868	0.4114 (0.3806) 542,254
3				0.3855 (0.3706) 32,959
	Panel B: All Pairs Excluding Those With an Index Fund			
All	0.3791 (0.3873) 33,450,827	0.4063 (0.3720) 4,354,996	0.3935 (0.3826) 6,287,290	0.3773 (0.3771) 2,034,885
1		0.4593 (0.3618) 149,966	0.4401 (0.3691) 590,782	0.4340 (0.3628) 287,644
2			0.4181 (0.3822) 322,569	0.4089 (0.3802) 525,117
3				0.3781 (0.3712) 31,443
	Panel C: Only Pairs in Which Both Funds are Index Funds			
All	0.6921 (0.3472) 11,313	0.6909 (0.3371) 599	0.6144 (0.3917) 1,654	0.7972 (0.2772) 893
1		0.6519 (0.3813) 6	0.6084 (0.3843) 51	0.7826 (0.3006) 30
2			0.4989 (0.4323) 60	0.5836 (0.4037) 131
3				0.9413 (0.0908) 13

Notes: Funds are initially included when they have one full calendar year of data. Correlations are only estimated when there exists over one year's worth of overlapping dates with which to conduct the calculation. Correlations are only included if both funds are in *separate* fund families. Fund introduction number refers to the order in which funds (in existence as of December 1999, and tracked by Morningstar) were introduced within a family. The oldest operating fund within a family is labeled 1, the second oldest 2, etcetera.

Cell entries: Top – average return correlation between each fund. Middle (in parentheses) – standard *deviation*. Bottom – number of observations.

Table 6: Beta Correlations Within Families

	Average	1 With 2	1 With 3	2 With 3
All Funds	0.2279	0.2212	0.2704	0.2116
	(0.4726)	(0.4514)	(0.4317)	(0.4541)
	239,201	1,759	340	1,423
Correlations ≠ ±1	0.2279	0.2212	0.2704	0.2116
	(0.4726)	(0.4514)	(0.4317)	(0.4541)
	239,191	1,759	340	1,423
Correlations ≠ ±1 & No Index Funds	0.2292	0.2220	0.2696	0.1995
	(0.4745)	(0.4535)	(0.4368)	(0.4309)
	231,727	1,710	318	1,366

Notes: Betas for each fund are calculated with a via a rolling tri-cubed kernel estimator. Window is set to five months before and after each date. The weighted regression excludes the month being estimated. For example, the estimated June beta for a fund uses a weighted average of the date from January to May and July to November. Funds are initially included when they have one full calendar year of data. Correlations are only estimated when there exists over one year's worth of overlapping dates with which to conduct the calculation.

Cell entries: Top – average correlation between the betas of each fund. Middle (in parentheses) – standard *deviation*. Bottom – number of observations.

Table 7: Beta Correlations Across Fund Family Funds

	Panel A: Fund Introduction Number Within a Family			
	All Pairs			
	All	1	2	3
All	0.1214 (0.6372) 34,712,391	0.0873 (0.5865) 4,470,118	0.1110 (0.6212) 6,496,828	0.1264 (0.5975) 2,122,379
1		0.0739 (0.5225) 152,157	0.0854 (0.5644) 450,435	0.0935 (0.5334) 141,992
2			0.1077 (0.6032) 331,868	0.1162 (0.5784) 210,386
3				0.1323 (0.5497) 32,959
	Panel B: All Pairs Excluding Those With an Index Fund			
All	0.1229 (0.6359) 33,450,827	0.0872 (0.5863) 4,421,938	0.1111 (0.6213) 6,390,393	0.1266 (0.5986) 2,064,742
1		0.0731 (0.5207) 149,966	0.0844 (0.5644) 440,816	0.0921 (0.5342) 137,678
2			0.1070 (0.6047) 322,569	0.1152 (0.5808) 202,548
3				0.1329 (0.5536) 31,443
	Panel C: Only Pairs in Which Both Funds are Index Funds			
All	0.3307 (0.7204) 11,313	0.3143 (0.7950) 599	0.2799 (0.6756) 1,654	0.4694 (0.6181) 893
1		-0.0102 (1.0844) 6	0.3540 (0.7552) 45	0.4223 (0.8341) 24
2			0.2591 (0.5822) 60	0.4218 (0.5809) 71
3				0.6110 (0.5423) 13

Notes: Betas for each fund are calculated via a rolling tri-cubed kernel estimator. Window is set to five months before and after each date. The weighted regression excludes the month being estimated. For example, the estimated June beta for a fund uses a weighted average of the data from January to May and July to November. Funds are initially included when they have one full calendar year of data. Correlations are only estimated when there exists over one year's worth of overlapping dates with which to conduct the calculation. Correlations are only included if both funds are in *separate* fund families. Fund introduction number refers to the order in which funds (in existence as of December 1999, and tracked by Morningstar) were introduced within a family. The oldest operating fund within a family is labeled 1, the second oldest 2, etcetera.

Cell entries: Top – average correlation between the betas of each fund. Middle (in parentheses) – standard *deviation*. Bottom – number of observations.

Table 8: Size Factor Beta Correlations Across Fund Family Funds

	Fund Introduction Number Within a Family – All Pairs Included											
	All			1			2			3		
	Large	Mid	Small	Large	Mid	Small	Large	Mid	Small	Large	Mid	Small
All	0.0919	0.0879	0.1290	0.074	0.070	0.0811	0.098	0.0943	0.1128	0.0968	0.0911	0.1193
	(0.6118)	(0.6270)	(0.6586)	(0.5591)	(0.5748)	(0.6099)	(0.5890)	(0.6060)	(0.6446)	(0.5636)	(0.5846)	(0.6190)
		34,712,391			4,470,118			6,496,828			2,122,379	
1				0.0751	0.0758	0.0688	0.0904	0.0862	0.0788	0.0878	0.0845	0.0874
				(0.5042)	(0.5212)	(0.5528)	0.5321	0.5505	0.5875	0.5090	0.5293	0.5612
					152,157			450,436			141,992	
2							0.0984	0.0943	0.1128	0.1105	0.1059	0.1103
							(0.5664)	(0.5858)	(0.6295)	0.5396	0.5607	0.5992
								331,868			210,386	
3										0.1194	0.1173	0.1351
										(0.5394)	(0.5589)	(0.5886)
											32,959	

Notes: Betas for each fund are calculated via a rolling tri-cubed kernel estimator. Window is set to five months before and after each date. The weighted regression excludes the month being estimated. For example, the estimated June beta for a fund uses a weighted average of the data from January to May and July to November. Funds are initially included when they have one full calendar year of data. Correlations are only estimated when there exists over one year's worth of overlapping dates with which to conduct the calculation. Correlations are only included if both funds are in *separate* fund families. Fund introduction number refers to the order in which funds (in existence as of December 1999, and tracked by Morningstar) were introduced within a family. The oldest operating fund within a family is labeled 1, the second oldest 2, etcetera.

Cell entries: Top – average correlation between the betas of each fund. Betas are for each fund's return regressed on the returns of CRSP's large(highest or first decile), mid (fifth decile), and small(ninth decile) capitalization stock portfolios. Middle (in parentheses) – standard *deviation*. Bottom – number of observations.