Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Projections Using Replicating Portfolios for Insurance Liabilities

Luuk Tegels

MSc Thesis 2015-048

Maastricht University

August, 2015

PROJECTIONS USING REPLICATING PORTFOLIOS FOR INSURANCE LIABILITIES

Author: Luuk Tegels

Supervisor: Prof. Dr. Antoon Pelsser

Second corrector: Drs. Anne Balter

Abstract

For insurers it is important to value and price their assets and liabilities in a market consistent manner. In this thesis the replicating portfolio technique will be applied for the fair valuation of three different insurance products: pure cash flow, inflation linked, and option embedded. To construct the replicating portfolios we make use of the ordinary least squares and constrained least square approach. The main concern is the maximum (relative) error of the replicating portfolio. We find that this maximum is highly dependent on the product which is evaluated, but even the worst case remains limited to deviation of at most 3,5% of the products true value. In the next step the replicating portfolios are used to make projections of the products. It is found that the errors of the projections are equal to the initial error with accumulated interest. We then argue that the distribution of these projection errors is equal to the errors of the fit of the replicating portfolio. Hence, we conclude that the fit of the replicating portfolio is of vital importance to the quality of the projections.

1. Introduction

For insurance liabilities, it is important that they are valued market-consistently, also known as fair value valuation. Since insurance liabilities as such are not (regularly) traded there are no reliable market prices to rely upon. The International Actuarial Association (IAA) (2000) made an overview of some fair value approaches to insurance liabilities. In this thesis I will illustrate how the market consistent- value can be found, using a replicating portfolio. This is a technique that was advocated by Schrager (2008) and is also on the IAA's list of techniques.

A replicating portfolio tries to match the cash flows of the insurance liability as closely as possible, given the replicating instruments or assets available in the market. In this way the portfolio matches the liabilities in such a way that changes in the market conditions (for example interest rates or inflation rates) cause a (almost) parallel shift in the fair value of both the asset and the portfolio. This concept will be further evaluated in section 2.3. However, it should be noted that the cash flows of the insurance liabilities are not deterministic as such and depend on mortality rates. To remain focused on the replicating portfolios themselves I will disregard mortality in the given examples. This is not necessarily a big simplification if one assumes that mortality rates and interest rates are independent. In that case the market value can still be calculated using the projected cash flows following from a different mortality table (Bouwknegt & Pelsser, 2002).

A replicating portfolio can have more applications than just pricing, in this thesis I will look at the possibility of using a replicating portfolio to make projections of the value of the liability in the future. The advantage of projecting a replicating portfolio as opposed to the liability itself is that when the replicating portfolio has been established considerably fewer calculations are needed to determine the projections.

The remainder of this thesis is organized as follows. First, a theoretic framework for the replicating portfolios will be given. This includes the pricing of assets and thus the replicating portfolio in general. Second, the question on how the replicating portfolio is constructed will be answered. This approach will be applied to three different products and the results of this test will be discussed. Third, these portfolios will be used to make projections to see how the portfolios behave at later time point. To conclude, the results will be summarized and discussed.

2. Theoretic Framework

Before looking into the theory about replicating portfolio themselves it is important to know more about the fundamental problem of pricing financial derivatives. Modern theory began in 1973 with the famous Black-Scholes paper about the theory of option pricing, Black and Scholes (1973), and Merton's (1973) extension of this theory.

2.1. Pricing

To explore all important features we discuss contingent claim pricing in its simplest setting and impose the following assumptions on the financial markets (Bingham & Kiesel, 2013):

Assumption	Implication
(1) No market frictions	No transaction cost, no bid/ask spread, no taxes, no margin requirements, no restrictions on short sales
(2) No default risk	Implies that the interest rate for borrowing and lending is equal
(3) Competitive market	Market participants act as price takers
(4) Rational agents	Market participant prefer more to less
(5) No arbitrage	More details see section 2.2

Table 1 General Assumptions

In practice any real market contains frictions; this assumption is therefore only made for simplicity. We look at the ideal –frictionless- market so we can focus on the essentials of the theory and to be able to give a first-order approximation of reality. In order to be able to understand markets with friction it is necessary to understand a frictionless market first.

The risk of failure of a company – default – is unavoidably present in a real life situation: just as with individuals, death is part of life. These risks appear everywhere, also at a national level: in recent decades we have seen default of interest payments of international debts, or at least the threat of it. A prominent contemporary example would be the euro crisis in Greece, but this was not the first, nor will it be the last. For the interested reader on this subject we refer to Jameson (1995) and Madan (1998).

The third assumption is closer to the real world: financial agents are assumed to be price takers, not price makers. This means that the trading of a security by one agent, even in large amounts, does not influence the security's price. Hence, an agent can trade as much as he likes of a certain security without changing the price. Considering that there are only a handful of institutions that would be able to exercise that much power we assume that the effect is negligible in the model.

It is very rational and realistic to assume that a market participant prefers more rather than less, or more precisely, agents will always accept an increase in consumption if there are no additional costs. This assumption is therefore very weak and can be considered a formality. Note that this is the only assumption that is made about the preferences of an agent and that an agent can have any other preference that is not conflicting with the preference of more to less.

The no-arbitrage assumption is the most important assumption that we consider. Consider that a theoretical price is developed under these assumptions and this price is not equal to the price observed. We would consider this as an arbitrage opportunity in our model and exploit it. This might lead to the relaxation of one of the assumptions and the procedure is restarted with again no-arbitrage assumed. The no-arbitrage assumption has a special status and it is the basis for the arbitrage pricing technique that we'll use. The *principle of no-arbitrage* is explained in more detail below.

2.2. No-Arbitrage

The *principle of no-arbitrage* is the formal definition of a very strong and straightforward economic argument: in any (complete) financial market it should not be possible to make a profit without any initial investment and without bearing any risk.

Definition: *An arbitrage opportunity is a portfolio strategy θ such that*

$$X_0^\theta = 0, \mathbb{P}\left[X_1^\theta \geq 0\right] = 1 \; and \; \mathbb{P}\left[X_1^\theta > 0\right] > 0$$

Or in words, an arbitrage opportunity is a strategy with zero initial costs which has a non-negative payoff in all states of the world and has a strictly positive pay-off in at least one state of the world (Touzi & Tankov, 2008). This would mean that such an opportunity guarantees a profit without being exposed to risk (since $\mathbb{P}[X_1^\theta < 0] = 0$). If such an opportunity exists, an arbitrageur (someone that exploits arbitrage opportunities) would exploit it, and use the market as a 'money-pump' to extract, arbitrarily large, quantities of riskless profit from the market. This would make it impossible for the market to be in equilibrium. For simplicity we restrict ourselves to markets that are in equilibrium, hence markets that do not violate the no-arbitrage assumption.

As we have seen from the example above a market *with* arbitrage opportunities would be disorderly - too disorderly - to model. However, the converse of this statement is also true: the absence of arbitrage opportunities combined with small enough transaction costs and an (almost) complete market is sufficient to be able to build a model of a financial market that is realistic enough to provide real insight and to handle the mathematics necessary to price standard contingent claims (Bingham & Kiesel, 2013). Allingham (1991) explains that this arbitrage condition is therefore sufficient to determine prices, the *arbitrage pricing technique.*

We have seen that, although the no-arbitrage principle seems so simple at first sight, a lot can be deduced such a straightforward concept (Schachermayer, 2002). For example Black, Scholes (Black & Scholes, 1973), and Merton (Merton, 1973) use it as the basis for their Nobel-prize winning paper on how to determine the price of any derivative security (e.g. a European Put option).

Furthermore, keeping the previous assumptions in mind, the absence of arbitrage is 'essentially' equivalent to the existence of a *risk-neutral measure* or an *equivalent martingale measure*, where the price of a random payment is the expected discounted value. This is also known as the *Fundamental Theorem of Asset Pricing* (Dybvig and Ross, 1987). Generally speaking, multiple equivalent martingale measures exist, but when considering a *complete* (and arbitrage-free) market the number of measures is reduced to one *unique* equivalent martingale measure. Harrison and Kreps (1979) and Harrison and Pliska (1981, 1983) were the first to make the connection between the no-arbitrage condition and the existence of an equivalent martingale measure. The basis for their results was the idea of risk-neutral valuation by Cox and Ross (1976). Dalang, Morton and Willinger showed in 1990 that in a finite discrete-time model the absence of arbitrage is equivalent to the existence of an equivalent martingale measure. Under less strict circumstances it is necessary to replace the term "equivalent to" with "essentially equivalent to" (Gerber & Shiu 1996). The details why are beyond the scope of this thesis, but some interesting results can be found in Artzner and Heath (1995), Christopeit & Musiela (1994) and Back (1991).

2.3. Replicating portfolio theory

In this thesis the no-arbitrage argument will be used as the basic idea behind replicating portfolios. A replicating portfolio is a portfolio that can contain different types of assets like: real assets with liquid market prices, real illiquid assets, and even theoretical assets with a theoretical price (Corrigan & Qin, 2011). The idea of a replicating portfolio is that it mimics the (scenario-dependent) cash flows of a certain asset or liability under any economic circumstances. One of the first applications of replicating portfolios was the replication of options like in Rubinstein & Leland (1981).

Consider the principle of no-arbitrage: if two instruments or portfolios have identical cash flows under all economic conditions, then the prices of the two instruments or portfolios must be equal. If this were not the case one could go long in the cheaper instrument and short in de more expensive instrument and the difference in price would be riskless profit. This works the same for the implementation of replicating portfolios. If the cash flows of a portfolio of assets (the replicating portfolio) and the liabilities are identical in all scenarios, then the market value of the liabilities must be equal to the value of the replicating portfolio (Bouwknegt & Pelsser, 2002). This type of replicating portfolio is called *static replication*. After a static replication is made, it is no longer necessary to make any changes in the portfolio to maintain the same cash flows. Static replication techniques for exotic options were introduced by Derman, Ergener and Kani (1995) and Carr, Ellis and Gupta (1998).

In contrast to static replication, dynamic replication is a type of replicating portfolio where the cash flows do not (need to) match, but the portfolio has the same "Greeks" as the reference asset. This means that for small changes in the market, the price of the reference asset and the replicating portfolio change in the same way. Hence, after any change in the market, the portfolio needs to be readjusted, because the dynamic portfolio and reference asset only behave similar in a single point (mathematically speaking, their partial derivatives are equal in that point). This also means that the value of the portfolio and reference asset are only equal at that point. Although this helps us price and hedge the product it has a big disadvantage: it is model dependent and it only gives the value for the reference asset and portfolio under a specific model. On the contrary, static replication is done independent of the model and the results can therefore be used in a more general environment or model. In addition, dynamic replication has some problems when we look at it from a realistic and practical viewpoint. Say we disregard our first assumption ("No market-frictions") for a moment and apply the dynamic hedging strategy to an insurance liability. Insurance liabilities have a long maturity. Considering dynamic hedging requires the portfolio to be continuously adjusted, a long maturity means a lot of transactions. Even with low transaction costs this will become expensive, whilst these costs would be avoided in static replication.

In this thesis we distinguish three methods for constructing a replicating portfolio: the balance sheet method, the annual cash flow method, and the aggregate cash flow method (Matson, 2010). The *balance sheet method* starts by constructing a portfolio of assets whose current market value matches the fair value of the liabilities on the balance sheet in a range of sensitivities. This portfolio can be used to update the balance sheet in a market-consistent manner. The drawbacks are that the value of these assets and liabilities already have to be known for a range of sensitivities. In addition, this might not capture all behavior unless very complex sensitivities are analyzed. The *annual cash flow method* tries to find a replicating portfolio that matches the cash flows of the liability as closely as possible in a set of scenarios. The drawback is that this potentially requires a lot of assets in the portfolio to be able to match the cash flow in each year. The last method, *aggregate cash flow method,* is the method that will be implemented in this thesis. This method starts by constructing a replicating portfolio that matches the *total future discounted cash flows* of the liabilities in a set of scenarios (Matson, 2010). We expect that the aggregation of cash flows decreases the number of assets required to get a decent fit.

Note that all methods state "construct a replicating portfolio", but that they do not explain how. In this thesis the portfolios are constructed using the "ordinary least squares" approach as will be explained in the section 2.4. Other examples of models for portfolio construction are the *generalized additive model (GAM)* and the *equity replication model* (Corrigan & Qin, 2011). Interested readers are referred to Yamada (2009, 2012) and Bamberg & Wagner (2000).

The insurance world is already broadly adopting the idea of replicating portfolios and is seeing huge potential for the future (Schrager, 2008). As the valuations of the replicating portfolios can be done in much less time than the complex actuarial calculations it allows them to be done more frequently. This gives management the opportunity to monitor capital positions more closely and intervene when necessary. In the beginning, the use of replicating portfolios will mainly lie in the field of regular risk reporting, the calculation of a market consistent balance sheet, and the ability to make realistic projections of the future of the balance sheet. When the technique will have matured, and analysts have become acquainted with the technique the possibilities are even bigger. Future possibilities include, but are not limited to: the reporting of fair value according to the IFRS norms, the analysis of hedging strategies, and the use of what-if analysis to test the impact of management decisions (Matson, 2010).

2.4. Construction of replicating portfolio

This thesis makes use of the ordinary least squares approach to construct a replicating portfolio. This approach is split into two phases. The first phase is to determine which replicating instruments will be included in the portfolio. It is important that these instruments are able to grasp the behavior of the product that is replicated. Deciding which instruments to choose is therefore often done by hand while keeping general economic principles in mind, however automated procedures also exist, for example the Dantzig selector which selects the best subset of replicating instruments by solving a very simple convex program, which can be made into a convenient linear program (Candes & Tao, 2007).

In the second phase the weights of the chosen instruments are determined: How much of each instrument do we need to get the best replication? The quality of the replication is based on the sum of squared errors between the aggregate discounted cash flow of the replicating portfolio and the aggregate discounted cash flow of the product. Although the aggregate cash flow method does not necessarily mimic every individual cash flow: it mimics the net present value (NPV) of all of the cash flows. This might seem as a big simplification, but the main interest will usually be the aggregate value (NPV). This enables us to decrease the number of replicating instruments. Less replicating instruments means less calculation time for the value of the replicating instrument.

To explain how phase two works exactly it is necessary to introduce some notation. First of all let *n* be the number of training scenarios used to calibrate the portfolio and m be the number of replicating instruments in the replicating portfolio. Secondly, let X be the $m \times n$ matrix that holds the NPVs of the replicating instruments in all scenarios. This X matrix looks like:

$$\begin{bmatrix} R_{0,1} & \cdots & R_{0,n} \\ \vdots & \ddots & \vdots \\ R_{m-1,1} & \cdots & R_{m-1,n} \end{bmatrix}$$

where $R_{i,j}$ is the NPV of replicating instrument *i* in scenario *j*. In addition we need two more vectors, Y, a $n \times 1$ vector that holds the NPVs of the product to be replicated in all scenarios, and β a $m \times 1$ vector that holds the weights of the replicating instruments.

The objective of the least squares approach is now to minimized the squared error between the NPVs of the product and the replicating portfolios. In this thesis this is either done using ordinary least squares or constrained least squares.

Ordinary least squares

The goal of the OLS approach is to minimize the sum of squared errors, this guarantees that we have an efficient portfolio (Hayashi, 2000). The replicating portfolio is mathematically defined by:

$$X\widehat{\beta} + \epsilon = Y$$

The sum of squared errors is defined as:

$$SSE = \epsilon'\epsilon = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$

The optimal $\hat{\beta}$ is then derived as follows:

$$\frac{\delta SSE}{\delta \hat{\beta}} = 2\left(Y - X\hat{\beta}\right)' X = 0$$
$$Y'X = \hat{\beta}'(X'X)$$
$$\hat{\beta} = (X'X)^{-1}(X'Y)$$

2.4.1. Constrained least squares:

The replicating portfolio is still defined as: $X\widehat{\beta} + \epsilon = Y$, however this time a constraint is added to the model. This constraint adds information about the NPV of the product under the current market conditions, also called the base scenario. Since these conditions are determined by the market they also follow the no-arbitrage rules and therefore the NPV of the product and replicating portfolio should be exactly equal under those conditions.

Now suppose this NPV is d_{base}, then the following constraint is added to the model $C^T \beta = d_0$, where C^T is a 1 x *n* vector which consists of the values of replicating instruments $R_{BASE,0}$, that is the values of the replicating instruments in the base scenario at t = 0. The analytical determination of the optimal value for β was done by Amemiya (1985) as follows:

$$minimize\ f(\beta) = \left||X\beta - Y|\right|^2$$
$$subject\ to\ c_i^T \beta = d_i, i = 1, \dots, p$$

Construct the Langrangian function with Langrange multipliers $z_1, \dots, z_p$

$$L(\beta, z) = f(\beta) + z_1(c_1^T \beta - d_1) + \cdots + z_p\left(c_p^T \beta - d_p\right)$$

The optimality conditions are

$$\frac{\delta L}{\delta \beta_i}\left(\widehat{\beta_{cls}}, z\right) = 0, \quad i = 1, \dots, n, \qquad \frac{\delta L}{\delta z_i}\left(\widehat{\beta_{cls}}, z\right) = 0, \quad i = 1, \dots, p$$

The second equation is not very interesting as this was already fixed by the constraint:

$$\frac{\delta L}{\delta z_i}\left(\widehat{\beta_{cls}}, z\right) = c_i^T\, \widehat{\beta_{cls}} - d_i = 0$$

The other *n* equations are more interesting:

$$\frac{\delta L}{\delta \beta_i}\left(\widehat{\beta_{cls}}, z\right) = 2\sum_{j=1}^{n}(X^T X)_{ij}\widehat{\beta_J} - (2X^T Y)_i + \sum_{j=1}^{p} z_j c_i = 0$$

Or in matrix form: $2(X^T X)\widehat{\beta_{cls}} - 2(X^T Y) + C^T z = 0$

To get the Karush-Kuhn-Tucker conditions we put the matrix together with $C\, \widehat{\beta_{cls}} = d$:

$$\begin{bmatrix} 2X^T X & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \widehat{\beta_{cls}} \\ z \end{bmatrix} = \begin{bmatrix} 2X^T Y \\ d \end{bmatrix}$$

Assuming that the matrix is invertible:

$$\begin{bmatrix} \widehat{\beta_{cls}} \\ z \end{bmatrix} = \begin{bmatrix} 2X^TX & C^T \\ C & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2X^TY \\ d \end{bmatrix}$$

To show the relation of $\hat{\beta}$ to $\hat{\beta}_{cls}$ blockwise inversion is applied to the matrix. This technique was reinvented multiple times and was eventually generalized and proven for correctness by Banachiewicz (1937). There are two formulas to apply blockwise inversion:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$$

Since in our case the matrix D is singular we make use of the first formula (the second formula requires inverting D, which is not possible due to singularity).

$$\begin{bmatrix} (2X^TX & C^T \\ C & 0 \end{bmatrix}^{-1} = \begin{bmatrix} (2X^TX)^{-1} + (2X^TX)^{-1}C^T(-C(2X^TX)^{-1}C^T)^{-1}C(2X^TX)^{-1} & -(2X^TX)^{-1}C^T(-C(2X^TX)^{-1}C^T)^{-1} \\ -(-C(2X^TX)^{-1}C^T)^{-1}C(2X^TX)^{-1} & (-C(2X^TX)^{-1}C^T)^{-1} \end{bmatrix}$$

Since only the $\widehat{\beta_{cls}}$ is interesting we look at the upper row of matrices, which determine the $\widehat{\beta_{cls}}$.

$$\widehat{\beta_{cls}} = ((2X^TX)^{-1}(2X^TY) + (2X^TX)^{-1}C^T(-C(2X^TX)^{-1}C^T)^{-1}C(2X^TX)^{-1})(2X^TY) - ((2X^TX)^{-1}C^T(-C(2X^TX)^{-1}C^T)^{-1}\, d$$

$$= \frac{1}{2}((X^TX)^{-1}2 * (X^TY) + \frac{1}{2}(X^TX)^{-1}C^T 2 * (-C(X^TX)^{-1}C^T)^{-1}\frac{1}{2}C(X^TX)^{-1}\, 2 * (X^TY)$$
$$- \frac{1}{2}((X^TX)^{-1}C^T\, 2 * \; (-C(X^TX)^{-1}C^T)^{-1}\, d$$
$$= ((X^TX)^{-1}(X^TY) + (X^TX)^{-1}C^T(-C(X^TX)^{-1}C^T)^{-1}C(X^TX)^{-1}(X^TY)$$
$$- ((X^TX)^{-1}C^T\, (-C(X^TX)^{-1}C^T)^{-1}\, d$$
$$\widehat{\beta_{cls}} = \hat{\beta} - (X^TX)^{-1}C^T(C(X^TX)^{-1}C^T)^{-1}(C\hat{\beta} - d)$$

2.4.2. Why is it acceptable to take the base scenario as a constraint?

The base scenario reflects the current state of the world and the associated economic conditions. The economic conditions emerge from, amongst others, the no-arbitrage (assumption) in the market. This means that this scenario reflects the expectations of market participants about future changes in the economic conditions. Hence, these economic conditions are the best-estimate at this point in time. If there would be a better 'best-estimate' traders would make use of that information and use it to make a risk-free profit, which violates the no-arbitrage assumption. Therefore, we also need to make use of this information. Since the base scenario is the best estimate (at this point in time) the replicating portfolio should be correct in this scenario. Especially for future projection purposes it is likely that this scenario reflects the real-world or is close to the real-world. In addition, it can be proven that the CLS estimator keeps a minimum distance to the OLS estimators (Hansen, 2015), thus the effect on efficiency is limited to a minimum.

2.5. Omitted variable error

In the previous section we have discussed the least squares model and the way a replicating portfolio is defined: $X\widehat{\beta} + \epsilon = Y$ (and possibly a constraint). The errors (ϵ) are usually noise in the process, especially when considering a stochastic process. However, in this case it is not a stochastic noise process because the real data-generating process is actually known and deterministic. The actuarial model can provide us with the cash flows of any insurance liability, given a certain mortality table. Considering that we are mainly interested in the behavior of the replicating portfolio itself we will ignore the mortality rate. This is, as already mentioned, not necessarily a big simplification assuming that mortality rates and interest rates are independent. In such case the market-consistent value can be calculated by using the projected cash flows from a different mortality table (Bouwknegt & Pelsser, 2002). Once the cash flows are known it is very simple to create a perfectly replicating portfolio by simply matching cash flows with maturity T with discount bonds with maturity T. This would mean that there are just as many different discount bonds as the number of cash flows.

However, for practical purposes this is not very useful. First of all, typical insurance product has cash-flows due in (much) more than 50 years, whilst the maturity of most traded bonds doesn't exceed 30 years. Second of all, a replicating portfolio with this many products becomes very complex to analyze (in our example we have monthly cash flows for 100 years, it would therefore consist of 1200 replicating instruments). Due to these difficulties and restrictions we choose to drop most of the replicating instruments and use only a small subset of these replicating instruments. Unless this subset is, in some way, perfectly correlated to the full set of replicating instruments, it is impossible to create a perfect replicating portfolio. The replicating portfolio will therefore inherently have errors in the estimated coefficient, this error is known as the omitted variable error.

What happens is the following: because we intentionally omit some relevant variables, the other variables try to compensate for this, which makes them biased. The biased coefficient is precisely what we are interested in, since we want to find the best replicating portfolio, using only a subset of the possible replicating instruments. Furthermore, the "true" coefficients of the full set of regressors are already known, these are simply the cash-flows at each time point (with the price of the discount-bond as regressor). It should be noted that the size and direction of the error is different for each coefficient and it fully depends on which variables were left out and what the correlation with them was (Clarke, 2005).

2.6. Quantifying the omitted variable error

We are in the extraordinary case were we actually *know* the true model. Therefore it is possible to quantify this error. Suppose that the correct specification of the model is

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4}$$

but we estimate the misspecified model

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \epsilon_i^*$$

where β_2 is the coefficient of interest and $\epsilon_i^* = \beta_4 X_{i4}$.

Under the assumption that the expected value of ϵ_i is zero (in our case we even *know* it is zero if all regressors are included), the expected value for $\widehat{\beta_2}$ is given by Hanushek and Jackson (1977)

$$\mathbb{E}(\widehat{\beta_2}) = \beta_2 + \beta_4 b_{42}$$

where

$$b_{42} = \frac{(\rho_{42} - \rho_{32}\,\rho_{43})}{1 - \rho_{32}^2}\sqrt{\frac{Var(X_4)}{Var(X_2)}}$$

b_{42} is equal to the regression coefficient on X_2 in the auxiliary regression of the excluded variable, X_4, on the included variables, X_2 and X_3. By writing it out explicitly it can be seen that the effect of omitting X_4 depends on the magnitude of the excluded coefficient β_4 and on the correlation (ρ_{ij}) between the included variables and the excluded variable.

The generalization of this result deals with the situation where a set of relevant variables is omitted from the estimated equations. Suppose that the true model is:

$$\boldsymbol{y} = \boldsymbol{X_1}\boldsymbol{\beta_1} + \boldsymbol{X_2}\boldsymbol{\beta_2}$$

where $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ are both vectors of relevant variables. Now we estimate

$$\boldsymbol{y} = \boldsymbol{X_1}\boldsymbol{\beta_1} + \boldsymbol{\epsilon}^*$$

where $\boldsymbol{\epsilon}^* = \boldsymbol{X_2}\boldsymbol{\beta_2}$. The expected value of $\boldsymbol{\beta_1}$, under the usual assumptions, is given by Greene (2003)

$$\mathbb{E}(\widehat{\boldsymbol{\beta_1}}) = \boldsymbol{\beta_1} + \boldsymbol{P_{1.2}}\boldsymbol{\beta_2}$$

where $\boldsymbol{P_{1.2}} = (\boldsymbol{X_1'}\boldsymbol{X_1})^{-1}\boldsymbol{X_1'}\boldsymbol{X_2}$ is the matrix of regression coefficients from the auxiliary regressions of the excluded variables, $\mathbf{X_2}$, on the included variables $\mathbf{X_1}$. In this form it becomes very explicit how much of the weight is actually there to compensate for each specific omitted variable.

It is important to note that the exact effect of omitting variables is not trivial, that means including or excluding more relevant variables can increase the error of one coefficient and decrease the error of another coefficient. In the end the error solely depends on the set of included and excluded variables and the correlations between them (Clarke, 2005).

This shows what the effect is of omitting variables on the error of the coefficient, however, for the replicating portfolio this is not a big issue as long as the performance is good. We are therefore more concerned what this does to the error of the regression itself, ϵ^*. With some assumptions on the distribution of the X's it is possible to see what the drivers are behind the error. Assume that the X's are multivariate-normal distributed and consider the following "true" model:

$$Y = X_1\beta_1 + X_2 B_2$$

And the misspecified model

$$Y = X_1\beta_1 + \epsilon^*$$

Then we are able to define the error as follows:

$$\begin{aligned}
&\mathbb{E}(\epsilon^*|Y = y) = \mathbb{E}(Y - \beta_1 X_1|Y = y) \\
&= y - \mathbb{E}(\beta_1 X_1|Y = y) \\
&= y - (\mu_{X_1} + \rho_{X_1Y}\frac{\sigma_{X_1}}{\sigma_Y}(y - \mu_Y) = (y - \mu_{X_1}) - (\rho_{X_1Y}\frac{\sigma_{X_1}}{\sigma_Y}(y - \mu_Y))
\end{aligned}$$

Then the error is zero for:

$$\frac{y - \mu_{X_1}}{y - \mu_Y} = \rho_{X_1Y}\frac{\sigma_{X_1}}{\sigma_Y}$$

This shows how the error would behave in such bivariate normal situation. We see that it is dependent on the individual variances and correlations. However, just like the omitted variable error of the coefficient it is not trivial how the error behaves. It only tells us some about the fact that the error is affected by the variances and the correlations, but it does not give an unambiguous answer in what direction it will be affected.

3. Method

In order to valuate a certain (insurance) product we try to construct a portfolio that mimics the product's behavior, or more specific, the product's cash flows in any scenario. Such a portfolio is called a replicating portfolio. A replicating portfolio can consist of many different instruments, e.g. stocks, derivatives, options or bonds. The choice of instruments heavily depends on the product and is therefore done by hand. The weights on each of these instruments are then computed by a linear regression.

Section 3.1 will describe the types of products that are considered, and addresses the question as to which types of instruments are needed to replicate these products. Section 3.2 shows how to calculate the weights of each replicating instrument in the replicating portfolio.

3.1. Products and replicating instruments

3.1.1. Pure cash-flow product

A pure cash-flow product pays out a fixed amount per policyholder under certain conditions, e.g. monthly payments while the policyholder is alive, a fixed payment when the policyholder reaches a certain age or a fixed payment when the policyholder dies. The amount is always fixed (per policyholder) and does not depend on external factors such as market return or inflation. The value of the product is therefore only dependent on the interest rates and the uncertainty of mortality, however we do not take the latter into account in this thesis.

To mimic these pure cash-flow products we make use of discount bonds since their value depends only on the maturity and interest rate, like the products at hand. The replicating portfolio consists of 20 replicating instruments (discount bonds) with different maturities, ranging from a maturity of 1 year to a maturity of 100 years.

3.1.2. Inflation linked product

Inflation linked products are very much like the pure cash-flow product, but there is one important difference: the amount per policyholder is not fixed anymore. Instead of paying a fixed amount the amount is now increased (yearly) with the inflation of that year with a minimum guarantee of 0,25%.

Simple discount bonds are (independently) not able to capture the effect of this inflation component. For example, consider two scenario's with the same interest rates but with (completely) different inflation rates. The inflation linked products will have different values under those scenarios whilst the value of the discount bonds (and hence the replicating portfolio) remains exactly the same. It is therefore necessary to add some sort of component to the replicating portfolio that takes into account the effect of inflation and the embedded guarantee.

To do this a (hypothetical) inflation bond was constructed. One could see it as a coupon bond which pays yearly coupons of $\max\{0{,}25\% \,; inflation\}$ and these coupons are then reinvested in the same coupon bond again. The payout therefore is actually principal yearly compounded with the maximum of the inflation rate and the minimum guarantee (0,25%). Let the inflation rate be 3% in year 1 and 0,15% in year 2 and the principal amount be €1. The payout at the beginning of year 3 then is: $€1 * (1 + 0{,}03) * (1 + 0{,}0025) = €1{,}0326$

For the construction of the replicating portfolio for the inflation linked product we use 10 of these inflation bonds and 10 regular discount bonds, all with different maturities ranging from 1 to 100 years.

3.1.3. Option embedded product

Whilst the inflation linked product always depends on the inflation rate, the general option embedded product can depend on virtually everything. In this case it is linked to the MY-rate, an internal measure of the insurers' financial performance. Whilst the inflation linked product only had a minimum guarantee the option embedded product has a complete structure for the payouts. This structure is based on the MY-rate and looks as follows:

$$yield = \begin{cases} 0 & if \quad MY < 3\% \ (I) \\ \max\{0{,}75 * MY - {,}0225; 0\} & if \ 3\% \leq MY < 5\% \ (II) \\ \max\{0{,}50 * MY - 0{,}01; 0\} & if \ 5\% \leq MY < 6\% \ (III) \\ \max\{0{,}125 * MY + 0{,}0125; 0\} & if \ 6\% \leq MY \leq 10\% \ (IV) \\ 0{,}03 & if \quad MY > 10\% \ (V) \end{cases}$$

This yield is again compounded in a yearly fashion. Hence, say we have an MY product expiring at the beginning of year 3, then let the MY-rate in year 1 be 4% and in year 2 3,5%, then at the start of year 3 the MY product will pay:

$$(1 + \max\{0{,}75 * 0{,}04 - 0{,}0225; 0\}) * (1 + \max\{0{,}75 * 0{,}035 - 0{,}0225; 0\}) \\ = (1 + 0{,}0075) * (1 + 0{,}00375) = 1{,}0113$$

In the previous products it was fairly trivial to decide what kind of instruments should be selected. For this product we consider two approaches on which kind of instruments should be used for the replicating portfolio. For the first approach five different path-dependent option instruments are constructed. These instruments accumulate their value in the same way as the product itself. The payoff depends on (the path of) the MY-rate and a certain threshold value. The thresholds correspond to the boundaries in the original product. This gives the following yields for the different replicating instruments:

I. $\max\{MY - 0{,}03; 0\}$
II. $\max\{MY - 0{,}05; 0\}$
III. $\max\{MY - 0{,}06; 0\}$
IV. $\max\{MY - 0{,}10; 0\}$
V. $0{,}03 \ if \ MY > 10\%$

Since this yield compounds yearly the maturity of these instruments is also of importance. In principle we try to include ten different maturities for each instrument to be able to capture the path[1]. In addition to these replicating instruments, discount bonds with different maturities will be included in both types of replicating portfolios. They are used to mimic the 'fixed' cash-flows that usually account for large portions of the product as well.

The second approach involves constructing a more complex replicating instrument, in fact, this replicating instrument uses the same yield structure as the original product. Again we include approximately ten different maturities of this instrument and approximately ten discount bonds.

Notice that both methods have advantages and disadvantages. The advantage of the first approach is that it is relatively easy and accurate to calculate the value of these instruments. However, these instruments obviously have less explanatory power than the instruments from the second approach

[1] Since the paths of some of the instruments might be very similar (especially for short maturities) it might be necessary to exclude a few instruments to prevent singular (non-invertible) matrices.

where the whole yield structure is mimicked. For the second approach this argument holds in the opposite order, such an instrument is very precise in following the yield but rather difficult to price.

3.2. Weights in replicating portfolio

In this section the method used to determine the weight of each replicating instrument in the replicating portfolio will be described. The method will, for simplicity, be illustrated with a pure cash-flow product, the simplest product on the market. However, this can easily be extended to allow for more complex products as long as an expected value (or a best-estimate) of the cash-flows is available. The following input variables are needed to construct a replicating portfolio:

- The (expected) value of the cash-flows is known for each point in time t = 0,1,…,T, where T is the time of the last payment.
- A set of *m* random scenario's. Depending on which product is evaluated one needs the interest rates, inflation rates and MY-rates for each scenario. For the pure cash-flow sample only the interest rates would be sufficient.
- One base scenario, the base scenario consists of the best-estimate of the interest rates. Under the no-arbitrage assumption this is equal to the current term-structure of interest rates. This scenario will be used as an out-of-sample benchmark.
- A set of *n* replicating instruments. Since the example considers only pure cash-flows a set of *n* discount bonds is used.

We then proceed as follows:

1. **Calculate the value of the product in all scenarios.**
   As already explained, it is favorable to keep the problem as small as possible. We therefore chose to use the aggregate cash flow approach because matching all cash flows exactly would involve too many replicating instruments. Furthermore we think that the NPV grasps the most important property of the product, the current value. Therefore the main concern is that the net present value (NPV) of the product in each scenario matches the NPV of the replicating portfolio as close as possible. The NPV of the product is simply the sum of the NPVs of the cash-flows. The NPV of a cash-flow can be calculated by multiplying the cash-flow with the appropriate discount factor. E.g. the cash-flow at time t is multiplied with the discount factor at time t. Discount factors are defined by:
   $$D_0 = 1$$
   $$D_1 = \frac{D_0}{1+r_1}$$
   $$D_t = \frac{D_{t-1}}{1+r_t} = \frac{1-\sum_{k=1}^{t-1}(D_k) * r_t}{1+r_t}$$
   where D_t is the discount factor at time t, and r_t is the interest rate at time t. The interest rate should be converted in such a way that the length of the periods match, that is if the cash-flows are monthly the interest rate should be monthly as well.

   The NPVs for each scenario are then aggregated in a *m x 1* vector $Y = [V_0,\dots V_m]^T$ where V_i is the NPV of the product in scenario i.

2. **Determine the value of the replicating instruments in all scenarios.**
   The value of a replicating instrument is determined in the same way as the NPV of the product: the sum product of the cash-flows and the appropriate discount factors. In this example the replicating instruments are solely discount bonds. A discount bond is a payment of 1 at a

certain point in time, therefore the value of such discount bond is equal to the appropriate discount factor for that point in time.

For each instrument this generates a column vector with *m* entries (quite similar to the vector Y). Aggregating these vectors in a matrix gives a *m x n* matrix X:

$$\begin{bmatrix} Df_{0,1} & \cdots & Df_{0,n} \\ \vdots & \ddots & \vdots \\ Df_{m-1,1} & \cdots & Df_{m-1,n} \end{bmatrix}$$

where $Df_{i,j}$ is the value of discount factor *j* in scenario *i*.

<u>Note:</u> the actual data-generating process of the product (the calculation of the NPV and the best-estimates of the cash-flows) is actually known in this case. Hence, the product could be perfectly replicated by adding a discount bond for each point in time to the replicating portfolio. However, we expect that we can get quite precise with only a small subset of these regressors. The error that remains is solely due to the fact that variables are omitted, resulting in error in coefficients see section 2.5.

3. **Determine the weight of the replicating instruments.**
   The weights are chosen such that they minimize the squared differences of the product's NPV and the NPV of the replicating portfolio. The first methods minimizes the overall difference, while the second method minimizes the squared differences under (a) certain constraint(s). In other words this means that the portfolios created are the best matching portfolios in these 1500 scenarios. We assume that these 1500 scenarios are representative for the real world, and therefore this is the best estimate for the best matching portfolio.

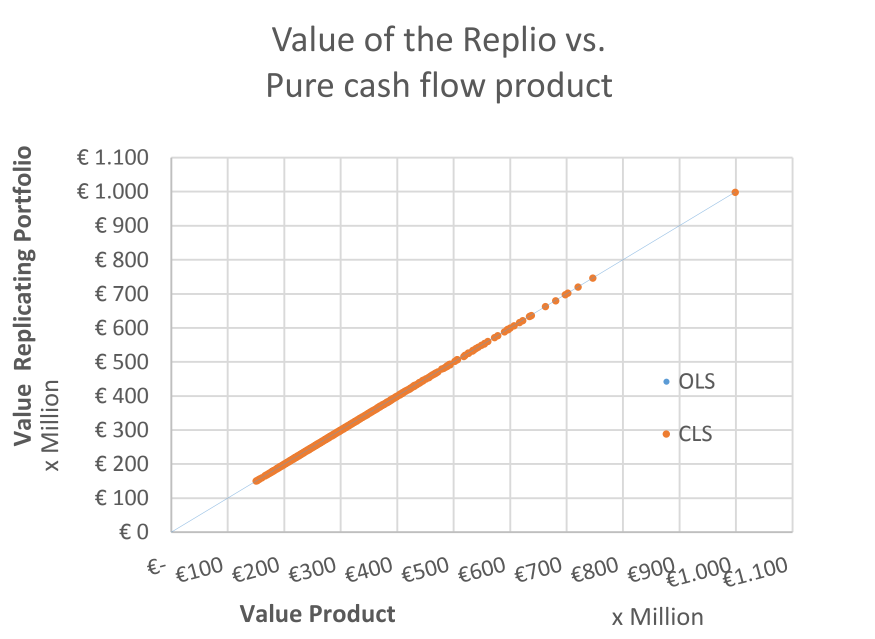
4. Results

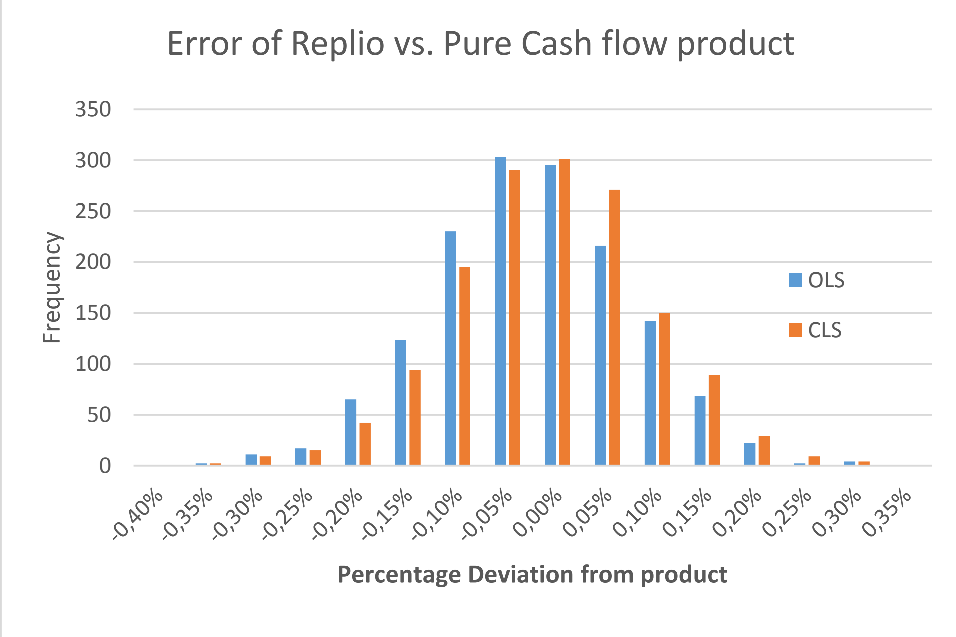
4.1. Results pure cash-flow replicating portfolio, using 20 discount bonds

Errors of replicating portfolio w.r.t. pure cash-flow product								
Quantile	**OLS (absolute value)**		**OLS (percentage)**		**CLS (absolute value)**		**CLS (percentage)**	
1,0000	€ -1.411.875,06	-0,286%	€ -1.022.452,40	-0,345%	€ -1.289.530,14	-0,262%	€ -1.017.158,59	-0,343%
0,9993	€ -1.268.348,50	-0,260%	€ -1.221.027,60	-0,319%	€ -1.266.848,98	-0,331%	€ -1.266.848,98	-0,331%
0,9987	€ -1.221.027,60	-0,319%	€ -1.411.875,06	-0,286%	€ -1.196.620,33	-0,245%	€ -903.208,68	-0,288%
0,9980	€ -1.146.073,46	-0,192%	€ -713.164,31	-0,286%	€ -1.188.889,41	-0,261%	€ -855.947,22	-0,268%
0,9973	€ -1.104.974,30	-0,243%	€ -747.606,46	-0,286%	€ -1.020.730,87	-0,265%	€ -699.850,36	-0,267%
0,9967	€ -1.095.529,56	-0,285%	€ -1.095.529,56	-0,285%	€ -1.017.158,59	-0,343%	€ -663.278,36	-0,266%
0,9960	€ -1.022.725,09	-0,192%	€ -997.291,09	-0,274%	€ -985.441,48	-0,165%	€ -1.020.730,87	-0,265%
0,9953	€ -1.022.452,40	-0,345%	€ -854.935,70	-0,273%	€ -955.665,51	-0,263%	€ -955.665,51	-0,263%
0,9947	€ -997.291,09	-0,274%	€ -769.687,36	-0,267%	€ -913.492,22	-0,171%	€ -1.289.530,14	-0,262%
0,9940	€ -970.471,85	-0,237%	€ -607.723,55	-0,265%	€ -910.318,50	-0,222%	€ -1.188.889,41	-0,261%
0,9933	€ -854.935,70	-0,273%	€ -845.822,06	-0,265%	€ -903.208,68	-0,288%	€ -733.155,94	-0,254%
0,9927	€ -845.822,06	-0,265%	€ -803.880,71	-0,263%	€ -855.947,22	-0,268%	€ -1.196.620,33	-0,245%

Ordinary Least Squares	
Mean	**294893061**
Total sum of squares	**1,031E+19**
Ssres	**1,322E+14**
R-squared	**0,9999872**

Constrained Least Squares	
Mean	**294893061**
Total sum of squares	**1,031E+19**
Ssres	**1,367E+14**
R-squared	**0,9999867**





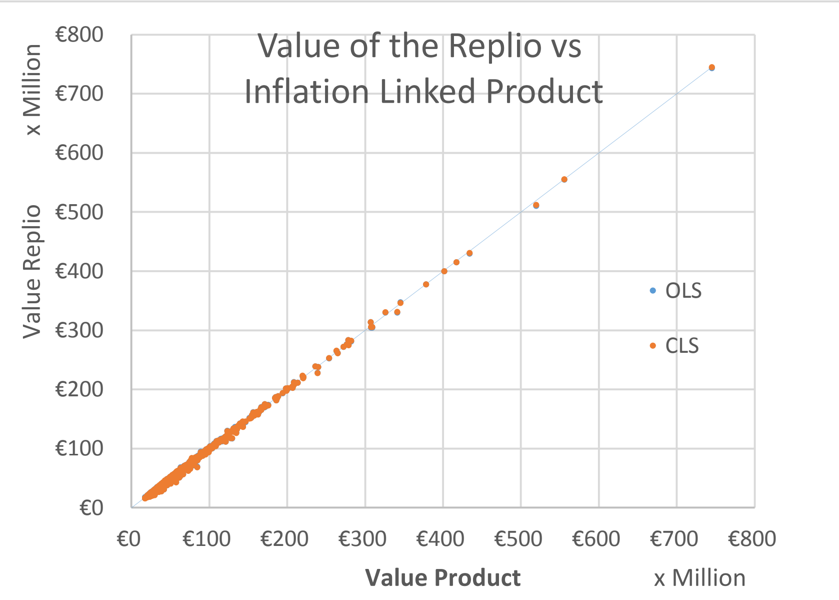
4.2. Results Inflation Linked Portfolio, using 10 discount bonds and 10 Inflation linked instruments

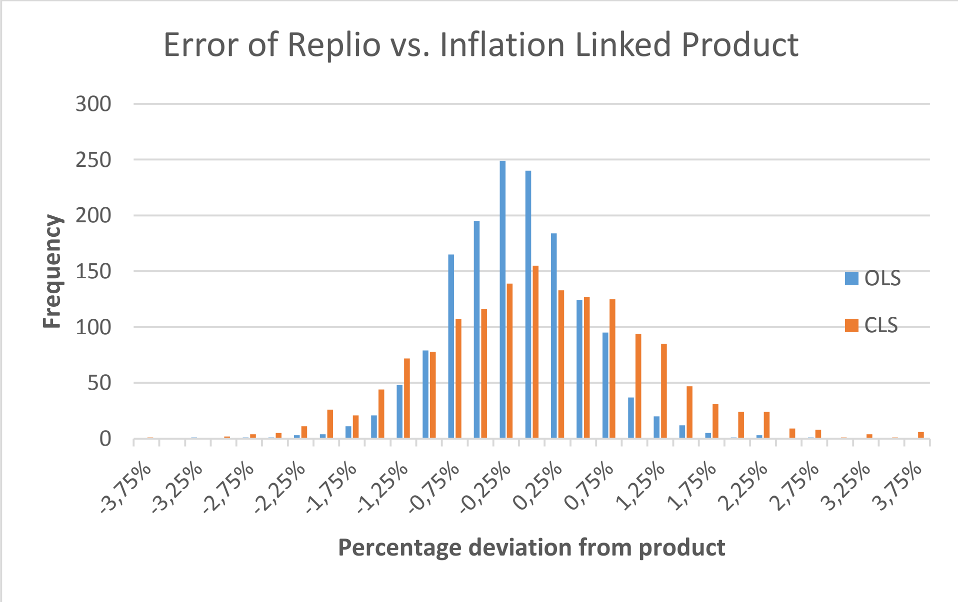
Errors of replicating portfolio w.r.t. inflation linked portfolio -Own products

Quantile	OLS (absolute value)		OLS (percentage)		CLS (absolute value)		CLS (percentage)	
1,0000	€ -8.897.845,39	-2,052%	€ -4.079.742,01	-3,020%	€ -8.545.053,75	-1,971%	€ -809.192,52	-3,524%
0,9993	€ -5.405.190,38	-2,585%	€ -5.405.190,38	-2,585%	€ -5.519.074,98	-2,640%	€ -630.111,77	-2,946%
0,9987	€ -4.616.103,85	-1,338%	€ -1.377.285,68	-2,318%	€ -4.793.145,15	-1,472%	€ -695.840,13	-2,801%
0,9980	€ -4.316.031,11	-1,325%	€ -2.620.288,10	-2,165%	€ -4.717.238,29	-1,367%	€ -3.699.967,12	-2,739%
0,9973	€ -4.079.742,01	-3,020%	€ -1.718.236,01	-2,150%	€ -3.699.967,12	-2,739%	€ -1.018.997,42	-2,684%
0,9967	€ -3.230.908,15	-1,624%	€ -8.897.845,39	-2,052%	€ -3.438.020,69	-1,458%	€ -5.519.074,98	-2,640%
0,9960	€ -3.222.506,56	-1,948%	€ -3.222.506,56	-1,948%	€ -3.326.033,97	-2,010%	€ -550.957,13	-2,520%
0,9953	€ -3.085.358,78	-1,308%	€ -983.633,39	-1,915%	€ -3.145.979,59	-0,422%	€ -723.723,79	-2,495%
0,9947	€ -2.859.821,89	-0,384%	€ -404.560,63	-1,891%	€ -3.033.320,22	-1,525%	€ -884.370,76	-2,426%
0,9940	€ -2.620.288,10	-2,165%	€ -1.670.566,31	-1,803%	€ -2.740.125,25	-1,606%	€ -689.578,63	-2,327%
0,9933	€ -2.606.942,56	-1,528%	€ -1.044.997,38	-1,742%	€ -2.577.875,32	-2,130%	€ -681.897,92	-2,267%
0,9927	€ -2.126.960,84	-0,692%	€ -1.466.598,53	-1,655%	€ -2.473.698,87	-1,472%	€ -2.098.284,12	-2,265%

Ordinary Least Squares	
Mean	62553608
Total sum squares	3,92E+18
Res sum squares	5,93E+15
R-squared	0,9984880

Constrained Least Squares	
Mean	62553608
Total sum squares	3,92E+18
Res sum squares	6,81E+15
R squared	0,9982642





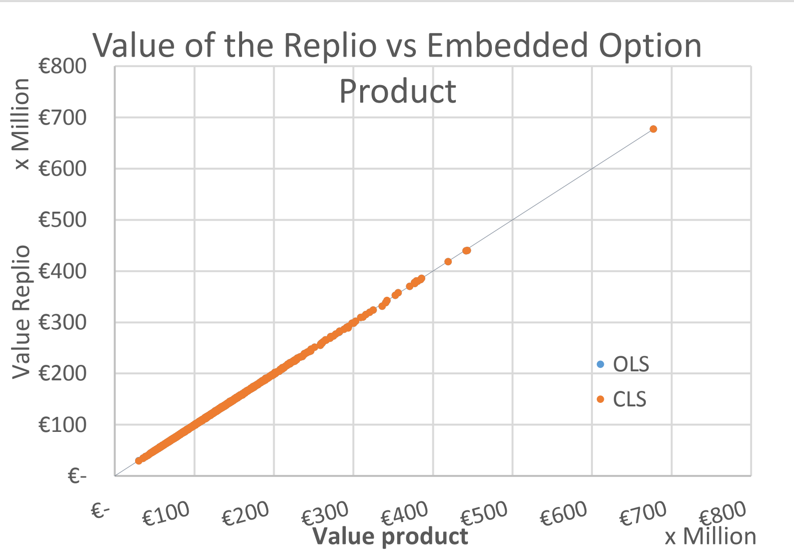
4.3. Results Embedded Option Portfolio, using 10 discount bonds and 10 MY-Staffel instruments[2]

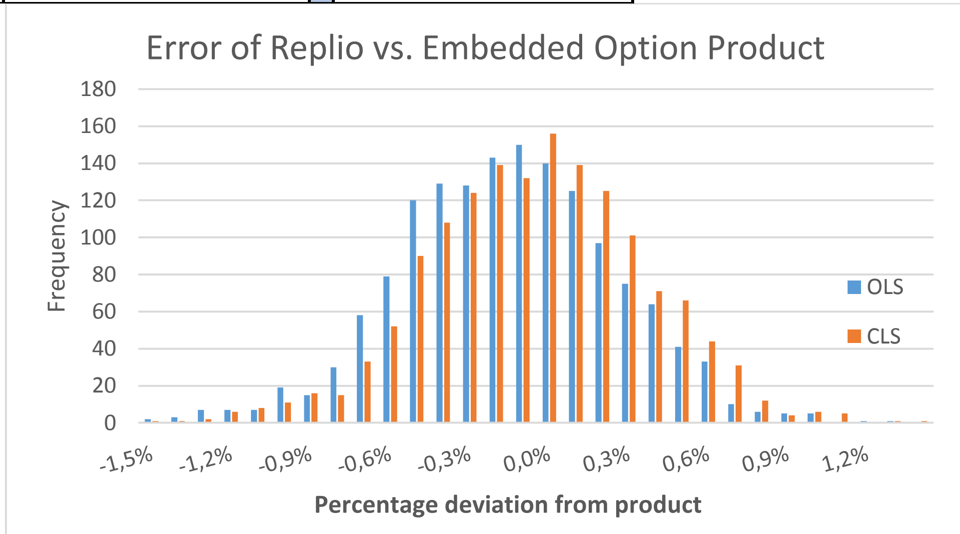
Upper quantile of the errors of the replicating portfolio w.r.t. the Embedded option product

Quantile	OLS (absolute value)		OLS (percentage)		CLS (absolute value)		CLS (percentage)	
1,0000	€ -3.987.823,99	-1,189%	€ -1.360.201,80	-1,346%	€ -3.954.245,68	-1,179%	€ -594.639,32	-1,460%
0,9993	€ -3.622.667,78	-1,236%	€ -538.603,73	-1,322%	€ -3.688.259,98	-1,258%	€ -1.345.649,75	-1,332%
0,9987	€ -2.276.650,72	-0,514%	€ -3.622.667,78	-1,236%	€ -2.281.754,00	-0,515%	€ -765.223,34	-1,289%
0,9980	€ -2.147.910,44	-0,833%	€ -727.429,01	-1,225%	€ -2.150.813,63	-0,914%	€ -3.688.259,98	-1,258%
0,9973	€ -2.101.700,20	-0,893%	€ -1.074.218,36	-1,223%	€ -2.148.655,13	-0,834%	€ -1.048.091,12	-1,193%
0,9967	€ -1.670.531,86	-0,679%	€ -3.987.823,99	-1,189%	€ -1.641.167,61	-0,667%	€ -3.954.245,68	-1,179%
0,9960	€ -1.651.207,45	-0,800%	€ -707.601,37	-1,154%	€ -1.633.436,55	-0,792%	€ -1.337.255,53	-1,177%
0,9953	€ -1.631.043,71	-1,018%	€ -1.001.189,76	-1,150%	€ -1.616.225,59	-1,009%	€ -1.001.762,84	-1,137%
0,9947	€ -1.552.302,70	-0,687%	€ -1.290.702,82	-1,136%	€ -1.575.231,37	-0,697%	€ -860.404,37	-1,111%
0,9940	€ -1.364.041,96	-0,662%	€ -997.846,85	-1,133%	€ -1.369.049,12	-0,665%	€ -678.487,66	-1,107%
0,9933	€ -1.362.613,59	-0,691%	€ -637.594,74	-1,118%	€ -1.345.649,75	-1,332%	€ -624.280,65	-1,095%
0,9927	€ -1.360.201,80	-1,346%	€ -859.001,56	-1,109%	€ -1.337.255,53	-1,177%	€ -952.371,90	-1,094%

Ordinary Least Squares	
Mean	62553607,71
Total sum of squares	3,92216E+18
Ssres	7,52838E+14
R-squared	0,9998081

Constraint Least Squares	
Mean	62553607,71
Total sum of squares	3,92216E+18
Ssres	9,48809E+14
R-squared	0,9997581





[2] Note that these are the results from the second approach for the Option Embedded portfolio, the results from the first approach were too unsatisfactory.

What can be derived from these graphs and tables (pages 16, 17 and 18)? Let us start with the results of the pure cash flow product and the corresponding replicating portfolio. Remember that a positive deviation is less of a problem than a negative deviation, this negative deviation is known as downside risk and is displayed in the big table (page 16). For both the OLS and the CLS approach the table summarizes the worst fits, either based on the worst absolute errors or the worst relative errors. The table shows that the relative errors are very small, at most -0,345%, a negligible percentage. The distribution of the errors is shown in the histogram (page 16), the distribution appears to be highly symmetrical for the OLS approach. The errors of the CLS approach seem to be symmetrical as well, but when taking a closer look it can be seen that the distribution is skewed to the right. This is actually a favorable property, it means there is less downside risk and more upside potential. This skewness in relative error is also evident from the table, the highest relative deviation is smaller for the CLS approach and this holds for most of the quantiles in the table.

To conclude, the overall fit (and thus R-Squared) of the OLS approach are better (by construction) than the CLS approach. However, if this replicating portfolio is to be used as a hedging instrument it is favorable to have positive errors, because that means that there is excess money in the replicating portfolio. From this point of view the CLS approach is therefore more useful. Other reasons to choose for the CLS approach are given in section 2.4.2 and 6.

The first thing that we notice about the results for the inflation linked product (page 17) is that the fit is worse. This replicating portfolio makes a bigger error of ~3% (OLS) or ~3,5%(CLS). Whilst this relative error might still be workable and is sufficiently small to be useful for calculations it is worse than for the cash flow product. The distribution of the relative errors is again quite symmetric, both for the OLS and CLS approach, however there are some clear differences between them. The errors from the CLS approach have a significantly lower kurtosis and fatter tails. In addition, one can notice that OLS errors are slight skewed to the left (of 0%) and the CLS errors are slightly skewed to the right.

This time it is harder to say which replicating approach is more attractive. Whilst the skewness of the distribution is definitely in favor of the CLS approach the fatter tails are definitely not. It is therefore more dependent on what the replicating portfolio will be used for pricing at time t = 0 the OLS approach might be favorable and for hedging the disadvantage of the fatter tails might be outweighed by the skewness to the right. Hence, there is not necessarily a better choice as to which approach is better. The choice will be dependent on the situation and the decision be taken at the users discretion.

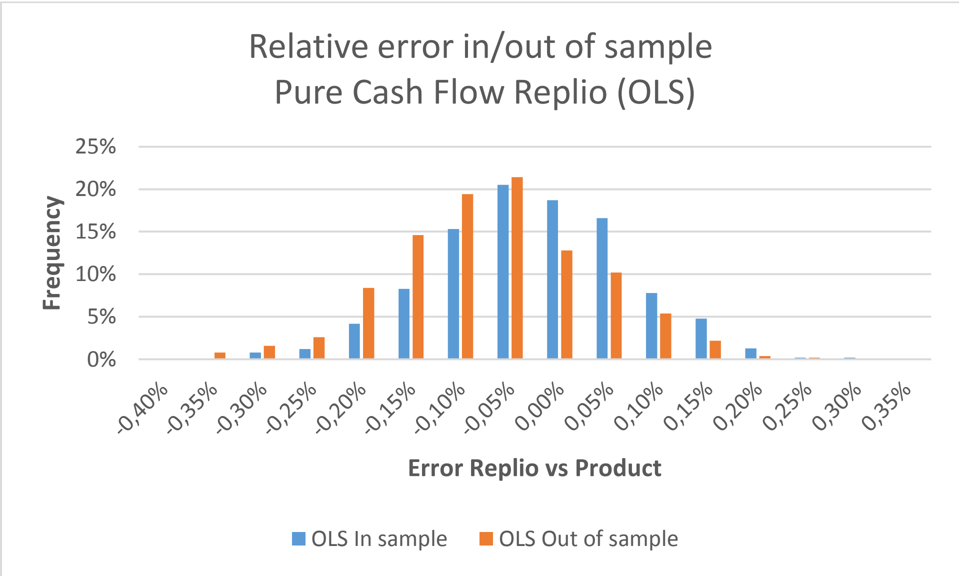
The last replicating portfolio was created for the embedded option product (page 18). The fit of this replicating portfolio is again very good. The worst case relative errors are less than 1,5% in both approaches and therefore significantly better than the errors of the inflation linked portfolio. The distribution of the relative error seems to be consistent with the previous observations: for both approaches highly symmetrical and the errors of the CLS approach are skewed to the right. In this case the CLS approach is again favorable.

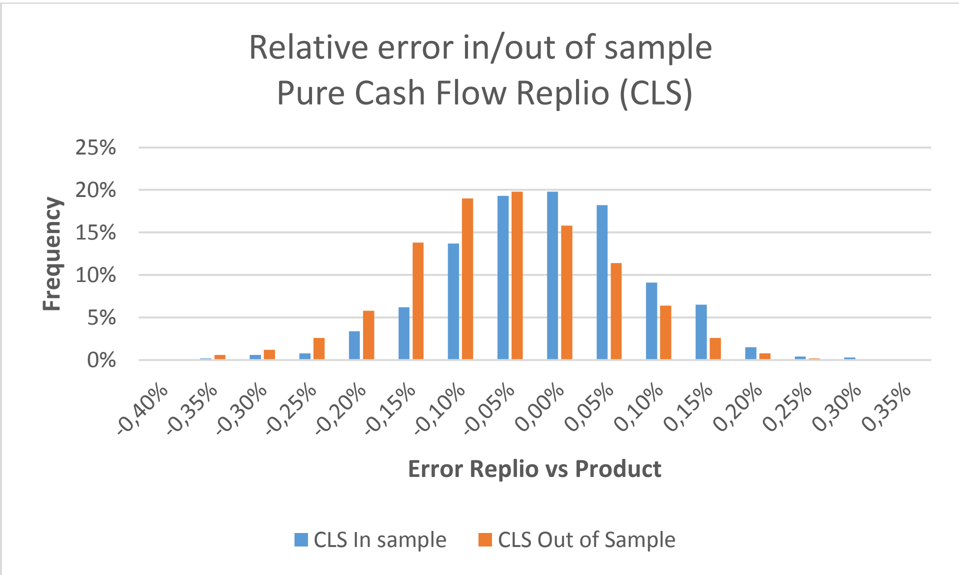
It is necessary to make a few comments about these results. Firstly note that these errors and results are based on the test scenarios used and that the real life possibilities are not limited by these test scenarios. However, we assume that these test scenarios are diverse enough to have explanatory power. It is therefore also acceptable to claim that these are the quantiles we should keep into account. Secondly we would like to comment on the results of the CLS approach; we saw that the (relative) errors were skewed more to the right than in the OLS approach. Obviously it depends on what constraint is used, apparently the current base scenario causes the relative error to skew to the right. This is most likely not the case for any base scenario and this conclusion should therefore be used cautiously.

4.4. In sample vs out of sample

To test whether the replicating portfolios are actually useful and not just a result from data mining we performed and in and out of sample test. For this test a total of 1500 scenarios was used. The first 1000 scenarios were used to estimate and calibrate a replicating portfolio. The weights of the replicating instruments in these portfolios can be found in Appendix B:. Now that a portfolio is found, it can be tested against the remaining 500 scenarios. From this we find the errors that were made in the set of 1000 scenarios, the in sample errors, and the errors made in the 500 scenarios, the out of sample errors. If these errors differ a lot from each other it means that the portfolio was only fitted for those specific 1000 scenarios and that it has less or no explanatory power for other scenarios. In contrast, when the errors are of similar magnitude it means that the replicating portfolio also has explanatory power beyond the initial set of scenarios.

4.4.1. Pure Cash Flow



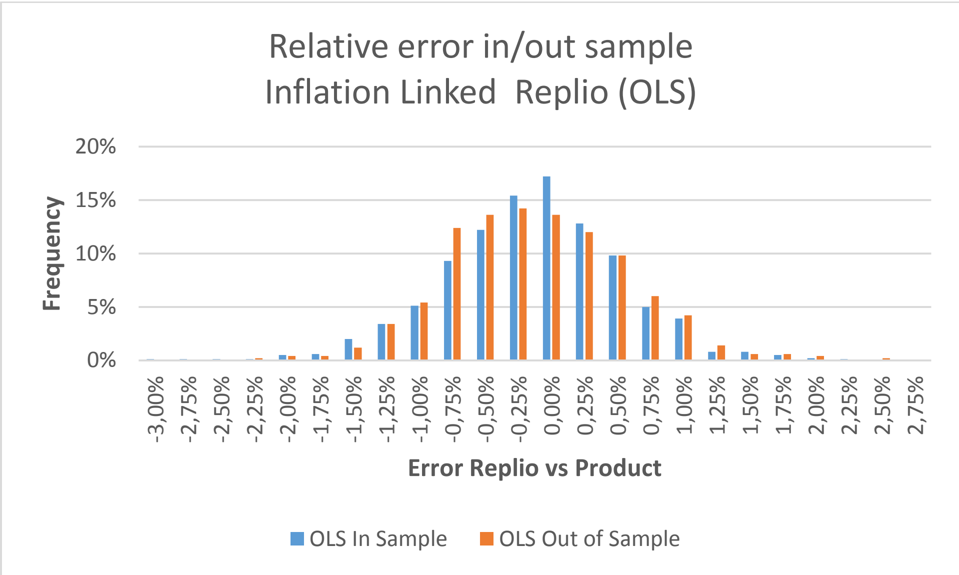


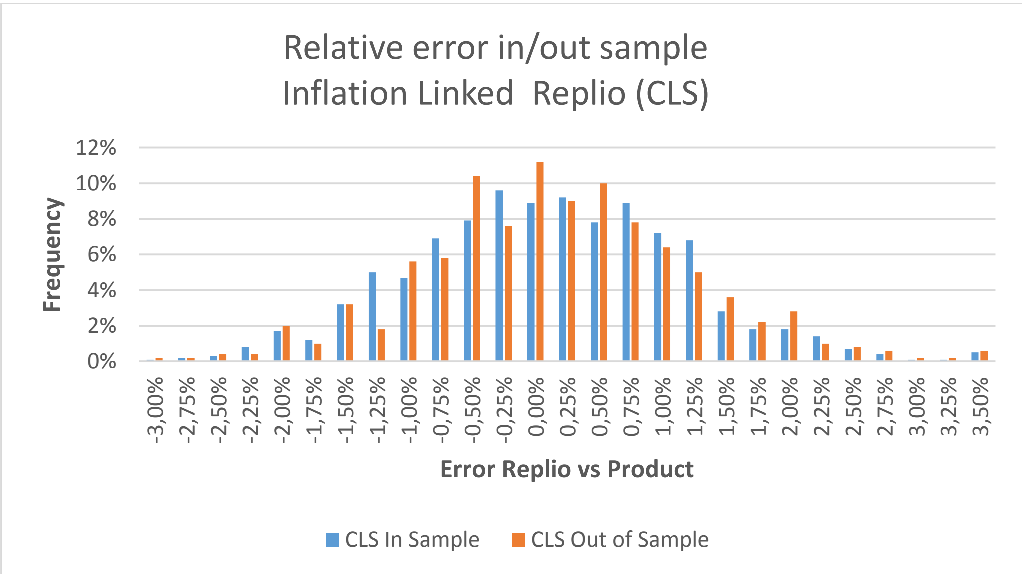
These graphs show that that there is a slight difference in the distribution of the relative errors. The out of sample error seems to be skewed a little to the left, although it is an undesirable property it is only a minor difference. Looking at the complete picture the magnitude of the errors is quite similar. This is also evident from Table 2 Maximum/minimum errors pure cash flowTable 2 where we see that the minimum error actually decreases and the maximum error slightly increases. The differences (positive and negative) between the in and out sample error is approximately 10-20%, which seems to be a workable amount, especially considering the fact that the interval of the errors has become smaller. From these observations we conclude that the replicating portfolio has explanatory power for the product. Furthermore, since the minimum error decreases it is most certainly useful for hedging purposes.

	In Sample		**Out of Sample**	
	OLS	*CLS*	*OLS*	*CLS*
Minimum Relative Error	-0,35%	-0,35%	-0,29%	-0,27%
Maximum Relative error	0,30%	0,32%	0,33%	0,34%

Table 2 Maximum/minimum errors pure cash flow

4.4.2. Inflation linked



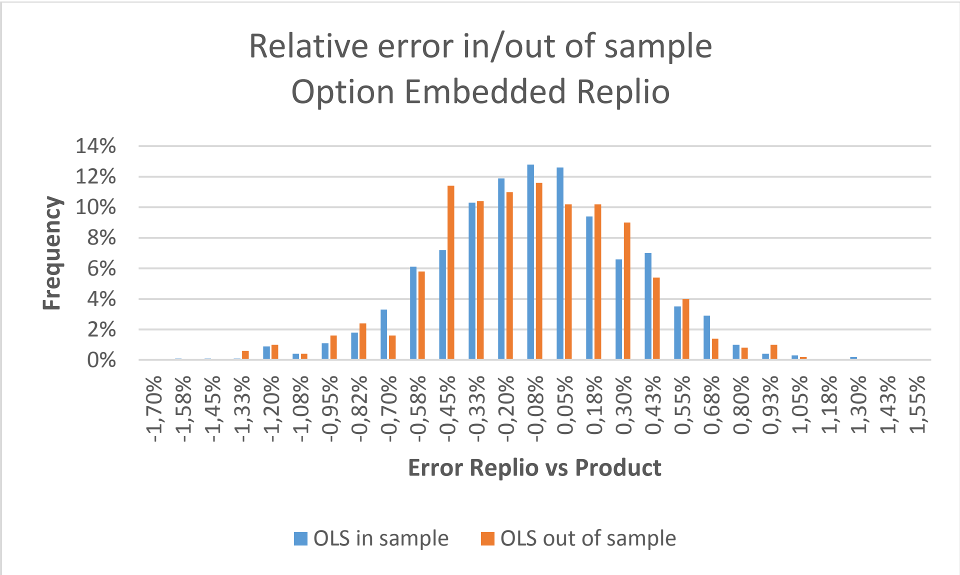


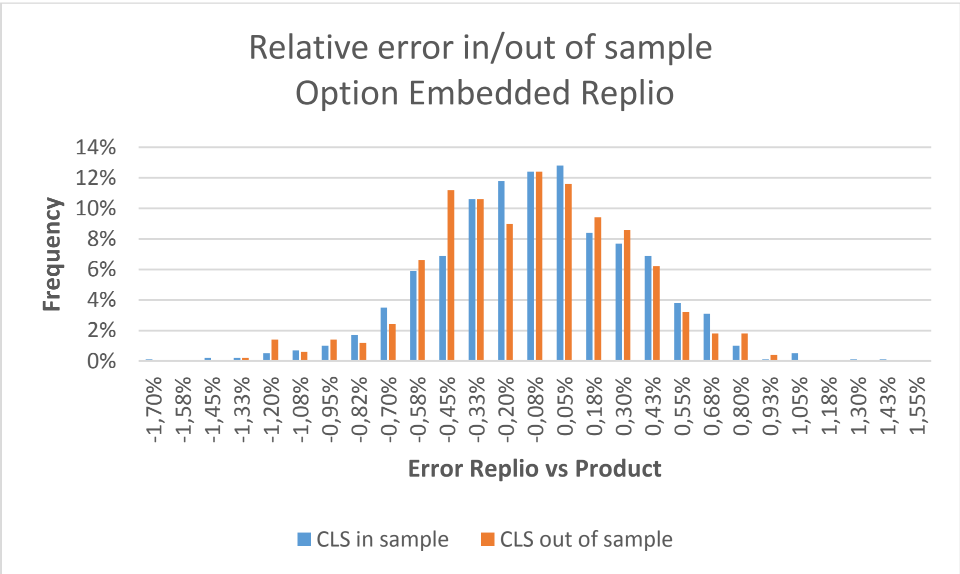
For the inflation linked replicating portfolio the graphs do not show any clear differences. Although at some specific levels the errors vary slightly, the general shape is almost the same. This is an indication that the replicating portfolio has good explanatory power of the product. When looking at the differences between the maxima and minima there are some differences. The most notable differences is the 33% decrease for the maximum error for the out of sample CLS approximation. Although a decrease in error is favorable it shows that there might be some differences between the in and out of sample errors. However, as we already saw from the graph this is probably an “unlucky” coincidence, since the overall shape of the errors is similar. The minimum and maximum in and out of sample errors again differ by at most 10-25%. Also, the interval between the maximum and minimum is decreased, just as was the case with the pure cash flow product. With these observations in mind we think that the replicating portfolio is able to explain the behavior of the real product.

	In Sample		Out of Sample	
	OLS	*CLS*	*OLS*	*CLS*
Minimum Relative Error	-2,96%	-2,79%	-2,18%	-3,36%
Maximum Relative error	2,26%	6,17%	2,59%	3,98%

Table 3 Minimum/maximum errors inflation linked

4.4.3. Option embedded





The differences between the relative errors in- and- out of sample are again very small. Apart from an occasional outlier, the general shape of the distribution is almost equal. This means that the replicating portfolio has (approximately) the same explanatory power of the product in the calibration sample as outside the calibration sample. Since the errors are sufficiently small it can be said in general that this portfolio has good explanatory power of the product (both the OLS and CLS portfolio). When looking at the maximum and minimum errors we notice the same phenomena as in the inflation linked replicating portfolio. The extrema of the out of sample errors are slightly smaller than the extrema in the sample. Although this smaller interval is favorable it is not clear why this is. Since the rest of the distribution is very similar we assume that this does not have any effects on our results.

	In Sample		**Out of Sample**	
	OLS	*CLS*	*OLS*	*CLS*
Minimum Relative Error	-1,48%	-1,67%	-1,25%	-1,21%
Maximum Relative error	1,38%	1,43%	1,12%	1,02%

Table 4 Minimum/maximum error option embedded

5. What can be said about the reliability of the errors?

We have seen that our replicating portfolio makes some errors since it is not a perfectly replicating portfolio. But what can we say about the reliability of these errors. How sure is it that the quantile of the errors that has been identified is the real quantile? When looking at the reliability we encounter these problems.

1. One of the assumptions for linear regressions is that the errors are independent and identically distributed (i.i.d.) distributed. Because our scenarios are generated by an i.i.d. autoregressive process the errors are i.i.d. by construction. If in addition the errors are normally distributed it becomes very easy to make inferences about the model or estimate the confidence intervals. However, we fail to accept the null hypothesis for normality, which means confidence intervals can't be constructed in the usual manner.
2. The true distribution of the errors is unknown. It is possible to determine the empirical distribution function, but this is most likely not exactly equal to the true distribution

Order statistics approach

Another approach at estimating the 99,5% worst case scenario is by using order statistics. In statistics the k^{th} order statistic is equal to the k^{th} smallest value. This order statistic are point estimates for the quantiles of the underlying distribution. As n tends to infinity this p^{th} quantile is asymptotically normally distributed, this is approximated by:

$$X_{([np])} \sim AN\left(F^{-1}(p), \frac{p(1-p)}{n\left[f\left(F^{-1}(p)\right)\right]^2}\right)$$

Where f is the density function, and F^{-1} is the quantile function for the distribution F. This approximation was proven in 1946 by Mosteller and improved by Bahadur in 1960. However, the problem is that the real distribution F is unknown, therefore the empirical counterpart F' is used. This means that we intentionally approximate for the sake of simplicity. We conjecture that this error, as n tends to infinity, converges to zero.

To find the 99,5% quantile of the errors at a 95% confidence interval we do the following:

1. Let X_i be the observations and $X_{(i)}$ the ordered observations. Let $U(i) = F_X(X_{(i)})$

2. Then for a large sample the p[th] sample quantile is asymptotically normally distributed

$$U_{[np]} \sim AN\left(p, \frac{p(1-p)}{n}\right)$$

$$p = 0{,}995$$
$$n = 1500$$

$$U_{(1493)} \sim AN\left(0{,}995; \frac{4{,}98 * 10^{-3}}{1500}\right)$$

3. Hence the confidence interval:

$$0{,}995 - 1.96 * \left(\sqrt{\frac{4{,}958 * 10^{-3}}{1500}}\right) < p < 0{,}995 + 1{,}96 * \left(\sqrt{\frac{4{,}98 * 10^{-3}}{1500}}\right)$$

$$0{,}9914 < p < 0{,}9985$$

Hence, with 95% certainty the 99,5% quantile is smaller than the empirical 99,85% quantile. To determine this quantile we use linear interpolation between the 1497[th] and 1498[th] order statistic. We then find that the best estimate for the 99,5% quantile is €-1.009.871, with a 95% confidence interval of (€ -808,396 --- € -1,210,103).

k[th] - Order Statistic	Quantile	Loss
1487	0.9913	€ -803.880,71
1488	0.9920	€ -834.857,00
1489	0.9927	€ -845.822,06
1490	0.9933	€ -854.935,70
1491	0.9940	€ -970.471,85
1492	0.9947	€ -997.291,09
1493	0.9953	€ -1.022.452,40
1494	0.9960	€ -1.022.725,09
1495	0.9967	€ -1.095.529,56
1496	0.9973	€ -1.104.974,30
1497	0.9980	€ -1.146.073,46
1498	0.9987	€ -1.221.027,60
1499	0.9993	€ -1.268.348,50
1500	1.0000	€ -1.411.875,06

Tabel 1: The order statistics and quantiles from the pure cash-flow product[3]

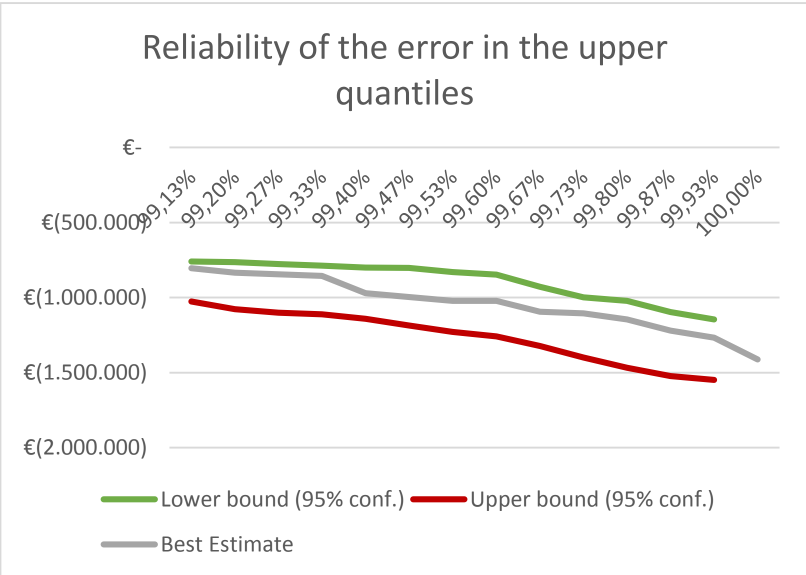


Figure 1 95% confidence interval of the upper quantiles

The graph shows the upper quantiles (>99%) with their corresponding 95% confidence bounds. Because the quantiles are so "high" we see that the upper bound is significantly further away than the lower bound.

[3] Note that the losses are ordered in a decreasing order instead of increasing order. This is done because the interests lies with the negative errors. Positive errors are not that important.

6. Projections

6.1. Projections towards the future

The replicating portfolio was constructed to match the product at time $t = 0$. It can therefore be seen as a hedge at time $t = 0$. For a hedge it is important that the value of the replicating portfolio is the same, or very close to the value of the product in all possible scenarios. The best possible situation would be a perfect hedge, where the value of the portfolio and the product would be the same in any scenario (even a scenario outside of our test set). However, since the replicating instruments are only a subset of the real data generating instruments, it is very unlikely that a perfect hedge can be found. We therefore aim at an efficient portfolio, a portfolio that minimizes the variance of the errors. In general the portfolio created by an OLS estimation will be an efficient portfolio given the set of replicating instruments. Whilst the CLS estimation will give the efficient portfolio given the set of constraints and the set of replicating instruments.

Although it is important that the portfolio is efficient at $t = 0$ we are also interested in the error over time. What happens to this error, will it increase, decrease or stay the same? Because the NPV takes the time value of money into account we expect this error to remain constant over time.

Firstly, considering that the value of the product, at any time point, could be determined with the recursive formula. Unfortunately a little more information is needed to determine the value of the replicating portfolio at some future time $t^* > 0$. The value of the replicating portfolio is build up out of two components: the value of the replicating instruments and the outgoing payments.

6.2. Valuing replicating instruments

Valuing discount bonds

First of all, remember that the replicating instruments "pay" at the beginning of the month and are therefore discount with the discount factor of that time-period, e.g. at time $t = 0$ (1-1-2015) the discount factor is 1.00. The next step is to calculate the NPV of these replicating instruments, which is simply the sum-product of the cash-flows from the replicating instruments and their appropriate discount factors. This gives the NPV of the replicating instruments at time $t = 0$. To see what it is worth in the future we make the following assumption:

> **Assumption 6:** *All cash received from a replicating instrument is immediately placed into a bank account and interest is accrued. If this cash position would be negative this amount would be borrowed from a bank at the same interest rate.*

This assumption makes sure that no money 'leaks' from our system. Under this assumption it is now very easy to roll the instruments into the future. The instruments can be rolled forward to time t^* by accruing interest over this time or mathematically, by dividing the NPV at time 0 by the discount factor D_{t^*} at time t^*.

Valuing inflation linked bonds

The value of an inflation linked bond is slightly different then the valuation of a discount bond. Instead of paying €1,- at maturity it now pay a, yet, unknown quantity at maturity. This quantity is dependent on the inflation rate and is compounded yearly. The cash flow it will generate can be computed as follows:

$$\text{cashflow}(T) = \text{principal}(0) * \prod_{i=0}^{T-1}(1 + \max\{0{,}0025; \pi_i\})$$

where T is the maturity of the inflation linked bond. The calculation of the NPV and rolling payments to a future time t^* is done in the same way as with the discount bond.

6.3. Value outgoing payments

The calculation is similar to the calculation for the discount bonds. However the payments of the outgoing premiums are at the end-of-month. To limit the complexity of the model such a payment is simply assumed to be paid on the first day of the next month:

> **Assumption 7:** *an end-of-month cash-flow in month i will be considered to be a begin-of-month cash-flow in month i+1.*

The payments can be rolled forward to time t^* in the same way as with the discount bond.

6.4. Value of the replicating portfolio

The value of the replicating portfolio is the sum of the NPV's of the replicating instruments. However, over time payments are made to policy holders, these payments need to be extracted from the replicating portfolio. Therefore the NPV of the payments that are already made, are subtracted to find the real value of the replicating portfolio. The portfolio is again rolled forward by dividing by the discount factor of the time to which it is rolled forward.

$$NPV_{replio}(t^*) = \frac{\overbrace{\sum_{i=0}^{T} D_{monthly}(i) * cashflow(i)}^{\text{Value replicating instruments}} - \overbrace{\sum_{j=0}^{t^*} D_{monthly}(j) * payment(j)}^{\text{Payments paid to policy-holders}}}{D_{monthly}(t^*)}$$

<u>Example:</u>

Time	$D_{monthly}(time)$	π_{time}	# of discount bonds	# of inflation linked bonds	Outgoing payment
0	1	.001	50		0
1	.9981	.015	20	30	36
2	.9965	.012		10	30
3	.9935	.021			28
4	.9919	.020			8

Discount bonds:

Value at time 0: $\sum_{i=0}^{T} D_{monthly}(i) * cashflow(i) = 1 * 50 + .9981 * 20 = 69.962$

Value at time 3: $\frac{1*50+.9981*20}{.9935} = 70.420$

Inflation linked bonds:

Value at time 0: $\sum_{t}^{T} \text{principal}(t) * \prod_{i=0}^{t-1}(1 + \max\{0.0025; \pi_i\}) * D_{monthly}(t)$

$$= \big(30 * (1 + 0.0025)\big) * 0.9981 + \big(10 * (1 + 0.0025) * (1 + 0.015)\big) * 0.9965 = 40.15$$

Value at time 3: $\frac{value\ at\ time\ 0}{D_{monthly}(3)} =$

$$\frac{\Big(\big(30 * (1 + .0025)\big) * 0.9981 + \big(10 * (1 + 0.0025) * (1 + 0.015)\big) * 0.9965\Big)}{.9935} = 40.42$$

Replicating portfolio

Value at time 0: $Value\ replicating\ instruments - extracted\ payments$

$$= \frac{(69.962 + 40.15)}{1} - \frac{0}{1} = 110.112$$

Value at time 3: $\frac{69.962+40.15}{.9935} - \frac{(36*.9981+30*.9965+28*.9935)}{.9935} = 110.83 - 94.26 = 16.57$

6.5. Comparing: Replicating portfolio vs. Product

The value of the replicating portfolio can then be compared to the 'real' value of the product at each point in time. This error is shown in figure 2 below. It can be seen that the error does not remain the same over time as expected. However, the behavior of the error can be explained, the error behaves as an exponential curve, and more precisely the exponential curve belonging to the term structure of interest rates. It can be noticed that the percentage difference is equal to the interest in that period. Hence, the error at any time is simply the error at t = 0 with monthly compounded interest up to that time point. Therefore, if we can force the error to be zero at t = 0 with respect to the current term structure of interest, then the error will remain zero over time (see figure 3), as long as the term structure remains the same.

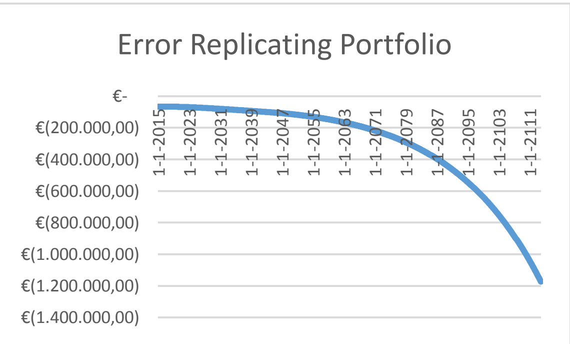


Figure 2 Difference between the replicating portfolio and the product over time.

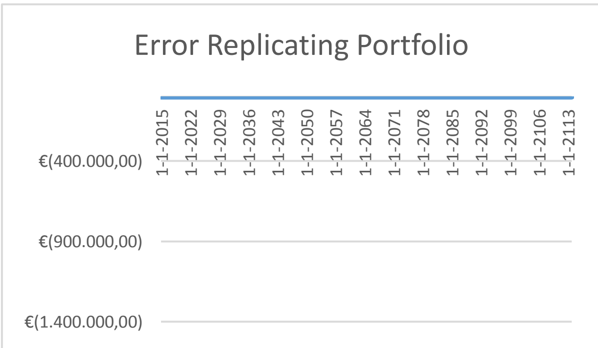


Figure 3 The error over time when the error at time t = 0 is forced to be zero

Theorem: *the error of the replicating portfolio with respect to the product at some future time t* is equal to the error at time t=0 with compounded interest until time t*. This only holds if the term structure of interest rates remains the same.*

Proof: This result can easily be proven mathematically. First we consider an example of a pure cash flow product and afterward we generalize the proof for all pure cash flow products.

Consider a cash flow product and a term structure of interest rates with the following characteristics and the corresponding replicating portfolio:

Time	Cash flows product	Cash flows replicating portfolio	Discount factor	Interest rate
0	x_0	y_0	$D_0 = 1$	r_0
1	x_1	y_1	D_1	r_1
2	x_2	y_2	D_2	r_2
3	x_3	y_3	D_3	r_3

Without loss of generality we assume the following:

> **Assumption 8**: *the cash flows at time t=0 do not have to be discounted. Hence the discount factor at time t=0* (D_0) *is assumed to be 1. Furthermore it is assumed that, without loss of generality, the inflation at time t=0* (π_0) *is equal to 0*

Firstly, the NPV of the product and the replicating portfolio are calculated at time t = 0. The value of the product is:

$$NPV_{prod,0} = x_0 + x_1 D_1 + x_2 D_2 + x_3 D_3$$

and of the replicating portfolio:

$$NPV_{replio,0} = y_0 + y_1 D_1 + y_2 D_2 + y_3 D_3$$

Hence, the error at time t = 0 is:

$$\begin{aligned}\epsilon_0 &= (y_0 + y_1 D_1 + y_2 D_2 + y_3 D_3) - (x_0 + x_1 D_1 + x_2 D_2 + x_3 D_3) \\ &= (y_0 - x_0) + (y_1 - x_1) D_1 + (y_2 - x_2) D_2 + (y_3 - x_3) D_3\end{aligned}$$

By definition of the discount factors this can be rewritten to:

$$\epsilon_0 = (y_0 - x_0) + \frac{(y_1 - x_1)}{1 + r_1} + \frac{(y_2 - x_2)}{(1 + r_1)(1 + r_2)} + \frac{(y_3 - x_3)}{(1 + r_1)(1 + r_2)(1 + r_3)}$$

Secondly, the NPV's are calculated for some later time point, say time t = 1.

$$NPV_{prod,1} = \frac{(x_0 + x_1 D_1 + x_2 D_2 + x_3 D_3)}{D_1}$$

$$NPV_{replio,1} = \frac{(y_0 + y_1 D_1 + y_2 D_2 + y_3 D_3)}{D_1}$$

Thus the error at time t = 1 is:

$$\epsilon_1 = \frac{(y_0 - x_0) + (y_1 - x_1) D_1 + (y_2 - x_2) D_2 + (y_3 - x_3) D_3}{D_1}$$

$$\epsilon_1 = \frac{\epsilon_0}{D_1} = \epsilon_0 * (1 + r_1)$$

From this it is obvious that the error at time 1, , only grows by the interest over ϵ_0 during the time period 0 to 1.

This result can very easily be generalized, suppose that $x_0, \dots, x_T$ are the cash flows of the product, and $y_0, \dots y_T$ are the cash flows of the discount bonds in the replicating portfolio and $r_0, \dots, r_T$ are the interest rates at each time point 0,…,T. Then the NPV's and the error are defined as follows:

$$NPV_{prod,t} = \frac{(x_0 + x_1 D_1 + \dots + x_T D_T)}{D_t}$$

$$NPV_{replio,t} = \frac{(y_0 + y_1 D_1 + \dots + y_T D_T)}{D_t}$$

$$\begin{aligned}\epsilon_0 &= (y_0 + y_1 D_1 + \dots + y_T D_T) - (x_0 + x_1 D_1 + \dots + x_T D_T) \\ &= (y_0 - x_0) + (y_1 - x_1) D_1 + \dots + (y_T - x_T) D_T\end{aligned}$$

$$\epsilon_t = \frac{(y_0 - x_0) + (y_1 - x_1)D_1 + \cdots + (y_T - x_T)D_T}{D_t} = \frac{(y_0 - x_0) + (y_1 - x_1)D_1 + \cdots + (y_T - x_T)D_T}{\frac{1}{\prod_{i=1}^{t}(1+r_i)}}$$

$$= \left((y_0 - x_0) + (y_1 - x_1)D_1 + \cdots + (y_T - x_T)D_T\right) * \prod_{i=1}^{t}(1+r_i) = \epsilon_0 * \prod_{i=1}^{t}(1+r_i)$$

Hence, for every point in time the error remains proportional to the compounded interest over ϵ_0.

The same property can be verified for the inflation linked product. In order verify this some extra notation is needed. Let $\pi_0, \ldots, \pi_T$ be the inflation rate and $z_1, \ldots, z_T$ be the principals of the inflation linked bonds included in the replicating portfolio at each time point t =0,…,T. And let G be the guaranteed return.

$$NPV_{replio,t} = \frac{\begin{array}{c}(y_0 + y_1 D_1 + \cdots + y_T D_T) + (z_0 + z_1 D_1(1+\max\{G;\pi_1\}) \\ +z_2 D_2(1+\max\{G;\pi_1\})(1+\max\{G;\pi_2\}) + \cdots + z_T D_T \prod_{i=1}^{T}(1+\max\{G;\pi_i\}))\end{array}}{D_t}$$

Let the inflation linked part be abbreviated:

$$Z = z_0 + z_1 D_1(1+\max\{G;\pi_1\}) + z_2 D_2(1+\max\{G;\pi_1\})(1+\max\{G;\pi_2\}) + \cdots$$
$$+ z_T D_T \prod_{i=1}^{T}(1+\max\{G;\pi_i\})$$
$$= z_0 + \frac{z_1(1+\max\{G;\pi_1\})}{1+r_1} + \frac{z_2(1+\max\{G;\pi_1\})(1+\max\{G;\pi_2\})}{(1+r_1)(1+r_2)} + \cdots + \frac{z_T D_T \prod_{i=1}^{T}(1+\max\{G;\pi_1\})}{\prod_{i=1}^{t}(1+r_i)}$$

Then the errors are as follows:

$$\epsilon_0 = \frac{(y_0 + y_1 D_1 + \cdots + y_T D_T) + Z)}{D_0} - \frac{(x_0 + x_1 D_1 + \cdots + x_T D_T)}{D_0}$$
$$= (y_0 - x_0) + (y_1 - x_1)D_1 + \cdots + (y_T - x_T)D_T + Z$$

$$\epsilon_t = \frac{(y_0 - x_0) + (y_1 - x_1)D_1 + \cdots + (y_T - x_T)D_T + Z}{D_t} = \frac{\epsilon_0}{D_t} = \epsilon_0 * \prod_{i=1}^{t}(1+r_i) \blacksquare$$

This shows that the error made at time 0 grows exponentially with the interest rate. Hence, if a projection is made with an initial error this error will be included in the projection, compounded accordingly. It is possible to minimize this error by using the CLS approach. The moment a projection is made, there is some current term-structure of interest rates (the base scenario). It has already been explained that it is a reasonable assumption or choice to fix this scenario as "true" at the start of our projection because of the no-arbitrage assumption. This forces the initial error to 0 and we prevent interest being compounded over this error. If interest is compounded over 0 it obviously remains zero and therefore no longer affects the projection.

This does not mean that there can be no projection errors anymore, but the errors that are made are only because the product behaves somewhat different from the replicating portfolio. Since a perfect replicating portfolio could not be found this error will always be there by construction. This means that also for the projection error the same results as presented in section 4 apply. Hence, it is not possible to make a perfect projection for the future, but we can estimate a projection of the value within some bounds. The distribution of these errors is namely identical to the distribution of the errors under the test scenarios. Questions about this may arise. The reason this is the case for the time = 0 case is trivial and already explained. The question is less trivial if one looks at a time t > 0. In the way that these replicating portfolios were modelled such change would mean that there is a completely new scenario. This would be an out of sample scenario, and it was tested that the distribution of these errors is similar to the errors of the in sample scenarios. Hence, the same distribution of errors applies. Note that these tests were done on a finite set of scenarios but we conjecture that as n tends to infinity the true and empirical distribution of these errors converge.

7. Conclusion

The main interest of this thesis was to look whether replicating portfolios could be useful for pricing of insurance products. The results show quite small maximum errors for the pure cash flow product and the embedded options product, respectively 0,345% and 1,46%. We argue that errors of this size should be workable in most typical applications. The inflation linked product was more difficult to model and gave a maximum error of 3,52%, this is significantly larger and could (in some cases) be troublesome. The results of all replicating portfolios can almost surely be improved by including some additional regressors. In case of the inflation linked product it might also be an option to include a product that would be able to grasp some of the yet unexplained behavior.

The consistency of the created replicating portfolios was also tested by comparing the in- and out of-sample results to one another. For all products these errors were very similar and no cause for concern. The replicating portfolio seems to perform quite similar in and out of the set of test scenarios. This indicates that the replicating portfolio's explanatory power is not dependent on the scenario. This is very important since this shows that the replicating portfolio performs well for all scenarios and not only for the test scenarios.

We were interested in the possibility of using a replicating portfolio to make a projection of the future. It was discovered that it is of vital importance that all cash flow, incoming and outgoing, are evaluated equally and thus compounding interest over them. When this is done the replicating portfolio is most certainly valuable for projection purposes, especially in the short- and medium run. It was shown that the projection error is essentially the error at time = 0 with compounded interest. Thus if the initial error can be held as small as possible this decreases the error further along the projection. Furthermore, it was argued that the projection error has the same error (distribution) as the initial error, since it is simply a new (out of sample) scenario. Thus the error of the forecast can be maximized by the error at time zero with compounded interest over the time that the projection lasts.

8. Literature

Allingham, M. (1991). *Arbitrage: Elements of financial economics*. Macmillan.

Amemiya, T. (1985). Advanced econometrics. *Harvard University press.*

Artzner, P., & Heath, D. (1995). Approximate completeness with multiple martingale measures. *Mathematical Finance*, *5*(1), 1-11.

Back, K. (1991). Asset pricing for general processes. *Journal of Mathematical Economics*, *20*(4), 371-395.

Bahadur, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* 37. 577–580.

Bamberg, G., & Wagner, N. (2000). Equity index replication with standard and robust regression estimators. *OR-Spektrum*, *22*(4), 525-543.

Banachiewicz, T. (1937). Zur Berechnung der Determinanten, wie auch der Inversen und zur darauf basierten Auflosung der Systeme linearer Gleichungen. *Acta Astronom. Ser. C*, *3*, 41-67.

Bingham, N. H., & Kiesel, R. (2013). *Risk-neutral valuation: Pricing and hedging of financial derivatives*. Springer Science & Business Media.

Black F., & Scholes M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy,* Vol. 81, pp. 637–659.

Bouwknegt, P., & Pelsser, A. (2002). Market value of insurance contracts with profit sharing. *The Journal of Risk Finance*, *3*(3), 60-64

Candes, E., & Tao, T. (2007). The Dantzig selector: statistical estimation when p is much larger than n. *The Annals of Statistics,* 2313-2351

Carr, P., Ellis, K., & Gupta, V. (1998). Static hedging of exotic options. *The Journal of Finance*, *53*(3), 1165-1190.

Christopeit, N., & Musiela, M. (1994). On the existence and characterization of arbitrage–free measure in contingent claim valuation. *Stochastic Analysis and Applications*, *12*(1), 41-63.

Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict Management and Peace Science*, *22*(4), 341-352.

Corrigan, J., & Qin, C. (2011). *Replicating Portfolios and Risk Management*. Presentation, Sydney.

Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of financial economics*, *3*(1), 145-166.

Derman, E., Ergener, D., & Kani, I. (1995). Static options replication. *The Journal of Derivatives*, *2*(4), 78-95.

Dybvig, P. H., & Ross, S. A. (1987). Arbitrage. *The new Palgrave: A dictionary of Economics*, *1*, 100-106.

Gerber, H. U., & Shiu, E. S. (1996). Actuarial bridges to dynamic hedging and option pricing. *Insurance: Mathematics and Economics*, *18*(3), 183-218.

Greene, W.H. (2003). Econometric analysis. *5th Ed.. Englewood Cliffs, NJ: Prentice Hall.*

Hansen, B.E. (2015). *Econometrics.* University of Wisconsin

Hanushek, E.A., & Jackson J.E. (1977). *Statistical methods for social scientists*. NewYork: Academic Press.

Harrison J.M., & Kreps D.M. (1979). Martingales and arbitrage in multiperiod securities markets. *J. Econ. Theory,* Vol. 20, pp. 381–408.

Harrison J.M., & Pliska S.R. (1981), Martingales and Stochastic integrals in the theory of continuous trading. *Stochastic Processes and Applications,* Vol. 11, pp. 215–260.

Harrison, J.M., & Pliska, S.R. (1983). A stochastic calculus model of continuous trading: complete markets. *Stochastic processes and their applications*, *15*(3), 313-316.

Hayashi, F. (2000). Econometrics. *Princeton University Press.*

IAA. (2000). "Insurance Liabilities - Valuation & Capital Requirements: General Overview of a Possible Approach". Retrieved from: http://www.actuaries.org/LIBRARY/Submissions/IASC_Insurance_Issues/Fair_Value_Overview.pdf

Jameson, R. (ed.) (1995). Derivative credit risk. *Risk Publications , New York London.*

Kreps D.M. (1981). Arbitrage and equilibrium in economies with infinitely many commodities. *Journal of Mathematical Economics.,* Vol. 8, pp. 15–35.

Madan, D. (1998). Default risk, in D. J. Hand, and S. D. Jacka, eds.: *Statistics in finance (Arnold, London)*.

Matson, T. (2010). *Basics of Replicating Portfolios*. Presentation.

Merton R.C. (1973). Theory of rational option pricing. Bell J. *Econom. Manag. Sci. Vol. 4,* pp. 141–183

Mosteller, F. (1946). On some Useful “Inefficient” Statistics. *Annals of Mathematical Statistics*, *17(4), 377-408.*

Rubinstein, M., & Leland, H. E. (1981). Replicating options with positions in stock and cash. *Financial Analysts Journal*, *37*(4), 63-72.

Schachermayer, W. (2002). No arbitrage: on the work of David Kreps. *Positivity*, *6*(3), 359-368

Schrager, D. (2008). Replicating portfolios for insurance liabilities. *AENORM 2008*.

Touzi, N., & Tankov, P. (2008). No-arbitrage Theory for Derivative Pricing. *Cours de l'Ecole polytechnique*.

Yamada, Y. (2009). Optimal Hedging of Illiquid Asset Derivatives Using Additive Models. Retrieved from http://www.econ.hit-u.ac.jp/~finmodel/events/09pdf/Yamada_thesis_Optimal%20Hedging%20of%20Illiquid%20Asset%20Derivatives.pdf

Yamada, Y. (2012). Properties of Optimal Smooth Functions in Additive Models for Hedging Multivariate Derivatives. *Asia-Pacific Financial Markets*,*19*(2), 149-179.

Appendix A: Calculation of the NPV of a Pure cash-flow product

In order to value the NPV of a pure cash-flow product we need the following input data:
For each month we need:

- Premiums Begin of Month (BoM)
- Outgoing payments BoM
- Outgoing payments End of Month (EoM)
- Zero rate

First we need to calculate the forward rates to be able to construct the monthly discount factors. The only input we need is the continuously compounded zero-rate.

1) **Convert to discount factors**

$$D(t) = e^{-z_t * t}$$

where t is in years and z_t is the continuously compounded zero rate at time t.

2) **Convert to 1Y forward rates**

From these discount factors we are able to derive the (1-year) forward rates. Let $f_{yearly}(t)$ be the 1-year forward rate between in year t-1 and t.

$$f_{yearly}(1) = \frac{1}{D(1)} - 1, \qquad for\ t = 1$$

$$f_{yearly}(t) = \frac{D(t-1)}{D(t)} - 1, \qquad for\ t > 1$$

Where t is in years

When considering the forward rate somewhere in the middle of the year we need to take certain weights into account:

$$f^*_{yearly}(1) = f_{yearly}(1), \quad for\ t = 1$$

$$f^*_{yearly}(t) = \left(1 + f_{yearly}(t-1)\right)^{\frac{rep\_month}{12}} * \left(1 + f_{yearly}(t)\right)^{\frac{12-rep\_month}{12}}, \quad for\ t > 1$$

Where $f^*_{yearly}(t)$ is the adjusted forward-rate for rates during the year and the nearest full month we are reporting of. This can be adjusted for days as well by replacing the powers with $\frac{report\ day}{365}$ and $\frac{365-report\ day}{365}$ respectively.

3) **Convert 1Y forward rates to monthly rates**

Because we are dealing with monthly cash-flows we would like to know the monthly forward-rates. To be able to do this we make an assumption:

Assumption: *the monthly forward rates are equal throughout the year.*

It is then possible to convert the yearly rates to monthly rates by doing:

$$f_{monthly}(t) = \left(1 + f_{yearly}(\left\lceil \frac{t}{12} \right\rceil)\right)^{\frac{1}{12}} - 1$$

where t is in months. The adjusted forward-rates can be constructed similarly.

4) Construct monthly discount factors

$$D_{monthly}(0) = 1$$

$$D_{monthly}(1) = \frac{1}{1 + f_{monthly}(1)}, for\ t = 1$$

$$D_{monthly}(t) = \frac{f_{monthly}(t-1)}{1 + f_{monthly}(t)}, for\ t > 1$$

$$t\ in\ months$$

Now that the monthly discount factors are known it is possible to value the (monthly) cash-flows of the product. For valuing the NPV of end-of-month cash-flows we assume the following

> **Assumption**: *an end-of-month cash-flow in month i will be considered to be a begin-of-month cash-flow in month i+1.*

Let *cashflow(t)* be the net cash-flow in time period t (t in months). The net cash-flow is defined as the sum of the incoming premiums and the outgoing payments in that period. Then the NPV at time 0 can be calculated by:

$$NPV = \sum_{t=0}^{T} cashflow(t) * D_{monthly}(t)$$

The NPV can also be calculated recursively, the advantage to this method is that it provides the NPV at any point in time.

$$NPV(t) = \frac{NPV(t+1) + cashflow(t+1)}{1 + f_{monthly}(t+1)}$$

Appendix B: Instruments and weights of the instruments per replicating portfolio

The replicating portfolios make use of five types of instruments, with the abbreviations in brackets:

1. **Constant (C)** – This means that a certain cash position is held at time t = 0. The NPV of this instrument is €1 in each of the scenarios.
2. **Discount bonds / Zero-coupon bonds (DF)** – This instrument pays €1 at maturity, the NPV depends on the term-structure of interest in each scenario. The NPV can be determined by simply discounting the cash-flow with the term-structure of interest rate.
3. **Inflation linked bonds (Inf) –** The payout of this instrument is linked to the inflation. Each year the payout is increased with 0,25% plus the inflation in that year (if there is deflation the payout is increased with 0,25% only).
4. **MY staffel bond (MY-Staffel)** – The payout of this bond is linked to the MY-rate, which is an internal measure of the insurers' financial performance.

The payouts of these instruments are explained in more detail in section 3.1 and the valuation in section 6.2.

The name of an instrument consist of two parts: firstly the abbreviation of the type of instrument, and secondly the maturity. E.g. DF-20 would be a discount bond with a maturity of 20 years and MY-Staffel-5 would be a MY staffel bond with a maturity of 5 years.

Replicating portfolio for the Pure Cash Flow product			
	Maturity (Y)	In Sample Weight OLS	Weight CLS
C	0	5.253.368,82	7.080.205,31
DF1	1	8.095.299,74	5.729.192,18
DF2	2	12.054.347,54	11.968.167,05
DF3	3	3.845.959,38	4.020.561,33
DF4	4	23.192.018,23	23.535.921,92
DF7	7	26.259.940,22	26.759.537,70
DF10	10	36.734.078,81	36.519.864,27
DF15	15	43.498.803,31	43.321.626,88
DF20	20	39.970.382,98	39.986.630,70
DF25	25	38.477.523,63	38.251.136,32
DF30	30	36.421.507,87	36.614.639,41
DF35	35	35.794.754,07	35.907.845,14
DF40	40	31.948.192,61	32.019.163,78
DF45	45	30.163.918,76	30.161.351,48
DF50	50	24.719.869,08	24.525.350,70
DF55	55	18.958.687,97	19.137.415,22
DF60	60	10.903.875,60	10.918.984,02
DF65	65	10.967.420,67	10.828.919,13
DF75	75	5.791.044,19	5.805.589,44
DF85	85	657.711,11	840.766,12
DF95	95	541.015,27	278.916,09

Appendix B: Continued

Replicating portfolio for the Inflation Linked product			
	Maturity (Y)	Weight OLS	Weight CLS
C	0	-4.004.277,41	-2.860.967,27
DF1	1	6.761.925,86	26.857.205,61
DF3	3	-1.889.386,28	19.595.993,25
DF7	7	-174.849,29	-2.242.307,65
DF15	15	308.843,93	-249.024,00
DF25	25	-497.845,59	-704.131,72
DF35	35	416.140,55	1.047.664,12
DF45	45	715.648,45	356.922,17
DF55	55	-1.422.648,19	-779.833,35
DF65	65	1.266.463,04	892.019,08
DF85	85	-518.776,51	-765.541,31
Inf2	2	1.376.335,96	-38.270.775,15
Inf4	4	2.456.109,40	3.441.883,94
Inf10	10	8.034.591,43	8.414.355,31
Inf20	20	6.819.856,81	6.918.153,81
Inf30	30	5.360.954,70	5.295.531,79
Inf40	40	4.155.375,20	4.158.669,34
Inf50	50	2.378.017,97	2.362.960,12
Inf60	60	3.650.232,72	3.661.060,79
Inf75	75	2.246.889,22	2.241.592,65
Inf95	95	86.740,71	87.712,036

Replicating portfolio for the Embedded Option product			
	Maturity (Y)	Weight OLS	Weight CLS
C	0	-1.383.055,39	-179.856,46
DF1	1	-4.112.032,29	-9.522.108,61
DF3	3	-10.790.887,02	-19.172.797,50
DF7	7	-8.008.695,64	-8.633.057,28
DF15	15	200.723,86	192.206,86
DF25	25	2.456.042,81	2.549.743,53
DF35	35	4.294.300,62	4.308.269,09
DF45	45	1.617.787,41	1.652.807,50
DF55	55	-680.578,06	-637.145,84
DF65	65	6.723.584,65	6.676.757,10
DF85	85	1.653.464,96	1.653.973,49
MY-Staffel2	2	4.502.762,94	13.358.646,28
MY-Staffel4	4	6.658.659,39	10.882.817,46
MY-Staffel10	10	18.875.798,89	19.049.864,14
MY-Staffel20	20	33.267.877,01	33.197.844,44
MY-Staffel30	30	43.092.024,74	43.038.746,39
MY-Staffel40	40	42.806.531,25	42.787.610,77
MY-Staffel50	50	36.058.352,90	36.025.429,23
MY-Staffel60	60	30.196.223,93	30.163.989,73
MY-Staffel75	75	23.620.890,37	23.673.596,17
MY-Staffel95	95	1.759.218,57	1.739.290,18