Mathematical Finance — Bachelier Congress 2000

Selected Papers
from the First World Congress
of the Bachelier Finance Society,
Paris, June 29—July 1, 2000

Springer
Future Possibilities in Finance Theory and Finance Practice

Robert C. Merton
Harvard Business School
Morgan Hall 397, Soldiers Field
Boston, MA 02163, USA

1 Introduction

The origins of much of the mathematics in modern finance can be traced to Louis Bachelier's 1900 dissertation on the theory of speculation, framed as an option-pricing problem. This work marks the twin births of both the continuous-time mathematics of stochastic processes and the continuous-time economics of derivative-security pricing. In solving his option-pricing problem, Bachelier provides two different derivations of the classic partial differential equation for the probability density of what later was called a Wiener process or Brownian motion process. In one derivation, he writes down a version of what is now commonly called the Chapman-Kolmogorov convolution probability integral in one of the earliest examples of that integral in print. In the other, he uses a limit argument applied to a discrete-time binomial process to derive the continuous-time transition probabilities. Bachelier also develops the method of images (or reflection) to solve for a probability function of a diffusion process with an absorbing barrier. All this in his thesis five years before the publication of Einstein's mathematical theory of Brownian motion.

However, for most of the century, the mathematical and finance branches from Bachelier's work evolved at different paces and independently of one another. On the mathematical side, Kiyoshi Itô was greatly influenced by Bachelier's work in his development in the 1940s and early 1950s of his stochastic calculus, later to become an essential tool in finance. Indeed, at the centennial celebration of Norbert Wiener's birthday at M.I.T. in 1994, Itô told me that Bachelier's thesis was far more influential on his work than Wiener's. Apparently, much the same story holds for the great probabilist A.N. Kolmogorov.

On the financial side, Bachelier's important work was largely lost to financial economists for more than a half century. As we have heard, Paul Samuelson via L.J. Savage brought Bachelier's thesis to the attention of the economics community, including arranging for its translation into English.

---

2 See Itô (1951, 1987).
3 The English translation by James Boness appears in Cootner (1964).
Furthermore, Samuelson’s theory of rational warrant pricing, published in 1965, was centrally motivated by that work.\textsuperscript{4} Henry McKean bridged both the mathematical and financial branches as a scientific collaborator with Itô on the mathematics of stochastic processes and with Samuelson on the finance application to warrant pricing.\textsuperscript{5} It was not, however, until the end of the 1960s and early 1970s that these two branches of research growing from Bachelier’s dissertation were actually reunited. Initially, Itô’s mathematics found its way into finance with the development of the continuous-time theory of optimal lifetime consumption and portfolio selection.\textsuperscript{6} This theory used diffusion processes to model asset price movements and applied the Itô calculus to analyze the dynamics of continuously traded portfolios. The connection between Itô’s work and option pricing was made when that same continuous-trading portfolio modeling tool was used to derive dynamic portfolio strategies that replicate the payoffs to an option, from which the famous Black-Scholes option pricing theory was born.\textsuperscript{7}

This unmistakable lineage from Bachelier’s dissertation to the Black-Scholes model in both its mathematics and its finance underscores the influence of his work on the development and growth of the listed options market, nearly three-quarters of a century later. Subsequently, that same influence greatly impacted the development, refinement and broad-based practical implementation of contracting technology. Derivative securities such as futures, options, swaps and other financial contractual agreements provide a prime example. Innovations in financial-contracting technology have improved efficiency by expanding opportunities for risk sharing, lowering transactions costs and reducing information and agency costs. The numbers reported for the global use of derivative securities are staggering, $70 trillion worldwide and there are a number of individual banking institutions with multi-trillion dollar off-balance-sheet derivative positions. These reported amounts are notional or face values (and often involve double counting), and thus, they do not measure the market value of the contracts themselves, which is much smaller. Nevertheless, it can surely be said that derivatives are ubiquitous throughout the world financial system and that they are used widely by non-financial firms and sovereigns as well as institutions in virtually every part of their financing and risk-managing activities. The significance of Bachelier’s contribution will continue to grow as improved technology, together with growing breadth and experience in the applications of derivatives, expands the scale and scope of

\textsuperscript{4} See Samuelson (1965, 1972) and Samuelson and Merton (1969). For an extensive review of Samuelson’s seminal contributions to the theory of option pricing as well as finance more broadly, see Merton (1983b).

\textsuperscript{5} See Itô and McKean (1964) and McKean (1965).


their use and both consumers and producers of derivatives move down the learning curve.8

Although the paper will address the practice of finance and the science of finance, it will not focus on the latest option-pricing models, nor is it my aim to introduce state-of-the-art computational tools, which might help implementation of these models. Instead I try my hand at providing a frame of reference for these models and tools by describing something of the interaction between those parts of the science of finance which have direct lineage from Bachelier's thesis and their influence on the practice of finance, and possibilities for future trends in each.

1.1 A Functional Perspective for Forecasting Institutional Change

There are two essentially different frames of reference for trying to analyze and understand innovations in the financial system. One perspective takes as given the existing institutional structure of financial-service providers, whether governmental or private-sector, and seeks what can be done to make those institutions perform their particular financial services more efficiently and profitably.

An alternative to this traditional institutional perspective - and the one I favor - is the functional perspective, which takes as given the economic functions served by the financial system and seeks what is the best institutional structure to perform those functions.9 The basic functions of a financial system are essentially the same in all economies, which makes them far more stable, across time and across geopolitical borders, than the identity and structure of the institutions performing them. Thus, a functional perspective offers a more robust frame of reference than an institutional one, especially in a rapidly changing financial environment. It is difficult to use institutions as the conceptual "anchor" for forecasting financial trends when institutional structures are themselves changing significantly, as has been the case for more than two decades and as appears likely to continue well into the future.

Finance theory, which is not institution based, is thus an apt tool for applying the functional perspective to forecast new trends. Indeed, during the last quarter century, finance theory has been a particularly good predictor of future changes in finance practice. That is, when theory seems to suggest that an institution, an instrument, or a service "should be there" and it is not, practice has tended to evolve so that it is. Placed in a normative context, current theory has been a fruitful idea source for subsequent innovations in finance practice.

The Black-Scholes option pricing theory is, of course, the most celebrated instance. However, it is surely not a singular case. The elementary state-

---

8 See Jin, Kogan, Lim, Taylor, and Lo (1997) for a live web site with extensive references documenting the wide range of applications of derivatives.
contingent securities developed as a theoretical construct by Kenneth Arrow (1953) to explain the function of securities in risk-bearing, were nowhere to be found in the real world until the broad development of the options and derivative-security markets. As we all know, it is now routine for financial engineers to use digital options and other Arrow-like derivative instruments in analyzing and creating new financial products.\textsuperscript{10} More broadly, Arrow's notion of "market completeness," long treated as a purely theoretical concept, is now seen as a (nearly) achievable long-run goal for real-world financial markets.\textsuperscript{11} Finance theory thus plays useful dual roles: as a \textit{positive} model for predicting the future direction of financial innovation, changes in financial markets and intermediaries, and regulatory design, and as a \textit{normative} model for identifying new product and service opportunities. Although framed in the positive context of "What \textit{will} the trends be?," the discussion to follow could apply equally in the normative context of "What \textit{should} the trends be?"

Just as the science of finance has helped shape the practice of finance, so practice in turn has helped shape the evolving theory. Financial innovation has generated a great variety of new institutions to serve financial functions, presumably more efficiently. Since theory is not institution based, those real-world innovations provide financial scientists a rich opportunity to understand the selective processes mapping institutions to functions. This strong interplay between research and practice is surely exemplified by this Congress where by my count there are roughly equal numbers of academics and practitioners presenting papers, including many individuals who qualify as both.

The view of the future of financial practices as elsewhere in the economic sphere is clouded with significant uncertainties. With this in mind, I will nevertheless try to apply finance theory, specifically the functional perspective, to talk about the possibilities for future trends in both financial products and services\textsuperscript{12} – giving examples from each of the four broad classes of customers for financial services – households, endowment institutions, non-financial firms, and governments.

\textsuperscript{10} For the theory of synthesis and production of Arrow securities from options and dynamic trading strategies, see Banz and Miller (1978), Breeden and Litzenberger (1978), Hakansson (1976) and Ross (1976b).

\textsuperscript{11} See Melnikov (1999, forthcoming) and Merton (1993a, 1995a) for the development of the financial innovation spiral as one model for the dynamic interplay between financial institutions and markets driving the evolution of financial instruments toward market completeness as the asymptotic long-run attractor.

\textsuperscript{12} See Claessens, Glaessner, and Klingebiel (2000) for an extensive survey on the impact of electronic and financial technologies on institutions and practices in the world financial system.
1.2 Financial Services for Households in the Future\textsuperscript{13}

As a result of major technological innovation and wide-spread deregulation, the household sector of users in the more fully developed financial systems have experienced a major secular trend of disaggregation...some call it disintermediation...of financial services. Households today are called upon to make a wide range of important and detailed financial decisions that they did not have to in the past. For example, in the United States, there is a strong trend away from defined-benefit corporate pension plans that require no management decisions by the employee toward defined-contribution plans that do. There are more than 7,000 mutual funds and a vast array of other investment products. Along with insurance products and liquidity assets, the household faces a daunting task to assemble these various components into a coherent effective lifetime financial plan.

Some see this trend continuing with existing products such as mutual funds being transported into technologically less-developed financial systems. Perhaps so, especially in the more immediate future, with the widespread growth of relatively inexpensive Internet access to financial "advice engines." However, the creation of all these alternatives combined with the deregulation that made them possible has consequences: Deep and wide-ranging disaggregation has left households with the responsibility for making important and technically complex micro financial decisions involving risk – such as detailed asset allocation and estimates of the optimal level of life-cycle saving for retirement – decisions that they had not had to make in the past, are not trained to make in the present, and are unlikely to execute efficiently in the future, even with attempts at education.

The availability of financial advice over the Internet at low cost may help to address some of the information-asymmetry problems for households with respect to commodity-like products for which the quality of performance promised is easily verified. However, the Internet does not solve the "principal-agent" problem with respect to more fundamental financial advice dispensed by an agent. That is why I believe that the future trend will shift toward more integrated financial products and services, which are easier to understand, more tailored toward individual profiles, and permit much more effective risk selection and control.

The integrated financial services in the impending future, unlike the disaggregated financial services of the recent past, will focus on the customer instead of the product as the prime unit of attention. That is, the service begins by helping the customer design a financial plan to determine his optimal life-cycle needs and then finds the products necessary to implement that integrated plan in a cost-efficient fashion. The past generation has seen explosive growth in asset management. Since 1974, mutual fund assets in the United States alone have increased 125-fold from $48 billion to around $6 trillion. In this time, the financial-service industry has made great strides

\textsuperscript{13} This section is an expanded version of Merton (1999).
in developing and improving portfolio-allocation and performance measurement. However, the central objective function employed, even in sophisticated practice, is still the same basic mean-variance efficient-frontier criterion developed by Markowitz (1952), Tobin (1958), and Sharpe (1964) in the 1950s and 1960s. This criterion, based on a static one-period model of maximizing the expected utility of end-of-period wealth, is simply not rich enough to capture the myriad of risk dimensions in a real-world lifetime financial plan.

The practical application of this status quo model is almost always limited to just the financial assets of the individual. Thus, in the models available to consumers today, there is no formal recognition of either the size or risk characteristics of human capital, which is the largest single asset for most people during much of their lifetime. In addition to taking into account the magnitude of human capital, advice models in the impending future should also capture the important element of its individual risk characteristics: a stockbroker, an automobile engineer, a baseball player, a surgeon, or a professor have very different risk profiles. The human capital of a stockbroker will surely be highly correlated with stock market returns. The human capital of the professor much less so. Without holding any equities among his financial assets, a stockbroker has a significant investment exposure to stock returns. Thus, between a stockbroker and a professor with the same total wealth and risk tolerance, the stockbroker should allocate a smaller part of his financial portfolio to equities. As we see, effective models of asset allocation cannot just focus on the expected levels of compensation, but must also consider its volatility and its correlation with other assets' returns.  

There are a number of other risks that are important to households besides the uncertainties about the future values of financial assets and about the returns to human capital. In addition to general inflation uncertainty, there is uncertainty about relative prices of individual categories of consumption goods such as local residential housing. There is even uncertainty about the menu of possible consumption goods that will be available in the future. There is uncertainty about one’s medical care needs and the age of death. There is uncertainty about one’s own tastes in the future, including the importance attached to bequests to transfer wealth to family and other heirs.

One particularly important intertemporal risk faced by households which is not captured in the traditional end-of-period wealth models of choice is uncertainty about the future investment opportunity set. That is, the unpredictable changes in the menu of expected returns and volatilities of returns

14 For an extension of the mean-variance one-period model to include human capital, see Mayers (1972). For intertemporal dynamic models of optimal consumption and portfolio selection that take into account uncertain human capital, see Bodie, Merton, and Samuelson (1992), Merton (1971, sections 7, 8, 1977b), and Williams (1978, 1979).
15 For an analytical development, see for examples Breeden (1979), Cox and Huang (1989), and Merton (1970, 1973b, 1977b, 1982).
Future Possibilities in Finance Theory and Finance Practice

available on investments in the future. To illustrate the point, consider the following choice question: Which would you rather have: $5 million or $10 million? The answer is obvious for all, take the $10 million, given all other variables are held fixed. However now consider that choice framed with further elaboration: Which would you rather have: $5 million in an environment in which the only investment available for the rest of time pays a risk-free real interest rate of 10 percent or $10 million in an environment in which the only investment available for the rest of time pays a risk-free real interest rate of 1 percent?

In a one-period model such as the Markowitz-Tobin one, the $10 million selection is still superior because one is presumed to consume all one's wealth at the end of the period and the future rate of return on investments are irrelevant. Note however that the $5 million selection can provide a $500,000 inflation-protected annual cash flow in perpetuity while the $10 million can only provide a $100,000 annual real cash flow in perpetuity. Thus, for anyone with a long enough future consumption horizon (approximately 10 years or longer in this case), the $5 million with a 10 percent interest is the better choice in terms of consumption standard of living.

Move from this simplified example to the general case when future investment rates on both risk-free and risky assets are uncertain. We see that for the household to maintain a stable consumption stream, it is necessary to plan its portfolio to hedge against unanticipated changes in interest rates. Thus, the household's portfolio is such that in future states of the world in which real interest rates are lower than expected, it has more wealth than expected and in states in which real interest rates are higher than expected, it accepts a lower than expected wealth because it doesn’t reduce its standard of living. The “natural” financial security to implement such hedging behavior is a long-term inflation-protected bond.

In addition to taking into account the various dimensions of risk, the household products and services of the future will be much more comprehensive and integrative. They will marry risk control and protection with optimal saving plans for lifetime consumption smoothing and bequests. To arrive at the necessary integrated lifetime consumption and asset-allocation decisions, more advanced financial models are required than have been used in past practice. The underlying analysis will have to combine the traditional efficient risk-return tradeoff for the tangible-wealth portfolio, accounting for human-capital risks and returns, hedging the risks of future reinvestment rates and relative consumption goods prices, incorporating mortality and other traditional insurance risks as well as income and estate tax risks.

Exemplifying my theme of “good practice evolving toward good theory” as a guide to predicting future financial innovation, the basic models for implementation can be found in the rich body of published academic research on optimal lifetime consumption and portfolio selection and intertemporal capi-
tal asset pricing that has developed since the classic Markowitz–Tobin–Sharpe work.17

In the new environment of these integrated retail products, success for financial-service providers will require much more than simply developing these decision models and performing an advisory role. They should also expect to undertake a principal intermediation role as either issuer or guarantor to create financial instruments that eliminate the "short-fall" or "basis" risk for households. One important category for such intermediation is hedging "targeted" expenditures, ones which are almost surely going to be made and the magnitude of which are not likely to depend on changes in the household's overall standard of living. A prime example is tuition, room and board for a child's college education. In the current investment product environment, the household must take the "basis" risk between the amount saved to provide for that education and the subsequent investment performance from those savings and the uncertain inflation rate for college tuition and housing. Basic finance theory suggests that a more efficient approach would be for an intermediary to issue to the household a contract for four years tuition, room and board delivered at a specified future date in return for a fixed price (which can be financed over time, if necessary, just like a car or house is). The intermediary would then bear the basis risk instead of the household.

To serve the households in the future efficiently, providers will find it advantageous to integrate the various risk-management products. To implement this integration will require bundling of some products that cut across traditional provider institutions and the unbundling of others. For example, by bundling long-term care insurance with retirement annuities, there is a prospect for an efficiency gain by reducing traditional selection bias problems for the mortality component. An unbundling of the accumulation for retirement in a pension fund from the life insurance feature of survivor benefits from that fund can provide a more efficient meeting of these two financial needs in different parts of the household life cycle. A third example would be comprehensive value insurance for the household residence, covering value risk from market price changes, fire and natural disasters.

Each of these integrated risk products combines traditional insurance risks with market risks. Effective implementation will require not only greater regulatory flexibility among banks, securities firms and insurance companies, but also that the rigid intellectual barriers between research in the fields of finance and actuarial science become more permeable and flexible as well. This permeability is already underway reflecting changes in real-world practice where insurance functions are extending well beyond traditional actuarial lines to include wide-ranging guarantees of financial performance by both private-sector institutions and governments. Examples are guaranteed income contracts, deposit insurance, pension-benefit guarantees, and guarantees of loan and other

17 See Merton (1992) for an extensive bibliography.
contractual obligations.\(^{18}\) The mathematical tools developed to evaluate risks of "nature" (mortality, weather, and fire) are not adequate to analyze those financial guarantees. Instead, the prototype insurance instrument for financial risks is the put option. As we have seen, the mathematical tools for option pricing are found in the finance literature evolving from Bachelier's thesis. Just as insurance is "moving" into the domain of finance, so finance is moving into the realm of insurance. Although only just at its beginnings, there is an effort for a major institutional shift to move much of the catastrophic risk insurance exposures outside insurance companies (or governments) and instead have them borne directly in the capital markets.\(^{19}\) Thus, cutting-edge research and practice in the future in either field will require a mathematical and substantive knowledge base that spans both fields.

The inadequacies of the current static model have been masked by the compound claim that classic one-period diversification across stocks handles the static risk of investing and that "time diversification" handles the intertemporal dynamic aspects of that risk. The false but oft-claimed belief that investing in stocks become less risky as the investment horizon becomes longer seemingly offers a practical short cut to addressing the multi-year investment and consumption problems of households in the real world. Unfortunately, that view of equities' risk is just plain flawed.\(^{20}\) A decade-long bull market may have kept the errors of this approach to multi-period investing from becoming apparent. But as we all know too well, that cannot continue indefinitely.

1.3 Production of Integrated Financial Products in the Impending Future

Production of the new brand of integrated, customized financial instruments will be made economically feasible by applying already existing financial pricing and hedging technology that permits the construction of custom products at "assembly-line" levels of cost.

Paradoxically, making the products more user-friendly and simpler to understand for customers will create considerably more complexity for their producers. The good news for the producers is this greater complexity will also

\(^{18}\) For discussion, see Cummins (1988), Kraus and Ross (1982), Merton (1977a), Merton and Bodie (1992), and Mody (1996).

\(^{19}\) For discussion and analysis, see Cummins and Geman (1995), Harrington, Mann, and Niehaus (1995), and Hayes, Cole, and Meiselman (1993).

\(^{20}\) Bodie (1995) makes this point quite dramatically. He shows that the premium for insuring against a shortfall in performance of stocks versus bonds is actually an increasing function of the time horizon over which the insurance is in force instead of a decreasing one, which would be expected with declining risk. A similar fallacy involving the virtues of investing to maximize the geometric mean return as the "dominating" strategy for investors with long horizons was addressed earlier by Samuelson (1971, 1972, 1979).
make reverse engineering and “product knockoffs” by second-movers more difficult and thereby, protect margins and create franchise values for innovating firms. Hence, financial-engineering creativity and the technological and transactional bases to implement that creativity, reliably and cost-effectively, are likely to become a central competitive element in the industry.

A key element for the success of these highly integrated, user-friendly products in the household sector will be to find effective organizational structures for ensuring product performance: that is, that the contingent payments promised by the products are actually paid by the issuing institution. The need for assurances on contract performance is likely to stimulate further development of the financial-guarantee business for financial institutions. It is encouraging to note that currently, credit risk analysis and credit-derivative contracting technologies are among the fastest growing areas of development in financial services. In general, the greater complexity in products combined with the greater need for contract performance will require more elaborate and highly quantitative risk-management systems within financial-service firms and a parallel need for more sophisticated approaches to external oversight. 21

All of these will significantly change the role of the mutual fund from a direct retail customer product to an intermediate or “building block” product embedded in the more integrated products used to implement the consumer’s financial plan. The “fund of funds” is an early, crude example. The position and function of the fund in the future will be much like that of individual traded firms today, with portfolio managers, like today’s CEOs, selling their stories of superior performance to professional fund analysts, who then make recommendations to retail “assemblers.” As we know, commercial marketing is very different from retail marketing, and some fund institutions may have difficulty making the transition. How and what institutional forms will perform the retail assembly and distribution functions is not clear. It does seem, however, that a fully vertically integrated fund complex of the usual kind that limits its front-end assembly operation to using only its own funds and products will be at a very distinct disadvantage, because it will not have the breadth of first-quality “building blocks” to assemble the best integrated products.

1.4 Financial Services for Endowment Institutions in the Future

Much the same story to the one on products and advice for households applies to serving endowment institutions. There are, however, significant enough substantive differences between the two to warrant separate attention here. 22

The standard approach to the management of endowment today is to treat it as if it were the only asset of the institution. As a result, investment advice

21 See Merton and Bodie (1992) for a discussion of the difference between customer-held and investor-held liabilities in terms of impact on a firm’s business and the various approaches to managing default risk.

22 This section is a revised version of a part of Merton (1993b).
and products for endowments are focused on achieving a mean-variance efficient portfolio with the appropriate level of risk and a prudent withdrawal or dividend rate. Thus, except for choosing the particular point on the risk-return frontier, the investment advice varies little across institutions. Of course, endowment is almost never the only asset of an institution. Specifically, institutions such as universities have a variety of other assets, both tangible and intangible, which are important sources of cash flow. In addition to tuition, there are gifts, bequests, publishing and other business income, and public and private-sector grants for research. Taking explicit account of those assets and their risk and return characteristics can cause the characteristics of the optimal endowment portfolio to change substantially. Although to be concrete the following discussion uses the context of a university, the same principles of analysis would apply to managing the endowments of museums, foundations, and religious organizations.

A procedure for selecting the investments for the endowment portfolio that takes account of non-endowment assets includes the following steps:

1. Estimate the market value that each of the cash flow sources would have if it were a traded asset. Also determine the investment risk characteristics that each of those assets would have as a traded asset.
2. Compute the total wealth or net worth of the university by adding the capitalized values of all the cash flow sources to the value of the endowment.
3. Determine the optimal portfolio allocation among traded assets, using the university’s total wealth as a base. That is, treat both endowment and cash flow-source assets as if they could be traded.
4. Using the risk characteristics determined in step 1, estimate the “implicit” investment in each traded-asset category that the university has as the result of owning the non-endowment (cash flow-source) assets. Subtract those implicit investment amounts from the optimal portfolio allocations computed in step 3, to determine the optimal “explicit” investment in each traded asset, which is the actual optimal investment allocation for the endowment portfolio.

As a simple illustration, consider a university with $400 million in endowment assets and a single non-endowment cash flow source. Suppose that the only traded assets are stocks and cash. Suppose further that the university estimates in step 1 that the capitalized value of the cash flow source is $200 million, with risk characteristics equivalent to holding $100 million in stock and $100 million in cash. Thus, the total wealth of the university in step 2 is $(400 + 200 =) \$600 million. Suppose that from standard portfolio-selection techniques, the optimal fractional allocation in step 3 is .6 in stocks and .4 in cash, or $360 million and $240 million, respectively. From the hypothesized risk characteristics in step 1, the university already has an (implicit) investment of $100 million in stocks from its non-endowment cash flow source.
Therefore, we have in step 4 that the optimal amount for the endowment portfolio to invest in stocks is $260 million, the difference between the $360 million optimal total investment in stocks and the $100 million implicit part. Similarly, the optimal amount of endowment invested in cash equals $(240 - 100 =) \$140$ million.

The effect on the composition of the optimal endowment portfolio induced by differences in the size of non-endowment assets can be decomposed into two parts: the wealth effect and the substitution effect. To illustrate the wealth effect, consider two universities with identical preference functions and the same size endowments, but one has non-endowment assets and the other does not. If, as is perhaps reasonable to suppose, the preference function common to each exhibits decreasing absolute risk aversion, then the university with the non-endowment assets (and hence larger net worth) will prefer to have a larger total investment in risky assets. So a university with a $400 million endowment as its only asset would be expected to choose a dollar exposure to stocks that is smaller than the $360 million chosen in our simple example by a university with the same size endowment and a non-endowment asset valued at $200 million. Such behavior is consistent with the belief that wealthier universities can “afford” to take larger risks with their investments. Thus, if the average risk of the non-endowment assets is the same as the risk of the endowment-only university’s portfolio, then universities with such assets will optimally invest more of its endowment in risky assets.

The substitution effect on the endowment portfolio is caused by the substitution of non-endowment asset holdings for endowment asset holdings. To illustrate, consider again our simple example of a university with a $400 million endowment and a $200 million non-endowment asset. However, suppose that the risk characteristics of the asset are changed so that it is equivalent to holding $200 million in stocks and no cash. Now, in step 4, the optimal amount for the endowment portfolio to invest in stocks is $160 million, the difference between the $360 million optimal total investment in stocks and the $200 million implicit part represented by the non-endowment asset. The optimal amount of endowment invested in cash rises to $(240 - 0 =) \$240$ million. If instead the risk characteristics of the asset had changed in the other direction to an equivalent holding of $0$ in stocks and $200$ million in cash, the optimal composition of the endowment portfolio would be $(360 - 0 =) \$360$ million in stocks and $(240 - 200 =) \$40$ million in cash.

Note that the changes in risk characteristics do not change the optimal deployment of total net worth ($360$ million in stocks and $240$ million in cash). However, the non-endowment assets are not carried in the endowment portfolio. Hence, different risk characteristics for those assets do change the amount of substitution they provide for stocks and cash in the endowment portfolio. Thus, the composition of the endowment portfolio will be affected in both the scale and fractional allocations among assets.

With the basic concept of the substitution effect established, we now apply it in some examples to illustrate its implications for endowment investment.
policy. Consider a university that on a regular basis receives donations from alums. Clearly, the cash flows from future contributions are an asset of the university, albeit an intangible one. Suppose that the actual amount of gift giving is known to be quite sensitive to the performance of the general stock market. That is, when the market does well, gifts are high; when it does poorly, gifts are low. Through this gift-giving process, the university thus has a "shadow" investment in the stock market. Hence, all else the same, it should hold a smaller portion of its endowment in stocks than would another university with smaller amounts of such market-sensitive gift giving.

The same principle applies to more specific asset classes. If an important part of gifts to a school that specializes in science and engineering comes from entrepreneur alums, then the school de facto has a large investment in venture capital and high-tech companies, and it should therefore invest less of its endowment funds in those areas. Indeed, if a donor is expected to give a large block of a particular stock, then the optimal explicit holding of that stock in the endowment can be negative. Of course, an actual short position may not be truly optimal if such short sales offend the donor. That the school should optimally invest less of its endowment in the science and technology areas where its faculty and students have special expertise may seem a bit paradoxical. But the paradox is resolved by the principle of diversification once the endowment is recognized as representing only a part of the assets of the university.

The same analysis and conclusion apply if alum wealth concentrations are in a different class of assets, such as real estate instead of shares of stock. Moreover, much the same story also applies if we were to change the example by substituting government and corporate grants for private donations and gift giving as the sources of cash flows. That is, the magnitudes of such grant support for engineering and applied science may well be positively correlated with the financial performance of companies in high-tech industries. If so, then the prospect of future cash flows to the university from the grants creates a shadow investment in those companies.

The focus of our analysis is on optimal asset allocation for the endowment portfolio. However, the nature and size of a university's non-endowment assets significantly influence optimal policy for spending endowment. For a given overall expenditure rate as a fraction of the university's total net worth, the optimal spending rate out of endowment will vary, depending on the fraction of net worth represented by non-endowment assets, the expected growth rate of cash flows generated by those assets, and capitalization rates. Hence, neglecting those other assets will generally bias the optimal expenditure policy for endowment.

In addition to taking account of non-endowment assets, our perspective on asset allocation differs from the norm because it takes account of the uncertainty surrounding the costs of the various activities such as education, research, and knowledge storage that define the purpose of the university. The breakdown of activities can of course be considerably more refined. For in-
stance, one activity could be the education of a full-tuition-paying undergraduate, and a second could be the education of an undergraduate who receives financial aid. The unit (net) cost of the former is the unit cost of providing the education less the tuition received, and the unit cost of the latter is the cost plus the financial aid given. An important function of endowment investments is to hedge against unanticipated changes in the costs of university activities.

Consider, for example, the decision as to how much (if any) of the university's endowment to invest in local residential real estate. From a standard mean-variance efficiency analysis, it is unlikely that any material portion of the endowment should be invested in this asset class. However, consider the cost structure faced by the university for providing teaching and research. Perhaps the single largest component is faculty salaries. Universities of the same type and quality compete for faculty from the same talent pools. To be competitive, they must offer a similar standard of living not just similar salaries. Probably the largest part of the differences among universities in the cost of providing this same standard of living is local housing costs. The university that invests in local residential housing hedges itself against this future cost uncertainty by acquiring an asset whose value is higher than expected when the differential cost of faculty salaries is higher than expected. This same asset may also provide a hedge against unanticipated higher costs of off-campus housing for students that would in turn require more financial aid if the university is to compete for the best students. The prescribed targeted investment in very specific real estate assets to hedge against an unanticipated rise in a particular university's costs of faculty salaries and student aid should not be confused with the often-stated (but empirically questionable) assertion that investments in real estate generally are a good hedge against general inflation. Inflation-indexed government bonds, such as Treasury Inflation-Protected Securities are the efficient instruments for that objective.

Similar arguments could be used to justify targeted investment of endowment in various commodities such as oil and natural gas to hedge against unanticipated changes in energy costs. Uncertainty about those costs is especially significant for universities located in extreme climates and for universities with major laboratories and medical facilities that consume large quantities of energy.

In the same fashion, the hedging analysis of whether selling tuition forward is risk reducing for the university cannot be made without understanding the interactive risk structures among both assets and liabilities.

In sum, one cannot properly evaluate the financial decisions of the institution without considering the risks and returns for the total wealth of the institution. Advice and products of the future will adopt this comprehensive perspective.
1.5 Financial Services for Non-Financial Firms in the Future

The optimal management of corporate pension assets follows closely the endowment-model prescription. Indeed, taking into account risks on both sides of the balance sheet is fundamental to providing effective financial services to non-financial firms in general. Enterprise risk management is one term for such a unified approach. The movement from tactical to strategic application of currency, interest rate, commodity, and equities hedging is already underway. The next major step is to integrate operational, market, credit and traditional insurance risk management. To implement such integration requires connecting the decisions on operations, on the use of contractual agreements to hedge targeted exposures, and on the choice of capital structure.23

A particularly promising area for further development is the management of factor risks, particularly labor. Firms can be leveraged with their explicit and implicit labor contracts in parallel fashion to more traditional financial leverage with debt. Both temporary-employee firms and consulting firms serve the function of “labor intermediaries” that allow more efficient management of the risks for both those who supply labor and those who demand it. Their rapid growth, both in the United States and abroad, is probably a good measure of the significance of these factor risks to enterprises. The point again, integrated risk management for firms.

1.6 Government and Financial Services in the Future

A consequence of all this prospective technological change will be the need for greater analytical understanding of valuation and risk management by users, producers, and regulators of financial services. Furthermore, improvements in these products and services will not be effectively realized without concurrent changes in the financial “infrastructure”—the institutional interfaces between intermediaries and financial markets, regulatory practices, organization of trading, clearing, settlement, other back-office facilities, and management-information systems. To perform its functions as both user and overseer of the financial system, government in the future will need to both understand and make use of new financial technology.

Government also serves a critical role as financial intermediary. We already see a major effort underway nearly world wide with respect to restructuring the intermediary roles played by government and the private sector in providing pensions benefits in the retirement segment of the life cycle. Even if the responsibility for retirement benefits shifts largely to the private sector, government must still assess the risks it is underwriting.24 These can be explicit guarantees as in the case of corporate pension insurance (in the United States

23 For a compact but comprehensive non-technical overview of modern-day integrated risk management for the firm, see Meulbroek (2000).
through the Pension Guarantee Insurance Corporation) and implicit ones in its role as the "guarantor of last resort" for a systemic shortfall in benefits that affects an entire generation of retirees. Government is almost surely the only viable provider in adequate size of long-dated, default-free inflation-indexed debt which can be used by private-sector financial intermediaries as the prime hedging asset for issuing life retirement annuity products that are protected against inflation. With all the current discussion in the United States about retiring large portions of the government debt, policy should ensure that an adequate supply of such debt is available for this intermediation purpose.

Application of new financial technology is critical to the future provision of risk-accounting standards, designing monetary and fiscal policies, implementing stabilization programs, and financial-system regulation. Many experts on monetary policy have expressed serious concerns about how financial innovation has been eroding the ability of central banks to conduct monetary policy through traditional channels. Much the same concern has been expressed about the effect of financial innovation on some fiscal and regulatory policies as well. Such concerns are manifestly valid to the extent that the effectiveness of these traditional channels rests on large frictions of transaction costs, institutional rigidities, and institutionally defined regulations. Indeed, policymakers who continue to depend on such channel frictions are effectively speculating against the long-run trend of declining transaction costs and growing flexibility in institutional design. However, financial innovation and improved technology also opens new opportunities for government policymakers to perform their financial functions more effectively.

To illustrate how governments of the future might use modern financial technology to pursue their policies more effectively, we present three examples: 1) "automatic" open-market operations for stabilization of interest rates or currencies; 2) providing international diversification to the domestic population under capital controls; 3) measuring and controlling country risks.

In the spring and summer of 1990, the German government issued a sizeable private placement of ten-year maturity Schuldentschein bonds with put-option provisions. They were just like standard government bonds, except they had the feature that the holders can put them back to the government for a fixed price.

26 The Federal Reserve has for some time used implied volatilities derived from prices of traded options on government bonds as an estimate of the market's current assessment of future interest rate uncertainty. See Nasar (1992). Other areas to consider applying option theory are in evaluating physical capital project alternatives (Trigeorgis, 1993) and education and training policy alternatives for human capital development (S.J. Merton, 1992).
27 For expanded discussion of this example, see Merton (1995b, 1997).
28 This example is taken from Merton (1990).
29 See also Merton (1999).
By issuing those bonds, the German government in effect introduced a pre-programmed dynamic stabilization policy. How is that? Suppose that it had issued a standard ten-year bond instead. Suppose further that afterwards interest rates start to rise, and therefore, bond prices fall. Normal ten-year bonds would fall in price in line with interest rate rises. But what happens to the bonds with the put option? The put bonds will not decline as much as the normal ten-year ones. Furthermore, the rate of decline in the put bonds becomes less and less, until they cease to decline at all. At that point, they will actually begin to behave just like a short-term money instrument. In terms of ‘hedge ratios’ or exposures relative to a normal ten-year bond, what is happening?

To answer, consider a single-factor interest-rate model with dynamics described by a diffusion process. If $B(t)$ denotes the price of a standard ten-year bond, then we can express the price of the puttable bonds as $F(B, t)$, where $F$ is derived from a replicating trading strategy using contingent-claims analysis. From that analysis, the puttable bonds are economically equivalent in exposure to a portfolio of $\partial F/\partial B$ units of the standard ten-year bond and $[F - B\partial F/\partial B]$ invested in the shortest-maturity Treasury bill. It is straightforward to show that $0 \leq \partial F/\partial B \leq 1$ and that $F$ is convex which implies $\partial^2 F/\partial B^2 > 0$. It follows that as the price $B$ falls, the equivalent number of units of $B$ represented by the puttable bonds, $\partial F/\partial B$, also falls.

In effect, because of the puts, the hedge ratio or equivalent number of ten-year bonds for each put bond gets smaller and smaller as the price of the ten-year bond falls. It is thus as if government were repurchasing normal bonds as in a regular open-market operation. In economic effect, the government is taking the interest-rate risk back from holders as if it was purchasing bonds even though it had not actually done so. If instead interest rates were to fall and bond prices rise, then the puts would become more out-of-the-money, the equivalent number of ten-year bonds per put bond rises, and the outstanding bond exposure held by investors would increase, which is effectively the same as the government issuing more bonds. Note that the decrease or increase in the equivalent bond exposure outstanding takes place immediately as interest rates change, without requiring that the bonds actually be put back to the government. So, in effect by issuing those put bonds, the German government put into place an automatic stabilizer to the extent that ‘stabilization’ means to ‘lean’ against market movements and buy bonds when bond price goes down and sell bonds when they go up. That is, the put bonds function as the equivalent of a dynamic, ‘open-market,’ trading operation without any need for actual transactions.

In comparison to traditional open-market activity, the put-option-bond automatic stabilizer will work well even over weekends, over non-trading days, and over crashes, especially in an environment with trading going on around the world because the central bank does not have to be on the scene to do the open-market operations. It automatically ‘kicks in’ as soon as events occur because it is built into the structure of the securities.
Turning from stabilization policy using open-market operations, we next examine the use of modern financial technology to reduce an unintended risk cost imposed on the domestic population from implementing capital controls. Numerous empirical studies of stock market returns have documented the gains in diversification from investing internationally. By diversifying across the world stock markets, there is significant improvement in the efficient frontier of risk versus expected return. As we know, the last decade has seen widespread implementation of such international diversification among investors in the large developed countries with the major stock markets. However, international diversification has not yet evolved in many smaller countries where indeed it may be relatively more important.

A major barrier to foreign stock market investment by citizens of some of these countries is capital controls, imposed by their governments to prevent flight of domestic capital. A common rationale for such restrictions is that they reduce the risk that the local economy will have inadequate domestic investment to promote growth. Another potential barrier is that the transaction cost paid by foreign investors to buy shares directly in these domestic stock markets can be so large that it offsets any diversification benefits that would otherwise accrue. The cost in lost welfare from less-efficient diversification affects both large-country and small-country citizens. However, the per capita magnitude of the cost is much larger for the latter, since the potential gains from international diversification are greatest for citizens of the smaller countries with domestic economies that are by necessity less well diversified. An additional cost may be that domestic physical investment is driven to become more diversified than would otherwise be efficient according to the principle of comparative advantage.

Of course, one (and perhaps the best) solution is to eliminate capital flow restrictions and open capital markets. However, with the capital controls taken as a given, a constrained solution exists that separates the capital-flow effects of investment from its risk-sharing aspects. Suppose that small-country domestic investors who already own the domestic equity (perhaps through domestic mutual funds or financial intermediaries) were to enter into “swap” agreements with large foreign investors. In the swap, the total return per dollar on the small country’s stock market is exchanged annually for the total return per dollar on a market-value weighted-average of the world stock markets. This exchange of returns could be in a common currency, dollars, as described or adjusted to different currencies along similar lines to currency-swap agreements. The magnitudes of the dollar exchanges are determined by the “notional” or principal amount of the swap to which per dollar return differences apply. As is the usual case with swaps, there is no initial payment by either party to the other for entering the agreement.

Without pursuing further the details of implementation, we see that the swap agreement effectively transfers the risk of the small-country stock market to foreign investors and provides the domestic investors with the risk-return pattern of a well-diversified world portfolio. Since there are no initial payments
between parties, there are no initial capital flows in or out of the country. Subsequent payments which may be either inflows or outflows involve only the difference between the returns on the two stock market indices, and no "principal" amounts flow. For example, on a notional or principal amount of $1 billion, if, ex post, the world stock market earns 10 percent and the small-country market earns 12 percent, there is only a flow of (.12 - .10) x $1 billion or $20 million out of the country. Furthermore, the small-country investors make net payments out precisely when they can "best" afford it: namely, when their local market has outperformed the world markets. In those years in which the domestic market underperforms the world stock markets, the swap generates net cash flows into the country to its domestic investors. Hence, in our hypothetical example, if the small-country market earns 8 percent and the world stock market earns 11 percent, then domestic investors receive (.11 - .08) x $1 billion = $30 million, a net cash inflow for the small country. Moreover, with this swap arrangement, trading and ownership of actual shares remain with domestic investors.

Foreign investors also benefit from the swap by avoiding the costs of trading in individual securities in the local markets and by not having the problems of corporate control issues that arise when foreigners acquire large ownership positions in domestic companies. Unlike standard cash investments in equities or debt, the default or expropriation exposure of foreign investors is limited to the difference in returns instead of the total gross return plus principal (in our example, $20 million versus $1.12 billion).

The potential exposure of foreign investors to manipulation by local investors is probably less for the swap than for direct transactions in individual stocks. It is more difficult to manipulate a broad market index than the price of a single stock. Even if settlement intervals for swaps are standardized at six months or one year, the calendar settlement dates will differ for each swap, depending upon the date of its initiation. Hence, with some swaps being settled every day, manipulators would have to keep the prices of shares permanently low to succeed. Furthermore, with the settlement terms of swaps based on the per-period rate of return, an artificially low price (and low rate of return) for settlement this year will induce an artificially high rate of return for settlement next year. Thus, gains from manipulation in the first period are given back in the second, unless the price can be kept low over the entire life of the swap. Since typical swap contract maturities range from two to ten years (with semi-annual or annual settlements), this would be difficult to achieve.

Note that this derivative-security innovation is not designed to circumvent the stated objective of the capital-control regulation, to prevent domestic capital flight. Instead, it is designed to eliminate (or at least reduce) the unintended and undesirable "side effects" of this policy on efficient risk bearing and diversification. Whether or not this proposed solution using a swap turns out to be an effective real-world solution is not the central point of the exercise here. Rather, it is to demonstrate how a modern financial technological
innovation by government could help reduce the social cost of using “blunt” traditional policy tools that affect a number of countries around the world.

A similar but considerably broader prospective application of modern financial technology by government is the measurement and management of country risk. That is, we ask “How do we explain different countries’ relative performance and the variations in performance across regions and what to do about it?”

That question prompts another one: How much of what we observe ex post as differences in performance is a consequence of ex ante different risk profiles versus different management and government policy decisions? For instance, Taiwan is heavily into electronics but produces no automobiles. More generally, few countries, if any, are well diversified when measured against the world market portfolio, the theoretically best-diversified portfolio if all assets including human capital, were traded or could be hedged and there were no transaction costs.

A non-traditional approach to address the performance issue and its implication for evaluating policy is to apply the technology of a well-studied problem in risk and performance measurement for investment management and financial firms. This is the problem of configuring all the decomposition and reintegration of risk-factor exposures that must be determined within a financial institution before the aggregate risk measures such as value-at-risk (VAR) can be applied. I believe that this technology, if properly adapted, can be used to measure country risk exposures.

In practice, measuring the differences in country exposures is not a simple task since many asset classes are not traded at all. But this is structurally the same problem faced in the risk measurement of non-traded assets and liabilities in financial institutions. In short, it is like the challenges of extending the VAR and stress-testing concepts to include the domain of non-traded assets and liabilities. But as with the application to financial institutions, I see this as a tough engineering problem, not one of new science... we know how to approach it in principle and what we need to model, but actually doing it is the challenge.

As with conventional private-sector applications, the country risk exposures give us important information about the dynamics of future changes that cannot be inferred from the standard “country” accounting statements, either the country balance sheet or the country income or flow-of-funds statements. That is, information not extractable from an accurate listing of the value of assets including foreign reserves, or from the trade flows or the capital flows, even if they are all mark-to-market numbers.

As we discover with more conventional applications of risk management systems, once we can measure the risk exposures we have, it is difficult to resist exploring whether we could improve economic efficiency and risk sharing by changing those exposures. Again, take the example of Taiwan. Suppose it decided to try to align its risk exposures more with the world portfolio. In
the past, that might lead to an industrial policy to develop an automobile industry... a truly inefficient solution!

However, as another application of the ubiquitous swap contract, it is now feasible to separate the risk exposure decisions from the investment decisions. Instead of physically building a new industry, we can imagine Taiwan implementing its risk policy by entering into swap contracts in which it is a payer of the returns on a world electronics portfolio and a receiver of the returns on a world automobile portfolio.

Would such a swap be feasible? It is certainly structurally attractive. On the ability to pay, Taiwan is a net payer when electronics outperforms autos and a net receiver when electronics underperforms. There is no moral hazard or major asymmetric information problem for the country's counterparts, because payments are not based on country-specific performance in an industry, but are instead based on its global performance.

For the same reason, it avoids the political economic issue that the country's government can be accused of "giving away" its best assets as sometimes happens when foreigners buy the shares of its industries, because the country gets to keep its "alpha." Expropriation risk is also minimized, both because there is no principal exposure and because (returning to the first structural point) the likely ability to pay is aligned with the liability. Finally, while the useful implementation of such a swap obviously requires a large-size market, there are natural counterparts: other countries seeking alignment of their risk exposures.

These points seem to mitigate the usual incentive and information asymmetry problems for transactions with sovereigns. The technical problems of building a set of surrogate portfolios to use as benchmarks for risk measurement and contract specification are well understood. Initially at least, using mixtures of traded indices as the underlying asset for swap purposes would make the liquidity much better and the settlement mechanics easier. Contract credit risk is important but here too we know a lot about designing solutions, whether by a combination of mark-to-market collateral, purchase of private-sector performance guarantees, or efforts involving government and quasi-government institutional guarantees.

While the benefits of country-risk management systems and the associated markets would be expected to accrue to all, those in smaller countries with developing financial systems have the greater potential to benefit. With more concentrated investment opportunities, they should gain disproportionately from developing global access for capital and, perhaps more importantly, from more efficient allocation of risk.

Moreover, if they design their financial system using the most-up-to-date financial technology, these countries can "leap-frog" existing systems in terms of efficiency. In so doing, they can dramatically reduce the cost of investment capital and thereby materially increase national wealth.

With the developed countries, Japan and EMU Europe in particular, and the emerging ones both working on major changes in their financial systems,
this may be an especially opportune time to explore country risk management. It is certainly an opportune time to be a finance professional, financial engineer, or financial architect.

As we take stock on this centennial of Bachelier's thesis, the influence of that work on 20th Century science and practice of finance is unmistakable. Its impact has been global and truly extraordinary. I feel secure with the forecast that when reviewed again at its bicentennial, Bachelier's influence and impact on both 21st Century finance science and finance practice will be even more extraordinary.

References


