The Evaluation of Target Date Funds

Recent changes to pension regulations in the United States (the Pension Protection Act of 2006, to be precise) have led to the promotion of target date funds as a default option (a so-called qualified default investment alternative, or QDIA) for DC pension plans. Target date funds are mutual funds that (as the name would suggest) are associated with a particular target date that roughly represents the retirement date of a typical investor in the fund. The exposure to equity is reduced over time, from a weight of 80% to 100% forty years before retirement to a weight of 40% to 60% at retirement. The allocation strategy is therefore dynamic and multi-period in nature. Traditional performance measures such as the Sharpe Ratio, Jensen's alpha, Treynor's measure, the Sortino ratio or the M^2 measure are based on single-period optimal portfolio theory, and are not appropriate for a multi-period strategy[1]. Indeed, since the classic work of Robert Merton [Merton, 1973], it has been known that the optimal multi-period strategy will generally deviate from a series of "rolled over" optimal single-period strategies (a so-called myopic multi-period strategy).

Within the mutual fund industry, several consultants and academics have proposed methodologies for the evaluation of target date funds:

Lipper: **Method #1** [Labovitz, 2006]: Block-bootstrapping portfolio returns along the glide path, to produce a wealth distribution at retirement that can be evaluated as to its moments, lower partial moments, VaR, CVaR, certainty-equivalent wealth, etc. **PROBLEM**: The method uses historical returns, and most in the investment industry believe that the premium paid for market risk will be lower going forward than it has been historically. The evaluation methodology may therefore be biased.

Method #2 [Clark, 2007]: Use of the Bradley-Terry statistic, which has been used (for example) to rank professional tennis players. (The method was actually invented by Zermelo in 1929, and re-discovered by Bradley and Terry in the 1950's.)

[1] They may be appropriate, however, to evaluate the value added by active management, since the relative value added by active vs. passive management tends to have a one year horizon.

PROBLEM: The payoff function for a pair comparison is symmetric (a zero-sum game), rather than penalizing the loser more severely, which reflects no adjustment for risk. The fund with the highest average return ex post will tend to win, and in a normal market with funds that are essentially efficient, this would be the fund with the highest market risk. Also, there is no attempt to tie the evaluation to a long-term wealth objective.

Method #3 [Turowski, 2007]: Use of paired-comparison digraphs (PCD's) to rank target date funds. This method is very similar to method #2 above.

PROBLEM: Same as for Method #2 above.

Morningstar: Three target date categories: 2000 – 2014, 2015 – 2029, and 2030+. Within a category, funds are ranked by total return over various horizons.

PROBLEM: In a normal market, a shorter-dated fund in a given category will tend to get a lower rank, by virtue of its reduced market risk. Also, even for funds with the same target date, there is no attempt to risk-adjust returns, so that in a normal market, the higher risk funds will tend to outperform. Again, there is no attempt to tie the evaluation of the fund to a long-term wealth objective.

Target Date Analytics (TDA) [Israelsen et al., 2008]: Aggressive, Moderate, Conservative and Defensive glide paths are proposed as four different solutions for investors of different levels of risk aversion.

PROBLEM: The methodology is only focused on the accumulation phase of the lifecycle. As a result, for the Conservative and Defensive glide paths, the equity weight immediately pre-retirement is at a level that cannot sustain a reasonable withdrawal rate (say, 4.5% to 5%) for a retirement of reasonable length (say, 25 years). Hence, for the full lifecycle, the strategy has a jump discontinuity at retirement, which disagrees with all academic lifecycle consumption and portfolio allocation models. Also, using historical returns, it can be shown that the Conservative and Defensive glide paths are second-order stochastically dominated by the Aggressive and Moderate glide paths, i.e. no rational, risk-averse investor would choose them. For the Conservative and Defensive glide paths to be reasonable choices for a rational investor, therefore, the forecasted equity risk premium must be much lower than has been the case historically.

Turnstone Advisory Group, LLC [Nagengast et al., 2007]: Use of a multidimensional scoring function for each target-date fund is recommended.

PROBLEM: This solution is structurally quite close to the "correct" one. The problem is that the five components of the scoring function are highly flawed. The first two address the agency conflict between the investor and the fund management company, with the first focused on the adverse selection problem and the second focused on the moral hazard problem. These are legitimate concerns, but the score itself is suspect (for adverse selection, for instance, no signal to communicate superior type is taken as a bad signal). The third component scores the glidepath, but is of a fuzzy, qualitative nature. The fourth and fifth components address performance and risk, respectively, but performance is restricted to the long-dated funds, and only outperformance of active management net of fees is considered; the fee is therefore being "double-counted" (it's incorporated in the second component as well). The risk score only considers the short-dated funds, and the score averages the average shortfall and the maximum shortfall relative to a passive benchmark, reflecting a high aversion to active risk (a.k.a. tracking error), penalizing active management. Note that there is no attempt to incorporate a long-term objective into the scoring function.

Dow Jones: Real (Dow Jones Real 2010 through 2045 Portfolios) and Nominal (Dow Jones 2010 through 2045 Portfolios) glide paths

PROBLEM: This form of benchmarking is the traditional, backward looking one. Each target date has a benchmark portfolio which is a weighted average of indexes; the real indices have a higher weight in real (as opposed to financial) assets. The fund ranking does not incorporate the long-term, forward looking fund objective.

Bodie and Treussard [Bodie and Treussard, 2008]: Optimize over weighted averages of the prospective target date fund and a risk-free target date fund, determine (for various levels of relative risk aversion and beta of income risk) the certainty equivalent wealth of the target date fund alone and the optimal combination, and express the former as a percentage of the latter.

PROBLEM: Creating a risk-free fund for different target dates is problematical, since inflation-indexed bonds (TIPS in the U.S.) only trade with at-most 30 year maturities, with real durations of roughly 20 years, so a retirement date more than 20 years in the future cannot be immunized against real interest rate risk. Even if one could do this, a synthetic CPI-U indexed T-bill maturing at retirement isn't really a risk-free investment for a prospective retiree (cf. [Huang and Milevsky, 2008] in this regard). Although the CPI-E index is more appropriate, the use of a single basket of goods to measure the price level for a retiree is based on the assumption that the investor has homothetic preferences, which is almost certainly ***false*** (preferences are homothetic if, when income goes up/down, the investor buys more/less of ***exactly the same*** bundle of goods). In reality, an investor's bundle is made up of **necessary** goods (whose fraction of the bundle decreases/increases when income goes up/down) and **luxury** goods (whose fraction of the bundle increases/decreases when income goes up/down[2]). The usual "excuse" here is to say that, unless this assumption is made, the problem is intractable. By optimizing, though, you get the right solution to the wrong problem. Also, the "traditional target date fund" is taken to implement a straight-line equity reduction over time, which is demonstrably sub-optimal (cf. [Cocco, Gomes, and Maenhout, 2004], p.521-526).

So, although the mutual fund industry has attempted to introduce a methodology that will rank target-date funds, none of the proposals to-date has really addressed the multi-period objective(s) of these funds, and how well these funds would be expected to do in attempting to meet these objectives.

GOALS: The evaluation of Target-Date Funds is essentially a problem in auction theory, in which a buyer must score competing bids according to a multidimensional scoring function, and then choose the bid with the highest score[3]. This is the same as for conventional mutual funds (although I have never seen mutual fund evaluation

[2] The homothetic preferences assumption is necessary, in order to utilize the notion of investor risk aversion, which is ubiquitous in the optimal consumption/portfolio choice literature. See [Kihlstrom and ,].
[3] The paper [Che, 1993] is one of the first papers of the rather small literature on auctions with multidimensional scoring functions.

approached from this perspective). In the case of conventional mutual funds, the time horizon is either implicitly or explicitly a single year, and the scoring functions are well known (the ones mentioned earlier – Sharpe ratio, Treynor measure, etc.). For multi-period strategies, the scoring function is less clear, since we must include not only scores based on past performance, but also on potential future performance. The first goal, therefore, is:

1) To devise a methodology (essentially, define a scoring function) that will enable a financial advisor, armed with an investor's level of risk aversion and income risk (and any other parameters deemed to be appropriate), to rank the available target date fund suites.

An individual investor, together with a financial advisor, will weight the component scores of the scoring function so as to arrive at a one-dimensional score to use in ordering the funds from best to worst. Note that the scoring function may need to include an evaluation of the agency conflict that exists between an investor and a fund management company (both adverse selection and moral hazard).

For target-date funds to be used as the default option for a defined contribution pension plan, there is the further problem that the DC pension plan sponsor must convert the multidimensional scoring function into a one-dimensional scoring function, not knowing the appropriate weights to use for each component of the scoring function, or possibly not even knowing one or more components of the scoring function (which may depend on specifying investor characteristics). The second goal, therefore, is:

2) To devise a methodology that will enable defined-contribution pension plan sponsors and pension consultants, without the luxury of specific information regarding investor risk aversion, income risk, etc., to rank the available target date fund suites for use as the default option in a DC pension plan.

With billions of dollars being invested in target date funds each year, much of this ***by default***, an objective methodology for the evaluation of these funds is desperately needed by the investment management industry (cf. the response [Smith, 2007] to [Bodie and Treussard, 2007] for an indication of the desire of financial advisors to resolve this issue).

REFERENCES

Bodie, Zvi and Jonathan Treussard, 2007. "Making Investment Choices as Simple as Possible, but Not Simpler," *Financial Analysts Journal* **63**(3), 42-47.

Che, Yeon-Koo, 1993. "Design Competition through Multidimensional Auctions," *Rand Journal of Economics* **24**(4), 668-680.

Clark, Andrew, 2007. "Ranking Target-Date Funds Using Past Performance," Lipper Research Series Fund*Industry* Insight Report, November 6, 2007.

Cocco, Joao F., Francisco J. Gomes and Pascal J. Maenhout, 2005. "Consumption and Portfolio Choice over the Life Cycle," *The Review of Financial Studies* **18**(2), 491.

Huang, Huaxiong and Moshe A. Milevsky, 2008. "Should Retirees Hedge Inflation or Just Worry About It?" Working Paper, IFID Centre, March 11, 2008.

Israelsen, Craig, Joe Nagengast and Ron Surz, 2008. "Pure Target Indexes from Target Date Analytics LLC (TDA)", posted on www.tdbench.com, January 10, 2008.

Kihlstrom and , 1984.

Labovitz, Mark L., 2006. "Targeting Target Maturity Funds," Lipper Research Series Fund*Industry* Insight Report, June 28, 2006.

Merton, Robert.C., 1973. "An Intertemporal Capital Asset Pricing Model," *Econometrica* **41**, 867-887.

Nagengast, Joseph C., Daniel Wong, Craig L. Israelsen and John M. Bucci, 2007. "Popping the Hood II: An Analysis of Target Date Fund Families," Turnstone Advisory Group LLC.

Smith, Jonathan, 2007. "Making Investment Choices as Simple as Possible, but Not Simpler: A Comment," *Financial Analysts Journal* **63**(4), 14.

Turowski, Hank, 2007. "Using Paired Comparison Digraphs to Rank Target-Date Funds," Lipper Research Series Fund*Industry* Insight Report, November 13, 2007.