The global financial crisis of 2007 and 2008 (Rajan, 2010; Roubini and Mihm, 2010; Shiller, 2008) provides an opportunity for all finance professionals to introspect about the changes that need to be made in their discipline. Economists in general and financial economists in particular have come in for a great deal of criticism after the crisis. One well-known book about the crisis was entitled “ECONned” (Smith, 2010), while a widely read columnist talked about the “Formula that Killed Wall Street” (Salmon, 2009).

It is difficult to deny that there were serious problems with finance as it was practised in the years before the crisis. Whether this was only a gap between theory and practice or whether there are fundamental problems in finance theory itself is a more difficult question. Even if the problems are only with the practice and not with the theory, finance academics must revisit how finance is taught so that these problems do not recur. If there are problems in finance theory itself, then finance academics must reflect on the directions that finance research should take to redress these problems.

This paper is the result of my own introspection about these issues. A lot of it has to do with how finance ought to be taught, but a significant part is also about how finance theory needs to change by drawing on insights from other disciplines. The paper first discusses what the crisis taught us about the 3 P’s of finance – individuals’ preferences, their assessment of probabilities, and the behaviour of market prices. This is followed by a discussion on the nature of changes that need to be made in the models that are used in modern finance and the need for finance to integrate insights from other source disciplines.

PREFERENCES, PROBABILITIES AND PRICES

How the Crisis Changed the 3 P’s

It is the interactions among the 3 P’s - preferences, probabilities, and prices - that give modern financial economics its richness and depth (Lo, 1999; 2004):

- Preferences (especially risk preferences) determine the choice between investment alternatives after their future returns have been estimated. Two investors who agree on the probability distributions of returns from two assets might still make
different choices because they have different levels of risk aversion.

- **Probabilities** enter the picture because the future cash flows of various investment alternatives are uncertain. In the case of deciding to switch from stock A to stock B, the investor has to make some judgements about the probability distribution of the earnings (or dividends) of the two companies. For some asset pricing models, the probability distribution can be summarized by its mean and variance, but in general the entire probability distribution is important.

- **Prices** represent investment opportunities available to each investor. For example, for an investor contemplating a switch from stock A to stock B, the price of A determines how much money she will get by selling the stock, while the price of B determines how many shares she can buy with that money. Prices are also the (equilibrium) outcome of the investment decisions of all the investors. Thus asset prices can be thought of as the starting point of the decision making process; equally, they can be viewed as the end result of the process. In some reductionist models, prices are all that we need to know because they impound all other relevant information.

The global financial crisis has given us sufficient reason to rethink many of our ideas about each of these three:

- If we consider preferences, it is necessary to reconsider whether risk aversion is stable (time invariant) or whether it changes during the course of a crisis.
- Turning to probabilities, we need to ask whether investors have sufficient relevant data to estimate statistical parameters with reasonable accuracy. Can investors be assumed to have homogeneous expectations or is the data so sparse that probabilities are inherently subjective and heterogeneous?
- When it comes to prices, the process of price formation needs to be re-examined. In particular, we need to pay attention to market microstructure theories in which traded prices are the outcome of a complex interaction of quotes and orders, and do not necessarily represent equilibrium prices at every instant of time.

In the following sections, I elaborate on these ideas regarding preferences, probabilities, and prices. Incorporating these ideas into finance teaching will hopefully lead to a richer and more nuanced understanding of the subject.

### Preferences: Risk Aversion may be Environment Contingent

Time varying risk aversion is usually frowned upon as a desperate attempt to reconcile a struggling theory with unfavourable evidence. It is true that time varying risk aversion can be abused to explain away many anomalies. For example, a stock market bubble can be explained away as a temporary decline in risk aversion; similarly, a temporary rise in risk aversion can explain away a market crash. Arbitrary time variation in risk aversion can thus ensure that many theories can never be falsified making them devoid of testable implications.

Many finance theorists have, however, gone to the other extreme of thinking of the risk aversion coefficient as an innate characteristic of a human being – almost as if there were a gene for risk aversion. It is true that there is a genetic element in risk aversion, but studies based on identical twins show that genetics explain only 20-30 per cent of the variation in risk aversion (Cesarini et al, 2009; 2010). Evolutionary biology provides a theoretical argument why a large part of risk aversion may not be purely genetic. Bell (2007) puts it very succinctly: “If a trait is heritable and linked to survival or reproductive success, then evolutionary theory tells us that variation will eventually disappear from the population.”

Evolutionary biologists explain risk aversion as resulting from different life-history strategies adopted in response to ecological pressures. In this sense, risk aversion is not so much an immutable trait as a (life-historical) strategic choice – for example, individuals with high future expectations (of evolutionary fitness) become more risk-averse than individuals with low expectations (Wolf, et al, 2007; Buss, 2009; Heilbronner et al, 2008). A neuroscience perspective views “preferences as transient state variables that ensure survival and reproduction” (Camerer, Loewenstein and Prelec, 2005).

This perspective of risk preferences being moulded by life experiences has received considerable attention in the behavioural sciences as well. Two decades ago, Sitkin and Pablo (1992) introduced the concept of risk propensity which goes beyond a general dispositional risk orientation to include two other important elements that are rooted in life history. First is the notion of inertia: “decision makers who have been risk averse in the past will tend to continue in their cautious ways, whereas previously risk-seeking decision makers will continue
to be more adventurous.” Second is the idea of outcome history: “successful risk-averse decision makers will become increasingly risk-averse, and successful risk-seeking decision makers will become increasingly risk-seeking.”

If the degree of risk aversion is a strategic choice – the choice of a strategy for solving recurrent adaptive problems – then drastic shifts in the environment – the distribution of such adaptive problems – could conceivably cause change in these strategies. Booms and busts could then lead to (evolutionarily) rational changes in aggregate risk aversion. To push the analogy with Heilbronner et al (2008) to the level of caricature, one may suspect that human investors may (quite rationally!) behave like chimpanzees in booms and like bonobos during market crashes.

In particular, exceptionally loose monetary policy during a boom could change aggregate risk aversion as a (evolutionarily) rational response to altered expectations of future rates of return. The empirically observed yield-seeking behaviour (shifting to higher risk assets to maintain portfolio yield levels) far from being irrational may in fact be ecologically rational when viewed as an environment contingent shift in life history strategies. Central banks may need to take this into account.

Probabilities: They are Always Subjective

Finance courses necessarily build on what has been covered in the statistics courses. A course on portfolio theory, for example, would assume knowledge of the meaning of covariance. Unfortunately, there is a problem with this division of labour – most statistics professors teach classical statistics. That is true even of those statisticians who prefer Bayesian techniques in their research work!

The result is that many finance students wrongly think that when the finance theory talks of expected returns, variances, and betas, it is referring to the classical concepts grounded in relative frequencies. Worse still, some students think that the means and covariances used in finance are sample means and sample covariances and not the population means and covariances.

In business schools like mine, where the case method dominates the pedagogy, these errors are probably less (or at least do less damage) because in the case context, the need for judgemental estimates for almost everything of interest becomes painfully obvious to the students. The certainties of classical statistics dissolve into utter confusion when confronted with messy “case facts,” and this is undoubtedly a good thing.

But if cases are not used or used sparingly, and the statistics courses are predominantly classical, there is a very serious danger that finance students end up thinking of the probability concepts in finance in classical relative frequency terms.

Nothing could be farther from the truth. To see how differently finance theory looks at these things, it is instructive to go back to some of the key papers that established and developed modern portfolio theory over the years.

Here is how Markowitz (1952) began his Nobel Prize winning paper more than half a century ago:

“The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio.”

Many finance students would probably be astonished to read words like observation, experience, and beliefs instead of terms like historical data and maximum likelihood estimates. This was the paper that gave birth to modern portfolio theory and there is no doubt in Markowitz’ mind that the probability distributions (and the means, variances, and covariances) are subjective beliefs and not classical relative frequencies.

Markowitz is also crystal clear that what matters is not the historical data but beliefs about the future – historical data is of interest only in so far as it helps form those beliefs about the future.

Unless finance professors are willing to spend time in the classroom discussing subjective probabilities, they must put pressure on the statistics professors to discuss probability from the subjective, Bayesian point of view. Finance students need to be confronted with probabilities that have no frequentist interpretation at all. For example, Borch (1976) tries to estimate the probability that the Loch Ness monster exists (and would be captured within a one year period) given that a large company
had to pay a rather high premium of 0.25 per cent to obtain a million pound insurance cover from Lloyd’s of London against that risk. This is obviously a question which a finance student cannot refuse to answer; yet there is no obvious way to interpret this probability in relative frequency terms.

**Expectations are Heterogeneous**

The passage quoted above from Markowitz (1952) seems to take it for granted that different people will have different beliefs about the parameters of the subjective probability distribution of future returns.

Perhaps the seminal paper in finance to introduce the assumption that all investors have the same expectations was by Sharpe (1964). To develop the Capital Asset Pricing Model (CAPM) that won him the Nobel prize, William Sharpe had to assume that all investors had the same beliefs so that they could determine the market equilibrium. But Sharpe (1964) made this assumption with great reluctance:

“... we assume homogeneity of investor expectations: investors are assumed to agree on the prospects of various investments – the expected values, standard deviations and correlation coefficients described in Part II. Needless to say, these are highly restrictive and undoubtedly unrealistic assumptions. However, ... it is far from clear that this formulation should be rejected – especially in view of the dearth of alternative models.”

(emphasis added)

While finance theory has been built on equilibrium models like the CAPM, the application of these models in investor decision making has always recognized the role of heterogeneous expectations. Treynor and Black (1973) interpreted the CAPM as saying that: “...in the absence of insight generating expectations different from the market consensus, the investor should hold a replica of the market portfolio.” They devised an elegant and widely used model of portfolio choice when investors had moved out of consensus beliefs:

“The viewpoint in this paper is that of an individual investor who is attempting to trade profitably on the difference between his expectations and those of a monolithic market so large in relation to his own trading that market prices are unaffected by it.”

Similar ideas can be seen in the popular Black Litterman Model. Black and Litterman (1992) started with the following postulates:

1. “We believe there are two distinct sources of information about future excess returns – investor views and market equilibrium.”
2. “We assume that both sources of information are uncertain and are best expressed as probability distributions.”
3. “We choose expected excess returns that are as consistent as possible with both sources of information.”

Heterogeneous expectations arise naturally when there is inadequate data to estimate the requisite parameters with high accuracy. Even if long time series is available and the apparent sample size is very large, parameter estimates would be very imprecise if there are frequent regime changes. The global financial crisis has highlighted the importance of regime changes, and therefore forced us to recognize the imprecision in statistical parameter estimates. Parameter estimates must therefore be subjective, and expectations will be heterogeneous.

**Estimation Must Almost Always be Bayesian**

The importance of Bayesian estimation of parameters can be illustrated nicely in terms of the CAPM beta, but the discussion is equally applicable to Fama-French multi-factor models, the Arbitrage Pricing Theory, and several other models in modern finance theory.

The derivation of the CAPM makes it clear that the beta is actually the ratio of a covariance to a variance and both of these are parameters of the subjective probability distribution that defines the market consensus. Statisticians instantly recognize that the ratio of a covariance to a variance is identical to the formula for a regression coefficient and are tempted to reinterpret the beta as such.

This may be formally correct, but it is misleading because it suggests that the beta is defined in terms of a regression on past data. That is not the conceptual meaning of beta at all. Rosenberg and Guy (1976) explained the true meaning of beta very elegantly in their paper introducing what are now called fundamental betas:

“It is instructive to reach a judgement about beta by carrying out an imaginary experiment as fol-
allows. One can imagine all the various events in the economy that may occur, and attempt to answer in each case the two questions: (1) What would be the security return as a result of that event? and (2) What would be the market return as a result of that event?"

This approach is conceptually revealing but is not always practical at this level of generality. The process of arriving at a usable estimate in practice may involve many sophisticated econometric procedures, but the purpose of all this econometrics is to provide a better foundation for our subjective belief about the true beta of a company based on at least the following inputs:

- The beta is equal to unity unless there is enough reason to believe otherwise. The value of unity (the beta of an average stock) provides an important anchor which must be taken into account even when there is other evidence. It is not uncommon to find that simply equating beta to unity outperforms the beta estimated by naive regression.
- What this means is that betas obtained by other means must be shrunk towards unity. An estimated beta exceeding one must be reduced and an estimated beta below one must be increased. One can do this through a formal Bayesian process (for example, by using a Bayes-Stein shrinkage estimator), or one can do it purely subjectively based on the confidence that one has in the original estimate.
- The beta depends on the industry to which the firm belongs. Since portfolio betas can be estimated more accurately than individual betas, this is often the most important input into arriving at a judgement about the true beta of a company.
- The beta depends on the leverage of the company and if the leverage of the company is significantly different from that of the rest of the industry, this needs to be taken into account by unlevering and relevering the beta.
- The beta estimated by regressing the returns of the stock on the market over different time periods provides useful information about the beta provided the business mix and the leverage have not changed too much over the sample period. Since this assumption usually precludes very long sample periods, the beta estimated through this route typically has a large confidence band and becomes meaningful only when combined with the other inputs.
- Subjective beliefs about possible future changes in the beta because of changing business strategy or financial strategy must also be taken into account.

Much of the above discussion is valid for estimating Fama-French betas and other multi-factor betas, volatility (used for valuing options and for computing convexity effects), and default correlations in credit risk models and many other contexts.

Good classical statisticians are quite smart and in a practical context would do many of the things discussed above when they have to actually estimate a financial parameter. In my experience, they usually agree that (a) there is a lot of randomness in historical returns; (b) the data generating process does not remain unchanged for too long. Therefore in practice there is not enough data to avoid sampling error and hence it is desirable to use a method in which sampling error is curtailed by fundamental judgement.

On the other side, Bayesians shamelessly use classical tools because Bayes theorem is an omnivore that can digest any piece of information whatever its source and put it to use to revise the prior probabilities. In practical terms, Bayesians and classical statisticians may end up doing very similar stuff.

The advantage of shifting to Bayesian statistics and subjective probabilities is primarily conceptual and theoretical. It would eliminate confusion in the minds of students on the ontological status of the fundamental constructs of finance theory.

**Prices: Market Microstructure has Macro Consequences**

At the microstructure level, there is no such thing as “the price.”

In discussions about price in market microstructure, the term price must be qualified to make clear what we are talking about. There is a bid price, an ask price, a mid price, a last traded price; and then, there is a volume weighted average price, but there is no such thing as “the price.”

For example, consider a market where the bid price (at which a small lot of shares can be sold) is 99.00 and the ask price (at which a small lot of shares can be bought) is 100.00. The mid price is 99.50 and is perhaps the closest that one can come to the concept of “the price.” To a
first approximation, it may be reasonable to regard 99.50 as the price if the focus is on small purchase or sale transactions.

For larger transactions, however, the price is quite different. A seller who wants to liquidate 1,000 shares may receive 99.00 for say, the first 250 shares, but may then receive progressively lower prices for subsequent blocks of shares and the volume weighted average price may be only 98.50.

A seller with 5,000 shares to sell might exhaust all the buy orders in the order book and would have to wait for the order book to refresh with new bids (latent orders) from value traders seeing a buying opportunity at the low prices induced by the large sell order. After several minutes, the sell order of 5,000 shares may finally get executed at a volume weighted average price of say 96.45 after having pushed the price down to 95.00. In other words, a large order can have a large “impact cost” and a long execution time.

There will often be an intervening period, when say 2,000 shares have been sold against the orders available in the order book, but value traders are still evaluating the situation and the latent orders have not materialized. At this point, the bid side of the order book is empty and the market is “ask only.” There is no bid price; there may be an ask price of 100.00; and there may be a last traded price of 98.00. The concept of “the price” is even more elusive at this point.

For microstructure theorists, there is nothing unusual in all this; on the contrary, this is the normal state of affairs. During crises, this phenomenon can occur at a bigger scale and over longer periods. For example, in the dollar/yen exchange rate, the bid and ask prices may normally be separated by only a couple of cents. But during the dramatic events of October 8, 1998, the bid ask spread widened to 200 cents (the yen was bid at ¥113.50/$ and asked at ¥111.50/$). A prominent hedge fund manager complained to his investors that: “The yen, which was as liquid as water, suddenly dried up like the Sahara” (Mallaby, 2010).

During the global financial crisis, this phenomenon was witnessed on an even larger scale with entire markets freezing for extended periods of time. From a microstructure perspective, what is new is not the phenomenon itself, but its scale, scope, and duration.

Is a Financial Crisis Simply Market Microstructure Writ Large?

Over the short time intervals of microstructure events (a few minutes), sharp and rapid price declines (market meltdowns) and the converse (melt-ups) happen all the time. For example, any sell order large enough to sweep through the whole or a major fraction of the bid side of the order book would cause a steep decline in prices within seconds (if not milliseconds). It might take several minutes for enough latent orders to enter the order book and reverse this meltdown. Conversely, a large buy order can send the price shooting upwards in the space of a few seconds or even milliseconds.

Over the short time intervals at which microstructure events take place, these “tail events” cause price movements that are several times the range that would be expected from a Gaussian distribution. In the earlier microstructure example, most of the time, as small sell and buy orders are filled at the current bid and ask prices respectively, the price may fluctuate between 99.00 and 100.00 leading to fluctuations of only 0.50 around the mid price of 99.50. An occasional large order could however cause the price to drop to 95.00. If the standard deviation is estimated from only the previous half hour of price movements, this drop of 4.50 from the mid price of 99.50 could be a 9 standard deviation move which is a near impossibility in a Gaussian distribution.

Such fat tails are very common in market microstructure, but microstructure theorists do not regard these markets as dysfunctional or irrational. On the contrary, what is important in this context is the self correcting ability of the market that restores equilibrium over the space of several minutes or hours. Taking into account the various frictions (search and information costs, transaction costs, and leverage restrictions), we should probably consider a market which experiences such microstructure meltdowns or meltups to be an efficient market.

During the crisis, booms and busts happened at a macro scale (over time frames of several months instead of minutes), but it is possible that the phenomena differed from microstructure events only in their scale and duration.

* I think it is a good idea to consistently use the term “Gaussian distribution” instead of “normal distribution” to avoid the risk of students inadvertently and subconsciously associating non-Gaussian distributions with some form of abnormality.
A financial crisis may simply be market microstructure writ large. Perhaps the complexities of “microstructure noise” persist at longer time scales as well, and the market is in a perpetual state of chaotic movement towards an ever changing equilibrium instead of being in a continuous state of equilibrium.

The hypothesis that financial crisis is simply market microstructure writ large implies that markets are messier and more complex than the ideal friction-free market. On the flip side, it means that we have the theoretical tools and techniques (of microstructure theory) to study crises.

At any rate, I think that all finance researchers must not only learn market microstructure theories, but also take them seriously as potential explanations for even macro scale phenomena.

MODELS

I now discuss the nature of changes that need to be made in the models that are used in modern finance.

Efficient Market Hypothesis: There is Still No Free Lunch

We must distinguish between two important aspects of the efficient markets hypothesis (EMH) because the global financial crisis has led to diametrically opposite conclusions regarding these two perspectives:

• The first perspective is summarized by the statement that there is no free lunch or that it is not possible to beat the market in risk-adjusted terms. If something is too good to be true, it is probably not true. The global financial crisis has strengthened this claim. All those apparently low-risk, high-return investments turned out to be high-risk.

• The second perspective is that prices are “right” in the sense that they reflect fundamentals. The global financial crisis has weakened this claim. Many prices were clearly not right.

It is easy to reconcile these two lessons from the global financial crisis by drawing on the limits to arbitrage literature. Limits to arbitrage imply that prices are not always “right,” but limits to arbitrage also tells us that the prices are wrong for a reason. The no free lunch argument remains true: there are anomalies, but no easily exploitable anomalies.

Another way of looking at it is that what appears like a free lunch is just the reward for a hidden tail risk. It is the unhedgeability of this risk (possibly a liquidity risk) that prevents arbitrageurs from correcting the anomaly. This apparent free lunch can be exploited only by those who can back their bets with a nearly infinite pool of liquidity and capital. The agents that meet this description best are the “too big to fail” (TBTF) banks with implicit sovereign support. TBTF banks may in fact go further and actively manufacture hidden tail risk and the resulting apparent free lunches that only they can exploit.

Unfortunately, regulators fail to understand the consequences. The EMH does not justify a light touch regulation of TBTF banks. On the contrary, the “no free lunch” form of the EMH ought to lead regulators to suspect that an incredibly profitable bank is an incredibly risky bank (with huge hidden tail risks) and therefore needs high levels of capital to mitigate the risk.

Finance courses need to teach more about the limits to arbitrage not just in terms of behavioural finance, but in terms of well specified market micro structure with proper attention paid to transaction costs, leverage, and collateral requirements. The important stream of literature (e.g., Brunnermeier and Pedersen, 2007) linking funding liquidity and market liquidity needs to be a part of the core courses in financial markets.

It is equally true that over reliance on the “prices are right” form of the EMH allowed much of modern finance to deviate too far from its micro foundations in terms of well-defined fundamentals. Derivative models allow us to compute implied volatility and implied correlations (and if necessary the entire implied risk neutral distribution). These models allow us to start valuing anything without any regard to fundamentals at all. Models then become over-calibrated to markets and under-grounded in fundamentals. For example, quite often derivative textbooks and courses do not encourage us to ask questions like: what is the fair value of an option if we assume that the underlying is 10 per cent overvalued in the marketplace.

Multi Factor Models are Unavoidable

We Must Go Beyond Size and Value

While most introductory courses in financial markets and corporate finance are grounded in the CAPM, this
is no longer the case in advanced courses in the MBA curriculum. Long ago, the Fama-French three-factor model replaced the CAPM as the workhorse for modeling asset returns when it comes to research in financial markets. Increasingly, this model is also the core model in elective courses in the MBA classroom. In recent years, the Fama-French model has begun to give way to the Carhart four-factor model which takes momentum into account. However, I think it is necessary to go beyond even this to consider liquidity as an explicit risk factor. Market, size, value, momentum, and liquidity are all essential to understand asset returns in the post-crisis world.

**Liquidity is a Systemic Risk**

Though liquidity was studied as far back as the mid-1980s (Amihud and Mendelson, 1986), finance academics began to develop sophisticated models for liquidity risk only after the LTCM crisis in 1998. Most of the important advances in this field came in the mid-2000s well before the global financial crisis.

First of all, it became clear that liquidity was a systematic risk and not a diversifiable risk. The commonality that was found in liquidity made it possible to talk about liquidity betas and the liquidity risk premium in a way completely analogous to the corresponding notions for market risk (or the size and value factors). Pastor and Stambaugh (2003) and Acharya and Pedersen (2005) have established liquidity as a risk factor that needs to be considered alongside the conventional risk factors.

The second key advance was the theoretical linkage that was established between market liquidity and funding liquidity (Brunnermeier and Pedersen, 2007). Market liquidity which is all about market microstructure is intimately related to funding liquidity which is all about macroeconomics – suddenly, micro has become macro!

An interesting application of including liquidity and other factors while evaluating the performance of an asset manager in a real world investment context is provided by Ang, Goetzmann and Schaefer (2009). Another interesting application is provided by Chen, Ibbotson and Hu, 2010.

Ang, Goetzmann and Schaefer (2009) evaluate the performance of the Norwegian sovereign wealth fund over a 12-year period from inception to September 2009. Pre-crisis, the fund had a consistent positive excess return (over the benchmark) of about 0.03 per cent per month cumulating to over 4 per cent at the onset of the crisis. This entire cumulative excess return was wiped out in 2008 with the cumulative excess return going negative before rebounding partly in 2009. Ang, Goetzmann and Schaefer fitted a multi-factor model including liquidity and volatility and found that the excess returns (both pre-2008 and in 2008) could be explained by the factor model: “What we find, perhaps surprisingly, is that using only data that would have been available at the time just prior to the period of very bad active returns in 2008, the very poor results following the collapse of Lehman could have been predicted to a significant extent conditional on the realizations of the factors.” More interestingly, they concluded that the fund’s liquidity and volatility exposures were appropriate as a means to earning factor premiums, but they needed to be communicated better.

**There is No Risk-free Rate**

The existence of a risk-free rate is not essential in most finance theories, but it is a very convenient simplification. Until the crisis, this simplification was largely harmless. With the onset of concerns about sovereign debt even in core developed markets, the assumption of a risk-free rate is no longer a harmless simplification.

In equity pricing theory, the notion of a risk-free rate was dispensed with as long ago as the zero beta model of Black (1972). Moreover, rising levels (and volatility) of inflation in the 1970s led to the realization that the nominal risk-free asset is not really risk-free.

In fixed income markets, the risk-free rate played a more important role as all bonds tended to be priced off the risk-free yield curve. In the early 2000s, this changed however and the swap yield curve displaced the sovereign yield curve as the pricing benchmark. During the crisis, the spread between Libor and government bonds reached stratospheric levels. The idea of the swap rate (which is tied to Libor) being risk-free became increasingly untenable.

At the same time, it was not possible to go back to the notion of government bonds being risk-free. During the crisis, credit default swap (CDS) premia for top-rated (AAA) sovereigns rose above 100 basis points. This implied that the annual risk neutral probability of default of these sovereigns was 1 per cent or more which is hardly
compatible with the idea of their bonds being risk-free. Post-crisis, OIS (overnight index swap) is regarded as the closest thing to a risk-free rate under the assumption that the probability of default of a highly creditworthy entity over a one-day horizon is negligible. This leads to the well known two-curve discounting model (Fujii, Shimada and Takahashi, 2009; 2010, Mercurio, 2009; Morini, 2009). I think the idea of a risk-free rate should be regarded as nothing more than a useful approximation.

Microstructure Needs Agent-based Modelling

Simple general equilibrium models relying on a representative investor typically assume homogeneous expectations. In these models, agents do not trade with each other – they only optimize against their budget constraints. There are no trades because the equilibrium price is defined as the price at which nobody wants to trade. The prices in these models are therefore more in the nature of shadow prices than real prices.

Of course, it is possible to build more complex models that allow two or three different types of agents with different information sets, but even these cannot capture the full range of heterogeneity that is apparent in real world financial markets.

Agent-based models, on the other hand, allow arbitrary number of heterogeneous players with different information sets, trading strategies, and objectives. Computer simulations are the principal tool because typically analytical solutions do not exist. The big advantage is that information aggregation and price discovery can be studied in detail, and the impact of alternative market microstructures can be quantified. A good example of this kind of work is Lee, Cheng and Koh (2010), who simulated a “flash crash” before it occurred on May 6, 2010 (in the US markets) using an agent-based model.

If it is true that microstructure theories are relevant for understanding phenomena at macro time scales, then it is necessary to embrace agent-based models in finance theory.

Tail Risk is the Only Real Risk

Way back in the 1960s, it was recognized that fat tails are pervasive in financial time series (Mandelbrot, 1963; Fama, 1963). But tail risk as the preponderant form of priced risk has gained ground only in this century with the influential papers of Barro (2006) and Barro and Ursua (2009). Before the global financial crisis (and the associated Great Recession), it was widely believed that depressions were impossible (at least in developed markets) and it was easy to brush aside the idea that the Equity Risk Premium is compensation for tail risks like the Great Depression risk. Post-crisis, this is a point of view that needs to be taken very seriously indeed.

Much of the tail risk of diversified portfolios comes not from the tail risks of the individual assets but from common jumps or other form of non-linear dependence. Empirically, only about half of the downside risk (fat tail) of a portfolio of stocks is due to the fat tails of its members (Langnau and Cangemi, 2011; Langnau, 2010). Thus tail risk would require us to embrace not just fat-tailed distributions, but non-Gaussian copulas as well.

Quantitative models based on non-Gaussian fat-tailed distributions with non-linear dependence structures (copulas) are hard from the point of view of teaching in the MBA classroom, but we must not shirk hard mathematics. Risk modelling using Value at Risk with Gaussian distributions and linear correlations is no longer defensible after all that we have seen during the global financial crisis (Varma, 2009).

Econometrics Must be Grounded in Financial History

The global financial crisis and its aftermath evoked parallels with

- The Great Depression of 1930s
- The Panic of 1907
- The sovereign defaults of 1890s and 1930s
- The financial (and sovereign debt) crises of 1830s and 1870s

From a long historical perspective, the financial crisis does not appear to be an aberration at all. On the contrary, it is the Great Moderation of the late 1990s and early 2000s that appears to be an aberration. For example, Haldane (2009) provides data for macro-economic volatility in the UK (Table 1).

A key mistake prior to the crisis was the assumption that the Great Moderation was a permanent structural change in the world economy that implied a permanently reduced volatility. The crisis has taught us that the statistical processes that we observe during any par-
ticular period should be viewed as just one of the several possible regimes. There is always a non-trivial probability of shifting to a different regime.

The “new normal” in this sense is that there is no unique and stable “normal.” Frequent regime changes imply that sample sizes (restricted to the current regime) are always small. More importantly, the possibility of future regime changes means that the parameters of the historical distribution estimated from such a sample are not reliably predictive of the future distribution. Since regime changes are relatively infrequent, the probability of a regime change is also estimated very imprecisely from past data. As a result, the historical data is never sufficient to completely dominate the subjective prior distribution. Heterogeneous expectations about the future distribution are therefore to be expected.

I see financial history as providing powerful inputs into the econometric procedures that we use. Since high quality data does not usually go back more than a few decades, we do not have the option of fitting econometric models directly to centuries of data. Yet, it is not sensible to limit the estimation process to only the limited sample duration that is available. What we need to do is to favour robust models that are qualitatively consistent with decades if not centuries of historical experience. Such models should not only provide a good fit to the high quality data of the recent past, but also allow us to extrapolate far beyond recent experience. Markov switching models using Bayesian priors are quite capable of doing this in a tractable and rigorous way.

Much simpler approaches may also be sufficient. For example, Barro and Ursua (2009) use nearly a century of macro-economic data and stock price data from 25 countries to estimate the linkage between stock market crashes (a drop of at least 25% in real terms) and economic depressions (a contraction of at least 10% in real terms). Much of the estimation involves only simple tabulation and counting to estimate switching probabilities. They also find that a simple power law distribution is sufficient to describe the size of the contraction. Clearly, the hard part is not the econometrics, but putting the data together and more importantly, recognizing the relevance of the data that is available.

In any case, a significant amount of financial history should be a part of the finance curriculum. Among the many excellent books that are available today, I would like to mention the ones by Reinhart, and Rogoff (2009), Homer and Sylla (2005), and Goetzmann and Rouwenhorst (2005).

### LEARNING FROM RELATED DISCIPLINES

I now examine the need for finance to integrate insights from other source disciplines.

#### Neuroscience and Sociology are as Important as Psychology

Behavioural finance is now so much a part of most standard finance that it is often difficult to distinguish between behavioural and neoclassical finance (Thaler, 1999; Berg and Gigerenzer, 2010). For example, the asset pricing models that include the momentum factor are clearly behavioural finance models, and even the Fama-French model has a strong behavioural interpretation. The limits to arbitrage literature are also often associated with behavioural finance. All these approaches have proved their worth during the crisis.

Yet, there are areas in which there is a need to re-emphasize the hard-nosed rational models. For example, the build-up to the subprime crisis was characterized by a reliance on credit history (FICO scores). The implicit assumption is that default is a behavioural trait that can be measured using past payment records. Rational models (Merton style models) assume that people default when it is rational to do so and focuses attention on modeling the fundamentals (for example, home prices). Clearly lenders would have been much better relying on rational models rather than presumed behavioural traits.

Unfortunately, during the lending boom, behavioural models held sway and these were supported by the short historical time series data that was then available. It is amazing but true that so much of what happened dur-

### Table 1: Volatility of UK Macroeconomic Variables during the Great Moderation compared with 150-year Average

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP growth (%)</td>
<td>0.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Earnings growth (%)</td>
<td>0.5</td>
<td>6.4</td>
</tr>
<tr>
<td>Inflation (%)</td>
<td>0.9</td>
<td>5.9</td>
</tr>
<tr>
<td>Unemployment (%)</td>
<td>0.6</td>
<td>3.4</td>
</tr>
</tbody>
</table>

Source: Haldane (2009) Annex Table 1.
ing 2007 and 2008 can be explained as the rational response of economic actors to altered fundamentals.

These divergent signals from the global financial crisis suggest that the right balance between behavioural models and rational models remains a major challenge for finance theory despite the growing synthesis of behavioural and neoclassical finance.

Continuing developments in neural imaging leads me to believe that neurosciences might have a lot to contribute to finance theory (Camerer, Lowenstein and Prelec, 2005; Bernheim, 2008). Drawing on the comprehensive survey in Camerer, Lowenstein and Prelec (2005), a few of the important insights from neuroeconomics can be summarized as follows:

- Neural imaging studies indicate that money is an intrinsic reward and has direct utility, unlike in standard finance models where the utility of money is a derived utility (derived from the utility of consumption). If this finding remains robust, it has vast implications for finance theory. It would challenge the entire consumption-based asset pricing theory. It would also require us to take money illusion (Shiller, 2007; Modigliani and Cohn, 1979) more seriously in asset pricing.

- Risk preferences and time preferences are state contingent and are not stable across situations.

- The importance of the reference point in risky choices may be partly due to neural sensitivity to changes and not to levels.

- Neuroscience suggests that risky choices sometimes reflect a tussle between the affective part of the brain (e.g., the amygdala which is responsible for the fear response) and the cognitive part of the brain (e.g., the frontal cortex). Often we may find that the observed choices and behaviour do not reveal any sign of fear, but what is going on is that the cortex is overriding the fear response emanating from the amygdala. Some animal studies show that, in this situation, a single new negative experience may be sufficient for a full blown fear response to manifest itself in behaviour. Theories that incorporate the affective dimensions of risk may be necessary to understand stock market booms and busts (Camerer, 2005).

Both behavioural and neo-classical finance theories assume the ability to calculate how other persons think and how other persons think about how other persons think and iterate these calculations to a higher order. In neoclassical economics, concepts of equilibrium like the Bayesian Nash equilibrium require the ability to iterate this to arbitrary order. In behavioural finance, we have Keynes’ famous beauty contest example: “It is not a case of choosing those which, to the best of one’s judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees.” (Keynes, 1936, Chapter 12). Neuroscience however suggests that “there may be no generic human capacity to iterate this kind of thinking beyond a couple of steps.” (Camerer, Lowenstein and Prelec, 2005). If this is so, both the Keynesian beauty contest (of behavioural economics) and the Bayesian-Nash equilibrium (of neoclassical finance) might lie beyond the cognitive capability of the average human mind.

The sociology of finance is another promising discipline which could improve our understanding of financial markets. In particular, the sociology of knowledge and the literature on market devices as socio-technical systems are particularly important:

- Sociologists argue that behavioural biases and the limitations of bounded rationality can be partly compensated by social cues. “[T]raditional economic theory (invoking the substantive rationality paradigm) succeeds wherever individual choice is strongly constrained by social and institutional scaffolding that has itself evolved subject to selective pressures to maximize rewards. Outside such highly constrained settings, genuine individual thought plays a greater role, and the psychological irrealism of the substantive rationality model takes its toll.” (Clark, 1997).

- Standard finance theory assumes that when the markets “calculate” a price, the “calculation” is simply an equilibrium arising out of the calculation (optimization) of individual agents. The sociology of finance suggests that the calculation is the result of “calculable goods, calculative agencies and calculated exchanges” (Callon and Muniesa, 2005). This in turn leads to the important insight that “Increasingly sophisticated devices allow for the proliferation of increasingly complex markets.” (Muniesa, Millo and Callon, 2007)
• “[F]inancial actors go back and forth between models, their understanding about what is being traded, and their ability to figure out what their competitors are doing. … calculation is social and … the social is technically mediated.” (Beunza and Stark, 2010).

• “Liquidity is, among other things, an issue in the sociology of knowledge” (Carruthers and Stinchcombe, 1999). In this context, MacKenzie (2010) emphasizes the critical role of “evaluation cultures” which he defines as “pockets of local consensus on how financial instruments should be valued.”

**Hard Mathematics Should Not be Eschewed**

The mathematical and statistical tools required in post-crisis finance are not new – they are the tools that have been widely used in finance theory during the last decade or more. What would be new would be their introduction into the MBA finance curriculum which has tended to be taught in a time warp of the 1980s or even the 1970s. The key elements of a modern curriculum in mathematical finance would include:

- **Fat-tailed distributions and power laws**: If we consider ten years of daily price movements, the normal distribution would predict that the largest price move is likely to be about 3-3½ standard deviations. In reality, we tend to observe price movements of 5-10 standard deviations. This is consistent with fat-tailed distributions like the student $t$-distribution with 5-8 degrees of freedom. The probability of large moves does not decline exponentially (as in the Gaussian distribution) but declines according to a power law (as in the student $t$ and other fat-tailed distributions).

- **Copulas and non-linear relationships**: The importance of copulas can be explained with a simple example of two assets which exhibit zero correlation implying that there is no linear dependence between them. Suppose we know that one asset has experienced a large (say three standard deviation) move. If we assume that the correlation captures the entire dependence relationship (so that there is no non-linear relationship between the two assets), then it is very unlikely that the second asset is also experiencing a large move. In practice, however, we will often find that the second asset is also having a large move at the same time. This can be consistent with a zero correlation if the large move in the second asset is sometimes in the same direction as the first and sometimes in the opposite direction. To someone who uses only linear correlations, it would appear as if the correlation has abruptly and unpredictably moved from 0 to ±1. A person using a non-linear copula (like the $t$-copula with zero correlation) would be prepared for this behaviour, and would not find it surprising at all.

- **Dragon kings and log periodic power laws**: While fat-tailed distributions with power law tails capture most of the non-Gaussian character of asset prices, several researchers have argued that they still do not account for the most extreme movements in asset prices. These “wilder” extremes (christened as dragon kings) are “associated with a neighbourhood of what can be called equivalently a phase transition, a bifurcation, a catastrophe (in the sense of René Thom), or a tipping point” (Sornette, 2009). Sornette argues that the presence of a phase transition means that it may be possible to diagnose in advance the symptoms associated with a coming dragon-king. Mathematical models based on log periodic power laws have been developed to attempt such predictions; Fantazzini and Geraskin (2011) provide a good and accessible review of these models.

- **Stochastic calculus for Levy processes and discontinuous semimartingales**: Many MBA programmes today have a course on stochastic calculus. Most of these courses go far enough to teach students the mathematical tools (essentially the stochastic calculus of processes driven by a Brownian motion) required to understand the derivation of the Black Scholes option pricing model. In line with the general shift away from the Gaussian distribution, derivative pricing today increasingly relies on processes more general than the Brownian motion in which all increments are Gaussian. Poisson jump processes have become popular as have processes based on general Levy processes which can be thought of as fatter tail generalizations of the Brownian motion. Many important financial models originally formulated with Brownian motions have been extended to Levy processes and semi-martingales. Good textbook treatments of the mathematics required for all this can be found in Applebaum (2004) and Rogers and Williams (2000).

- **Bayesian models**: Bayesian statistics is extremely important for several reasons as explained earlier in this paper – estimation of betas and other asset pricing parameters, estimation of models with regime shifts, dynamic learning models in which investors are try-
ing to estimate parameters that are changing over time and so on. In the past, Bayesian models were difficult to use in practice because of the high computational burden. Increases in computational power and efficient implementation methods like MCMC (Markov Chain Monte Carlo) methods have changed this and made Bayesian models computationally feasible.

Network Models: Integrating Financial Markets and Institutions

Over the years, finance teaching (and to some extent finance theory) has come to be segmented between financial institutions and financial markets. This segmentation is increasingly untenable as the dividing line between institutions and markets gets blurred.

The repo market is a good example of this blurring of lines. Pre-crisis, courses on fixed income markets placed a lot of emphasis on the repo market as a critical component of the bond market. But neither the markets courses nor the banking courses looked at the repo market as being akin to a bank or a financial institution. After the crisis, Gorton and Metrick (2009) have taught us that the repo market is a “shadow bank” vulnerable to old-fashioned “bank runs” (see also Pozsar et al, 2010).

Similarly, courses in financial markets did not focus much on the fact that banks were investing in securitized instruments through Special Investment Vehicles (SIVs) that funded themselves with Asset Backed Commercial Paper (ABCP). This fragile form of maturity transformation had devastating implications for certain securitization markets during the crisis.

I believe that network theory has a great deal of potential in unifying institutions and markets because it teaches us to focus less on the individual components (markets, instruments, institutions) and more on their interconnections. Many of the properties of the networks depend on these interconnections. Contagion of crises and many other emergent phenomena can perhaps be best understood as network effects.

Easley and Kleinberg (2010) provide a powerful framework for looking at markets as networks. There has been a great deal of effort to understand the interdependence of banks by using network models (Garratt, Mahadeva and Svirydenka, 2011). I believe that network theory would become an integral part of the toolkit of finance theory and it is time for these tools to enter the finance curriculum.

At another level, it is perhaps true that we teach too much of ephemeral institutional detail. Many of the details which we taught to our students during the last 3-5 years have been rendered obsolete by changes in the market structure. Investment banks are gone; the Libor market is barely recognizable; and risk-free government paper is no longer risk-free. When we are preparing students for a career and not for their first job, we must emphasize functions and not institutions; concepts and not context. Again network theory could provide a unifying theme. In some sense, the financial sector is just a vast network that connects providers of capital at one end to users of capital at the other.

CONCLUSION

The global financial crisis has revealed serious problems with the finance that is taught in a typical MBA programme. One major reason for this is that the coverage in the finance courses has not kept pace with the developments in finance theories in the last decade or more. For example, the global financial crisis demonstrated that risk modelling using Value at Risk with Gaussian distributions and linear correlations is a terrible idea. However, finance theory had moved far beyond this naïve model since the late 1990s. In other words, while a lot needs to change in finance teaching, much less may need to change in finance theory itself.

Another important conclusion is that many ideas that are well understood within certain subfields in finance need to be better assimilated into mainstream models. For example, many concepts in market microstructure cannot remain niche ideas, but must become part of the core toolkit of finance.

Finally, finance theory itself is constantly evolving and needs to draw on insights from several other disciplines to enrich itself. Behavioural finance has succeeded in integrating several models from psychology into mainstream finance, but the global crisis has demonstrated that many phenomena have their roots in sociological factors. Apart from sociology, finance must learn from evolutionary biology, neurosciences, financial history, and the multidisciplinary field of network theory. Above all, the increasingly complex world of finance needs more sophisticated mathematical models and statistical tools.
REFERENCES


Langnau, Alex and Cangemi, Daniel (2011). “Marking Systemic
Jayanth R Varma is currently a Professor in the Finance and Accounting Area at the Indian Institute of Management, Ahmedabad. He has been the Dean of the Institute for three years. He obtained a post-graduation (PGP) and doctorate in management from the Indian Institute of Management, Ahmedabad (IIMA) and was awarded a Gold Medal for scholastic performance in his PGP. He was a full-time Member of the Securities and Exchange Board of India (SEBI) for a year and a part-time member of SEBI for three years. He has been the Chairman of the Secondary Markets Advisory Committee and several other committees of SEBI, and also of several committees set up by the Forward Markets Commission and the Ministry of Finance and Company Affairs. He was a member of the Raghuram Rajan Committee on Financial Sector Reforms and is currently a member of the Financial Sector Legislative Reforms Commission. He is currently a director on the Board of Infosys BPO Limited and of Onmobile Global Ltd. He has also been a Director on the Board of Punjab National Bank. He has carried out extensive research in the field of Indian financial markets and finance theory and has authored extremely popular books on portfolio management and Indian securities scam. His book on the securities scam was a national best-seller.

e-mail: jrvarma@iimahd.ernet.in