The History of Finance

An eyewitness account.

Merton H. Miller

At five years, the German Finance Association is not very old as professional societies go, but then neither is the field of finance itself. Finance in its modern form really dates only from the 1950s. In the forty years since then, the field has come to surpass many, perhaps even most, of the more traditional fields of economics in terms of the numbers of students enrolled in finance courses, the numbers of faculty teaching finance courses, and above all in the quantity and quality of their combined scholarly output.

The huge body of scholarly research in finance over the last forty years falls naturally into two main streams. And no, I don't mean "asset pricing" and "corporate finance," but instead a deeper division that cuts across both. The division I have in mind is the more fundamental one between what I will call the *business school* approach to finance and the *economics department* approach. Let me say immediately, however, that my distinction is purely "notional," not physical — a distinction over what the field is really all about, not where the offices of the faculty happen to be located.

In the United States, the vast majority of academics in finance teach in business schools, not economics departments, and always have. At the same time, in the elite schools at least, a substantial fraction of the finance faculties have been trained in — that is, have received their Ph.D.s from — economics departments. Habits of thought acquired in graduate school have a tendency to stay with you.

MERTON H. MILLER is Robert R. McCormick distinguished service professor emeritus at the University of Chicago (IL 60637).

The characteristic business school approach tends to be what we would call in our jargon "micro normative." That is, a decision-maker, whether an individual investor or a corporate manager, is seen as maximizing some objective function, be it utility, expected return, or shareholder value, taking the prices of securities in the market as given. In a business school, after all, that's what you're supposed to be doing: teaching your charges how to make better decisions.

To someone trained in the classical traditions of economics, however, the dictum of the great Alfred Marshall stands out: "It is not the business of the economist to tell the brewer how to make beer." The characteristic economics department approach thus is not micro, but macro normative. The models assume a world of micro optimizers, and deduce from that how market prices, which the micro optimizers take as given, actually evolve.

Note that I am differentiating the stream of research in finance along macro versus micro lines, and not along the more familiar normative versus positive line. Both streams of research in finance are thoroughly positivist in outlook in that they try to be, or at least claim to be, concerned with testable hypotheses. The normal article in finance journals over the last forty years has two main sections: the first presenting the model, and the second an empirical section showing that real-world data are consistent with the model (which is hardly surprising, because had that not been so, the author would never have submitted the paper in the first place, and the editors would never have accepted the article for publication).

The interaction of these two streams, the business school stream and the economics department stream — the micro normative and the macro normative — has largely governed the history of the field of finance to date. I propose to review some of the highpoints of this history, taking full advantage of a handy organizing principle nature has given us: to wit, the Nobel Prizes in Finance.

Let me emphasize that I will not be offering a comprehensive survey of the field — the record is far too extensive for that — but rather a selective view of what I see as the highlights, an eyewitness account, as it were, and always with special emphasis on the tensions between the business school and the economics department streams.

After my overview, I offer some very personal views on where I think the field is heading, or at least where I would be heading were I just entering the field today.

MARKOWITZ AND THE THEORY OF PORTFOLIO SELECTION

The tension between the micro and macro approaches was visible from the very beginning of modern finance — from our big bang, as it were — which I think we can all agree today dates to the year 1952 with the publication in the *Journal of Finance* of Harry Markowitz's article, "Portfolio Selection." Markowitz in this remarkable paper gave, for the first time, a precise definition of what had hitherto been just vague buzzwords: risk and return.

Specifically, Markowitz then identified the yield or return on an investment with the expected value or probability-weighted mean value of its possible outcomes; and its risk with the variance or squared deviations of those outcomes around the mean. This identification of return and risk with mean and variance, so instinctive to finance professionals these days, was far from obvious then. The common perception of risk even today focuses on the likelihood of losses — on what the public thinks of as the "downside" risk — not just on the *variability* of returns.

Markowitz's choice of the variance as his measure of risk, counterintuitive as it may have appeared to many at the time, turns out to have been inspired. It not only subsumes the more intuitive view of risk — because in the normal or at least the symmetric distributions we use in practice the downside risk is essentially the mirror image of the upside — but it also has a property even more important for the development of the field. By identifying return and risk with mean and variance, Markowitz makes the powerful algebra of mathematical statistics available for the study of portfolio selection.

The immediate contribution of that algebra is the famous formula for the variance of a *sum* of random variables; that is, the weighted sum of the variance *plus* twice the weighted sum of the covariances. We in finance have been living on that formula, literally, for more than forty years now. That formula shows, among other things, that for the individual investor, the relevant unit of analysis must always be the whole portfolio, not the individual share. The risk of an individual share cannot be defined apart from its relation to the whole portfolio and, in particular, its covariances with the other components. Covariances, and not mere numbers of securities held, govern the risk-reducing benefits of diversification.

The Markowitz mean-variance model is the perfect example of what I call the business school or micro normative stream in finance. And this is somewhat ironic, in that the Markowitz paper was originally a thesis in the University of Chicago's economics department. Markowitz even notes that Milton Friedman, in fact, voted against the thesis initially on the grounds that it wasn't really economics.

And indeed, the mean-variance model, as visualized by Markowitz, really *wasn't* economics. Markowitz saw investors as actually applying the model to pick their portfolios using a combination of past data and personal judgment to select the needed means, variances, and covariances.

For the variances and covariances, at least, past data probably *could* provide at least a reasonable starting point. The precision of such estimates can always be enhanced by cutting the time interval into smaller and smaller intervals. But what of the means? Simply averaging the returns of the last few years, along the lines of the examples in the Markowitz paper (and later book) won't yield reliable estimates of the return *expected* in the future. And running those unreliable estimates of the means through the computational algorithm can lead to weird, corner portfolios that hardly seem to offer the presumed benefits of diversification, as any finance instructor who has assigned the portfolio selection model as a classroom exercise can testify.

If the Markowitz mean-variance algorithm is useless for selecting optimal portfolios, why do I take its publication as the starting point of modern finance? Because the essentially business school model of Markowitz was transformed by William Sharpe, John Lintner, and Jan Mossin into an economics department model of enormous reach and power.

WILLIAM SHARPE AND THE CAPITAL ASSET PRICING MODEL

That William Sharpe was so instrumental in transforming the Markowitz business school model into an economics department model continues the irony. Markowitz, it will be recalled, submitted his thesis to an economics department, but Sharpe was always a business school faculty member, and much of his earlier work had been in the management science/operations research area. Sharpe also maintains an active consulting practice advising pension funds on their portfolio selection problems. Yet his capital asset pricing model is almost as perfect an example as you can find of an economists' macro normative model of the kind I have described.

Sharpe starts by imagining a world in which every investor is a Markowitz mean-variance portfolio selector. And he supposes further that these investors all share the same expectation as to returns, variances, and covariances. But if the inputs to the portfolio selection are the same, then every investor will hold exactly the same portfolio of risky assets. And because all risky assets must be held by somebody, an immediate implication is that every investor holds the "market portfolio," that is, an aliquot share of every risky security in the proportions in which they are outstanding.

At first sight, of course, the proposition that everyone holds the same portfolio seems too unrealistic to be worth pursuing. Keep in mind first, however, that the proposition applies only to the holdings of risky assets. It does not assume that every investor has the same degree of risk aversion. Investors can always reduce the degree of risk they bear by holding riskless bonds along with the risky stocks in the market portfolio; and they can increase their risk by holding negative amounts of the riskless asset; that is, by borrowing and leveraging their holdings of the market portfolio.

Second, the idea of investing in the market portfolio is no longer strange. Nature has imitated art, as it were. Shortly after Sharpe's work appeared, the market created mutual funds that sought to hold all the shares in the market in their outstanding proportions. Such index funds, or "passive" investment strategies, as they are often called, are now followed by a large and increasing number of investors, particularly by U.S. pension funds.

The realism or lack of realism of the assumptions underlying the Sharpe CAPM has never been a subject of serious debate within the profession, unlike the case of the Modigliani and Miller propositions to be considered later. The profession, from the outset, wholeheartedly adopted the Friedman positivist view: that what counts is not the literal accuracy of the assumptions, but the *predictions* of the model.

In the case of Sharpe's model, these predictions are striking indeed. The CAPM implies that the distribution of expected rates of return across all risky assets is a *linear* function of a single variable, namely, each asset's sensitivity to or covariance with the market portfolio, the famous beta, which becomes the natural measure of a security's risk. The aim of science is to explain a lot with a little, and few models in finance or economics do so more dramatically than the CAPM.

The CAPM not only offers new and powerful theoretical insights into the nature of risk, but also lends itself admirably to the kind of in-depth empirical investigation so necessary for the development of a new field like finance. And its benefits have not been confined narrowly to the field of finance. The great volume of empirical research testing the CAPM has led to major innovations in both theoretical and applied econometrics.

Although the single-beta CAPM managed to withstand more than thirty years of intense econometric investigation, the current consensus within the profession is that a single risk factor, although it takes us an enormous length of the way, is not quite enough for describing the cross-section of expected returns. Besides the market factor, two other pervasive risk factors have by now been identified for common stocks.

One is a size effect; small firms seem to earn higher returns than large firms, on average, even after controlling for beta or market sensitivity. The other is a factor, still not fully understood, but that seems reasonably well captured by the ratio of a firm's accounting book value to its market value. Firms with high book-to-market ratios appear to earn higher returns on average over long horizons than those with low book-to-market ratios after controlling for size and for the market factor.

That a three-factor model has now been shown to describe the data somewhat better than the single-factor CAPM should detract in no way, of course, from appreciation of the enormous influence of the original CAPM on the theory of asset pricing.

THE EFFICIENT MARKETS HYPOTHESIS

The mean-variance model of Markowitz and the CAPM of Sharpe et al. are contributions whose great scientific value was recognized by the Nobel Committee in 1990. A third major contribution to finance was recognized at the same time. But before describing it, let me mention a fourth major contribution that has done much to shape the development of the field of finance in the last twenty-five years, but that has so far not received the attention from the Nobel Committee I believe it deserves.

I refer, of course, to the efficient markets hypothesis, which says, in effect, that no simple rule based on already published and available information can generate above-normal rates of return. On this score of whether mechanical profit opportunities exist, the conflict between the business school tradition in finance and the economics department tradition has been and still remains intense.

The hope that studying finance might open the way to successful stock market speculation served to support interest in the field even before the modern scientific foundations were laid in the 1950s. The first systematic collection of stock market prices, in fact, was compiled under the auspices of the Alfred Cowles Foundation in the 1930s.

Cowles had a lifelong enthusiasm for the stock market, dimmed only slightly by the catastrophic crash of 1929. The Cowles Foundation, currently an adjunct of the Yale University economics department, was the source of much fundamental research on econometrics in the 1940s and '50s.

The Cowles indexes of stock prices have long since been superseded by much more detailed and computerized data bases, such as those of the Center for Research in Security Prices at the University of Chicago. And to those computer data bases, in turn, goes much of the credit for stimulating the empirical research in finance that has given the field its distinctive flavor.

Even before these new computerized data bases came into widespread use in the early 1960s, however, the mechanical approach to above-normal investment returns was already being seriously challenged. The challenge was delivered, curiously enough, not by economists, but by statisticians like M.G. Kendall and my colleague, Harry Roberts — who argued that stock prices are essentially random walks. This implies, among other things, that the record of past stock prices, however rich in "patterns" it might appear, has no predictive power for future stock returns.

By the late 1960s, however, the evidence was accumulating that stock prices are not random walks by the strictest definition of that term. Some elements of predictability *could* be detected, particularly in long-run returns. The issue of whether publicly available information could be used for successful stock market speculation had to be rephrased — a task in which my colleague, Eugene Fama, played the leading role — as whether the observed departures from randomness in the time series of returns on common stocks represent true profit opportunities after transaction costs and after appropriate compensation for changes in risk over time. With this shift in focus from returns to cost- and risk-adjusted returns, the efficient markets debate becomes no longer a matter of statistics, but one of economics.

This connection with economics helps explain why the efficient markets hypothesis of finance remains as strong as ever, despite the steady drumbeat of empirical studies directed against it. If you find some mechanical rule that seems to earn above-normal returns — and with thousands of researchers spinning through the mountains of tapes of past data, anomalies, like the currently fashionable "momentum effects," are bound to keep turning up — then imitators will enter and compete away those above-normal returns exactly as in any other setting in economics. Above-normal profits, wherever they are found, inevitably carry with them the seeds of their own decay.

THE MODIGLIANI-MILLER PROPOSITIONS

Still other pillars on which the field of finance rests are the *Modigliani-Miller propositions* on capital structure. Here, the tensions between the micro normative and the macro normative approaches were evident from the outset, as is clear from the very title of the first M&M paper, "The Cost of Capital, Corporation Finance and the Theory of Investment." The theme of that paper, and indeed of the whole field of corporate finance at the time, is capital budgeting.

The micro normative wing was concerned with finding the "cost of capital," in the sense of the optimal cutoff rate for investment when the firm can finance the project either with debt or equity or some combination of both. The macro normative or economics wing sought to express the aggregate demand for investment by corporations as a function of the cost of capital that firms are actually using as their optimal cutoffs, rather than just the rate of interest on long-term government bonds.

The M&M analysis provided answers, but ones that left both wings of the profession dissatisfied. At the macro normative level, the M&M measure of the cost of capital for aggregate investment functions never really caught on, and, indeed, the very notion of estimating aggregate demand functions for investment has long since been abandoned by macro economists. At the micro level, the M&M propositions imply that the choice of financing instrument is irrelevant for the optimal cutoff. Such a cutoff is seen to depend solely on the risk (or "risk class") of the investment, regardless of how it is financed, hardly a happy position for professors of finance to explain to their students being trained, presumably, in the art of selecting optimal capital structures.

Faced with the unpleasant action consequences of the M&M model at the micro level, the tendency of many at first was to dismiss the assumptions underlying M&M's then-novel arbitrage proof as unrealistic. The assumptions underlying the CAPM, of course, are equally or even more implausible, as noted earlier, but the profession seemed far more willing to accept Friedman's "the assumptions don't matter" position for the CAPM than for the M&M propositions.

The likely reason is that the second blade of the Friedman positivism slogan — what *does* count is the descriptive power of the model itself — was not followed up. Tests by the hundreds of the CAPM fill the literature. But direct calibration tests of the M&M propositions and their implications do not.

One fundamental difficulty of testing the M&M propositions shows up in the initial M&M paper itself. The capital structure proposition says that if you could find two firms whose underlying earnings are identical, then so would be their market values, regardless of how much of the capital structure takes the form of equity as opposed to debt.

But how do you find two companies whose earnings are identical? M&M tried using industry as a way of holding earnings constant, but this sort of filter is far too crude. Attempts to exploit the power of the CAPM for testing M&M were no more successful. How do you compute a beta for the underlying real assets?

One way to avoid the difficulty of not having two identical firms, of course, is to see what happens when the *same* firm changes its capital structure. If a firm borrows and uses the proceeds to pay its shareholders a huge dividend or to buy back shares, does the value of the firm increase? Many studies have suggested that it does. But the interpretation of such results faces a hopeless identification problem.

The firm, after all, never issues a press release saying "we are just conducting a purely scientific investigation of the M&M propositions." The market, which is forward-looking, has every reason to believe that the capital structure decisions are conveying management's views about changes in the firm's prospects for the future. These confounding "information effects," present in every dividend and capital structure decision, render indecisive all tests based on specific corporate actions.

Nor can we hope to refute the M&M propositions indirectly by calling attention to the multitude of new securities and of variations on old securities that are introduced year after year. The M&M propositions say only that no gains could be earned from such innovations if the market were in fact "complete." But the new securities in question may well be serving to complete the market, earning a first-mover's profit to the particular innovation. Only those in Wall Street know how hard it is these days to come by those innovator's profits.

If all this seems reminiscent of the efficient markets hypothesis, that is no accident. The M&M propositions are also ways of saying that there is no free lunch. Firms cannot hope to gain by issuing what looks like low-cost debt rather than high-cost equity. They just make the cost of higher-cost equity even higher. And if any substantial number of firms, at the same time, seek to replace what they think is their high-cost equity with low-cost debt (even tax-advantaged debt), then the interest costs of debt will rise, and the required yields on equity will fall until the perceived incentives to change capital structures (or dividend policies for that matter) are eliminated.

The M&M propositions, in short, like the efficient markets hypothesis, are about *equilibrium* in the capital markets — what equilibrium looks like, and what forces are set in motion once it is disturbed. And this is why neither the efficient markets hypothesis nor the Modigliani-Miller propositions have ever set well with those in the profession who see finance as essentially a branch of management science.

OPTIONS

Fortunately, however, recent developments in finance, also recognized by the Nobel Committee, suggest that the conflict between the two traditions in finance, the business school stream and the economics department stream, may be on the way to reconciliation.

This development, of course, is the field of options, whose pioneers, recently honored by the Nobel Committee, were Robert Merton and Myron Scholes (with the late Fischer Black everywhere acknowledged as the third pivotal figure). Because the intellectual achievement of their work has been commemorated over and over — and rightly so — I will not seek to review it here. Instead, in keeping with my theme, I want to focus on what options mean for the history of finance.

Options mean, among other things, that for the first time in its close to fifty-year history, the field of finance can be built, or as I will argue be rebuilt, on the basis of "observable" magnitudes. I still remember the teasing we financial economists, Harry Markowitz, William Sharpe, and I, had to put up with from the physicists and chemists in Stockholm when we conceded that the basic unit of our research, the expected rate of return, was not actually observable. I tried to parry by reminding them of their neutrino — a particle with no mass whose presence is inferred only as a missing residual from the interactions of other particles. But that was eight years ago. In the meantime, the neutrino has been detected.

To say that option prices are based on observables is not strictly true, of course. The option price in the Black-Scholes-Merton formula depends on the current market value of the underlying share, the striking price, the time to maturity of the contract, and the risk-free rate of interest, all of which are observable either exactly or very closely. But the option price depends also, and very critically, on the variance of the distribution of returns on the underlying share, which is not directly observable; it must be estimated.

Still, as Fischer Black always reminded us, estimating variances is orders of magnitude easier than estimating the means or expected returns that are central to the models of Markowitz, Sharpe, or Modigliani-Miller. The precision of an estimate of the variance can be improved, as noted earlier, by cutting time into smaller and smaller units — from weeks to days to hours to minutes. For means, however, the precision of estimate can be enhanced only by lengthening the sample period, giving rise to the well-known dilemma that by the time a high degree of precision in estimating the mean from past data has been achieved, the mean itself has almost surely shifted.

Having a base in observable quantities — or virtually observable quantities — on which to value securities might seem at first sight to have benefited primarily the management science stream in finance. And indeed, recent years have seen the birth of a new and rapidly growing specialty area within the profession, that of financial engineering (and the recent establishment of a journal with that name is a clear sign that the field is here to stay). The financial engineers have already reduced the original Black-Scholes-Merton formula to Model-T status.

Nor has the micro normative field of *corporate* finance been left out. When it comes to capital budgeting, long a major focus of corporate finance, the decision impact of what have come to be called "real" options — even simple ones like the right to close down a mine when the output price falls and reopen it when it rises — is substantially greater than that of variations in the cost of capital.

The options revolution, if I may call it that, is also transforming the macro normative or economics stream in finance. The hint of things to come in that regard is prefigured in the title of the original Black-Scholes paper, "The Pricing of Options and Corporate Liabilities." The latter phrase was added to the title precisely to convince the editors of the *Journal of Political Economy* — about as economics a journal as you can get — that the original (rejected) version of the paper was not just a technical *tour de force* in mathematical statistics, but an advance with wide application for the study of market prices.

And indeed, the Black-Scholes analysis shows, among other things, how options serve to "complete the market" for securities by eliminating or at least substantially weakening the constraints on high leverage obtainable with ordinary securities. The Black-Scholes demonstration that the shares in highly leveraged corporations are really call options also serves in effect to complete the M&M model of the pricing of corporate equities subject to the prior claims of the debtholders. We can go even further: *Every* security can be thought of as a package of component Arrow-Debreu state-price contingent claims (options, for short), just as every physical object is a package of component atoms and molecules.

RECONSTRUCTION OF FINANCE?

I will speculate no further about these and other exciting prospects for the future. Let me close rather with a question: What would I advise a young member of the German Finance Association to specialize in? What would I specialize in if I were starting over and entering the field today?

Well, I certainly wouldn't go into asset pricing or corporate finance. Research in those subfields has already reached the phase of rapidly diminishing returns. Agency theory, I would argue, is best left to the legal profession, and behavioral finance is best left to the psychologists. So, at the risk of sounding a bit like the character in the movie "The Graduate," I reduce my advice to a single word: options.

When it comes to research potential, options have much to offer both the management science/business school wing within the profession *and* the economics wing. In fact, so vast are the research opportunities for both wings that the field is surely due for a total reconstruction as profound as that following the original breakthrough by Harry Markowitz in 1952.

The shift toward options as the center of gravity of finance that I foresee should be particularly welcomed by the members of the German Finance Association. I can remember when research in finance in Germany was just beginning and tended to consist of replication of American studies using German data. But when it comes to a relatively new area like options, we all stand roughly equal at the starting line. And this is an area in which the rigorous and mathematical German academic training may even offer a comparative advantage.

It is no accident, I believe, that the Deutsche Termin Borse (or Eurex, as it has now become after merging with the Swiss exchange) has taken the high-tech road to a leading position among the world's futures exchanges only eight years after a great conference in Frankfurt where Hartmut Schmidt, Fischer Black, and I sought to persuade the German financial establishment that allowing futures and options trading would not threaten the German economy. Hardware and electronic trading were the key to DTB's success, but I see no reason why the German scholarly community cannot duplicate that success on the more abstract side of research in finance as well.

Whether it can should be clear by the time of the twenty-fifth annual meeting. I'm only sorry I won't be able to see that happy occasion.

ENDNOTE

This is a slightly modified version of an address delivered at the Fifth Annual Meeting of the German Finance Association in Hamburg on September 25, 1998.