

# RECORD, Volume 29, No. 3\*

---

Orlando Annual Meeting  
October 26–29, 2003

## Session 129PD Investment Performance Attribution

**Track:** Investment

**Moderator:** JOSEPH KOLTISKO  
**Panelists:** CATHERINE E. EHRLICH  
BASIL P. RABINOWITZ  
FRED SHEN†

*Summary: This session provides an overview of how insurers, pension plans and other institutional investors measure and report manager performance.*

*Topics include:*

- *Investment benchmarks—how to use and abuse them*
- *Yield and total return—target or constraint?*
- *Incentives and accountability*
- *Standards of professionalism for the investment reporting process*
- *Report design*

*Attendees take home examples of best practices for reporting investment performance.*

**MR. JOSEPH KOLTISKO:** First up will be Catherine Ehrlich from Swiss Re Life and Health America. She's going to help us understand the process of deciding what's important for investment performance and for measuring progress against benchmarks. Basil Rabinowitz is the senior vice president for investments at Fortis Inc. He's going to discuss some of the pitfalls and opportunities in using investment benchmarks for insurers. We also have Fred Shen, vice president for risk analytics

---

\* Copyright © 2004, Society of Actuaries

† Mr. Fred Shen, not a member of the sponsoring organizations, is vice president for risk analytics at the Bank of Montreal in Quebec, Canada.

**Note:** The chart(s) referred to in the text can be found at the end of the manuscript.

at the Bank of Montreal. He's going to show how to measure performance using economic capital. He will show us a balanced perspective that applies to both banking and insurance. So we're going to talk about measurement and attribution, benchmarks, ground rules, pitfalls and capital at risk.

**MS. CATHERINE EHRLICH:** Performance attribution is an important process, but it can be difficult. Knowing what securities to pick, what sectors to be in and what duration is a lot of work. Attribution helps you figure out if your given strategy is appropriate for what you're trying to do as well as if your manager is doing a good job. But before you can get into the process of performance attribution you need to make certain decisions. You need to figure out how you're going to define your performance. You need to figure out what you're going to compare it to. We'll go through some of the different types of benchmarks people can use with their portfolios. Finally, you need to figure out if you have broken your return down to the proper components so it actually makes sense.

Now, the first thing is very straightforward. I'm just trying to figure out how my portfolio did and how I'm going to measure this performance. However, at insurance companies it seems that nothing is a very easy decision. There are a lot of issues involved. Basically you need to decide if you're going to be looking at results on a total-return basis or on a book-yield basis. Total return has a lot of very wonderful qualities about it, and most of the work I've done on performance attribution has been on a total-return basis. Total return reflects the economic environment of the period you're measuring. It reflects the decisions that we've made during that period. There really is no historical component to it, but it's a good metric.

However, in an insurance company your financial statements are not on a total-return basis. Your crediting-rate methodology and your dividend-rate methodology are not on a total-return basis. So managing your portfolio through the total-return benchmark when you have these book-yield considerations can be suboptimal. There are other constraints that are put on a portfolio. That is something that really needs to be worked through as a company. Next to this choice the other decisions aren't so fundamental. You have to decide if you're going to look at things on a time-weighted basis or a dollar-weighted basis. A time-weighted basis gets rid of the effects of cash flow. It's the return you would achieve if you invested \$1 in that portfolio at the beginning of the period and no other cash flows happened. Dollar-weighted basis takes into account the cash flows during the period and lets you know what the internal rate of return of that portfolio was over the period. So from a manager's perspective, if they are not generating the cash flows and getting the cash flows into the portfolio, they would probably prefer to be measured on a time-weighted basis. That's going to measure their performance rather than what happened to them, because you've got a large amount of money coming in or out of the portfolio at a certain time. Finally, you need to decide your frequency of reporting. Is it going to be annual or quarterly, monthly or daily? It may depend on whatever your systems can support.

The measures we talked about so far didn't really incorporate anything for risk and returns. Without any adjustment for risk, the portfolios can be very misleading. If you have a portfolio manager who achieves a 10 percent return by having a lot of risky assets in that portfolio, did they do a better job than the manager who achieved a 6 percent return but has low-risk assets in the portfolio? A straight return measure doesn't really get at that. On the equity side, we've had a number of measures that do incorporate risk.

The first measure is called the Treynor measure, which will look at the excess of return of the portfolio over the risk-free rate and divide it by the beta for the portfolio. That measure assumes you have diversified away all unsystematic risk in your portfolio.

The Sharpe ratio looks very similar. It takes the excess return of the portfolio over the risk-free rate, but divides it by the standard deviation of the current portfolio. So it does incorporate the whole volatility of the actual portfolio, and not just the systematic volatility.

Then there's the Jensen measure, which really looks at the portfolio's "alpha": it tells us the return over and above what can be expected for the data in the portfolio.

The final measure that we're looking at here is the information ratio, which is the excess return of the portfolio over a benchmark portfolio divided by the standard deviation of that excess return. If you have asked the manager to manage an index portfolio for you, and there is a lot of variation of return from what that portfolio achieved versus what the benchmark or the index achieves, then you've got some idea that something's wrong.

Recently there are some other performance measures that people have looked at. One is risk-adjusted return on capital (RAROC), which is defined as the gain or loss, adjusted for risk, and then divided by capital. RAROC can be used as a forward-looking measure in the risk-management process and capital allocation. It can also be calculated on an ex post basis, so you can use it for performance attribution. In that case the profit would include your income on the risk capital, and it would not include realized losses. Tracking error, which actually is the generator of the information ratio, is quite commonly used for fixed-income portfolio managers who are trying to track an index.

Once you've finished determining what you're going to call your performance, then you need to describe what you're going to compare it to. In insurance companies the most obvious bogies are the liabilities. What you're trying to do in an insurance company is make sure you have enough money to pay off those liabilities. So if your asset portfolio outperforms your liabilities, you did well, and if it underperforms it, you didn't do so well. However, unless your liability is a static

stream, it's going to be very hard for the portfolio manager to know what to do with this. They have a very deep understanding of asset markets, but not a very deep understanding of the liabilities. As actuaries we need to help the portfolio managers understand what they're trying to do. Just telling them to beat the liabilities doesn't do it for them.

There are several types of benchmarks that have been used. The first type is the liability. They're not easily understood by the portfolio managers. The next type is the model portfolio or normal portfolio. That's a universal security that's been specifically selected to match the characteristics of the liabilities. A model portfolio takes some work to set up. You need to make sure it looks like the liabilities and behaves the way you want it to behave. Also, it needs to be adjusted over time.

The next is a commercial benchmark, which is an index that's made available by a third party, like Lehman Brothers or Citigroup. That index could be a basket of securities that have certain inclusion rules such as investment grade bonds between one and 10 years. They are very easily understood by portfolio managers. It's really easy to get returns on them. They don't often do a good job of mimicking the liabilities. Commercial benchmarks can be customized, and I call this the Frankenstein benchmark. That's basically the weighted average of certain commercial benchmarks. You might not want your benchmark to look like the Lehman Brothers aggregate. You may want it to be a specialized average or combination of a mortgage index and an asset-backed index, or a treasury index, and you want to select the weights for those. You can specify that and add it all up and call it whatever you want. Those are also very commonly used.

The final type of benchmark I will include is the peer group universe, which is really not much of a benchmark. It says, after the fact, let me compare my portfolio manager to all these other portfolio managers and see how the return compares. This is not fair to the managers because you're not specifying in advance what you're really trying to do.

Every CFA candidate learns the qualities of an effective benchmark. From the portfolio manager's perspective, a useful benchmark is unambiguous, investable, measurable, appropriate, reflective of current investment opinions and specified in advance. The first characteristic is that it's unambiguous, which means the names and the weights for securities that comprise the benchmark are clearly defined. Second is that it is investable. The manager could simply hold the securities that are in the benchmark as an option. Third is that it's measurable, and the returns can be calculated on a reasonably frequent basis. The next is that it is appropriate, and the benchmark is consistent with the stated investment strategy or guidelines. Next, that it reflects current investment opinions: the manager has current investment knowledge of securities that make up the benchmark. Finally, that it is specified in advance. The benchmark should be constructed prior to the start of an evaluation period.

	Unambiguous	Investable	Measurable	Appropriate	Current	Specified in Advance
Liabilities				X		X
Model	X	X	X	?	X	X
Commercial	X	?	X	?	X	X
Frankenstein	X	?	X	?	X	X
Universe						

This chart is kind of a summary. I took the five types of benchmarks that we discussed and six qualities of an effective benchmark and tried to see which one came out the best. I really couldn't find a winner. I found a clear loser. It appears that the universe is criticized as having none of the qualities of a good benchmark. There's no one that can really invest in what his or her peer group is automatically investing in, and it's not specified in advance. It's rarely appropriate, and it has none of the good qualities.

The liabilities, I decided, clearly are appropriate and are specified in advance. They have none of the other qualities that would be helpful to an investment manager. You can't just invest in the liabilities and hold them. It's not a basket of securities that someone can look at. The model portfolio does pretty well, although I have a question mark for appropriate because it's often hard to find a model portfolio that exactly mimics your liabilities. There are certain characteristics in liabilities that just can't necessarily be replicated in the financial markets. The commercial and the Frankenstein are in the same areas. I did mention they might not be appropriate. I didn't talk about the fact that they might not be investable.

Commercial available fixed-income indexes are made up of thousands of securities that do not trade every day and often do not trade for a long period of time. It is hard to believe that any portfolio manager could actually hold the appropriate weights of all the securities in the large index. The quality and characteristics of the commercial industries can change over time because they're based on the market at the time.

Chart 1 shows how something like the Lehman Brothers aggregate might have changed over time. This is not based on that; this is just outstanding public and private debt. So if I were going to make an index that was based on outstanding public and private debt in the marketplace starting in 1985, I would have had approximately 50 percent of my index in treasuries. By the year 2003 it would be down to 20 percent in treasuries. You can imagine that the quality of the index would have fallen dramatically over that period of time. You can see that the amount of mortgage-related securities increases, so then you would have different convexity characteristics of your benchmark than you did at the start of the period. Basically you don't have too much control over what that looks like, it's just what's out there in the marketplace.

So if you were able to specify what your measurement metric was going to be and what your benchmark was, then you could actually move on and do some performance attribution. On a very macrobasis for your total portfolio, you can look at your attribution and break it down into two major different components of return. One is an allocation effect, and the other is selection effect.

An allocation effect is the over- or underperformance of my portfolio relative to my index because of what sectors I was in or whether I decided to invest in equities when I should have been in bonds or vice versa, compared to the benchmark. It doesn't try to say whether I was in the right bonds or the right equities. For example, if the benchmark had a 15 percent allocation to equities and the actual portfolio had a 25 percent allocation to equities, we would take that 10 percent overweight and multiply it by the overperformance of equities and the total benchmark return. If the equities and the benchmark returns are 10 percent with a total benchmark return of 7 percent, then that 10 percent overweight would be multiplied by the 3 percent overperformance, which is kind of a percentage on a percentage.

To get at the selection effects and determine if we had equally weighted or properly weighted our return for the securities that we selected versus what the benchmark had in turn is the second component in this selection effect. So if in our example the portfolio's equity return was only 5 percent because of poor security selection, the 25 percent weight would be multiplied by a negative 5 percent for the underperformance. So in this instance we'd see a positive allocation effect. The manager was good at predicting that the equities were going to outperform, but the negative selection effect shows they were bad at figuring out which of those equities were going to outperform.

Fixed-income performance attribution has many more components. Fixed-income investing tends to be more complex than equity investing. I'm a little biased because that's what I've worked on a lot. There are a lot of different ways that you can take your total fixed-income return and split different components. The hierarchy that I use is very similar to what you might see coming out of the BondEdge system because that's what I used for a lot of years. On a high level, we split the return into an income return and a price return. The income return is just the coupon plus the change in accrued that you would get. Time marches on, and you get the income return no matter what happened in the market. The price return, on a total-return basis, is the return that's generated by a change in the underlying price of the security. That's a bit more interesting. The principal components analysis of bond returns shows that the most important component in telling you how your bond is going to return is what happened to the overall level of interest rates.

So we have an interest rate effect, and we break that into components of level and shape. The level explains if the whole curve moved in a parallel fashion, how would

my price have changed and what my return would be. Now, in reality the curve didn't move in a parallel fashion but actually changed shape a little bit. Shape is my next most important component of the return. How much return is there because the shape of the yield curve changed?

The following effect is the sector quality effect, which says how much price return I achieved because the option-adjusted spreads (OASs) in the sector quality grouping that I'm invested in changed. That's usually the next most important effect. Next is the selection effect. The issue of my individual bond may have some good news or bad news come out about the company. Perhaps my bond did a little different from any other bond in its sector quality grouping. That's the selection effect.

The final thing is the residual, which is what you get when you add up all your different components and compare it to what the total return is. If they don't exactly match, that's a residual. When you do a performance attribution analysis, you really want the residual to be very small. That's when you say, "I can't explain it, although I know what's going on." It generally happens because of data issues. Maybe you traded securities more frequently than you've measured return, or something hasn't been captured properly. Having a small residual is fairly normal. It should offset over time. Having a large unexplained residual indicates that you haven't really done a good job of specifying all the different components of return. There's something else that's happened there that needs some explaining, and you're missing a component.

To help illustrate the fixed-income performance attribution, Chart 2 shows an example of a one-bond portfolio that I bought and held for one month. It shows what that would look like for fixed-income performance attribution. Oftentimes you can get months where absolutely nothing happens. It's really hard to say that in front of a room full of people. You spend a lot of time talking about basis points of return. I tried to pick a month where something happened. Over the summer we had a pretty good month to look at, and I took something that went from the end of June to the end of July. I have five-year bonds showing 60 months here. This is my yield curve from zero to 60 months. The treasury coupon curve at the end of June rose by the end of July. At the five-year point it rose by 92 basis points, but at the short end not much happened. It's even more dramatic up further on the curve.

Table 1 shows the specifics of my performance attribution example. For simplicity I used a flat OAS spread over the whole curve. For the sector quality grouping the spreads tightened by 17 basis points during this period. So the average yield for the five-year corporate increased by only 75 basis points in a month. The Treasuries went up by 92, but because the spreads tightened a bit, the impact on corporate was a little bit smaller. For my one bond this shows where it was at the end of June and where it was at the end of July. I had a little bit of accruing time. You can see my price went down. For the price return I added the yield curve roll effect, which I haven't talked about. It's small and positive, and it is the effect due to the bond

aging and rolling down the yield curve. I didn't want that to end up in a residual and confuse things, so I split it as three basis points.

Table 1

	Market Value	Accrued Income	Price	Duration	Spread Duration	Sector OAS	Bond OAS
30-Jun	1,000.00	-	100.00	4.43	4.43	1.16%	1.38%
31-Jul	978.56	3.13	97.54	4.32	4.32	0.99%	1.10%

<b>Income Return</b>	0.31%
<b>Price Return</b>	
Yield Curve Roll	0.03%
Interest Effect	
<i>Level</i>	-4.08%
<i>Shape</i>	0.15%
Sector Effect	0.75%
Selection Effect	0.70%
<b>Total</b>	<b>-2.46%</b>

As for the interest effect, most of this return is because for the overall level of the yield curve I used as a proxy for the whole yield curve, I just grabbed the five-year maturity, and that went up by 92 basis points. My duration is a little over four years, so I had -4.08 percent as my return because the whole yield curve moved up. As for the shape effect, I actually gained back 15 basis points because the short end didn't move up, but the long end did. The sector effect is because I had a tightening in the market or in that sector quality group. I gained 75 basis points on my return for that and my selection effect because my bond tightened more than an average bond in the sector quality grouping, even though it still has a wider spread; the change is really what I'm interested in. I had another 70 basis points of return because of that.

To do good performance attribution takes a tremendous amount of work for capturing the data, for analyzing the results and for making sure that everything is happening correctly. Very often companies get paralyzed by the fact that there's too much to do and there's too much noise in their data. There are systems out there that can help companies do that. The companies I've worked with that have gone through the whole process have really found it useful for their risk-management work.

I'm going to close with a few comments about performance standards for global investments. In the investment world one of the most important questions is how returns should be communicated to the investor—that is, the external client. What we have been generally talking about so far is how to measure and analyze returns internally in a company. So the Global Presentation Standards (or GPS, from the Association for Investment Management and Research [AIMR]) wouldn't apply to an insurance company, but understanding what an investor needs to know can help



you figure out what is the right way to do it at your firm. GPS is a set of ethical standards for performance reporting, based on fair representation and full disclosure. Doing your performance reporting with the result of having fair representation and full disclosure often means that you need to do more than the minimum required by these standards. They require certain calculations for presentation methods and disclosures. Many of these things concern how you show fees, how many years of performance you're going to show and how you calculate your composites. In a perfect world you and internal management wouldn't really need to worry about these things too much.

They also rely heavily on the integrity of the input data. They apply to methods you use to present the performance of assets managed by a third party. So none of the standards are really only for performance attribution. There are portions that are more relevant and useful for you. I don't think the requirements are terribly difficult hurdles conceptually. For example, "All areas necessary for performance must be captured and maintained": that can be difficult in practice. They ask that performance be shown on a market-value basis, but, then again, these are not the internal accounts that we're talking about. Performance needs to be measured at least monthly, and accrual accounting for fixed-income securities is necessary. So there is not anything that's really surprising. In conclusion, the AIMR promotes professional standards for reporting investment performance. While they largely represent common sense and basic business ethics, it can be worthwhile to work with your investment professionals to compare actual practices to these standards.

**MR. KOLTISKO:** In the example you have of total-return management: do you find that it's possible to use the same sort of information with the investment manager and with the liability manager? Can both get what they need from the same business process? In the chart you had income returns and total returns together. Is it possible to have the same set of reports that both sides can use, or is that a real challenge?

**MS. EHRLICH:** It is a real challenge. The companies that I have worked with that have done performance attribution have gone through it regardless of what else they need, how many different reports they show on book yield and how they measure people's performance. When they have gone through the work of doing performance attribution, they have done it on a total-return basis. If you want to fully understand the impact of the economic environment and the changes in the portfolio and capture it in the period in which it occurred, then I think you have to parallel it with the total return.

**MR. BASIL RABINOWITZ:** Cathy has very nicely laid the groundwork for what I'm going to do. I'm going to discuss pitfalls with investment benchmarks. I've gotten burned by a lot of things—that's how I found out about them—and I'm not going to have a solution for everything, but sometimes knowing about the problem ahead of time can help you avoid the pitfalls. I'll try to point those kinds of things out, but please don't think they reflect on the current company that I work for, which is

Fortis Inc. I did work in consulting before that, and I take a lot of examples from that work.

Benchmarks give you a yardstick for measuring relative performance. We'll discuss why this is appropriate. Why else would you benchmark? Well, you may want a model portfolio that could help you to perform your modeling and determine your sensitivities. You don't want to work with a huge disconnect between the model portfolio and your actual portfolio. It's another way of showing risk control.

The next point is the communication tool. Investment folk typically do not speak the same language as actuarial folk. Many years ago when I first started in this, the problem was much worse than it is today. I think today we're much better off. In the old days it was really a communication problem, and there was distrust between both sides. I think we're far better off today than we were then, but giving model portfolios helps with communication. It puts the issue in investment language.

What considerations are important when we develop a benchmark? What are the ground rules? It must be appropriate for the liabilities the portfolio supports. We need to be careful what we're measuring ourselves against. We should do the appropriate asset/liability management (ALM) studies to make sure that we're setting up the right benchmark to support the liabilities. I worked for a number of years for Thomas Ho at Global Advanced Technologies, where he developed something called ABCD. It's not quite as simple as A, B, C, D, which stands for arbitrage-free bond canonical decomposition. Basically we were able to decompose a liability into assets that look exactly like the liabilities, just using treasuries, caps and floors, swaps and swaptions. If you could take the liabilities and cast them into an asset framework that reflects the same cash flows, you could then go to the investment folks and say you don't understand the liabilities at all, but you understand this in terms of caps, floors and zero treasury bonds, and we could work from there. So you set up a framework that you can use to translate things into instruments the traders already know and love.

The other thing that's appropriate is establishing standards for tradeoffs between risk/reward, and you have to be very careful with that. Risk/reward is very important because you have to know the risk fully.

Clear rules before the period are so critical, and we fall down all the time because there's always something we overlook. Let's talk about benchmark transition rules, for example. Let's say you have a portfolio that has a duration of four, and you do an asset liability study, and you want to change it. Let's assume you're at four and want to go to five over the period of one month. How do I take a benchmark and mirror it against the actual when I'm transitioning from a duration of four to five? The AIMR would probably tell you to take it off the clock. Make the change and exclude that month from the bonus calculation. Management doesn't like to give free reign over a month. So what would be the fair thing to do?. The fair thing is to measure the benchmark at four, measure the new benchmark at five over the

period and take the average between the two. I think that's a fair deal. If you were the portfolio manager, you would probably be satisfied. That shift in the benchmark sounds like it will mimic what we actually want the traders to do.

Now, what about this scenario: the managers change the portfolio from duration four to five over the month, and nothing happens, until one day before the month ends, when they are at five, and the market rates rise extremely. For the first part of a month (when they were at four) the traders were disadvantaged by averaging in the benchmark at five. At the end of the month they were disadvantaged by averaging the benchmark at four, even though we told them to do it. When you're longer and rates go up, you get hurt more; whereas I'm measuring two portfolios, one at five and one at four, and the four portfolio didn't get as much of a hit as the five, and I'm taking an average. It's pretty painful, and that actually happened. You have to try and find as many of the loopholes beforehand. Remember, if you're paying people based on this, you're messing with their rice bowl. That's something you have to be really careful with.

The measuring objective has to be clearly specified. What are you measuring—yield, total returns or customized returns? You wouldn't believe what some of these customized indexes are. We need to know how to measure benchmark portfolios fairly, and that includes knowing what's being measured and over what period of time.

We must have unambiguous benchmark constituents that are very clearly defined. Alignment with the business objectives is also important. If your actuarial staff and other senior management are being paid on the bottom line, that is, some GAAP accounting or whatever measure, and they're very critically driven by income, and you pay your investment folks totally on return, then you have misalignment in the company, and distrust becomes a problem. You want to make sure that you align the investment folk with the businesses, but you have to do it in a way that you don't hurt your investments either. That is not easy to do.

Make sure the accounting measures are unambiguous. You have to write things down. The rules on this have changed a lot in the last few years. If you have an asset that's fallen 20 percent below book, then you have to write it down and take that hit. The rule is that it has to be below that for over a year to be other than temporary impairments, so you get rid of these market aberrations. We have professionals out there managing money, trying to worry about the net investment income budget, and we keep changing the rules on how we're counting this. It's like telling the surgeon to operate while the patient is moving on a boat. It's extremely difficult to have the rules keep changing in the middle of the game.

What about compensation? It's a function of performance versus benchmark that is very critical. It's hard to beat the market. It's nice to be average, but if you ever think about it, 50 percent of the people are below average. It's hard to beat the market average. You want to set up compensation so that they're driven to do the

right thing for the company. It's so important how you structure it. I found that a lot of times portfolio managers whom I worked with had a lot of integrity, and they brought things to my attention that hurt them but were for the company's good. I don't want to say they just do it for compensation. I want to make it very clear I've seen a lot of that, and it's been very heartwarming.

To every thrust there is a parry, but in economics and in the market it's very hard to have that parry in place before you get the thrust. You usually find it out in retrospect.

Now let's talk about realistic constraints. You have to put a lot of constraints on these benchmarks. Some of them are external, some of them are internal, and some of them are inherent in what you've got. You may have an internal loss budget. You may have liquidity problems, and you want to make sure you have enough liquidity. You may have internal risk limits that you've set that could be by sector or by issuer. I don't want to hold more than 2 percent of any one issuer. If he goes belly up, I want to lose only no more than 2 percent of the portfolio. You may even say 1 percent is the most per issuer. Also, I don't want to be exposed too much to lower credits, and I have a duration target that I want. I want some good ALM, and I'm happy if I'm within some portion or some corridor around that, so I could have a duration target with some permissible deviation. Convexity, or prepayment risk, is another thing to think about. So you want to make sure your benchmark reflects all these constraints that you would truly require for regulatory compliance and risk control.

Your benchmark must be measurable and practical. There should be no question as to what the measurement was, and usually publicly available results are preferable. It had better be compliant. There's no use giving a benchmark to your portfolio managers that isn't compliant, and by that I mean you've got to satisfy the rating agencies. They doing a good job, and they come up with interesting things. They might ask you what your mortgage-to-surplus ratio is and tell you that they don't want more than X percentage. So you have to make sure you factor such limits into your benchmarking. You want to make sure you satisfy the formal legal regulators too. South Dakota has a rule that states you can have a 40 percent maximum in corporates. You have to make sure you satisfy whatever state you are in. Internal limits are another thing to look at when benchmarking. Your risk committees in the company set up internal limits, and you have to reflect that in the benchmark.

Additionally, the benchmark must be investable. In Canada, for example, the market is not deep in noncorporate issuance and long mortgages, so it's no use putting a huge amount in that. Even though you need that duration, you can't get there. You can't build a benchmark like that because you just can't get it invested. So doing the theoretical benchmark doesn't help you if you don't consult the people who are actually involved in investments and understand the market. They tell you that what you've done there is just not investable.

I am now going to discuss yield versus total return. When you see this debate, avoid it. You don't want to get caught in this debate. It reminds me of a famous chess game that the master Nimzovich played during the Second World War where two famous German players took him on. At one point in the game they felt that they couldn't do anything and that it was a totally lost game. They were going to resign. So he offered to switch sides with them, at which point they accepted. I think there was some bet on who was going to win. Fifteen minutes later he offered to switch sides again and they accepted again, and they lost. If you want to debate yield versus total return, we'll debate it. If you want to switch sides, we'll debate again and we'll both lose, but there are good reasons for both sides.

Paradoxically we do know that a large part of total return does come from yield. A lot of return on an asset comes from the yield. If you start pushing for yield, the managers will want to play the spread game. They will want to load up as much as possible on whatever they can on high-spread product, and they're going to keep pushing up the risky assets. They're going to invest in stuff that's on the cusp. You clearly don't want that. By putting in total-return constraints you keep one honest because you want to protect capital.

Buy and hold versus active management is another area of discussion. Buy and hold is seen as being the yield side, whereas active management is seen as total return. I would say that people who operate in the true buy-and-hold fashion expect the portfolio not to churn very much. If rates are dropping, they don't expect the portfolio to go toward market rates quickly. If you have a four-duration portfolio, you can see 15 to 20 percent of your portfolio maturing every year without any active management. When you add new money coming in from coupons to that, hopefully adding to cash flow, you will see a yield go down even without any active management. So that argument doesn't hold completely. Keep it simple. Things tend to become obfuscated. We lose effectiveness when simplicity isn't there.

Small customized benchmarks versus indexes is another topic to look at. With small benchmarks that are customized with a few bonds in them, you learn that if one of those bonds goes south, your whole index doesn't look so good. So you want a lot of bonds in there. You can't invest in the index. You've got a lot of little positions that you cannot hold, but you have to take some kind of tradeoff. You want to customize the asset allocations and weight the pieces correctly. Asset allocation accounts for the majority of a long-term return. Are there appropriate indexes for these things?

What real estate benchmark would you use? The debt ratio target of your company, the appetite for leverage and the way you structure the deals affect it. Are there good indexes you can use for commercial mortgages? What if you don't have an index? Bench it against itself at worst.

What about alternative investments? Not just hedge funds, but the interesting miscellaneous opportunities that come up, for instance, premium tax credits. You

don't want to avoid such investments completely, but they don't easily fit into the process. To whom should that credit go? I'm getting invested in a bond that doesn't pay any dividends, or interest, but on the premium tax side I'm getting a credit. So whom does that reflect on? Do I get credit for making that asset allocation that's going to show up as a zero on me?

I can't use a single index, like the Shearson-Lehman. This thing has 37 percent treasuries. If you need income, you don't want 40 percent of your portfolio in treasuries, 19 percent in agencies and a duration of 5.5. It may not select what you want at all, so you customize this. Cut it up and subindex it by different maturities, weight the different slices; cut it up by different maturity or yield curve, or diversify by credit. But the more you cut it up and the more you customize it, the less the return is a public number, and the more you have to calculate it yourself. Customize the duration but use subindexes.

Have compliance tolerance corridors around the weights that you set, so that the portfolio can deviate from the benchmark. You have to encourage them to take a slight deviation. If the market allows them an opportunity, they should be able to take that opportunity when it arises. You don't do benchmarks every month. They're usually done annually at best. Again, diversify across a yield curve. The challenge is to put in the benchmark what you clearly need, for risk control, but not to overspecify it so there is no room for the trader to add value. If they want to put in a strategy of barbell/bullet because they think the curve's going to move, that's the portfolio manager's strategy. Your benchmark is what the duration should be.

Now I'm now going to talk about large issuers. Sometimes Ford is 3 percent of corporate bond issuance. You don't want exposure to that. You want to trim down those large exposures. If you trim Ford down, then some of the other smaller ones who are below the 3 percent suddenly pop up because now the trimmed-down version is more, and you have to trim them down as well. It becomes circular. You have to be careful not to violate your own issuer constraints. For the small issuers, as I mentioned earlier, that violates the investable constraints. You can't invest in all of them, and so we go around in sort of a circle.

Customize the credit exposure. You want a credit-convexity tradeoff. What do I mean by that? Prepayment in mortgage backs, because mortgage backs have good credit, so it's not bad to take on prepayment risk when you're diversifying your risk away from credit risk.

Next let's talk about liquidity. At the end of every year the market becomes kind of illiquid in the corporate bond market. They lock down, and they're happy with the returns they got. At the end of the year you can't trade too well, so you get aberrations. Sometimes it's good to look at longer-term rather than just one-year kind of measures. When long-term capital blew up and you had the whole Russian-Asian crisis, you couldn't get bond prices; they weren't trading. At one point we couldn't get a quote on a Treasury. It was unbelievable. So at those points you

can't measure that performance. In fact, even if you would get a quote, it was just an artificial quote because no one would buy it for you. It loosened up after a month or two, but at that point it was very difficult to measure things.

Finally, attribution: Cathy spoke a lot about it. You want to know where value was created. Her examples make sense. Using the benchmark, you need to communicate to management and the traders about what was done to create or destroy value, be it sector bets, security selection, asset allocation and timing.

Classification can be a source of other problems you run into. This is when you have a material impact because of simplistic reporting practices. Sometimes spare cash is parked in money market mutual funds, and it shows up as equities. The traders have no allocation to equities, but short term, the best place to park that was in the money market mutual funds. Worse than that, you've got a period that you've put it in for, and you put it in after the period started and took it out before the period ended, so if you look at your average assets, you had nothing to start with and nothing to end with and zero assets in the asset pool, but you've got some income. What's your return? It's infinite. By looking at averages over a period, you run into the same problem. So you also have to be aware of ad hoc management decisions and accounting rule changes. Also, the rating agency focus keeps changing. This has been an overview of the practical issues you will face in using benchmarks.

**MR. FRED SHEN:** I came from the insurance side. I used to work for ManuLife, which is now the sixth largest insurer in the United States after its acquisition of John Hancock, and I can sympathize with what Basil and Cathy said in their presentations because I was responsible for performance measurement for Manulife. I then went on to banking, and did it work out any better? Well, maybe; you can see for yourself. My objectives are to introduce traditional performance measures, introduce Bank of Montreal (BMO) measures, review capital measures, introduce the components of capital measures, examine the effects on decision making, describe some applications at BMO and give the advantages of using a capital-at-risk measure.

There are different kinds of benchmarks, such as peer group universe. The Standard and Poor's (S&P) 500 index is another kind of benchmark. Typically we define superior performance as something that beats the benchmark, which is something very simple. You may beat it at a basis point above the performance of an index or above medium in a peer universe. However, I'm actually in finance; I'm not an actuary.

Finance theory tells you that reward and risk go hand-in-hand. A lot of people don't understand that. There are many different ways to measure rewards. At the insurance company we went with economic value added as our measure. Basically it's the number of basis points by which you beat the benchmark times the assets under management. That's a very simple dollar value that you can show in terms of value added to the company.

What about risk? We talk about returns and we talk about benchmarks, but what about risk? Cathy already mentioned why we need risk-adjusted measures. Let me give you a few more reasons. Peer group measures are usually not risk adjusted. You think your peer group has the same risk, but you really don't know. Cathy mentioned that Treynor and Jensen typically apply to equity investing. There's really no equivalent measure in a bond or fixed-income world. Sharpe measures the volatility, and that's a better measure, but again you have to ask yourself if the volatility is the correct measure of risk. You have duration-of-bond portfolios, tracking error and a whole bunch of different elements of risk that are missing. No one ever talks about statistical significance. Why is that? Does beating a benchmark by five basis points rank it as significant? Did you really outperform, or was it by chance? You are actuaries. Statistics plays a big role in your world, such as mortality tables and all of that. Is it really relevant, especially when the S&P 500, for example, has a volatility around 20 or 25 percent? How much of a chance was it that you beat the index by five basis points and you consider yourself a superior performer? That's a good question.

Let me go to something that we use at the bank. We use something called capital at risk (CAR); it's the BMO term. Basically we have certain value-at-risk banking principles embedded from the top down at the bank. All of that drives the strategy, capital allocation, performance reporting and reward programs. Investment communication is very important in all of these measures. You really want to be sure that the strategy that you choose has enough capital allocated for you to execute successfully. If you're successful, or even if you're not successful, how do you measure it, because very often people don't measure what happens after an acquisition. There's no postperformance measurement or postacquisition measurement.

Why do we use CAR? We think that it's comprehensive and captures all aspects of risk. If you think you know risk, maybe I can change your mind. We quantify the risk that an entity will take. An entity could be a portfolio, a liability or a subsidiary. The strategies that we employ now actually have the cost of that risk taken into account. How do we define CAR? What we look at it is the maximum unexpected loss in value that a line of the business can incur over a particular time horizon and a defined confidence level because of any and all types of risk. For example, one year is a standard timeline for the bank. A typical defined confidence level could be 99.5 percent. One reason for that is because BMO is an AA-rated bank, and so for us to maintain that AA rating, we need capital that supports the risk that we take at this risk level. If we go down to a 95 percent confidence interval, we hold less capital, which means we are less able to protect our investors from an adverse shock to our assets. That means our credit rating will be downgraded by the rating agencies. A graphical form of CAR can be seen in Chart 3. I've been told that I actually reversed what a loss distribution looks like. They're right, but to get the graph to go from left to right, I had to do it this way. It's for simplicity.



Why do we need capital? Let me give you some perspective. There's regulatory capital, whether it's an insurance company or whether it's a bank. In Canada you're governed by the Minimum Continuing Capital and Surplus Requirements. I'm sure you're governed by the same in the United States. What the regulators do is unfortunate: they target an average bank. They don't recognize that you have more mortgages and fewer trading activities, and therefore you are less risky, so you can have less capital. They say, "You're an average bank, here's what an average bank capital should have, and these are the numbers that you will use to apply to your regulatory capital." It's a very crude measure, but it is a measure that we have to face.

Rating agency perspectives are probably a little bit better. They look at certain factors in their income statement and balance sheets and decide, based on those factors, how much capital you need. From our perspective, we use CAR. We think it provides the correct economic signals to the portfolio manager, senior management and outside world about what our proper risk measures are.

I want to give you a definition of CAR: capital at risk is the internal capital measure used at the bank. It measures the future potential variability of the value of a business. For a given target debt rating, this is also a measure of the capital required to sustain the business—that is, the amount of capital that the bank must hold to cover its risks and ensure that debt holders are not exposed to potential credit rating downgrades at a frequency greater than that implied by the bank's current debt rating.

What components of risk are we measuring when we use CAR? We all have proxies. I'm going to give you the components of CAR that we use at BMO. First and foremost, of course, is credit capital. As you know, one of the biggest things the banking institutions do is lend money. As a portfolio manager investing in fixed income, one of the biggest things that you should be doing is looking at credit and ensuring that the credit of your portfolio meets the particular targets that the liabilities holders think and the investment officer thinks you should have. So credit capital in our case is an appropriate measure of the credit risk of a portfolio of assets. It doesn't matter whether the assets are equities, whether they're fixed income or whether they're loans. We can measure credit capital.

We also have market risk capital. We measure market capital much like CAR because, as you've seen, prices fluctuate up and down. You can lose a lot of money in a very short time. BMO also runs a trading portfolio; at ManuLife we ran a trading portfolio. We actually ran a trading portfolio of minihedge funds. That takes on a lot of risk. So you have to be able to measure the losses from activities in the market because of changes in interest rates, points change, spreads, commodities prices and so on. At BMO we actually have two sources of market risk. There is one that relates to our trading activities—we took a bet in one direction or another—and a second one that relates to our structural book, which is something that you're probably more familiar with. Our asset liability management group says how much

structural market risk we should have, because as a bank we have liabilities that are depositors and note holders, and we have assets that we invest to meet those liabilities.

The other two aspects of risk that people often don't remember or don't recall would be operational and business risk. I am going to give you a quick sense of operational risk and why you might need to include it even if you are a portfolio manager. We did this at ManuLife, for example. When we bought a life insurance company in Tokyo in 1988, one of the things that we had a problem with was liability guarantees. The yield targets were anywhere from 100 to 150 basis points over five years with no penalty for termination or early withdrawal. The problem was with assets in Japan. Five-year JGBs were yielding 50 basis points, so you had to go down the credit curve. If you don't know the market, how are you going to get the yield? One of the things that we tried to do was buy U.S. dollar bonds and swap them back to yen. Why? We did it because we know U.S. dollar markets better than the Japanese markets. If you ever operate in the Japanese market, you will see that it is a maze of confusion. At least to me it was.

If you operate in the U.S. market, you know your credit. You know Nortel: Nortel bonds at that time were yielding around 300 basis points above comparable treasuries. That is a good yield. It is a great yield when you swap it back to yen. You get the yield required to meet the liability targets. So you did the transaction. Were there additional risks in that transaction besides the market and the credit that you have encountered? The answer is yes. When you use currency swaps, you have a derivative group that now has extra transactions it must get into. You have counterparty risk that you may not have taken into account. So there are other risks that exist by employing different types of strategies to boost your asset yields, such as operational risks. Even the asset-backed market or the mortgage-backed securities market has risks that you don't think of. You just think of yield.

Finally there is business capital. Why do we need business capital? When you think about it for a moment, the last two or three years have been really bad for equity markets. Most of you have probably heard of Charles Schwab, a discount brokerage dealer. Most of you probably trade through it. Therefore, when the markets crashed, bonds plunged. That's business, right? Competitors come in. TD Waterhouse is another discount brokerage, competing directly against Charles Schwab. So competition, markets or anything that drives volumes down in your business is something that we consider as a business risk.

Let's quickly discuss the different methodologies and how we look at market risk. We take a value-at-risk approach. We do scaling, and we take a look at stress testing. Stress is important because we're talking about events that occur in the tail of the distribution. One thing that we do know about the tail of distribution is that we don't know very much about it because the events are very rare. So we apply extreme value theories to model the tail. That's part of what I do.

With components of CAR for credit, we recognize collateral. If we hold collateral, then we will offset that against the credit capital for a particular holding. We have the data input, a modeling tool and certain assumptions. We have a distribution, and we can calculate very simply credit capital.

Because business risk is driven by competition, it's how your volumes and margins change over time. We model it as changes over time. The way we look at operational risk capital at this point in time is to look at event frequency versus severity. Event frequency we model as a Poisson distribution, so it comes in regular cycles. Severity we model as a lognormal distribution. We get a final distribution that we simulate for operational risk for BMO.

Now I will very quickly go over the mechanics. There is a very simple measure of return on capital. Cathy talked about RAROC measures, which are exactly what we use. Essentially the return on capital is net income minus the expected loss minus the cost over the appropriate capital measure. It's a very simple calculation. We use CAR in all of our decisions, and we hold every portfolio manager and every line of business accountable for the amount of capital that we need to allocate to the line of business. We do that because we're able to say to them, "The capital measure that you have is really how much risk you're going to run in your business. We have the fully allocated cost to your business, and for us to make the right decision we still have to calculate a correct measure that compares to alternative investments because capital is scarce." So that's how we can hold the line of business accountable.

Chart 4 is an example of the highest level the bank sees in terms of CAR. We call it shareholder value added, and we identify which businesses contribute most to the value of BMO, all the way down to businesses that don't add any value and are businesses that we should be looking to divest. So we actually have these kinds of reports that go up to the board of directors that show where value is being added.

Another way of looking at it, as shown in Chart 5, is to examine our return on equity less the cost of equity. We call that the net economic profit approach. We can say that 33 percent of our capital generates three-quarters of the profits of the firm, and 33 percent of the capital at this end loses money for the firm. It gives us a good indication of where we should be going with respect to marginal investments.

Where do you see it at BMO? At the highest levels we report on a quarterly basis capital that is segmented by risk and segmented by business. As an example of the segments, 30 percent of its capital comes from credit, 26 percent is due to market and 23 percent is because of operational risk. So the board gets a very good oversight as to where the risks are being generated, and that's our capital viewpoint.

Chart 6 is another example of reporting to the board. This is my means of telling them where we are compliant with the principles and the policies set down and

where we're not. So what you see in the middle tones of gray means this is a great methodology and you should trust the number. The light gray indicates that it is an okay methodology. It is saying this is still pretty well state of the art, but there may be some questionable data issues, and that these are still pretty good numbers that you can rely on. The dark tone definitely means it is a bad number, a bad methodology or bad data. So the board sees in an instant who complies and who doesn't, and this is a real attention grabber. Where are your eyes focused? They are focused on the dark areas. Those are elements of areas that we have to focus on in terms of getting our risk exposures aligned.

I want to now go over some of the advantages of CAR. It's a very precise measurement, and it's quantified. It actually helps us in our RAROC; we use that in our loan decision process. So when a client comes to us and says he or she wants a loan, we calculate the RAROC. Does it meet the minimum hurdle that we require for return on capital in the bank? If not, that client is rejected. By the way, there should not be any arbitrage in a sense that if a client comes to our corporate group wanting a loan and the corporate group turns them down, the client shouldn't be able to go to the commercial group, which is in the retail part of the bank, and get the loan because of a more relaxed capital measure. That's not acceptable. We have instituted that RAROC measures and our measures across the bank so that they are all consistent.

One good thing about this measure is it allows us to add up all the risks. It's a dollar measure. You add up total dollar market at risk, you add up total dollar CAR, you add up all the dollar measures, and that's the total amount that you need to have to protect yourself. One of the things I wish to emphasize about CAR measures is they have to be driven, and they require a culture that is driven from the top of the house. CAR is definitely not driven from the portfolio managers and not driven from the lines of business. It is something that the top of the house must believe in, and it's driven down to all the lines of business and the portfolio manager.

So because of all of these good things that capital has, we are able to do that. The key is that the highest levels of the bank agree with our approach. Not only that, the highest levels of the bank have said risk belongs not to the risk-management group where people think risk belongs, but to the people that generate the risk. The people that generate the risk must understand where it comes from, and they must take responsibility for it. Once that's understood, they realize they need to know their risk and that they had better go calculate those capital numbers appropriately.

BMO has gone to approximately 96 percent reliability, which means we have good numbers. We're actually going through a dynamic measure of capital where we can actually measure it monthly, quarterly or whatever cycle that you choose. That's a requirement of a good performance measure.

Chart 1

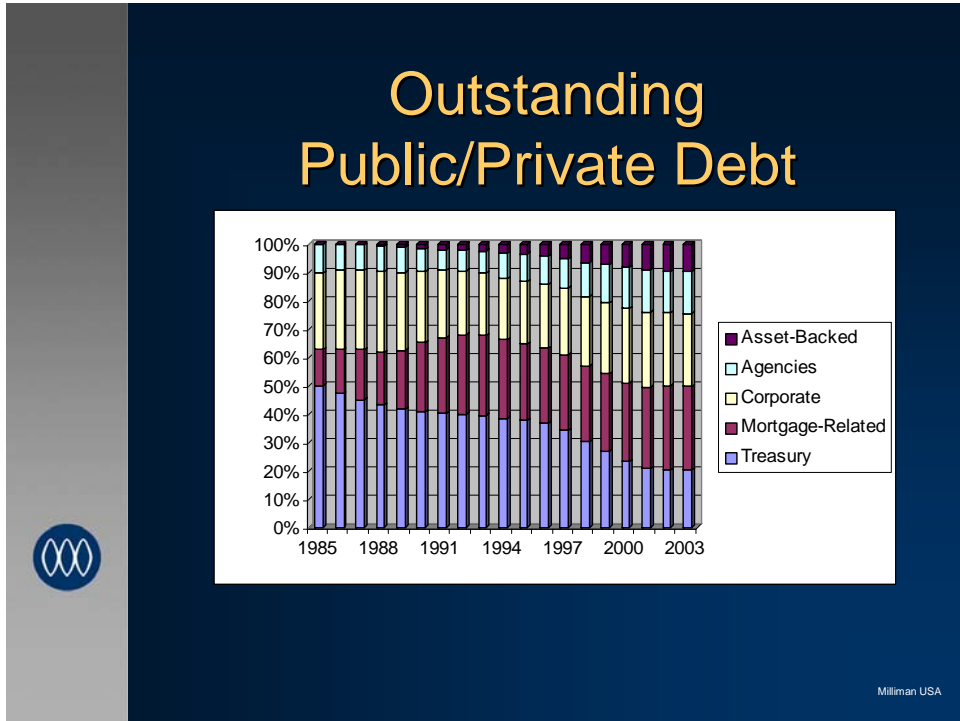


Chart 2

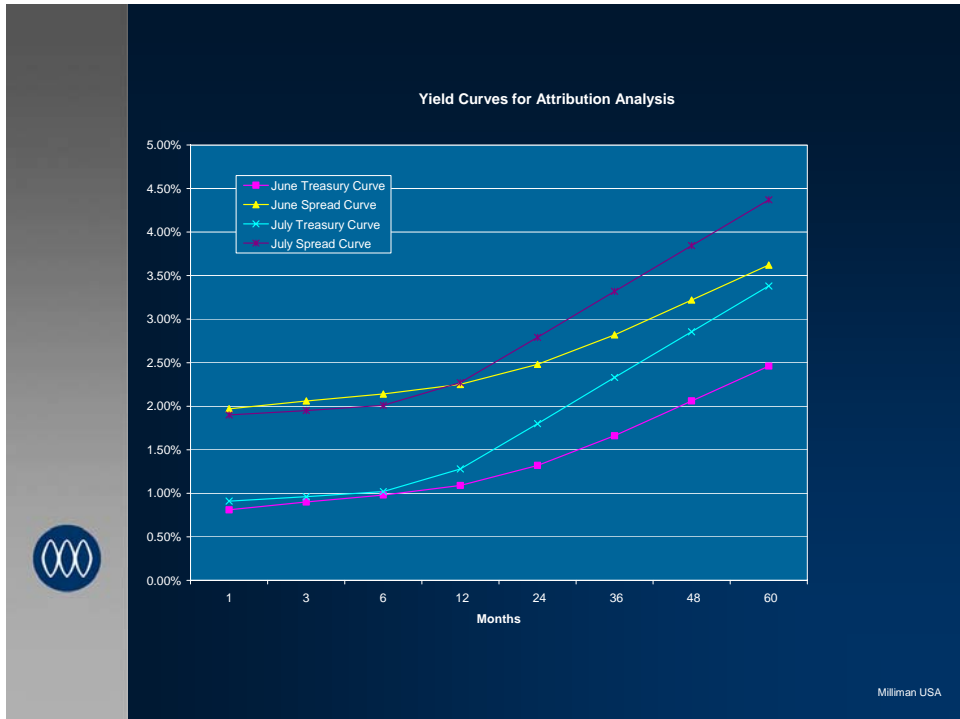


Chart 3

### What is Capital at Risk ?

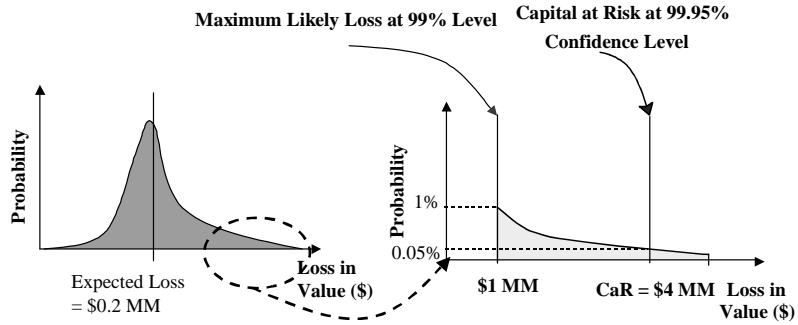


Chart 4

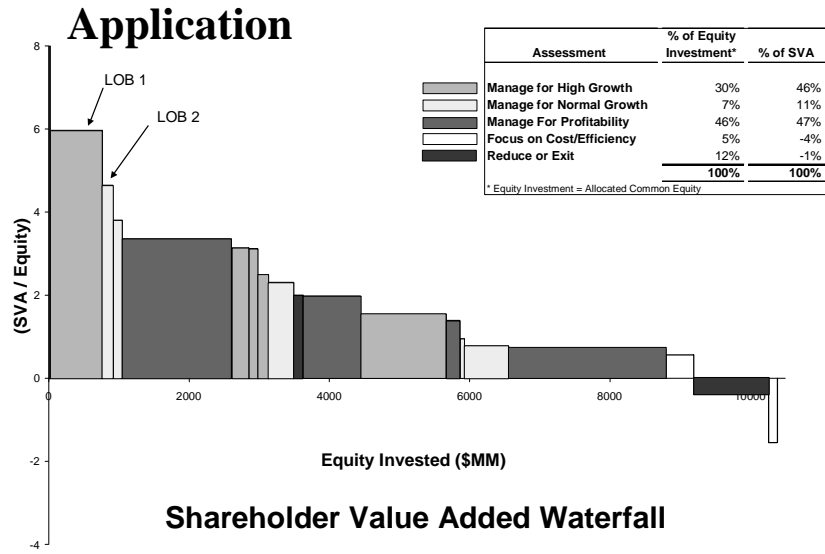


Chart 5

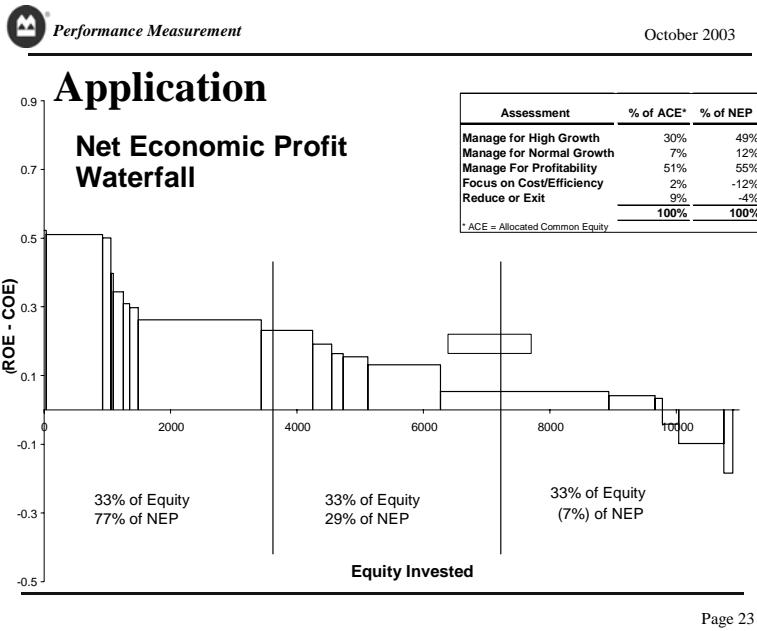


Chart 6

**Performance Measurement** October 2003

### Application: Capital at Risk Reporting

	Credit	Market	NDMR	Operational	Business	Equity	Insurance	Total
Auto	29		9	6	4			48
Capital Mgmt	45		13	20	1			79
Direct Lending	182		49	27	42			300
<b>Consumer Lending</b>	<b>256</b>		<b>71</b>	<b>53</b>	<b>47</b>			<b>427</b>
<b>Mortgages</b>	<b>17</b>	<b>150</b>	<b>40</b>	<b>31</b>	<b>37</b>			<b>275</b>
Electronic Banking	189		56	80	36			361
Processing			10	30	18	73		131
Corporate Cards	5		3	5	3			16
Miscellaneous	2		1					3
<b>Cards</b>	<b>196</b>		<b>70</b>	<b>115</b>	<b>57</b>	<b>73</b>		<b>511</b>
Commercial Lending	186		30	20	9	1		246
Small Business	167		65	61	66	48		407
<b>Consolidated</b>	<b>353</b>		<b>95</b>	<b>81</b>	<b>75</b>	<b>49</b>		<b>653</b>
Insurance	2	18	29	84	25		29	187
Everyday Banking			19	51	45			115
Broker Services			6	28				34
Direct Banking			5	24				29
Distribution			86	431				517
<b>Canada</b>	<b>824</b>	<b>168</b>	<b>421</b>	<b>898</b>	<b>286</b>	<b>122</b>	<b>29</b>	<b>2748</b>
U.S. Banking	95		81	210	68			454
U.S. Retail	40		7					47
<b>U.S.</b>	<b>135</b>		<b>88</b>	<b>210</b>	<b>68</b>			<b>501</b>
<b>Group Total</b>	<b>959</b>	<b>168</b>	<b>509</b>	<b>1108</b>	<b>354</b>	<b>122</b>	<b>29</b>	<b>3249</b>

Page 25