Using Trading Costs to Construct Better Replicating Portfolios

Curt Burmeister
Risk Solutions, Algorithmics Inc.

Helmut Mausser
Quantitative Research, Algorithmics Inc.

Oleksandr Romanko
Quantitative Research, Algorithmics Inc.

March 2010

Abstract

Regularization, by means of trading restrictions, is an effective way to obtain sparse replicating portfolios. By including only the most relevant replicating instruments, the resulting portfolio is more efficient computationally, easier to interpret and better able to approximate the liability on an out-of-sample basis. The instruments selected depend largely on their respective trading costs. This paper evaluates a number of alternative methods for specifying trading costs. We find that trading costs based on simple statistics of the instrument and liability cash flows are an effective choice in practice.

Keywords: replicating portfolio, regularization, trading restriction, trading cost

1 Introduction

Calculating economic capital is a task that lies at the core of enterprise risk management (ERM). Using ERM to improve operational results is premised on (a) the accurate quantification of the risks faced by an organization and (b) the ability of risk to be quantified in a timely manner, thereby enabling risk-informed decision-making. For insurers, the growing trend towards market-consistent valuation as the basis for computing economic capital presents unique computational challenges. In particular, the reliance on Monte Carlo methods for both liability pricing and obtaining an empirical loss distribution for risk assessment (Solvency II, for example, requires insurers to hold an amount of capital sufficient to withstand 99.5% of the possible losses over a one-year horizon) leads to nested stochastic simulations that are extremely time-consuming. A replicating portfolio lessens this computational burden by acting as a proxy for the liability during risk calculations; effectively, the liability is replaced by a set of standard financial assets/instruments that can be priced quickly using analytical methods. Replicating portfolios are constructed by using optimization to find the asset positions that best match certain characteristics of the liability. In doing so, however, the resulting portfolios often contain an excessive number of assets, which degrades computational performance, obscures economic interpretation, and reduces the effectiveness of the replicating portfolio out-of-sample. Regularization is a technique that mitigates these problems by encouraging the selection of only the most important replicating assets during optimization. This paper investigates the effectiveness of various forms of trading costs for regularization purposes when constructing replicating portfolios with linear and quadratic optimization algorithms.

Consider computing economic capital for a portfolio of liabilities at some time horizon τ (e.g., $\tau = 1$ year for Solvency II). Using Monte Carlo simulation, this calculation might proceed as follows:

1. Generate a set of S_{rw} real-world scenarios representative of future economic conditions.
2. Compute v_s^τ, the fair market value of the portfolio of liabilities at time τ in scenario $s = 1, 2, \ldots, S_{rw}$.
3. Compute r_s^τ, the fair market value of the assets at time τ in scenario $s = 1, 2, \ldots, S_{rw}$.
4. From the S_{rw} sampled losses ($v_s^\tau - r_s^\tau$), construct an empirical loss distribution and estimate the economic capital as a given quantile of this distribution.

The assets contained in a typical insurance company's investment portfolio can be priced efficiently in Step 3 using analytical models. However, obtaining each v_s^τ in Step 2 requires further stochastic simulation, over a set of S_{pr} scenarios, to accurately capture policyholder behavior, management decisions, embedded options and guarantees, and any other such scenario-dependent cash flows. Thus, generating the liability cash flows used for economic capital calculation is expensive, as it requires running complex actuarial models across a total of $S_{rw} \times S_{pr}$ scenarios.

Because S_{rw} needs to be large in order to estimate the economic capital[1] the resulting computational cost of Step 2 may be prohibitive. In practice one is faced with the prospect of either using fewer real-world scenarios, which increases the quantile estimation error, or reducing the time spent computing v_s^τ, which may introduce pricing errors. Replicating portfolios are representative of the second approach, in that they replace the liability with a proxy portfolio of financial assets that can be priced quickly using analytical models. In other words, instead of the liability value v_s^τ we use the value of the replicating portfolio $\hat{v}_s^\tau$, where the latter greatly reduces the time spent in Step 2. If $\hat{v}_s^\tau$ suitably approximates v_s^τ, it is possible to choose S_{rw} large in order to improve the accuracy of the quantile estimate.

The question then becomes how to find a replicating portfolio that satisfies $\hat{v}_s^\tau \approx v_s^\tau$ for all s. The approach usually adopted in practice is motivated by the no-arbitrage principle, which states that two assets that pay identical cash flows in all states of the world must have the same price. This suggests that an acceptable replicating portfolio may be found by matching the liability cash flows over some relatively small set of S replicating scenarios.

An effective replicating portfolio should exhibit certain characteristics. First, in order to be a valid substitute for the liability, the replicating portfolio's value must closely match that of the liability over a wide range of market conditions. Second, the practical computation of economic capital dictates that the time needed to price the replicating portfolio be extremely short. Finally, from an ERM perspective, the replicating portfolio should illuminate the nature of the risk inherent in the liability. Clearly, small size is advantageous in meeting the second and third objectives: the simpler the replicating portfolio, the more quickly it can be priced and the more readily it can be interpreted

[1] In practice, S_{rw} is typically 50,000 to 500,000 scenarios

economically. Moreover, to a certain degree, smaller replicating portfolios also provide a better fit to the liability in unanticipated market conditions (i.e. out-of-sample.)

Construction of a replicating portfolio begins with the selection of financial assets that are good candidates to mimic the liability. Optimization is then used to find the collective asset positions that best match the liability cash flows over the set of replicating scenarios. Intuitively, replicating portfolios rely on the fact that complex liability cash flows can be decomposed into simpler cash flows that correspond to vanilla instruments. For example, fixed cash flows can be represented as zero coupon bonds, minimum guarantees for variable annuities as puts on market indices, and fixed annuity options as swaptions with physical settlement.

In simple terms, constructing a replicating portfolio can be viewed as a regression problem; the liability and asset cash flows represent dependent and independent variables respectively, while the position sizes correspond to the regression coefficients. However, while conceptually straightforward, in practice finding an effective replicating portfolio poses a number of challenges. Identifying a suitable set of candidate assets is often difficult due to the number and diversity of cash flows that may be generated by a liability during its lifetime. For example, a 50-year liability portfolio with quarterly cash flows yields 200 possible maturity dates for replicating instruments. The assets necessary to cover these cash flow dates, and with sufficient variation in type, strike and underlying terms to span the liability’s response to alternative economic scenarios, can easily number in the thousands. Given the long time horizons of many liability products and the ensuing need to extrapolate the behavior of economic factors such as interest rates, replicating instruments are often of a “theoretical” nature; their strikes and/or maturities may not correspond to instruments actually traded in the market (e.g. a 50-year zero coupon bond.)

While having a large number of potential replicating instruments increases the likelihood of obtaining a good match to the liability cash flows in the replicating scenarios (i.e., in-sample), this typically comes at the expense of overall performance. Specifically, without some form of intervention, optimization methods will tend to take positions in an excessive number of assets, thereby producing a large, complex replicating portfolio that over-fits the liability in-sample. Moreover, such portfolios often contain offsetting long and short positions in similar assets, making it hard to interpret the results.

One way to address this problem is through the careful pre-selection of a small set of replicating instruments. This typically involves an in-depth analysis of the liability to identify one or more candidate sets of assets, followed by experimentation to determine the best of these alternatives (e.g., Daul and Gutiérrez Vidal (2009)). Bucketing may also be used to reduce the required number of candidate assets, in this case by aggregating all cash flows within a specified time interval so as to decrease the effective number of cash flow dates.

An alternative approach retains the large set of candidate assets, but modifies the optimization process so that only the most important ones are included in the replicating portfolio. The simplest technique, consistent with all-subsets regression, merely places an upper limit on the number of instruments that can be selected. For example, one might specify that the replicating portfolio contain no more than one hundred out of one thousand candidate assets. However, this type of restriction, known as a cardinality constraint, transforms what might otherwise be a convex optimization problem (i.e., a linear or quadratic program) into a mixed integer program that is much harder to solve. In fact, cardinality-constrained replicating portfolio problems of practical size typically preclude the possibility of finding an optimal solution with integer programming algorithms. In light of this, such problems are sometimes solved approximately by specialized heuristic methods, such as the genetic algorithm proposed by Ogrodzki (2007).

Recently, Burmeister and Mausser (2009) showed that adding trading restrictions, in the form of linear constraints or penalties, to a problem is an effective way to obtain small replicating portfolios. Rather than placing an explicit limit on the number of instruments in the replicating portfolio, this approach imposes trading costs that are proportional to the size of the asset positions, thereby preventing the replicating portfolio from becoming too large. Since the resulting problem remains convex, it can be solved efficiently using standard linear or quadratic optimization algorithms that are widely available. This approach is an example of regularization, a well-known variable selection technique often used in regression analysis to obtain sparse statistical models (see Hesterberg et al (2008) for an overview). In this case, the usual objective of minimizing the regression residuals is augmented with a constraint or a penalty term that discourages the inclusion of unnecessary predictors. Regularization has also been used to obtain sparse portfolios in the context of financial optimization, e.g., Brodie et al (2008), DeMiguel et al (2008), Gotoh and Takeda (2009).

Burmeister and Mausser (2009) considered only a simple form of regularization that puts an upper limit on the total size of positions in the replicating portfolio. This type of constraint effectively imposes a trading budget for constructing a replicating portfolio when all candidate assets have an identical trading cost of one dollar per unit. In this sense, it is similar to the Least Absolute Shrinkage and Selection Operator (LASSO) technique originally proposed by Tibshirani (1996) for ordinary least squares regression, which penalizes all predictors equally. As discussed in Hesterberg et al (2008), enhanced versions of the lasso allow the penalty to vary by predictor and/or by iteration. This suggests that a more general form of regularization, using instrument-dependent trading costs, may be more effective when constructing replicating portfolios.

This paper evaluates, on the basis of out-of-sample performance, a number of methods for determining instrument-dependent trading costs. Section 2 formulates optimization problems that can be used to construct replicating portfolios and various forms of trading costs are derived in Section 3. Section 4 contains computational experiments that use the trading costs to obtain replicating portfolios for a liability portfolio of variable annuities. Concluding remarks appear in Section 5.

2 Constructing Replicating Portfolios

Consider the problem of replicating the cash flows of a liability with a set of N candidate replicating instruments. Since cash flows can occur at any time, for practical purposes we partition the time horizon into a set of T time buckets (intervals) and consolidate all cash flows paid by an instrument within a bucket into a single cash flow (cash flows may optionally be adjusted by discount factors to reflect their actual time of occurrence). Specifically, we let c_{ij}^t denote the consolidated cash flow per unit of instrument j, where j = 0 is the liability and j =1, 2, .., N are replicating assets, that occurs in bucket t in scenario i.[2] Depending on its associated terms and conditions, instrument j generates a cash

[2] The amount constituted by a "unit" of an instrument is somewhat arbitrary. In the case of a stock, for example, a unit typically refers to a single share but, more generally, it could be defined to equal any number of shares. The liability, which is not tradable, is most readily viewed as comprising a single unit.

flow in some subset Q_j of the T buckets, i.e., $Q_j = \{t \mid \text{instrument } j \text{ generates a cash flow in bucket } t\}$. The liability is assumed to generate cash flows in all buckets, i.e., $Q_0 = \{1, 2, \ldots, T\}$.

For $j = 1, 2, \ldots, N$, let x_j denote the position size (i.e., the number of units) of instrument j in the replicating portfolio. Our goal is to find positions $\boldsymbol{x}$ such that the replicating portfolio matches, as closely as possible, the liability cash flows. Note that the larger the value of T, the more emphasis is placed on matching the timing of the liability cash flows; if the goal is simply to replicate the aggregate present or terminal values of the liability cash flows then it is sufficient to set $T = 1$.

A replicating portfolio is obtained by minimizing some measure, denoted $f(\boldsymbol{x})$, of the discrepancy between the liability cash flows and those of the replicating portfolio. One common measure of the replication error, motivated by least squares regression, is the weighted sum of the squared differences

$$f(\boldsymbol{x}) = \sum_{t=1}^{T} \sum_{i=1}^{S} w_i^t \left(\sum_{j=1}^{N} c_{ij}^t x_j - c_{i0}^t \right)^2 \tag{2.1}$$

where w_i^t is a priority weight that indicates the relative importance of matching c_{i0}^t.

Another measure, consistent with least absolute deviation regression, is the weighted sum of the absolute differences

$$f(\boldsymbol{x}) = \sum_{t=1}^{T} \sum_{i=1}^{S} w_i^t \left| \sum_{j=1}^{N} c_{ij}^t x_j - c_{i0}^t \right| \tag{2.2}$$

Note that if w_i^t equals the probability of scenario i then Equations (2.1) and (2.2) minimize the expected squared deviation and the expected absolute deviation, respectively, of the cash flows. For ease of reference, hereafter we will refer to Equations (2.1) and (2.2) as quadratic mismatch (QM) and linear mismatch (LM), respectively.

The replicating portfolio is given by the set of positions $\boldsymbol{x}$ that minimizes the replication error. Solving the unconstrained problem

$$min \quad f(\boldsymbol{x}) \tag{2.3}$$

often produces a replicating portfolio with non-zero x_j for most j, effectively overfitting the liability cash flows in the set of S scenarios. However, it is possible to modify Problem (2.3) in a way that reduces the number of non-zero x_j and that improves out-of-sample performance.

First, define the "cost" of obtaining the positions $\boldsymbol{x}$ to be

$$g(\boldsymbol{a}, \boldsymbol{x}) = \sum_{j=1}^{N} a_j \left| x_j \right| \tag{2.4}$$

where $a_j > 0$ for all j. Note that $g(\boldsymbol{a}, \boldsymbol{x})$ increases as the positions in the replicating portfolio become larger, regardless of whether they are long ($x_j > 0$) or short ($x_j < 0$). For example, if $\boldsymbol{a} = \mathbf{1}$ then $g(\boldsymbol{a}, \boldsymbol{x})$ is the total size of all positions in the replicating portfolio. More generally, a_j can represent a trading cost specific to instrument j.

To put Equation (2.4) in a form readily handled by optimization solvers, we make the substitution $x_j = \overline{x}_j - \underline{x}_j$ where $\overline{x}_j$ and $\underline{x}_j$ are non-negative variables that denote purchases and sales, respectively, of instrument j. The cost then becomes

$$g(\boldsymbol{a}, \boldsymbol{x}) = \sum_{j=1}^{N} a_j \left(\overline{x}_j + \underline{x}_j \right) \tag{2.5}$$

Consider the two problems

$$\begin{aligned} & \min \quad f(\boldsymbol{x}) \\ & s.t. \\ & \qquad g(\boldsymbol{a}, \boldsymbol{x}) \le b \end{aligned} \tag{2.6}$$

and

$$\min \quad f(\boldsymbol{x}) + \lambda\, g(\boldsymbol{a}, \boldsymbol{x}) \tag{2.7}$$

where $\lambda > 0$. The regularized problems (2.6) and (2.7) include a trading budget and a trading penalty, respectively. Problems (2.6) and (2.7) are effectively equivalent in the sense that b and λ both parameterize the same set of “efficient” replicating portfolios, i.e., those that provide the lowest replication error for a given cost. Decreasing b in Problem (2.6) or, equivalently, increasing λ in Problem (2.7), gives smaller replicating portfolios. Since standard optimization algorithms can easily solve the regularized problems (for linear or quadratic measures such as Equations (2.2) and (2.1), respectively), in practice it is possible to generate a large number of efficient replicating portfolios. Typically, an examination of the efficient set will identify a better replicating portfolio than the one found by solving Problem (2.3).

3 Alternative Trading Costs

The problem of choosing a small number of replicating assets is identical to the variable selection problem in linear regression, i.e., how to select an appropriate set of predictor variables from a large number of candidates. Like replicating portfolios, statistical models benefit from sparsity; when a model includes only the most important predictors, interpretation is easier and overfitting is less likely to occur, producing a more stable model with improved out-of-sample prediction accuracy. Variable selection is particularly relevant for so-called “high-dimensional” problems where the number of predictors (independent variables) exceeds the number of observations. In this case, which occurs frequently in signal processing and genomics applications, there are infinitely many models that yield zero error since the system is underdetermined. The development of regularization methods for regression has, therefore, been the subject of considerable research by statisticians, and such results have potential applications to replicating portfolios as well.

In regression analysis, a regularization method ideally eliminates those independent variables that are unnecessary (by setting their coefficients, or betas, to zero) without unduly reducing the coefficients of the selected variables (see, for example, Johnstone and Titterington (2009)).[3] Clearly, similar behavior is desirable when constructing replicating portfolios; the portfolio should contain relatively few assets but their position sizes should not be excessively small as a result of their trading costs.

The LASSO (Tibshirani 1996) has proved to be an effective technique for regularizing least squares regression. Given standardized variables having mean zero and variance one, the LASSO places an upper limit on the sum of the absolute regression coefficients, which corresponds to setting $\boldsymbol{a} = \mathbf{1}$ in Equation (2.4). Burmeister and Mausser (2009) used a similar approach to construct replicating portfolios that matched a liability's aggregate present-valued cash flows ($T = 1$). They were able to improve the out-of-sample performance by approximately 55% when minimizing absolute differences (Equation (2.2)) and by 90% when minimizing squared differences (Equation (2.1)), even when the cash flows were not standardized.

Fan and Li (2001) argued that the form of the penalty cost used by the LASSO is not appropriate; while it is able to identify the correct variables in the model, the coefficients of the selected variables tend to be too small (i.e., the LASSO imparts a negative bias on the estimated coefficients). They showed that this deficiency could be overcome by using a non-linear cost function, but this makes the resulting optimization problems difficult to solve in practice.

Zou (2006) proposed an alternative approach, called the Adaptive LASSO, that preserves the linear cost structure but varies the size of the penalty assigned to each variable (which corresponds to making the trading costs, a_j, in Equation (2.4) instrument-dependent). Ideally, larger costs should be assigned to those variables that are not selected and lower costs to those that appear in the model. To determine the penalties, Zou (2006) suggested solving an initial unpenalized ordinary least squares (OLS) regression to obtain coefficients $\tilde{\beta}^{OLS}$ and then setting

$$a_j = \frac{1}{\left|\tilde{\beta}_j^{OLS}\right|^{\gamma}} \tag{3.1}$$

for some $\gamma > 0$. A related version of the Adaptive LASSO, proposed by Wang et al (2007) for least absolute deviation (LAD) regression, uses the unpenalized LAD coefficients $\tilde{\beta}^{LAD}$ in Equation (3.1).

This way of obtaining penalty costs is ineffective for high-dimensional problems since the unpenalized coefficient estimates are not consistent. Thus, Zou (2006) proposed using an initial regularized regression in this case. As a simpler alternative, Huang et al (2008) suggested basing the costs on the marginal regression coefficients. [4]

[3] A method that is both able to select the correct variables and estimate their regression coefficients precisely (i.e., as well as if the correct model was known in advance) is said to possess the "oracle property".
[4] Huang et al (2008) showed that the resulting Adaptive LASSO estimator possesses the oracle property if a "partial orthogonality" condition is satisfied. In their computational results, satisfactory performance was obtained even when this condition was violated.

Before applying these regularization methods to the construction of replicating portfolios, it is worthwhile to briefly discuss the concept of standardization in this context. Standardization is often used in regression analysis to express variables in common units, namely standard deviations, with the intent of having the sizes of regression coefficients better reflect the relative importance of the variables.[5] Standardization also confers a degree of “fairness” when all variables are assigned the same penalty cost during regularization, as is the case for the LASSO (if variables are not standardized then those whose units of measurement result in smaller regression coefficients receive a favorable bias). As is readily shown, however, the same effect can be achieved by simply re-scaling the penalty costs themselves.

Consider, for example, the following regularized problem with unstandardized cash flows that constructs a replicating portfolio by minimizing the linear mismatch

$$\begin{aligned} & \min \quad \sum_{t=1}^{T}\sum_{i=1}^{S} w_i^t \left| \sum_{j=1}^{N} c_{ij}^t x_j - c_{i0}^t \right| \\ & s.t. \\ & \quad \sum_{j=1}^{N} a_j \left| x_j \right| \le b \end{aligned} \tag{3.2}$$

Suppose that $\theta_j > 0$ is a constant scaling factor for instrument *j*. The resulting problem with standardized cash flows is

$$\begin{aligned} & \min \quad \sum_{t=1}^{T}\sum_{i=1}^{S} w_i^t \left| \sum_{j=1}^{N} \frac{c_{ij}^t}{\theta_j} x_j - \frac{c_{i0}^t}{\theta_0} \right| \\ & s.t. \\ & \quad \sum_{j=1}^{N} a_j \left| x_j \right| \le b \end{aligned} \tag{3.3}$$

Upon substituting $\hat{x}_j = \frac{\theta_0}{\theta_j} x_j$ in Problem (3.3), one obtains

$$\begin{aligned} & \min \quad \frac{1}{\theta_0}\sum_{t=1}^{T}\sum_{i=1}^{S} w_i^t \left| \sum_{j=1}^{N} c_{ij}^t \hat{x}_j - c_{i0}^t \right| \\ & s.t. \\ & \quad \sum_{j=1}^{N} a_j \frac{\theta_j}{\theta_0} \left| \hat{x}_j \right| \le b \end{aligned} \tag{3.4}$$

Observe that Problem (3.2) will yield the same optimal solution as the standardized Problem (3.4) if the trading cost a_j is replaced with $a_j \frac{\theta_j}{\theta_0}$ (the constant factor θ_0^{-1} in the objective function of Problem

[5] The merits of standardization are subject to some debate. See, for example, Greenland et al (1991).

(3.4) has no effect on the solution). A similar argument applies when minimizing the quadratic mismatch except that the objective function is scaled by θ_0^{-2} in this case.

Thus, we see that the effect of standardizing cash flows (in the usual regression sense) can be achieved by making an instrument's trading cost proportional to the standard deviation of its cash flows. Since the scaling factor θ_j is completely arbitrary, one might also consider alternatives to the standard deviation.

In light of the preceding discussion, we now specify several forms of trading costs that will be used in Equation (2.4) during our computational experiments.

3.1 Unconstrained

Following Zou (2006) and Wang et al (2007), when the number of cash flows to be matched exceeds the number of replicating instruments (i.e., $N < ST$), instrument-dependent trading costs can be derived from the optimal position sizes in the unconstrained problem (2.3). Using the quadratic mismatch yields the following trading cost for instrument *j*

$$a_j^{QM} = \frac{1}{\left|\tilde{x}_j^{QM}\right|} \tag{3.5}$$

where $\tilde{x}_j^{QM}$ is the optimal position size of asset *j* as obtained by minimizing Equation (2.1), while the linear mismatch yields

$$a_j^{LM} = \frac{1}{\left|\tilde{x}_j^{LM}\right|} \tag{3.6}$$

where $\tilde{x}_j^{LM}$ is the optimal position size of asset *j* as obtained by minimizing Equation (2.2). If the denominator is zero for some instrument *j* then we simply fix the decision variable x_j equal to zero in the regularized problem (consistent with the trading cost being infinite).

3.2 Smallest Perfect Match

If the number of replicating instruments exceeds the number of cash flows to be matched (i.e., $N > ST$) then problem (2.3) has an infinite number of solutions with optimal value zero. In this case, we propose to derive the trading costs from the smallest replicating portfolio, in terms of the total size of positions, which perfectly matches the liability cash flows. Thus, we find $\tilde{\boldsymbol{x}}^{SPM}$ that solves

$$\begin{aligned} &\min \quad g(\boldsymbol{1}, \boldsymbol{x}) \\ &s.t. \\ &\qquad f(\boldsymbol{x}) \le 0 \end{aligned} \tag{3.7}$$

and set the trading cost for instrument *j* equal to

$$a_j^{SPM} = \frac{1}{\left|\tilde{x}_j^{SPM}\right|} \tag{3.8}$$

Note that the solution of Problem (3.7) is the same regardless of whether Equation (2.1) or Equation (2.2) appears in the constraint.

3.3 Marginal Unconstrained

Following Huang et al (2008), the trading cost based on the marginal optimal position size is

$$a_j^{MQM} = \frac{1}{\left|\tilde{x}_j^{MQM}\right|} \tag{3.9}$$

where $\tilde{x}_j^{MQM}$ is the optimal position size of asset *j* as obtained by minimizing Equation (2.1) when the positions of all instruments other than *j* are fixed to zero, i.e., the replicating portfolio contains only instrument *j*.

Similarly, using linear mismatch yields the trading cost

$$a_j^{MLM} = \frac{1}{\left|\tilde{x}_j^{MLM}\right|} \tag{3.10}$$

where $\tilde{x}_j^{MLM}$ is the optimal position size of asset *j* as obtained by minimizing Equation (2.2) when the positions of all instruments other than *j* are fixed to zero.

3.4 Standard Deviation Scaling

Consistent with the usual way of standardizing variables in regression, we use a trading cost based on the ratio of the standard deviations of the instrument and the liability cash flows. Since instrument *j* only contributes cash flows in certain time buckets, namely those in the set Q_j, the standard deviations are computed accordingly.

Specifically, let $\#(Q_j)$ represent the cardinality of the set Q_j. The mean cash flow of instrument *j* in those buckets in which instrument *k* generates a cash flow is

$$\mu_j^{(k)} = \frac{1}{\#\left(Q_k\right)\times S} \sum_{t\in Q_k} \sum_{i=1}^{S} c_{ij}^t \tag{3.11}$$

Similarly, the standard deviation of the cash flows of instrument *j* in those buckets in which instrument *k* generates a cash flow is

$$\sigma_j^{(k)} = \sqrt{\frac{1}{\#\left(Q_k\right)\times S} \sum_{t\in Q_k} \sum_{i=1}^{S} \left(c_{ij}^t - \mu_j^{(k)}\right)^2} \tag{3.12}$$

The resulting trading cost for instrument *j* is

$$a_j^{\sigma} = \frac{\sigma_j^{(j)}}{\sigma_0^{(j)}} \tag{3.13}$$

3.5 Mean Scaling

If all instruments are assigned equal trading costs then a favorable bias exists for instruments having large per-unit cash flows (the size of the position required to produce a cash flow of given size tends to be smaller in this case and so a smaller total cost is incurred). To counteract this bias, it is reasonable to assign larger trading costs to instruments that have larger cash flows. Making the trading cost proportional to the standard deviation of cash flows is consistent with this view provided that the standard deviation increases with the size of the cash flows themselves (this is a reasonable assumption in practice; a counterexample would be an instrument that pays a large, constant cash flow in all scenarios, such as a zero coupon bond). As an alternative to the standard deviation, we also consider a trading cost based on the relative magnitudes of the mean instrument and liability cash flows

$$a_j^{\mu} = \frac{\left|\mu_j^{(j)}\right|}{\left|\mu_0^{(j)}\right|} \tag{3.14}$$

3.6 Correlation Scaling

Intuitively, one might expect an instrument whose cash flows are highly correlated with those of the liability to be more important for replication purposes. Thus, it is reasonable to consider a trading cost that is inversely proportional to the magnitude of this correlation, namely

$$a_j^{\rho} = \frac{1}{\left|\rho_{j0}^{(j)}\right|} \tag{3.15}$$

where, in general,

$$\rho_{jl}^{(k)} = \frac{1}{\#\left(Q_k\right) \times S} \sum_{t \in Q_k} \sum_{i=1}^{S} \frac{\left(c_{ij}^t - \mu_j^{(k)}\right)\left(c_{il}^t - \mu_l^{(k)}\right)}{\sigma_j^{(k)} \sigma_l^{(k)}} \tag{3.16}$$

is the correlation between the cash flows of instrument *j* and *l* in those time buckets in which instrument *k* generates a cash flow.

3.7 Beta Scaling

To account for the effects of both correlation and standard deviation, we also consider a trading cost given by the product $a_j^{\rho} a_j^{\sigma}$, namely

$$a_j^{\beta} = \frac{\sigma_j^{(j)}}{\left|\rho_{j0}^{(j)}\right| \sigma_0^{(j)}} \tag{3.17}$$

Upon inspection, it is apparent that a_j^{β} equals the inverse of the regression coefficient, β_j, obtained by regressing (in a least squares sense) cash flows of the liability on instrument j in Q_j, i.e.,

$$Y = \beta_j X + \beta_0 \tag{3.18}$$

where Y and X denote the liability and instrument j cash flows.

4 Case Study

To assess the performance of the trading costs described in the previous section, we now use them to construct replicating portfolios for a liability portfolio comprising a block of variable annuities. Replicating portfolios are obtained by solving Problem (2.6) with both quadratic (Equation (2.1)) and linear (Equation (2.2)) mismatches. In each case, the trading budget is varied to produce a set of efficient replicating portfolios and their replication errors are evaluated on an out-of-sample basis.

As a benchmark, we also consider a constant trading cost, $\boldsymbol{a}^{EQ} = \mathbf{1}$, and to facilitate comparisons, we normalize the trading costs so that the total per-unit costs are the same in each case. Specifically, given a set of trading costs $\boldsymbol{a}$, we enforce the budget constraint

$$\sum_{j=1}^{N} \hat{a}_j \left| x_j \right| \le b \tag{4.1}$$

where

$$\hat{a}_j = \frac{K a_j}{\sum_{j=1}^{N} a_j} \tag{4.2}$$

for some constant K (in our experiments, we use $K = 1000$).

In the following, we refer to

$$\sum_{j=1}^{N} \hat{a}_j \left| x_j \right| \tag{4.3}$$

as the total cost of the portfolio $\boldsymbol{x}$. For convenience, hereafter we let $\boldsymbol{x}[f, Y, b]$ denote the optimal replicating portfolio obtained using measure f, trading costs $\boldsymbol{a}^{Y}$ and budget b, e.g., minimizing the quadratic mismatch with equal trading costs and budget 50 yields the portfolio $\boldsymbol{x}[QM, EQ, 50]$.

4.1 Problem Description

Consider a block of 15,000 variable annuity policies, all of which are in the accumulation phase, with a mixture of minimum guaranteed death benefits (ROPs, roll-ups and ratchets) sold over ten year period. The total annual cash flows comprising of both benefits paid and fees collected are projected over a 20-year time horizon. The economic scenario set used for the projection consists of 500 stochastic scenarios, which are a mixture of risk neutral and real world scenarios on the eight market indices listed in Table 1 plus the USD interest rate curve. [6]

Market Indices	**Liability Cash Flows**
S&P 400 Midcap (MID)	guaranteed minimum death benefit
Russell 1000 (RUX)	commissions expenses
S&P 500 (SPX)	mortality/expense charge
Nasdaq 100 (NDX)	revenue sharing
MSCI EAFE (MSDUEAFE)	surrender charges
MSCI EM (MSEUEGF)	per policy fees
MSCI US REIT (RMS)	
Lehman US Aggregate (USAGG)	

Table 1: Market Indices and Liability Cash Flows

Based on the characteristics of the liability, a set of 883 replicating instruments is identified consisting of zero coupon bonds, index forwards and European index options on the indices in Table 1, and physical-settlement swaptions. All instruments pay a single cash flow upon maturity except for the swaptions, which may trigger a stream of cash flows spanning several years. Intuitively, the European index options will match any minimum guaranteed death benefit payments, the index forwards will pick up the fees collected and the swaptions will reflect any annuity options. The maturities of the instruments are distributed evenly throughout the entire 20-year time period. Instruments with optionality are selected to emphasize at-the-money (ATM) strikes (i.e., there are more instruments that have strikes set close to ATM than strikes set far out-of-the-money). Annual settlement cash flows for the replicating instruments are projected over a 20-year time horizon for each of the 500 economic scenarios used in the liability cash flow projection. For replication purposes, all cash flows are present-valued (PV) using scenario-dependent discount factors, i.e., cash flows are expressed in current dollars.

The scenarios are split into two sets of 250 scenarios denoted *A* and *B*, such that both sets have equal numbers of risk neutral and real world scenarios. *A* is used for finding the optimal replicating portfolio while *B* is reserved for out-of-sample evaluation. Figure 1 shows the range, mean and the 10^{th} and 90^{th} percentiles of the discounted liability cash flow profiles for the in- and out-of-sample scenarios. The results indicate that the out-of-sample scenarios include a small number of extreme cash flows that are absent in-sample.

[6] We include real world scenarios that are representative of "stressed" economic conditions in this illustrative example. The methodology described in the paper, however, is applicable regardless of the particular scenarios used for replication.

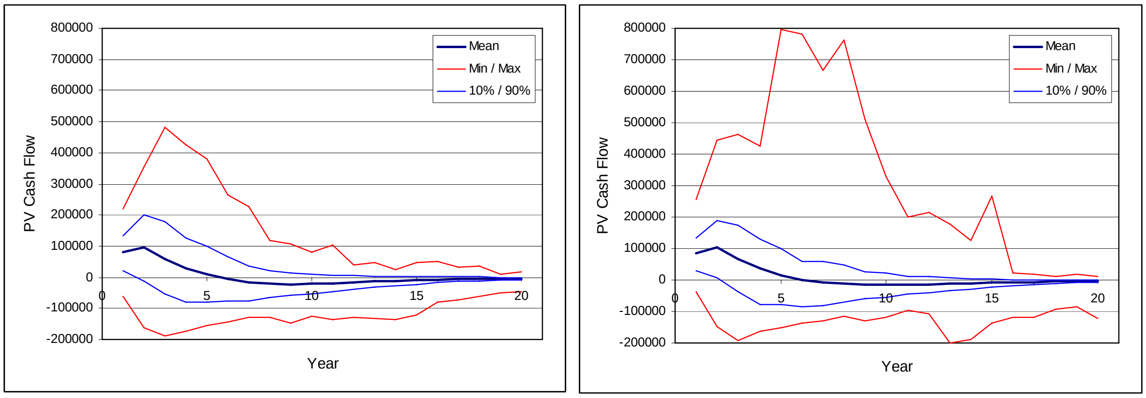

Figure 1: In-sample (left) and out-of-sample (right) discounted liability cash flows.

4.2 Single Time Bucket

We first consider a single time bucket spanning the entire 20-year horizon, so that the replicating portfolio is constructed to match the sum of the discounted liability cash flows in each scenario. Since the number of replicating instruments (883) exceeds the number of cash flows to be matched (250), the problem is underdetermined and multiple portfolios can perfectly replicate the liability in this case.

For each of the relevant trading cost schemes, Table 2 shows the replicating portfolio that perfectly matches the liability at minimal cost. All portfolios contain 250 instruments and yield zero replication error in-sample but, aside from those in the first two columns, the composition of the portfolios is different.[7]

Table 2: Single time bucket, minimum cost perfect replicating portfolios.

	Trading Cost							
	EQ	*SPM*	*MLM*	*MQM*	σ	μ	ρ	β
Total Cost	7,913	19	1,227	638	3,599	4,269	1,760	1,122
Cardinality	250	250	250	250	250	250	250	250
Total Position Size (Units)	6,987	6,987	13,693	13,097	438,514	22,860	16,360	22,147
LM (Out-of-Sample)	166,262	166,262	159,058	151,564	146,062	128,120	133,360	157,958
QM (Out-of-Sample)	4.1E+11	4.1E+11	3.9E+11	2.8E+11	2.9E+11	1.1E+11	2.5E+11	3.9E+11

Let us examine the smallest of the replicating portfolios in Table 2, **x**[-, *EQ*, 7913], which contains total positions of 6,987 units. Figure 2 plots the cash flows of **x**[-, *EQ*, 7913] against those of the liability. While R^2 = 1.00 in-sample, the out-of-sample fit is extremely poor for some scenarios, resulting in a much lower R^2 = 0.57. Thus, there is evidence that the replicating portfolio overfits the liability in this case, raising serious questions about its suitability as a proxy for the liability in economic capital calculations.

[7] **x**[-, *EQ*, 7913] and **x**[-, *SPM*, 19] are the same in this case because $\boldsymbol{a}^{SPM}$ is computed as one divided by the position sizes in **x**[-, *EQ*, 7913]. If an instrument does not appear in **x**[-, *EQ*, 7913], its SPM cost becomes infinite and so its position size will be fixed to zero. Thus, the same 250 instruments appear in both solutions.

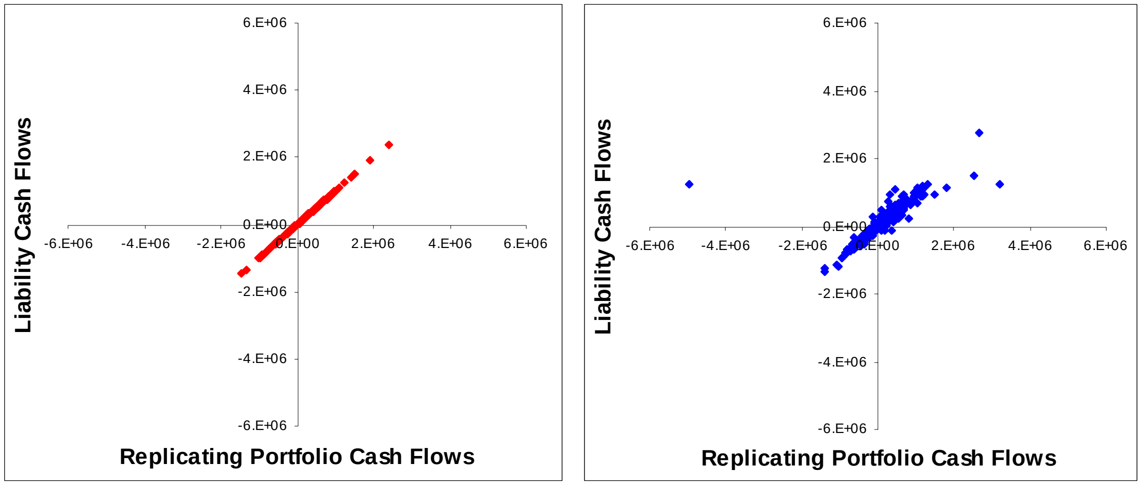

Figure 2: In-sample (left) and out-of-sample (right) cash flows for *x*[-, *EQ*, 7913].

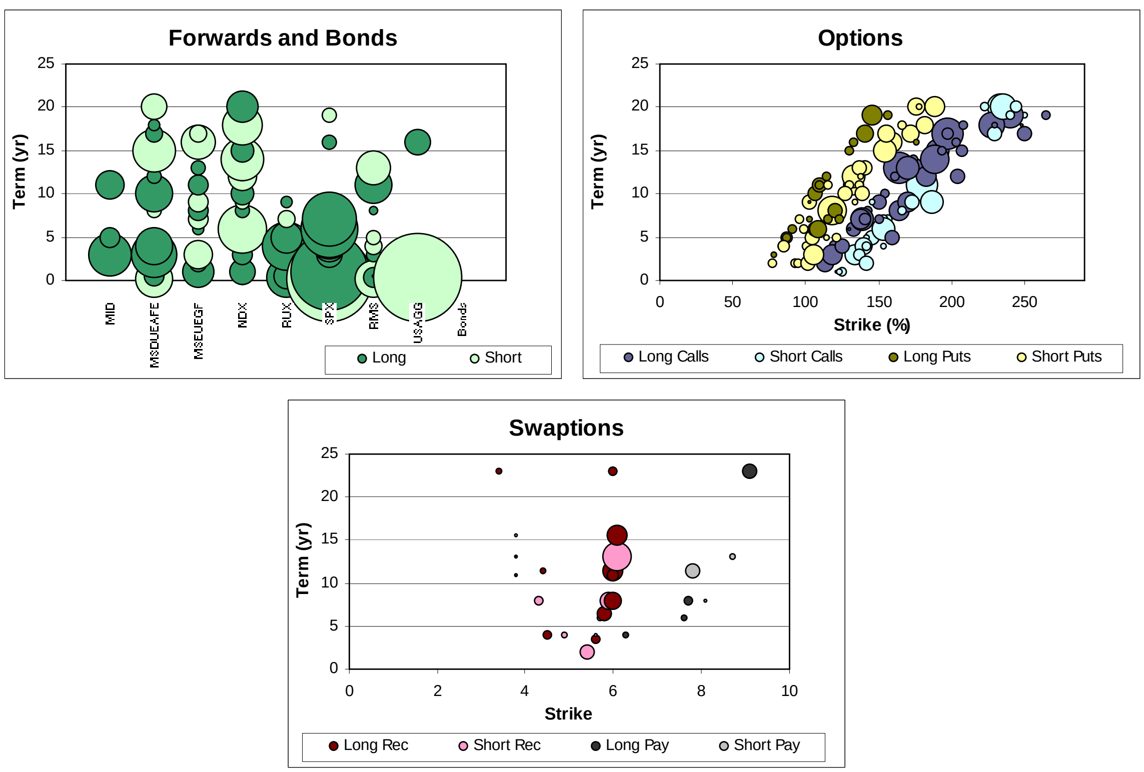

Figure 3: Composition of *x*[-, *EQ*, 7913].

To investigate further, we examine the composition of the replicating portfolio to see if it fits the liability at an intuitive level. Figure 3 plots the dollar values of the constituent positions, distinguishing between those that are long or short. In this case, an abundance of offsetting positions in similar instruments largely "cancel out", and make it difficult to discern any dominant structure. Thus, it is impossible to reconcile the replicating portfolio and the liability from this perspective.

The preceding analysis suggests that ***x***[-, *EQ*, 7913] will perform poorly as a replicating portfolio. In an effort to find a better candidate, we minimize the linear mismatch while gradually reducing the trading budget. As a result, each of the trading cost schemes produces an efficient set of progressively smaller replicating portfolios. Figure 4 plots the out-of-sample replication error against the cardinality of the resulting portfolios (the portfolios that obtain the lowest out-of-sample error for each cost scheme, identified by solid white circles, are summarized in
Table 3).

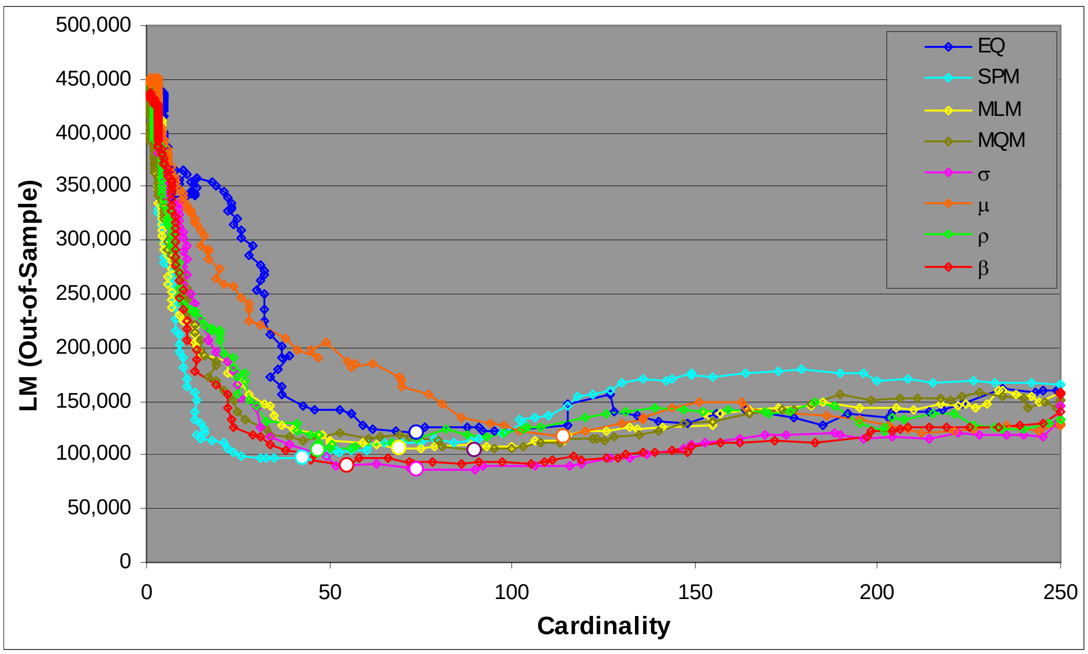

Figure 4: Efficient portfolios obtained by minimizing linear mismatch, single bucket.

Table 3: Best out-of-sample replicating portfolios (single time bucket, minimize LM).

	Trading Cost							
	EQ	*SPM*	*MLM*	*MQM*	σ	μ	ρ	β
Total Cost	1,772	2	203	127	953	1,598	249	156
Cardinality	74	43	69	90	74	114	47	55
Total Position Size (Units)	1,565	2,306	2,769	3,221	10,863	7,937	5,749	4,483
LM (In-Sample)	53,168	50,896	35,556	31,571	31,219	25,169	56,829	42,787
LM (Out-of-Sample)	120,924	96,440	105,382	104,844	86,642	117,651	104,523	90,147
Improvement (Out-of-Sample)	27.27%	42.00%	33.75%	30.83%	40.68%	8.17%	21.62%	42.93%

The results demonstrate that regularization is able to reduce the size of the replicating portfolio while also improving out-of-sample performance, regardless of the cost scheme used. Standard deviation scaling yields the lowest out-of-sample replication error in this case, followed closely by beta scaling and SPM, while equal costs and mean scaled costs are considerably less effective.

We now repeat the previous analysis with one of the more promising replicating portfolios, $\boldsymbol{x}[LM, \beta, 156]$. While its out-of-sample error is slightly higher than that of $\boldsymbol{x}[LM, \sigma, 953]$, its smaller size (55 positions) is an attractive feature.

As shown in Figure 5, $\boldsymbol{x}[LM, \beta, 156]$ maintains a high-quality fit ($R^2 = 0.98$) in-sample and achieves a much better out-of-sample fit ($R^2 = 0.89$) than $\boldsymbol{x}[-, EQ, 7913]$. In particular, it eliminates the extreme mismatches observed previously.

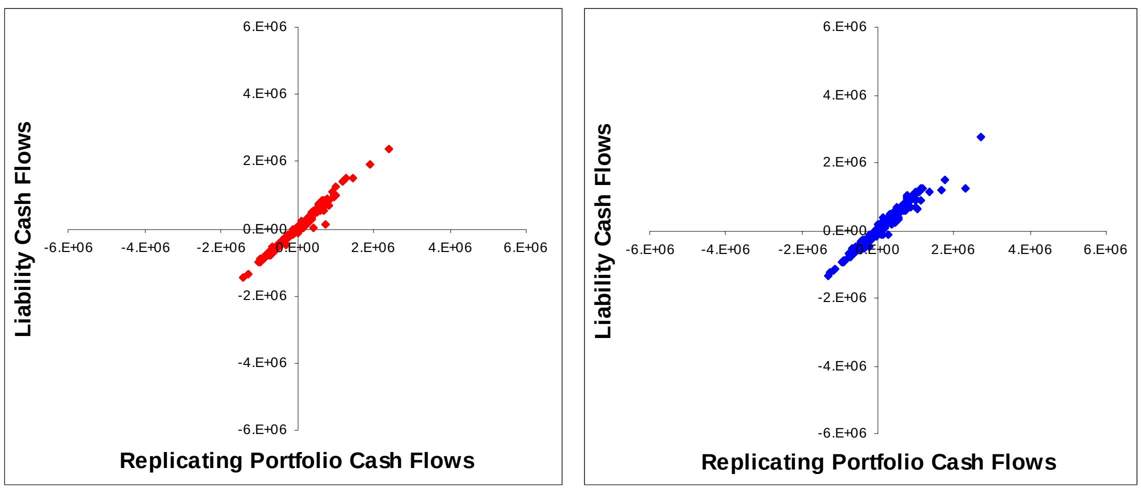

Figure 5: In-sample (left) and out-of-sample (right) cash flows for $\boldsymbol{x}[LM, \beta, 156]$.

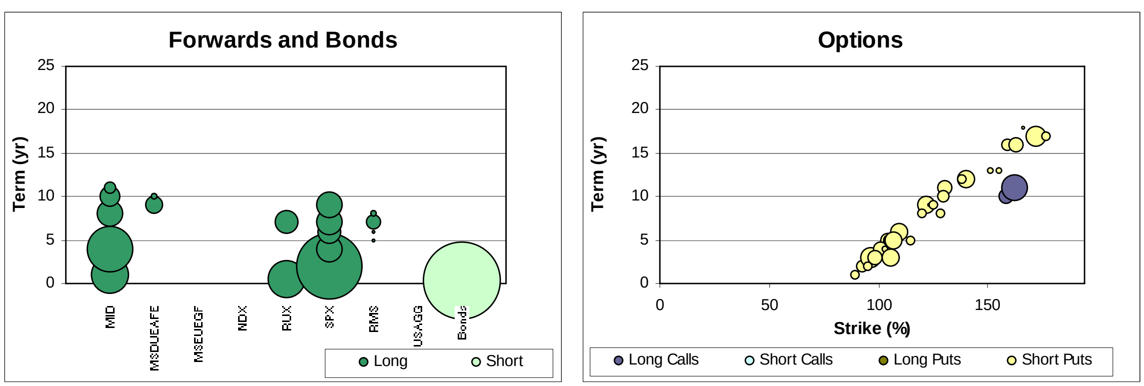

Figure 6: Composition of $\boldsymbol{x}[LM, \beta, 156]$.

Furthermore, the composition of $\boldsymbol{x}[LM, \beta, 156]$ (Figure 6) is much more readily interpreted than that of $\boldsymbol{x}[\text{-}, EQ, 7913]$ (Figure 3). In this case, the portfolio has four dominant positions:

- long positions in equity forwards are consistent with performance-based fee income
- a short bond position reflects lump sum payments
- short equity puts correspond to guaranteed death benefits (which, in effect, are put options on equity performance held by policyholders)
- a small long call option position is consistent with surrender charges (policyholders are more likely to opt out when equity markets perform well, making the guarantees less valuable)

Note that there are no large swaption positions in the replicating portfolio. Given that policies are in the accumulation phase, this may be due to limited annuitization occurring during the 20-year horizon.

Figure 7 plots the efficient replicating portfolios produced when the process is repeated with quadratic matching. In this case, trading costs based on mean scaling are unable to improve out-of-sample performance, while the improvements obtained by the other cost schemes are even more significant than in the linear matching case (Table 4). Once again, trading costs derived by SPM, beta scaling and standard deviation scaling produce the best results.

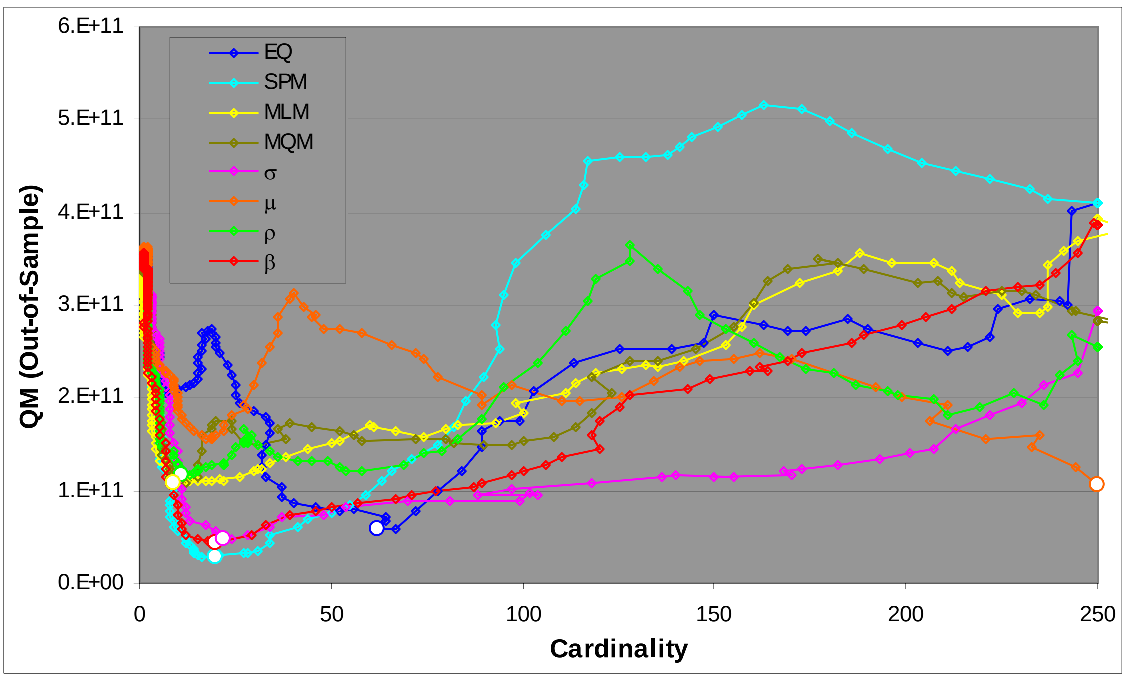

Figure 7: Efficient portfolios obtained by minimizing quadratic mismatch, single bucket.

In general, the best-performing replicating portfolios contain fewer instruments than those obtained by linear matching. For example, $\boldsymbol{x}[QM, \beta, 109]$ has only 20 positions while $\boldsymbol{x}[LM, \beta, 156]$ has 55, although the fit of $\boldsymbol{x}[QM, \beta, 109]$ is of only slightly lower quality (in-sample $R^2 = 0.96$, out-of-sample

$R^2 = 0.88$). As shown in Figure 8, $\mathbf{x}[QM, \beta, 109]$ contains long equity forwards, a short bond and short puts, but it lacks a position in long calls.

Table 4: Best out-of-sample replicating portfolios (single time bucket, minimize QM).

	Trading Cost							
	EQ	*SPM*	*MLM*	*MQM*	σ	μ	ρ	β
Total Cost	1,507	1	61	34	593	4,269	54	109
Cardinality	62	20	9	9	22	250	11	20
Total Position Size (Units)	1,331	1,723	1,419	1,521	3,142	22,860	3,822	3,359
QM Error (In-Sample)	9.4E+09	1.7E+10	6.9E+10	7.1E+10	2.8E+10	0	8.9E+10	1.8E+10
QM Error (Out-of-Sample)	5.8E+10	2.8E+10	1.1E+11	1.1E+11	4.6E+10	1.1E+11	1.2E+11	4.2E+10
Improvement (Out-of-Sample)	85.96%	93.14%	72.66%	61.61%	84.11%	0.00%	54.06%	89.06%

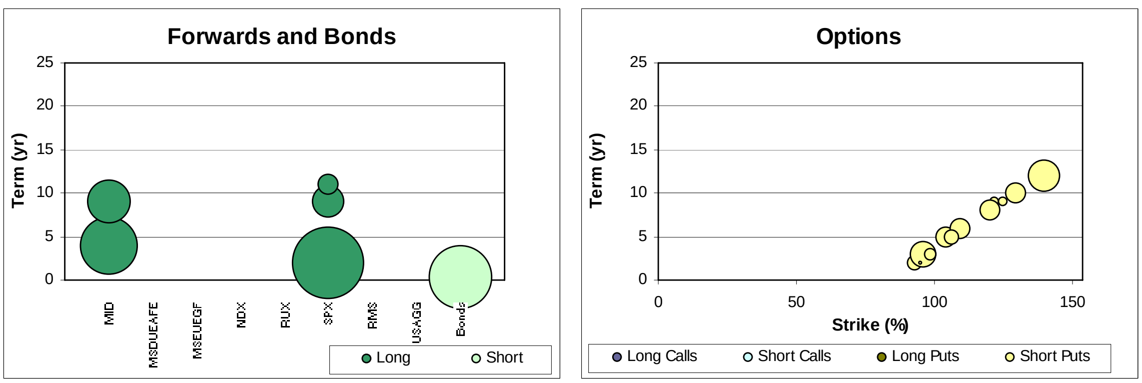

Figure 8: Composition of $x[QM, \beta, 109]$.

4.3 Annual Time Buckets

We now partition the time horizon into 20 annual buckets, so that the replicating portfolio is constructed to match the discounted liability cash flow in every bucket of each scenario. Thus, there are a total of 5,000 cash flows to be matched, which exceeds the number of replicating instruments. Since the problem is overdetermined in this case, one might expect the effects of regularization to be less significant than when all cash flows are consolidated into a single bucket.

With linear matching, the unique, unconstrained replicating portfolio comprises 883 instruments and a total position size of 78,512 units. The in- and out-of-sample LM replication error is 128,386 and 212,404, respectively.

The efficient portfolios produced by regularized LM minimization (Figure 9) do not dramatically improve out-of-sample performance. As shown in Table 5, all cost schemes deliver relatively modest reductions of 5% - 8% in out-of-sample replication error with approximately 300 – 400 instruments. The reason for the large number of instruments is evident in Figure 10, which shows the composition of $\boldsymbol{x}[LM, \beta, 102]$; since cash flows must be matched in 20 time buckets, replicating instruments are distinguished by the timing of their cash flows (in constrast, when cash flows are consolidated into a single bucket, the timing of the cash flows is largely immaterial).

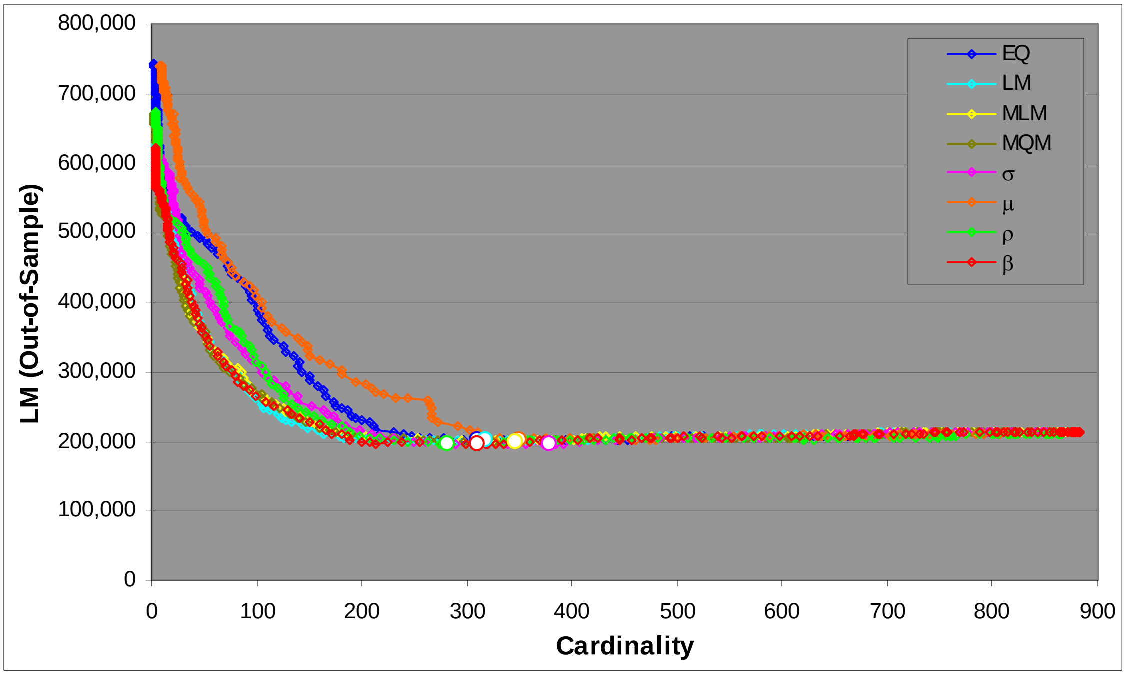

Figure 9: Efficient portfolios obtained by minimizing linear mismatch, multiple buckets.

Table 5: Best out-of-sample replicating portfolios (multiple time buckets, minimize LM).

	Trading Cost							
	EQ	*LM*	*MLM*	*MQM*	σ	μ	ρ	β
Total Cost	2,419	151	8	97	454	1,003	256	102
Cardinality	310	317	346	347	378	350	281	311
Total Position Size (Units)	2,136	12,766	6,359	7,709	23,109	6,201	3,640	6,742
LM Error (In-Sample)	154,832	143,085	145,558	143,442	140,891	150,543	153,121	147,543
LM Error (Out-of-Sample)	201,292	200,545	199,653	198,491	195,531	201,703	196,228	195,637
Improvement (Out-of-Sample)	5.23%	5.58%	6.00%	6.55%	7.94%	5.04%	7.62%	7.89%

This additional time dimension can be useful when validating replicating portfolios and/or using them to understand better the nature of the liability. For example, Figure 10 shows a gradual reduction in both fee income and payouts over the time horizon.

When using quadratic matching, the unique, unconstrained replicating portfolio comprises 883 instruments and a total position size of 71,165 units. The in- and out-of-sample QM replication error is 3.2 billion and 18 billion, respectively.

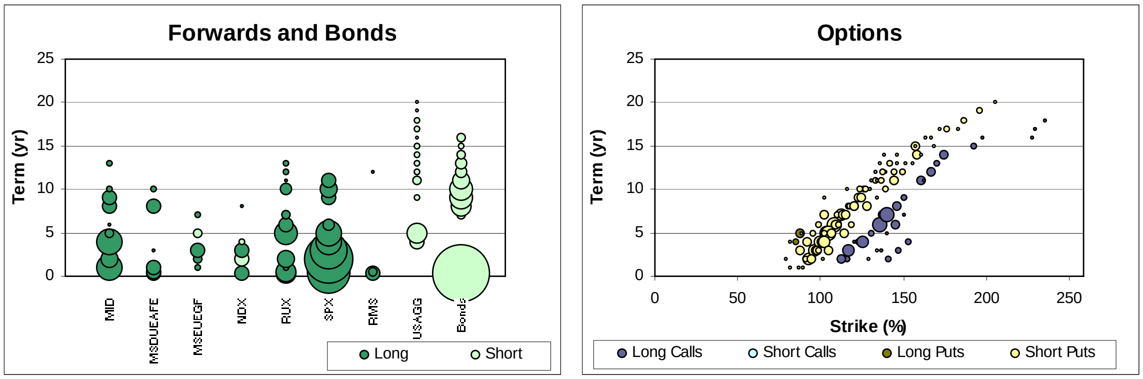

Figure 10: Composition of *x*[*LM*, *β*, 102].

Regularized quadratic matching produces more significant out-of-sample improvements than what is observed for its linear counterpart (Figure 11). Surprisingly, the QM costs are noticeably ineffective in this case, yielding only a 3% reduction in out-of-sample replication error compared to the upwards of 20% reductions delivered by standard deviation scaling, correlation scaling and beta scaling (Table 6). While the portfolios in Table 6 contain fewer instruments than those in Table 5, a comparison of ***x***[*QM*, *β*, 58] (Figure 12) and ***x***[*LM*, *β*, 102] (Figure 10) shows that they share the same dominant structure.

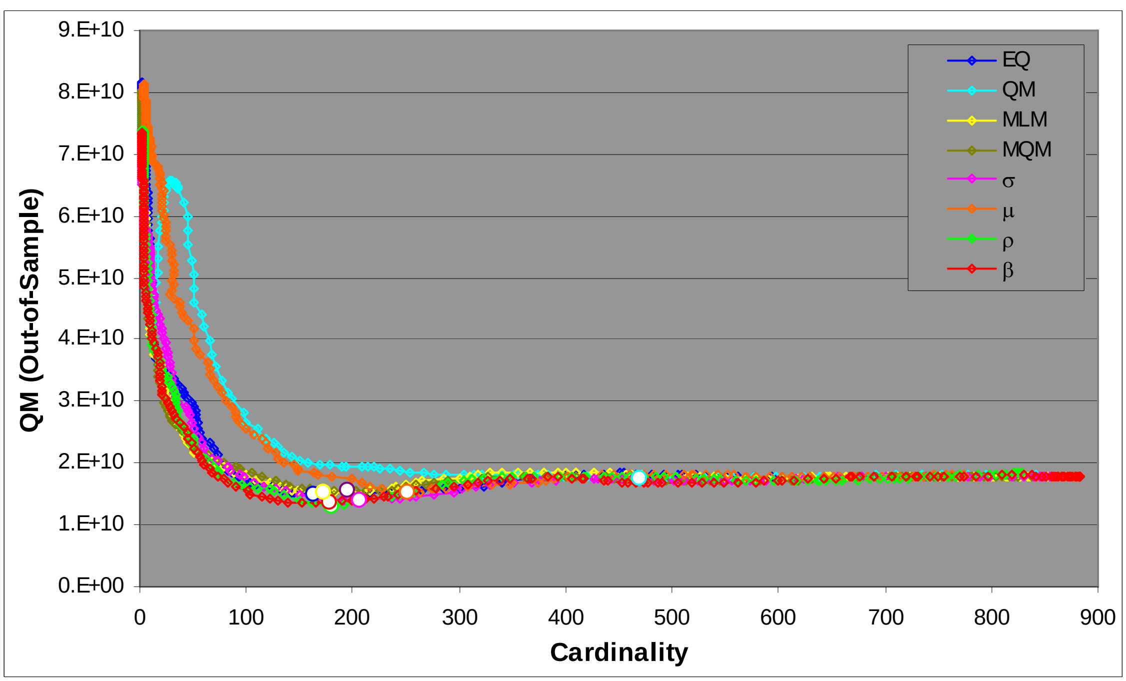

Figure 11: Efficient portfolios obtained by minimizing quadratic mismatch, multiple buckets.

Table 6: Best out-of-sample replicating portfolios (multiple time buckets, minimize QM).

	Trading Cost							
	EQ	*QM*	*MLM*	*MQM*	σ	μ	ρ	β
Total Cost	1,507	498	4	51	248	749	196	58
Cardinality	163	469	172	196	207	252	180	178
Total Position Size (Units)	1,331	36,669	3,507	3,896	8,120	5,059	2,939	3,870
QM Error (In-Sample)	6.3E+09	3.3E+09	4.7E+09	4.4E+09	4.5E+09	4.7E+09	4.9E+09	4.6E+09
QM Error (Out-of-Sample)	1.5E+10	1.7E+10	1.5E+10	1.5E+10	1.4E+10	1.5E+10	1.3E+10	1.4E+10
Improvement (Out-of-Sample)	17.56%	3.06%	15.68%	14.17%	22.01%	15.72%	27.38%	24.29%

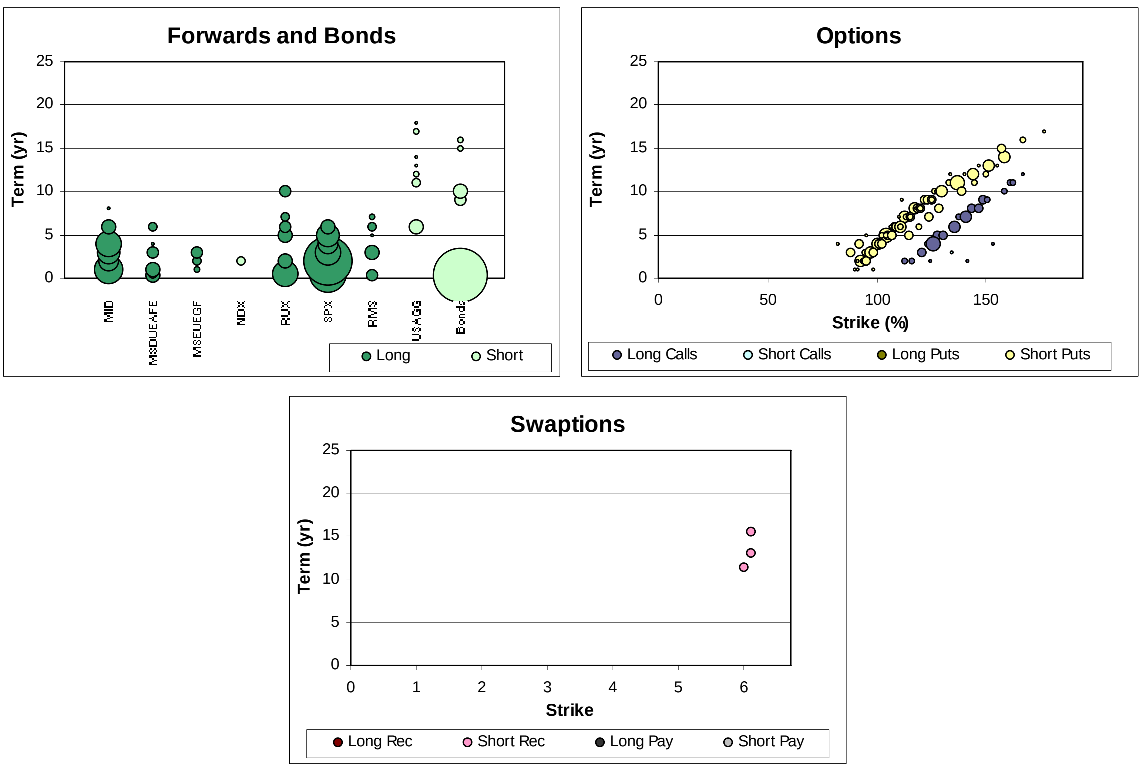

Figure 12: Composition of *x*[*QM*, *β*, 58].

4.4 Practical Considerations of Trading Costs

Based on the preceding examples, it is not possible to identify a single best way of specifying trading costs when constructing replicating portfolios with regularization. However, the equal cost approach (without standardized cash flows) and mean scaling appear to be rather ineffective. Among the other methods, computational considerations may further determine their usefulness. Notably, costs based on SPM, LM, QM, MGM and MLM require solving optimization problems (in particular, the marginal approaches require a separate problem to be solved for every instrument). In constrast, standard deviation scaling, correlation scaling and beta scaling allow trading costs to be computed analytically from cash flow data. Thus, the latter methods have an added appeal on a practical level.

It is important to recognize, however, that trading cost regularization presents an enormous computational advantage over imposing cardinality constraints in the optimization. For example, we formulated cardinality-constrained versions of both the single-bucket (50 and 100 instruments) and multiple-bucket (100, 200 and 300 instruments) problems with both linear and quadratic matching. We attempted to solve the problems with a state-of-the-art mixed-integer optimizer, within a 12-hour time limit (over 1000 times greater than the time required to solve any of the regularized problems considered in this paper). Of the ten problems attempted, we obtained feasible, but not optimal, solutions for eight of them (problems that minimize quadratic mismatches with a single time bucket gave no result whatsoever). In only one case, namely when minimizing linear mismatches for a single bucket with 100 instruments, was the out-of-sample replication error lower than that of the best regularized solution with the same number of instruments.[8] In contrast, producing an entire frontier of more than 200 efficient solutions using regularization takes only a matter of minutes.

5 Concluding Remarks

Replicating portfolios are computational tools that facilitate the calculation of economic capital for complex insurance liability portfolios. For practical reasons, a replicating portfolio needs to closely match the liability value with a minimal number of standard financial instruments. Trading restrictions are an effective way of regularizing optimization problems so that they produce small replicating portfolios. Making the trading costs in such restrictions instrument-dependent yields better results than using the same cost for all instruments. We investigated a number of different methods for obtaining effective trading costs, based on techniques used in regularized regression as well as several simpler approaches that have intuitive appeal. Experiments with a portfolio of variable annuities suggested that costs based on simple statistics provide good performance with minimal computational effort, making them an attractive choice in practice.

6 References

Brodie, J., Daubechies,I., De Mol, C., Giannone, D. and I. Loris (2008), "Sparse and stable Markowitz portfolios," European Central Bank Working Paper Series No. 936 (available at www.ecb.europa.eu)

Burmeister, C. and H. Mausser (2009), "Using Trading Restrictions in Replicating Portfolios," *Life & Pensions* (Nov), 36-40.

DeMiguel, V., Garlappi, L., Nogales, F.J. and R. Uppal (2008), "A Generalized Approach to Portfolio Optimization: Improving Performance By Constraining Portfolio Norms." London Business School.

Daul, S. and E. Gutiérrez Vidal (2009), "Replication of Insurance Liabilities," *RiskMetrics Journal* 9:1, 79-96.

[8] However, there exist regularized solutions with more than 100 instruments whose out-of-sample error is less than that of the cardinality-constrained solution

Fan, J. and R. Li (2001), "Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties," *Journal of the American Statistical Association*, 96, 1348-1360.

Gotoh, J.-Y. and A. Takeda (2009), "On the Role of the Norm Constraint in Portfolio Selection," Department of Industrial and Systems Engineering Discussion Paper Series ISE 09-03, Chuo University.

Greenland, S., Maclure, M., Schlesselman, J.J., Poole, C. and H. Morgenstern (1991), "Standardized Regression Coefficients: A Further Critique and Review of Some Alternatives," *Epidemiology* 2:5, 387-392.

Hesterberg, T., Choi, N.H., Meier, L. and C. Fraley (2008), "Least angle and l_1 penalized regression: A review," *Statistics Surveys* 2, 61-93.

Huang, J., Ma, S. and C.-H. Zhang (2008), "Adaptive Lasso for Sparse High-Dimensional Regression Models," *Statistica Sinica* 18, 1603-1618.

Johnstone, I.M. and D.M. Titterington (2009), "Statistical Challenges of High-Dimensional Data," *Philosophical Transactions of the Royal Society A*, 367, 4237-4253.

Markowitz, H. (1952), "Portfolio Selection," *Journal of Finance* 7, 77-91.

Ogrodzki, J. (2007), "Replicating portfolios," *Insurance and financial services review* (Autumn), 4-7 (available at www.watsonwyatt.com)

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B*, 58, 267-288.

Wang, H., Li, G. and G. Jiang (2007), "Robust regression shrinkage and consistent variable selection through the LAD-lasso," *Journal of Business & Economic Statistics*, 25:3, 347-355.

Zou, H. (2006), "The Adaptive Lasso and its Oracle Properties," *Journal of the American Statistical Association*, 101, 1418-1429.