NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Victor Lapshin

MODEL RISK IN BOND PORTFOLIO HEDGING

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: FINANCIAL ECONOMICS
WP BRP 88/FE/2022

SERIES: *FINANCIAL ECONOMICS*

Victor Lapshin[1]

MODEL RISK IN BOND PORTFOLIO HEDGING[2]

Abstract

Empirically testing a bond portfolio hedging model is usually carried out when proposing a new model or to compare several existing models using real data. However, there are many methodological choices to be made during such exercise, which are usually made either implicitly or without sufficient discussion. We review the empirical literature and highlight differences in testing methodology. We then carry out a massive numerical experiment to assess how each of these differences influences the testing outcome thus quantifying the model risk associated with these model features. The model risk measure values we report are easily interpretable and offer insight on the discrepancies in the results of existing empirical studies and ways to avoid such discrepancies in future work.

JEL: G12, E43.

Keywords: immunization, mark-to-market, mark-to-model, empirical test, bond portfolio.

[1]HSE University, Moscow, Russia; vlapshin@hse.ru

[2]Support from the Basic Research Program of the National Research University Higher School of Economics is gratefully acknowledged.

1 Introduction

We consider the problem of hedging the interest rate risk of a portfolio of bonds in a trading book. The source of this risk is the repricing of the bonds in the portfolio due to changes in the term structure of interest rates. This is a case of the bond portfolio immunization problem which is widely discussed in the literature.

The problem of immunizing a bond portfolio has traditionally been stated as follows—given an investment horizon H and a financial obligation, find a suitable investment strategy to meet the obligation with the least possible risk. As a rule, only the interest rate risk is considered even though the investment might be subject to credit or other risks. The usual setting consists in having to meet a single fixed financial obligation at time H. This problem has been the subject of extensive research, both theoretical and empirical, over recent decades. However, this informal problem statement can be formalized in various ways—thus numerous solutions have accumulated in the literature. Testing and comparing the performance of these solutions also requires a considerable number of modelling choices all of which could possibly impact the results. Empirical papers on comparing the performance of several immunization strategies usually made these choices implicitly—and differently—thus making the aggregation of comparison results over empirical papers impossible.

Our main contribution is methodological—we review the empirical literature and highlight the methodological steps (assumptions, goals, data, etc.) in which they implicitly differ one from another. Then we conduct a massive numerical experiment in which we try all the possible combinations of modelling choices and assess how much these modelling choices influence the resulting immunization performance estimates, thus separating the important choices from the less important.

2 Literature Review

The immunization problem dates back to Redington (1952) who introduced duration-based immunization together with a periodic rebalancing of the portfolio to maintain the duration-based hedge. The research was mostly theoretical—with a few numerical examples but no empirical study. There was no clearly defined mathematical objective, however, from the equations one might deduce that the objective was to maintain a steady actuarial valuation of the business over time, not to some predefined final reporting date.

Since then, the research has followed several separate directions. The first considered various kinds of shocks to the term structure. Duration-based hedging was shown to be equivalent to hedging against parallel additive shifts of the term structure. Bierwag (1977) considered additive and multiplicative shocks and their combinations. Cooper (1977) considered parametric durations—sensitivities to small changes in the parameters defining the term structure curve. Willner (1996) considered these parametric durations with respect to the Nelson-Siegel term structure parameters. Boyle (1978); Ingersoll et al. (1978); Cox et al. (1979); Beekman and Shiu (1988) considered stochastic processes for the instantaneous interest rate and the patterns of term structure shifts implied by various stochastic process specifications. Au and Thurston (1995) inferred durations as bond price sensitivities within a one-factor HJM model.

At roughly the same time, another branch of research set out to avoid having to assume a specific functional form of either the term structure or the stochastic process governing the short rate. Bierwag et al. (1983) introduced the so-called stochastic process risk (caused by the uncertainty in determining the stochastic process governing the short rate) and tied it to the discrepancy in cash flow times. This may be considered as a direct precursor to the M-absolute measure introduced later. Fong and Vasicek (1984) considered arbitrary instantaneous forward rate changes with bounded derivatives and showed that a lower bound for the loss in the portfolio value is proportional to the M^2 measure of cash flow discrepancy:

$$M^2 = \frac{\sum_{i=1}^{N}(t_i - H)^2 CF_i d(t_i)}{\sum_{i=1}^{N} CF_i d(t_i)},$$

where CF_i, $i = 1..N$ is the size of the total portfolio cash flow expected at time t_i, $d(t)$ is the discount factor for term t to maturity, and H is the investment horizon.

Shiu (1988) considered so-called convex shifts—such instantaneous forward rate shifts $\varepsilon(t)$ that the corresponding multiplicative discount factor change, $f(t) = e^{-\int_0^t \varepsilon(\tau)d\tau}$, is a convex function of the term t. This branch was later refined by Uberti (2000) and Hürlimann (2002) to include a larger set of possible forward rate shifts. Unfortunately, these results are not applicable to the bond portfolio immunization problem as is—Uberti (2000) stated that this was a direction for further research, which however, does not seem to have been carried out.

Bierwag et al. (1987a) considered immunizing with respect to the second-order duration measure D^2 along with the traditional duration. Chambers et al. (1988) introduce the duration vector D consisting of weighted j^{th} powers of cash flow times:

$$D^j = \frac{\sum_{i=1}^{N}(t_i)^j CF_i d(t_i)}{\sum_{i=1}^{N} CF_i d(t_i)}.$$

The actual formula used in the paper was a bit different as the authors considered the expected duration vector at the nearest rebalancing moment rather than today. Immunization was carried out as to make the first 1 to 7 elements of the portfolio duration vector match those of the liability. This is another example of a nonparametric approach to immunization via the Taylor expansion.

Reitano (1990) and Ho (1992) introduced the so called key rate durations—partial derivatives with respect to a change in the rate for one single key term to maturity (linearly interpolated between key terms if necessary). That approach has since become the de facto industry standard in hedging. Dattatreya and Fabozzi (1995) advanced this approach further by considering sensitivities to quotes of real-world financial instruments used for hedging the obligation in question—as opposed to some theoretical interest rates of Reitano (1990) and Ho (1992).

Nawalkha and Chambers (1996, 1997) proposed first the M-absolute and then the M-vector immunization approaches. They aimed to reduce the cashflow discrepancy measured by

$$M^j = \frac{\sum_{i=1}^{N}(t_i - H)^j CF_i d(t_i)}{\sum_{i=1}^{N} CF_i d(t_i)}.$$

Nawalkha et al. (2003) further elaborated the concept by considering the generalized M-vector:

$$M^j = \frac{\sum_{i=1}^{N}(g(t_i) - g(H))^j CF_i d(t_i)}{\sum_{i=1}^{N} CF_i d(t_i)},$$

where $g(t)$ is a strictly increasing weight function, e.g., $g(t) = t^\alpha$ or $g(t) = \ln t$.

Tark (1990) proposed estimating the most plausible term structure movements from empirical data using principal component analysis. D'ecclesia and Zenios (1994); Barber and Copper (1996); Hill and Vaysman (1998) put this approach into practice.

A different perspective was considered by Bierwag and Khang (1979); Prisman (1986). They employed a game-theoretical approach with the investor playing against the market and inferred portfolio immunization as the max-min strategy yielding the best possible results for the worst-case term structure change. The M^2 approach by Fong and Vasicek (1984) can also be considered a part of this branch since they showed that minimizing M^2 maximizes the worst-case portfolio yield. Significant new results in this direction were obtained by Balbás and Ibáñez (1998) who showed that the $\tilde{N}$ measure (equivalent to M-absolute) maximized the worst-case portfolio return for some specific problem statement (no short positions and requiring a maximum guaranteed return over a fixed time horizon). Barber and Copper (1998) used the clever trick of integrating by parts to minimize the sensitivity to any L_2 term structure shock and recovered the classical duration and M^2 immunization approaches as particular cases. Balbás et al. (2002b); Balbás and Ibáñez (2002); Balbás and Romera (2007) provided deeper theoretical foundations for the max-min fixed income investment approach but formulating an immunization approach based on these results did not turn out to be straightforward—and to the best of our knowledge has not been done yet.

Some of theoretical studies cited above also provided empirical tests, others only offered single-case illustrations or no empirical testing at all. There also were purely empirical papers comparing several immunization approaches using various criteria.

Unfortunately, it is not easy to arrange empirical immunization studies into one or several progressing lines of methodological development. Most empirical papers aimed at comparing various immunization methods did not discuss methodological aspects either just stating the choices made or citing previous studies as justification. The effects this or that particular methodological choice had on the empirical results were not discussed as a rule.

The review of empirical literature below is organized aspect-wise, not article-wise. We identify several important aspects of an empirical test of bond portfolio immunization and describe the various choices made by researchers over time.

Investment Horizon

Apart from obvious data limitations, we have found no discernible pattern in the way the investment horizon was chosen in immunization studies:

- 1 or several days (Wu, 2000),
- 1 week (Litterman and Scheinkman, 1991; Ortobelli et al., 2018),
- 1 month (Guletkin and Rogalski, 1984; Elton et al., 1990; Ilmanen, 1992; Carcano and Foresi, 1997),
- 2 months (Lacey and Nawalkha, 1993),
- 3 months (Guletkin and Rogalski, 1984; Bierwag et al., 1987b; Chambers et al., 1988; Carcano and Foresi, 1997),
- 6 months (Guletkin and Rogalski, 1984; Reitano, 1992),
- 1 year (Cooper, 1977; Guletkin and Rogalski, 1984; Nawalkha et al., 2003; Soto, 2004; Agca, 2005; Carvalhal and Daumas, 2010; Bravo and Fonseca, 2012; Oliveira et al., 2014; Zhu et al., 2018),
- 2 years (Lee and Cho, 1992; Bravo and da Silva, 2006),
- 2.5 years (Bierwag et al., 1987a),
- 3 years (Balbás et al., 2002a; Soto, 2004; Diaz et al., 2008; Díaz et al., 2009; Bravo and Fonseca, 2012; Oliveira et al., 2014),
- 4 years (Nawalkha and Chambers, 1996, 1997; Bravo and da Silva, 2006),
- 5 years (Fisher and Weil, 1971; Bierwag et al., 1987a; Fooladi and Roberts, 1992; Bierwag et al., 1993; Nawalkha et al., 2003; Mato, 2005; Agca, 2005; Oliveira et al., 2014),
- 10 years (Fisher and Weil, 1971; Bierwag et al., 1981; Agca, 2005),
- 20 years (Fisher and Weil, 1971).

Rebalancing Period

Most immunization strategies require rebalancing the portfolio once in a while. There are two popular choices for the rebalancing strategy in the literature—periodic rebalancing, e.g., once a year, or event-based rebalancing when the current market situation calls for it. Some papers assumed no rebalancing while most papers used periodic rebalancing with various frequencies. Studies with investment period of 1 year and shorter usually assumed no interim rebalancing while longer investment periods were usually accompanied by rebalancing one or several times a year. Additionally, some papers rebalanced the portfolio every time a coupon was paid on any bond in the portfolio.[3] These are the choices from the literature:

- no rebalancing (Cooper, 1977; Guletkin and Rogalski, 1984; Bierwag et al., 1987b; Chambers et al., 1988; Elton et al., 1990; Litterman and Scheinkman, 1991; Ilmanen, 1992; Reitano, 1992; Lee and Cho, 1992; Lacey and Nawalkha, 1993; Carcano and Foresi, 1997; Wu, 2000; Mato, 2005; Zhu et al., 2018),

- every time a coupon is paid (Balbás et al., 2002a; Soto, 2004; Diaz et al., 2008; Díaz et al., 2009; Oliveira et al., 2014),

- weekly (Ortobelli et al., 2018),

- monthly (Agca, 2005; Oliveira et al., 2014),

- quarterly (Lee and Cho, 1992; Soto, 2004; Theobald and Yallup, 2010),

- 2 times a year Bierwag et al. (1987a); Lee and Cho (1992); Bierwag et al. (1993); Bravo and da Silva (2006),

- once a year Fisher and Weil (1971); Bierwag et al. (1981); Lee and Cho (1992); Nawalkha and Chambers (1996, 1997); Nawalkha et al. (2003); Bravo and da Silva (2006); Bravo and Fonseca (2012).

Testing Intervals

When it comes to choosing the time intervals for testing, the two possible choices are overlapping and non-overlapping intervals. Overlapping intervals allow for more data points but at the cost of the observations becoming dependent thus rendering statistical comparisons more complicated. With longer investment horizons one usually does not have a choice other than to use overlapping intervals.

- Overlapping intervals were used by: Fisher and Weil (1971); Bierwag et al. (1981, 1987a); Fooladi and Roberts (1992); Bierwag et al. (1993); Nawalkha and Chambers (1996, 1997); Balbás et al. (2002a); Nawalkha et al. (2003); Soto (2004); Mato (2005); Bravo and da Silva (2006); Diaz et al. (2008); Díaz et al. (2009); Bravo and Fonseca (2012); Oliveira et al. (2014).

- Non-overlapping intervals were used by: Cooper (1977); Bierwag et al. (1987b); Chambers et al. (1988); Elton et al. (1990); Litterman and Scheinkman (1991); Ilmanen (1992); Lacey and Nawalkha (1993); Carcano and Foresi (1997); Wu (2000); Ortobelli et al. (2018); Zhu et al. (2018).

[3] The portfolios in these papers usually consisted of a very small number of bonds—two or three, thus coupon payments did not occur very often.

Nature of the Data

When it comes to the data, a very important distinction should be made. While some papers use real market bond prices, other empirical works use term structure data as an input with bond prices determined via discounting the promised cash flows. The bonds in question may be real or hypothetical—with any desired characteristics. This is easier as term structure data is more readily available. It also removes the necessity to deal with various issues like low liquidity, missing quotes, etc. This term structure can be either real or simulated via a dynamic term structure model fitted to real data.

- Real term structure data was used by: Fisher and Weil (1971); Cooper (1977); Bierwag et al. (1981, 1987a,b); Elton et al. (1990); Fooladi and Roberts (1992); Nawalkha and Chambers (1996, 1997); Carcano and Foresi (1997); Balbás et al. (2002a); Nawalkha et al. (2003); Soto (2004); Bravo and Fonseca (2012).

- Simulated term structure data was used by: Lee and Cho (1992); Agca (2005).

- Real bond prices / quotes were used by: Guletkin and Rogalski (1984); Chambers et al. (1988); Litterman and Scheinkman (1991); Ilmanen (1992); Lacey and Nawalkha (1993); Wu (2000); Mato (2005); Bravo and da Silva (2006); Diaz et al. (2008); Díaz et al. (2009); Carvalhal and Daumas (2010); Theobald and Yallup (2010); Oliveira et al. (2014); Ortobelli et al. (2018); Zhu et al. (2018).

Newer works seem to prefer working directly with bond price data and not with term structures. However, term structure pricing or using real bond data could be preferred in appropriate circumstances as discussed by Lapshin (2021).

Transactional Costs

Most papers assume no transactional costs in restructuring portfolios. When working with bid/ask quote data, transactional costs can be easily incorporated by modeling sales at bid prices and purchases at ask prices (Oliveira et al., 2014). For other types of data, transactional costs have to be specifically imposed via some sort of assumption e.g., fixed proportional costs (Fisher and Weil, 1971; Lee and Cho, 1992; Ortobelli et al., 2018) or a fixed term structure of bid-ask spreads (Agca, 2005). All the other papers assumed no transactional costs.

Strategies Tested

Most empirical papers tested one or several 'traditional' immunization strategies and possibly one or two specific strategies introduced by the authors. Widely tested traditional strategies are:

- naïve strategy: just invest in some long bonds and possibly roll over the portfolio;

- maturity strategy: invest in the shortest bond maturing after the investment horizon;

- duration strategy: invest in two (often) or more (rarely) bonds so that the combined duration of the immunizing portfolio be equal to the duration of the obligation being hedged. Most popular subtypes are bullet—the two bonds are chosen to have maturities closest to the investment horizon—and barbell—the two bonds are chosen to be the furthest apart in terms of maturity.

Other strategies which were tested more than occasionally and not entirely by their respective authors:

- higher order duration hedging (convexity and sometimes higher order): Chambers et al. (1988); Lacey and Nawalkha (1993); Carcano and Foresi (1997); Soto (2004); Agca (2005); Ortobelli et al. (2018),

- immunization based on the M-absolute measure (or $\tilde{N}$): Nawalkha and Chambers (1996); Soto (2004); Mato (2005); Diaz et al. (2008); Carvalhal and Daumas (2010),

- immunization based on the M-squared measure: Bierwag et al. (1993); Mato (2005); Carvalhal and Daumas (2010),

- other M-family measures: Nawalkha and Chambers (1997); Balbás et al. (2002a); Nawalkha et al. (2003); Bravo and da Silva (2006),

- partial (parametric) duration hedging: Cooper (1977); Ilmanen (1992); Soto (2004); Bravo and Fonseca (2012); Zhu et al. (2018),

- external factor hedging (e.g., PCA-based): Litterman and Scheinkman (1991); Ilmanen (1992); Soto (2004).

The majority of other hedging strategies were only tested by their respective authors and/or with significant methodological drawbacks (e.g., with sample sizes of several data points)—this particularly applies to empirical appendices to mainly theoretical papers.

Comparison criteria

Surprisingly enough, there is no consensus on the objective function for the immunization problem. Moreover, even on the conceptual level some researchers focus on various measures of expected return from the strategy (average realized return or the fraction of scenarios in which the strategy produced not enough funds to meet the immunized obligation) while others compared various risk measures (RMSE of realized returns, Value-at-Risk, etc.). Many papers employed unique comparison criteria based on the objective on the paper. Almost all papers calculated all measures relative to the promised return, i.e., the expected return calculated via the term structure of interest rates observed at the time of forming the immunizing portfolio. The most common comparison criteria are:

- mean absolute deviation of the realized return from the promised return (Fisher and Weil, 1971; Bierwag et al., 1987a; Chambers et al., 1988; Fooladi and Roberts, 1992; Nawalkha and Chambers, 1996, 1997; Nawalkha et al., 2003; Diaz et al., 2008; Bravo and Fonseca, 2012; Oliveira et al., 2014),
- standard deviation of the realized return from the promised return (Fisher and Weil, 1971; Guletkin and Rogalski, 1984; Bierwag et al., 1987a; Chambers et al., 1988; Lee and Cho, 1992; Lacey and Nawalkha, 1993; Bierwag et al., 1993; Carcano and Foresi, 1997; Wu, 2000; Soto, 2004; Bravo and da Silva, 2006; Diaz et al., 2008; Bravo and Fonseca, 2012; Oliveira et al., 2014; Ortobelli et al., 2018),
- mean absolute deviation or root mean squared deviation computed only from negative deviations—measures of average immunization shortfall (Fooladi and Roberts, 1992; Nawalkha and Chambers, 1996, 1997; Bravo and Fonseca, 2012),
- unexpected earnings (realized return minus the promised return) (Fisher and Weil, 1971; Guletkin and Rogalski, 1984; Reitano, 1992; Lacey and Nawalkha, 1993; Ortobelli et al., 2018; Zhu et al., 2018),
- R^2 from regressing the returns on factors (Cooper, 1977; Litterman and Scheinkman, 1991; Ilmanen, 1992),
- percentage of times when the strategy in question produced the highest return compared to either competitors or the promised return (Bierwag et al., 1981, 1993; Ortobelli et al., 2018),
- percentage of times when the strategy in question produced the return closest to that promised (Fooladi and Roberts, 1992; Balbás et al., 2002b; Nawalkha et al., 2003; Bravo and da Silva, 2006; Diaz et al., 2008),
- percentage of times when the strategy produced the return within a fixed tolerance around the promised return (Bierwag et al., 1981; Agca, 2005),
- Value-at-Risk and / or Conditional VaR, i.e., some lower quantile of the strategy result or the average of the worst $x\%$ of the outcomes (Mato, 2005; Carvalhal and Daumas, 2010; Ortobelli et al., 2018).

Other optimality measures were used occasionally.

3 Methodology

In this section we outline our methodological framework for testing and comparing immunizing approaches and briefly discuss the rationale behind some choices. As noted in the introduction, our main interest lies with the problem of hedging the interest rate risk of a bond portfolio in the trading book.

We consider a company having an obligation on its balance sheet. The company is not worried about having to meet this obligation in future—that is, we assume its business strategy is sound. The company is however worried about the excess volatility that this obligation introduces into its financial or internal reporting—and therefore chooses to hedge (immunize) it. Thus, we adopt the view of a risk management unit, not a business unit. Note that this is different from the classical bond portfolio immunization problem as we are mainly interested in reducing the interest rate risk of the portfolio and not in maximizing the financial outcome. This setup motivates most of the choices listed below.

Immunization horizon. Various kinds of reporting occur with different frequencies from 1 day to 1 year. To study the effects of this choice, we test the horizons of 1 day, 1 and 2 weeks, 1, 3, 6, and 12 months. Longer horizons do not make sense within our setting as most financial and management reporting is done at least annually. Since we use relatively short horizons, non-overlapping intervals is a valid choice which we adopt for the sake of improved statistical properties.

Portfolio composition. We use a variant of leave-out-one cross-validation. We choose each bond traded on a given business day as the original obligation—except for the longest and the shortest bond. Having the obligation outside the maturity range of the hedging instruments would introduce yet another degree of complexity—term structure extrapolation—which we would like to avoid at this stage. For each immunization method we form an immunizing portfolio and assess its performance at the end of the immunization period. Since the original obligation is modeled by a traded bond, we are able to use its market price if needed. If this market price is not available for the end of the immunization horizon we drop this observation.

Risk measures. We report the three most popular risk measures in assessing immunization performance, namely mean absolute deviation, root mean squared deviation and 95% Value-at-Risk of the financial result R_t. The financial result R_t is defined as the difference between the actual future portfolio value V_{t+H} and the expected portfolio value $\widetilde{V_{t+H}}$, where t is the calendar time and H is the immunization horizon. The expected portfolio value is obtained from the initial portfolio value V_t via the initial risk-free rate $r_t(H)$ for the term corresponding to the immunization horizon:

$$R_t = V_{t+H} - \widetilde{V_{t+H}} = V_{t+H} - e^{r_t(H)H} V_t.$$

Note that we report the returns per dollar of the original obligation, not per dollar of the immunizing portfolio as the immunizing portfolio is not required to be self-financing.

Portfolio rebalancing. We do not consider interim rebalancing. Instead we introduce short horizons into the consideration as this offers more flexibility in setting up the computational experiment. An immunization problem with a long horizon and several rebalancing moments is actually equivalent to several successive one-step immunization problems with no interim rebalancing—with their outcomes aggregated in a special way. Since all the risk measures already imply a kind of aggregation across calendar time, we believe that a two-stage aggregation (first within a single multi-step immunization problem and then across several immunization problems) is redundant. Therefore, we limit ourselves to testing only one-step immunizations with no interim rebalancing over a broad spectrum of horizons.

We do not rebalance the portfolio when the bonds in the portfolio pay coupons because some of our immunization portfolios consist of all the bonds available in the market—this would defeat the idea of testing long rebalancing periods as coupon payments are bound to happen very often in this case. Instead, we assume that all proceedings are kept in cash until the end of the period subject to a risk-free interest rate. We tested various reinvesting setups, however the details turned out to be insignificant in terms of the contribution to the resulting volatility.

Portfolio valuation. International and most national financial reporting standards allow for two distinct modes of reporting the value of a financial instrument—mark-to-market and mark-to-model. The exact details are not important at this stage. However, the distinction itself is very important. We consider the following three combinations:

- *both the obligation and the immunizing portfolio are marked to a model*, i.e., priced as the sum of the discounted cash flows via an estimate of the term structure of interest rates:
$$V_t = -\sum_{i=1}^{N_0} CF_{0,i} e^{-r_t(t_{0,i})t_{0,i}} + \sum_{k=1}^{K} w_k \left(\sum_{i=1}^{N_k} CF_{k,i} e^{-r_t(t_{k,i})t_{k,i}} \right),$$
where $CF_{0,i}$ are the cash flows of the original obligation expected at terms $t_{0,i}$, $CF_{k,i}$ are the cash flows of the hedging instruments expected at terms $t_{k,i}$, weights w_k define the immunization portfolio composition. $r_t(\cdot)$ is the term structure estimate at the calendar time t to be discussed later;

- *the obligation is marked to a model, but the immunizing portfolio is marked to the market*—this routinely happens when hedging an illiquid or bespoke obligation with liquid market instruments:
$$V_t = -\sum_{i=1}^{N_0} CF_{0,i} e^{-r_t(t_{0,i})t_{0,i}} + \sum_{k=1}^{K} w_k P_{t,k},$$
where $P_{t,k}$ is the observed market price of the hedging instrument k for the calendar time t;

- *both the obligation and the immunizing portfolio are marked to the market*—this scenario can also serve as an approximation to unanticipated transactional costs. The additional volatility introduced by the market prices can be viewed as a substitute for the uncertainty caused by unanticipated costs.

$$V_t = -P_{t,0} + \sum_{k=1}^{K} w_k P_{t,k},$$

where $P_{t,0}$ is the market price of the original obligation, which is observable by the choice of the latter.

Data issues. Market prices of some bonds were not observed at some points. For model-based pricing this is less of a problem; however we cannot use market prices which are unavailable. We consider two approaches to deal with this issue:

- *look-ahead*—at the time of forming the immunizing portfolio we restrict our attention only to the bonds for which the prices are available at the end of the immunization horizon;
- *fill in the missing data*—if at the end of the immunization period the prices of some bonds in the portfolio are not observed, we use their model prices instead.

Another closely related question is whether to include newly issued bonds—if at the end of the immunization period we observe a new bond which was not available in the beginning of the period, should we use this information to infer the term structure at the terminal moment?

- *Keep new bonds*—use all available information at the end of the immunization period even if the bonds in question were not available in the beginning. Including new bonds might significantly change the term structure estimate especially if the new bond is shorter or longer than all other bonds in the dataset.
- *Drop new bonds*—only use the bonds which were there at the time of forming the immunization portfolio. This reduces the term structure volatility by reducing one of its sources—the variability of the instrument set.

We expect these choices to influence the performance of all immunization strategies with mark-to-model valuation as the sudden appearance or disappearance of bonds is an additional source of term structure volatility which might or might not be considered by an immunization model.

Immunization strategies. We test several popular immunization models as well as several less popular. These strategies are individually described below.

Term structure estimation. For mark-to-model pricing we consider the following popular methods of estimating the term structure of interest rates from bond prices:

- *a Nelson-Siegel model fitted to bond prices via nonlinear least squares*—given prices

P_k for bonds promising $CF_{k,i}$ at times t_i, fit the parameters $\theta = (\beta_0, \beta_1, \beta_2, \tau)$ as

$$\sum_{k=1}^{K} \left(P_k - \sum_{i=1}^{N} CF_{k,i} e^{-r_\theta(t_i)t_i} \right)^2 \to \min_\theta, \tag{1}$$

for

$$r_\theta(t) = \beta_0 + \beta_1 \frac{1 - e^{-\frac{t}{\tau}}}{t/\tau} + \beta_2 \left(\frac{1 - e^{-\frac{t}{\tau}}}{t/\tau} - e^{-\frac{t}{\tau}} \right); \tag{2}$$

- *a smoothing spline with a penalty on the first derivative (piecewise linear)*:

$$\sum_{k=1}^{K} \left(P_k - \sum_{i=1}^{N_k} CF_{k,i} e^{-r(t_{k,i})t_{k,i}} \right)^2 + \gamma \int_0^T [r'(x)]^2 \, dx \to \min_{r(\cdot)}; \tag{3}$$

- *a smoothing spline with a penalty on the second derivative (piecewise cubic)*:

$$\sum_{k=1}^{K} \left(P_k - \sum_{i=1}^{N_k} CF_{k,i} e^{-r(t_{k,i})t_{k,i}} \right)^2 + \gamma \int_0^T [r''(x)]^2 \, dx \to \min_{r(\cdot)}. \tag{4}$$

Three separate term structures are estimated for each day and for each bond playing the part of the original obligation—because the obligation is excluded from the estimation dataset. Here and in what follows we suppress the dependence on the calendar time t to avoid cluttering. These models are used for pricing and for immunization strategies where appropriate. Even if an immunization method assumes a specific term structure model, we test it against all term structure estimation models to assess the model risk arising from term structure model misspecification.

Portfolio composition. Almost all immunization strategies impose only a few linear restrictions on the portfolio composition leaving a considerable number of degrees of freedom (e.g., duration-based immunization imposes only one linear constraint). The usual approach is to minimize the sum of squares of individual instrument weights in the portfolio: $\sum_{k=1}^{K} w_k^2 \to \min$. This is supposed to minimize the overall portfolio variance. Motivated by preliminary results, we also test another approach based on the work of Barber and Copper (1998). Their immunization approach is formulated as a quadratic programming problem and is thus directly applicable as a substitute to minimizing the sum of squares of the individual weights.

We test these two functionals in two formulations—minimizing the quadratic functional of choice, subject to all other constraints imposed by the chosen immunization strategy or adding it as a regularizer. The first (conditional) approach uses the regularization functional only to select one of the infinitely many portfolios satisfying the immunization conditions. The immunization conditions are always satisfied exactly; and if only one portfolio satisfies these conditions, then no regularization is performed. On the contrary, the second approach regularizes the immunization problem itself—the immunization conditions are now satisfied only approximately but with smaller immunizing portfolio weights. This might prevent some immunization methods from financing immunization via incredibly large short positions in the presence of a budget constraint. The equations are as follows:

- for the least-squares conditional:

$$\begin{cases} \frac{1}{2}w^T w \to \min; \\ Aw = b; \end{cases}$$

- for the least-squares regularized:

$$\frac{1}{2}w^T(\alpha I + A^T A)w - (A^T b)^T w \to \min;$$

- for Barber and Copper (1998) conditional:

$$\begin{cases} \frac{1}{2}w^T L^T T L w - (L^T T L_0)^T w \to \min; \\ Aw = b; \end{cases}$$

- for Barber and Copper (1998) regularized:

$$\frac{1}{2}w^T(\alpha L^T T L + A^T A)w - (A^T b + \alpha L^T T L_0)^T w \to \min.$$

Here w is the vector of portfolio weights, $Aw = b$ is the system of linear equality constraints imposed by the selected immunization strategy, I is the identity matrix, $T_{i,j} = \min(t_i, t_j)$, $L_{k,i} = CF_{k,i}e^{-r(t_{k,i})t_{k,i}}$ for $k = 0..K$, and α is the regularization parameter. We test two values of the regularization parameter: 10^{-7} as an example of a subtle regularization which should not influence the portfolio composition and 10^{-3} as an example of more aggressive regularization capable of reducing the portfolio weights closer to 0. Note that this is not the final problem formulation as portfolio constraints can introduce further modifications.

Portfolio constraints. Some empirical papers assumed no short positions; other allowed them. We found no clear tendency in this matter, so we test both approaches. Additionally, we consider a budget constraint—the immunizing portfolio should cost no more than the present value of the obligation. This makes sense if the proceeds from the sale of the obligation are to be invested in the immunization portfolio. The additional constraints are formulated as follows:

- no additional constraints,
- $w \geq 0$—no short positions,
- $P^T w \leq P_0$—budget constraint; here P is the vector of current bond prices and P_0 is the current obligation price,
- $w \geq 0$ and $P^T w \leq P_0$—both budget constraints and no short positions.

For some immunization strategies linear constraints are infeasible with some types of portfolio constraints. This only happens in the conditional problem formulation—in the regularized problem we just do our best to satisfy the hedging equations given the constraints.

Transactional costs can be classified as anticipated or unanticipated. Anticipated transactional costs are determined at the time of choosing the immunizing portfolio and are a function of its composition—current bid-ask spreads are a good example of anticipated costs. One can easily include this kind of cost in the optimization functional when forming the immunizing portfolio. Note that both current and future bid-ask spreads can be considered anticipated costs—current spreads are deterministic and will be included in the 'expected return' part of the optimization functional while future spreads are random and can be included in the 'volatility' part of the functional—assuming we have a decent estimate of future spread volatility. Unanticipated costs are by definition those which are not included in the functional.

In our one-step setting, bid-ask spreads and fixed transactional costs are anticipated and can be regarded as a kind of regularization. Since we test quite a few types of regularization variants, we refrain from treating this kind of transactional cost separately. Also note that since we aim to measure the residual risk as opposed to the investment profit, anticipated costs should not be considered. Unfortunately, there is no easy and universal way to model unanticipated costs. For the mark-to-market valuation approaches, the (random) market price can be considered as a kind of a proxy for unanticipated costs.

Numerical issues. Some immunization strategies could yield prohibitively poorly conditioned optimization problems. For the formulation of the conditional problem we use the 10^{-7} singular value threshold for removing the approximately dependent constraints. On the contrary, in the regularized case we just scale $Aw = b$ to make its largest singular value 1 and then let the regularization do the job of forcing singular values that are too small out of play.

3.1 Immunization strategies

For ease of notation we assume that the original obligation and all hedging instruments have a common set of cash flow times t_i. If this is not the case, zero cash flows $CF_{k,i}$ can be introduced when necessary.

3.1.1 No hedging

This is the reference strategy with all immunization weights equal to 0. Note that it produces different outcomes depending on whether we value the obligation at its market price or at its model price. The relative performance indicators we report are normalized by the performance of this strategy.

3.1.2 Duration-vector hedging

Implemented after Chambers et al. (1988). For a bond k with the cash flow vector $CF_{k,i}$ and cash flow times t_i, its j-th duration $D_{j,k}$ is given by

$$D_{j,k} = \frac{\sum_{i=1}^{N}(t_i)^j CF_{k,i} e^{-r(t_i)t_i}}{\sum_{i=1}^{N} CF_{k,i} e^{-r(t_i)t_i}},$$

where $r(t)$ is the current term structure estimate. Since the duration vector is linear in portfolio weights w, the optimization constraints $Aw = b$ can be written as

$$Dw = D_0,$$

where D_0 is the duration vector of the obligation calculated similarly. The length of the duration vector is arbitrary. We test duration-vector strategies for dimensions 1 (ordinary duration hedging) to 7. However, for dimensions larger than 5, the system becomes ill-conditioned. One would probably be better off reformulating these immunization problems in terms of Chebyshev polynomials to avoid numerical instability.

3.1.3 M-vector hedging including M-absolute

Implemented after Nawalkha and Chambers (1996, 1997). This is identical to the duration-vector hedging except that instead of the duration matrix D we use the M-matrix M defined by

$$M_{j,k} = \frac{\sum_{i=1}^{N}(t_i - H)^j CF_{k,i} e^{-r(t_i)t_i}}{\sum_{i=1}^{N} CF_{k,i} e^{-r(t_i)t_i}}.$$

The optimization constraint is $Mw = M_0$, where M_0 is the M-vector of the obligation. We test M-vector hedging strategies for dimensions 2 to 7, Instead of 1-dimensional M-vector, we use the M-absolute measure:

$$M_{\text{abs},k} = \frac{\sum_{i=1}^{N}|t_i - H| CF_{k,i} e^{-r(t_i)t_i}}{\sum_{i=1}^{N} CF_{k,i} e^{-r(t_i)t_i}}.$$

3.1.4 Generalized M-vector hedging

Implemented after Nawalkha et al. (2003). Here the generalized M-matrix is formed as

$$M_{j,k}^{gen} = \frac{\sum_{i=1}^{N}(g(t_i) - g(H))^j CF_{k,i} e^{-r(t_i)t_i}}{\sum_{i=1}^{N} CF_{k,i} e^{-r(t_i)t_i}}$$

for a monotonically increasing function $g(x)$. We choose $g(x) = \ln x$ motivated by the results of Nawalkha et al. (2003). We test generalized M-vector hedging for dimensions 2 to 7. As before, we use the absolute value instead of 1-dimensional hedging.

3.1.5 Nelson-Siegel parametric hedging

Implemented after Willner (1996). We assume that the bonds are valued at their model prices PV_k which are calculated as

$$PV_k = \sum_{i=1}^{N} CF_{i,k} e^{-r_\theta(t_i)t_i},$$

where $r_\theta(t)$ is given by Eq. (2) and $\theta = (\beta_0, \beta_1, \beta_2, \tau)$ is the vector of parameters changing from day to day. For each bond we calculate the vector of its parametric durations $\frac{\partial PV_k}{\partial \theta} = \left(\frac{\partial PV_k}{\partial \beta_0}, \frac{\partial PV_k}{\partial \beta_1}, \frac{\partial PV_k}{\partial \beta_2}, \frac{\partial PV_k}{\partial \tau}\right)$ and since the parametric durations are linear in portfolio weights, the optimization constraint is easily written as

$$\left(\frac{\partial PV}{\partial \theta}\right)^T w = \left(\frac{\partial PV_0}{\partial \theta}\right)^T,$$

where PV_0 is the sum of discounted cash flows for the obligation.

3.1.6 Spline parametric hedging

Implemented after Lapshin (2019). We assume that the value PV_k of bond k is calculated as

$$PV_k = \sum_{i=1}^{N} CF_{i,k} e^{-r(t_i)t_i},$$

where the term structure estimate $r(\cdot)$ is found from the current observed bond prices P_k by solving Eq. (3) or Eq. (4). Here we treat P_k as parameters and basically calculate the parametric duration matrix $\left(\frac{\partial PV}{\partial P}\right)^T$. Since the parametric duration is linear in portfolio weights, the optimization constraints can be written as

$$\left(\frac{\partial PV}{\partial P}\right)^T w = \left(\frac{\partial PV_0}{\partial P}\right)^T.$$

The solution of this system is

$$w = \left(B^T \Omega^{-1} B\right)^{-1} B^T \Omega^{-1} B_0,$$

where $\Omega = BB^T + \sum_{k=1}^{K} (PV_k - P_k) \frac{\partial^2 PV_k}{\partial r^2} + \alpha J^T J$ with $J^T J$ defined by

$$r^T J^T J r = \inf_{f(\cdot)\,|\,f(t_i)=r_i} \int_0^T f''(x)^2 \, dx$$

or a similar expression with the first derivative.

Note that even though we observe the market bond price P_k, here we assume that for the purpose of portfolio valuation we use our term structure model to discount the cash flows (mark-to-model).

3.1.7 Mixed mark-to-market and mark-to-model immunization

Implemented after Lapshin (2021). Assuming that the original obligation is marked-to-model while the hedging instruments are marked-to-market immediately yields the following formulas for the immunization weights:

$$w_k = \frac{\partial PV_0}{\partial P_k} = \left(\frac{\partial PV_0}{\partial r}\right)\left(\frac{\partial r}{\partial P_k}\right),$$

where we have split the obligation model price sensitivity $\frac{\partial PV_0}{\partial P_k}$ into the sensitivity of the model price to changes in the term structure $\frac{\partial PV_0}{\partial r}$ and the sensitivity of the term structure estimate to changes in the observed bond prices $\frac{\partial r}{\partial P_k}$. The first is determined by the choice of the obligation while the second by the chosen term structure estimation model.

For the Nelson-Siegel parametric term structure Eq. (2) we get

$$w = \left(\frac{\partial PV_0}{\partial P}\right)^T = B^T Q^T (QBB^T Q^T + A_\theta)^{-1} QB_0,$$

where $B = \left(\frac{\partial PV}{\partial r}\right)^T$ is the matrix of bond price sensitivities to the term structure changes, B_0 is the same for the original obligation, $Q = \left(\frac{\partial r}{\partial \theta}\right)$ is the matrix of term structure sensitivities with respect to its parameter vector θ, $A_\theta = \sum_{k=1}^{K} (PV_k - P_k) \frac{\partial^2 PV_k}{\partial \theta^2}$.

For the nonparametric term structure Eq. (3) or Eq. (4) we get

$$w = B^T (BB^T + A + \alpha J^T J)^{-1} B_0$$

using the notation above.

3.1.8 Classical key rate duration hedging

Implemented after Ho (1992). We choose the usual set of key rates $t_s^* = \{1, 3, 5, 7, 10, 15\}$ years and consider the sensitivity of the bond price PV_k to piecewise linear changes in the term structure $c_i(t)$ described by

$$c_i(t) = \begin{cases} 1, t = t_i^*; \\ 0, t = t_{j\neq i}^*; \\ c_i(t_a^*)\frac{t-t_a^*}{t_b^*-t_a^*} + c_i(t_b^*)\frac{t_b^*-t}{t_b^*-t_a^*}, t_a^* < t < t_b^*; \\ c_i(t_1^*), t < t_1^* = 1; \\ c_i(t_6^*), t > t_6^* = 15. \end{cases}$$

The i-th key rate duration $KRD_{i,k}$ of bond k is given by the directional derivative

$$KRD_{i,k} = \left. \frac{\partial PV_k(r + x \cdot c_i)}{\partial x} \right|_{x=0},$$

where with a slight abuse of notation $PV_k(r)$ is the price of bond k with the function $r(\cdot)$ playing the part of the term structure. The optimization constraints are then written as

$$KRD \cdot w = KRD_0,$$

where KRD_0 is the key rate duration vector of the obligation.

Note that due to the peculiarities of the dataset, for some days the key rate duration matrix KRD becomes singular (e.g., if there are no bonds over 10 years to maturity). In these cases we drop the zero or dependent rows of the key rate duration matrix and continue with the reduced set of key rates.

3.1.9 Adapted key rate duration hedging

As discussed by Lapshin (2019), classical key rate duration hedging can be viewed as parametric hedging when the term structure linearly interpolates the observed zero-coupon yields for fixed set of key terms to maturity. Therefore, it is natural to suggest that for a practical situation we choose the key terms to coincide with the observed bonds' terms to maturity rather than the standard set of 1, 3, 5, 7, 10, 15 years. We choose the key terms to maturity to be the maturities of all bonds in the dataset. Presumably this approach will provide better hedging than the standard key rate duration hedging due to being tailored to the data specifics.

3.1.10 Worst-case L2 immunization

Barber and Copper (1998) derive an upper bound for the sensitivity of a bond portfolio defined by its cash flows to any change in the term structure with $\int (\Delta r(x))^2\, dx = 1$. Minimizing this upper bound in a kind of worst-case immunization results in solving the following quadratic program:

$$(Lw - L_0)^T T (Lw - L_0) \to \min_w,$$

where L is the matrix of discounted cash flows $L_{i,k} = CF_{k,i} e^{-r(t_i)t_i}$, L_0 is the same for the obligation, and $T_{i,j} = \min(t_i, t_j)$. The immunization conditions can be restated in the form of equality constraints, however there is no practical need in this as the constraints identify all portfolio weights thus making the conditional regularization redundant. Instead, since this approach is formulated in terms of minimizing a quadratic functional, we use it instead of the least squares to regularize other methods.

3.2 Data description

We use a dataset of Spanish government bonds prices from 1996 to 2019 obtained from Bloomberg. We use only bonds with known coupon payments and no embedded options. Daily observations ranged from 11 to 46 bonds. On 1 Aug 2014, the ex-dividend date for the bonds changed from 3 business days before the coupon to 2 business days before the coupon.

We removed the bonds having less than 1 month to maturity from consideration. That is, if a bond has less than 1 month to maturity at the beginning of the immunization period, we do not consider it at all. If it has less than 1 month to maturity at the end of the period, we consider it unobserved at that time—and either drop it or mark it to the model.

We also remove bonds with coinciding cash flow times. Some bonds in the dataset have coinciding coupon payment and maturity dates. Even with slightly different coupon rates such bonds disrupt many of the hedging methods discussed, so in such cases we use only one.[4]

4 Empirical results

There are 10 dimensions to our numerical experiment—we performed 10,000 simulations for every combination of: risk measure, portfolio constraints, term structure fitting model, immunization horizon, immunization method, rule for including newly issued bonds, rule for dealing with missing bonds, portfolio valuation approach, regularization functional, and regularization parameter. Of these, we call 'features' everything except the immunization method.

We assume that an analyst has one of the two motives below.

1. Comparing various immunization methods by their performance in reducing the risk, possibly to choose one for implementation or to assess whether a change of immunization algorithm is in order. The quantity of interest in this case is the relative ranking of various strategies in terms of their performance.

2. Assessing whether immunization can deliver a necessary risk reduction compared to no hedging. The quantity of interest in this case is the performance of a chosen immunization method.

To estimate the magnitude of the influence of modelling choices on these quantities of interest, we calculate several measures of model risk implied by different model features as follows.

General model risk of a feature i is calculated as

$$\mathrm{Risk}_i = \mathrm{Std}_{f,o,m}\big[\mathrm{Res}_{f,o,m} - \mathrm{Mean}_f\left(\mathrm{Res}_{f,o,m}\right)\big],$$

where f is the value of feature i, o is a 8-dimensional value vector for all other features except i, and m denotes the immunization method, so $\mathrm{Res}_{f,o,m}$ is the quantity of interest for method m with feature i set to f and all other features set to the corresponding entries in o. With a slight abuse of notation, the subscript in $\mathrm{Std}_{f,o,m}$ and Mean_f denotes the variable(s) across which the standard deviation or average is to be taken.

As discussed above, we use two quantities of interest: the relative ranking of a given immunization method among its competitors $\mathrm{Res}_{f,o,m} = \mathrm{Rank}_m[P_{f,o,m}]$, where $P_{f,o,m}$ is the hedging

[4] In such cases, we leave the bond with the smallest ID number, which is usually the one issued the earliest.

performance measured as the relative change in the chosen performance measure:

$$P_{f,o,m} = \frac{\text{Risk Measure with Hedging}_{f,o,m}}{\text{Risk Measure without Hedging}_{f,o,m}},$$

and Rank_m denotes the rank of $P_{f,o,m}$ across the index m, i. e., among all values $P_{f,o,\cdot}$, which gives us the rank-based general model risk R_i:

$$R_i = \text{Std}_{f,o,m}\left[\text{Rank}_m[P_{f,o,m}] - \text{Mean}_f\left(\text{Rank}_m[P_{f,o,m}]\right)\right].$$

The other possible quantity of interest is the relative performance of a given immunization method in reducing the risk $\text{Res}_{f,o,m} = \ln(P_{f,o,m})$, which gives the log-based general model risk L_i:

$$L_i = \exp\left\{\text{Std}_{f,o,m}\left[\ln P_{f,o,m} - \text{Mean}_f\left(\ln P_{f,o,m}\right)\right]\right\} - 1.$$

The transformation $e^x - 1$ is added merely for ease of interpretation to undo the effect of the log and revert to the original scale.

Note that the choice of the risk measure is determined by one of the features—it can be the standard deviation of the financial result, its mean absolute deviation or 95% Value-at-Risk depending on the feature 'Risk Measure'.

For example, if the rank-based general model risk for feature i, $R_i = 2$, we can say that variations in feature i can typically shuffle the immunization methods leaderboard by $\pm$ 2 places for given values of the other features. On the other hand, if the log-based general model risk for feature i, $L_i = 0.15$, we can say that variations in feature i can affect the performance of an immunization method in reducing the chosen risk measure by $\pm$15% in a given setting.

For the log-based model risk we take one further step to split it into systemic and individual components as follows.

Systemic log-based model risk L_i^{sys} of feature i is calculated as

$$L_i^{sys} = \text{Std}_f\left[\text{Mean}_{o,m}\left(\text{Res}_{f,o,m}\right)\right].$$

Systemic model risk measures whether any single value of feature i on average improves (or reduces) the performance of all random immunization methods for all other feature values. Near-zero values of systemic model risk for some feature i mean that a change in this feature is just as likely to increase the performance of a random method with random other feature values as to decrease it. On the contrary, positive values indicate that some values of feature i are generally better than alternatives regardless of the chosen method and the other parameters.

$L_i^{sys} = 0.15$ can be interpreted as follows: by tuning the feature i one could achieve an average gain of about 15% in reducing the chosen risk measure regardless of the chosen

Feature i	R_i, places	P_i, points	L_i, %	L_i^{sys}, %	L_i^{ind}, %
Constraints	4.1	1.6	69	21	63
Term Structure	1.8	1	49	16	45
Valuation	2.4	1.3	77	32	64
Horizon	2.6	1.1	56	29	44
Risk Measure	2.3	1	69	43	46
Regularization Functional	3.7	1.3	91	44	70
Regularization Parameter	4.1	1.4	78	29	69
New Bonds	0.3	0.2	3.9	0.8	3.8
Missing Data	1.4	0.6	25	2.1	25

Table 1: Measures of model risk associated with various model features.

immunization method and the other parameters. Note that it is not sensible to calculate the systemic component of rank-based model risk since the average rank is not supposed to change.

Individual log-based model risk L_i^{ind} of feature i is calculated as

$$L_i^{ind} = \mathrm{Std}_{f,o,m}\left[\mathrm{Res}_{f,o,m} - \mathrm{Mean}_f\left(\mathrm{Res}_{f,o,m}\right) - \mathrm{Mean}_{o,m}\left(\mathrm{Res}_{f,o,m}\right)\right].$$

This measures the unpredicted effect of changing setting i, which depends on the choice of the immunization method and of other parameters. $L_i^{ind} = 0.15$ could mean that changes in feature i affect the immunization performance by $\pm 15\%$ on average, however this effect is positive for some methods and some parameters and negative for other methods and/or other parameters.

Points-based model risk P_i is calculated as a kind of robustness check—we acknowledge that for rank-based model risk estimates, perturbations at the top of the leaderboard are usually much more important than those at the bottom. Therefore, instead of calculating the average change in places, we assign points according to the following scheme: the first place is worth 10 points, the second place 7 points. The 3rd to the 7th places earn 5, 4, 3, 2, and 1 point respectively. All other places earn no points (the maximum rank is 25 as there are 25 hedging methods tested including no hedging).

Table 1 reports these measures. We now examine a few two-dimensional slices of this table in more detail. Figure 1 shows the model risk estimates in two dimensions. We can summarize the model risk of our 9 features as follows.

- The rule for including the newly issued bonds virtually does not matter for the purposes of estimating the immunization performance of a single method and comparing various immunization methods. This is probably due to the fact that new bonds are issued relatively rarely.

- The rule for dealing with missing data is a little bit more important but not much. We expect this to change with the overall liquidity of the market in question.

- The choice of the term structure fitting model is not very important in general—although in certain circumstances it might significantly favor methods which are designed with this particular term structure model in mind.

- The immunization horizon not only changes the performance of the methods, but also shuffles their relative positions by $\pm$ 2–3 places on average. This suggests that the choice of the best immunization method should be tied to a specific immunization horizon.

- The choice of the performance measure (Value-at-Risk, mean absolute deviation, standard deviation) and the portfolio valuation approach (mark-to-market or mark-to-model) cause significant performance discrepancy but moderate leaderboard changes. This can happen if the performances of various methods are affected similarly. Note that since we deal with relative performance change, the fact that Value-at-Risk is normally greater than the deviation does not influence the results as all changes are thus considered relative to the corresponding hedge-free risk measure.

- Portfolio constraints greatly impact the leaderboard with moderate changes to the performance. This implies that, for example, top performing methods with no portfolio constraints might well exhibit quite modest performance in the presence of trading or budget constraints.

- Finally, both the regularization method and the regularization parameter produce the highest model risk in terms of both performance and ranking. This suggests that immunization studies employing regularization techniques should devote more attention to the technical issues, i.e., to choosing the regularization functional and the regularization parameter.

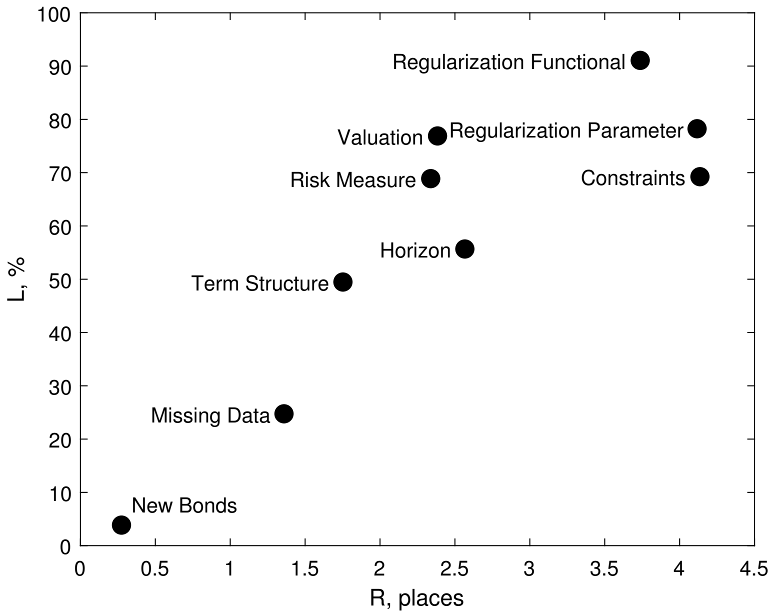

Figure 1: Relation between rank-based model risk R and log-based model risk L.

Our robustness check is visualized in Fig. 2. We can see that the relationship between the rank-based model risk R and the points-based model risk P is almost linear. We might say that the choices of the valuation principle and of the term structure impact the best-performing methods a little bit more than average while the choices of portfolio constraints and regularization tend to affect the best-performing methods a little bit less than average.

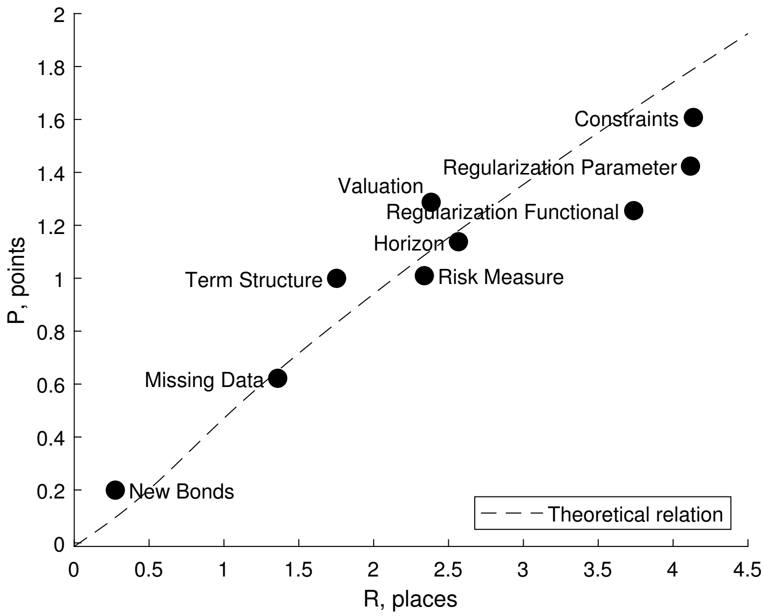

Figure 2: Robustness check: rank-based model risk R vs. points-based model risk P.

5 Conclusion

The design of the empirical bond portfolio hedging exercise means making a lot of choices. Some of these choices, e.g., immunization horizon, could be informed by the financial problem being solved while some others, e.g., choosing the regularization parameter or functional, are usually chosen arbitrarily.

We have considered two model problems—assessing the performance of a single hedging strategy and comparing several strategies. For each of these modeling choices (features), we have assessed the extent to which the answer depends on this choice thus estimating the model risk.

Some of the features turned out to be insignificant while some others give rise to significant model risk. Moreover, these significant features cannot be inferred from the model formulation in an obvious way and hence admit a great deal of voluntarism in their choice. There are two direct consequences of this result. First, future empirical immunization studies are advised to take greater care in designing the experiment. Second, existing empirical studies are most likely not directly comparable due to significant methodological differences—this applies to the studies proposing a new immunization method while comparing it to the existing best practice approach and to the studies comparing several existing immunization methods.

References

S. Agca. The Performance of Alternative Interest Rate Risk Measures and Immunization Strategies under a Heath-Jarrow-Morton Framework. *Journal of Financial and Quantitative Analysis*, 40 (3):645–669, 2005. doi: 10.1017/s0022109000001903.

K. T. Au and D. C. Thurston. A new class of duration measures. *Economics Letters*, 47(3-4): 371–375, 1995. doi: 10.1016/0165-1765(94)00553-E.

A. Balbás and A. Ibáñez. When can you immunize a bond portfolio? *Journal of Banking and Finance*, 22(12):1571–1595, 1998. doi: 10.1016/S0378-4266(98)00070-3.

A. Balbás and A. Ibáñez. Maxmin Portfolios in Models Where Immunization is Not Feasible. In *Computational Methods in Decision-Making, Economics and Finance. Applied Optimization, vol 74.*, pages 139–165. Springer, Boston, MA, 2002. doi: 10.1007/978-1-4757-3613-7_8.

A. Balbás and R. Romera. Hedging interest rate risk by optimization in Banach spaces. *Journal of Optimization Theory and Applications*, 132(1):175–191, 2007. doi: 10.1007/s10957-006-9124-6.

A. Balbás, A. Ibáñez, and S. López. Dispersion measures as immunization risk measures. *Journal of Banking and Finance*, 26(6):1229–1244, 2002a. doi: 10.1016/S0378-4266(01)00168-6.

A. Balbás, A. Ibáñez, and R. Romera. Shadow Risk-Free Returns When Hedging the Interest Rate Risk. Technical Report January, Universidad Carlos III de Madrid, 2002b.

J. R. Barber and M. L. Copper. Immunization using principal component analysis: An immunization strategy based upon historical term structure shifts. *Journal of Portfolio Management*, 23 (1):99–105, 1996. doi: 10.3905/jpm.1996.409574.

J. R. Barber and M. L. Copper. A minimax risk strategy for portfolio immunization. *Insurance: Mathematics and Economics*, 23(2):173–177, 1998. doi: 10.1016/S0167-6687(98)00028-6.

J. A. Beekman and E. S. Shiu. Stochastic models for bond prices, function space integrals and immunization theory. *Insurance Mathematics and Economics*, 7(3):163–173, 1988. doi: 10.1016/0167-6687(88)90073-X.

G. O. Bierwag. Immunization, Duration, and the Term Structure of Interest Rates. *Journal of Financial and Quantitative Analysis*, 12(5):725–742, 1977.

G. O. Bierwag and C. Khang. An Immunization Strategy is a Minimax Strategy. *The Journal of Finance*, 34(2):413, 1979. doi: 10.2307/2326980.

G. O. Bierwag, G. G. Kaufman, R. Schweitzer, and A. L. Toevs. The art of risk management in bond portfolios. *The Journal of Portfolio Management*, 7(3):27–36, 1981. doi: 10.3905/jpm.1981.408806.

G. O. Bierwag, G. G. Kaufman, and A. Toevs. Immunization Strategies for Funding Multiple Liabilities. *The Journal of Financial and Quantitative Analysis*, 18(1):113, 1983. doi: 10.2307/2330807.

G. O. Bierwag, G. G. Kaufman, and C. M. Latta. Bond Portfolio Immunization: Tests of Maturity, One- and Two-Factor Duration Matching Strategies. *Financial Review*, 22(2):203–219, 1987a. doi: 10.1111/j.1540-6288.1987.tb00761.x.

G. O. Bierwag, G. G. Kaufman, C. M. Latta, and G. S. Roberts. Duration: Response to Critics. *The Journal of Portfolio Management*, 13(2):48–52, 1987b. doi: https://doi.org/10.3905/jpm.1987.409098.

G. O. Bierwag, I. Fooladi, and G. S. Roberts. Designing an immunized portfolio: Is M-squared the key? *Journal of Banking and Finance*, 17(6):1147–1170, 1993. doi: 10.1016/0378-4266(93)90018-9.

P. P. Boyle. Immunization Under Stochastic Models of the Term Structure. *Journal of the Institute of Actuaries*, 105(2):177–187, 1978. doi: 10.4324/9781315145976-15.

J. M. V. Bravo and C. M. P. da Silva. Immunization using a stochastic-process independent multi-factor model: The Portuguese experience. *Journal of Banking and Finance*, 30(1):133–156, 2006. doi: 10.1016/j.jbankfin.2005.01.006.

J. M. V. Bravo and J. Fonseca. Parametric Immunization in Bond Portfolio Management. In *Proceedings of the 9th AFE International Conference on Applied Financial Economics*, pages 562–571, 2012.

N. Carcano and S. Foresi. Hedging against interest rate risk: Reconsidering volatility-adjusted immunization. *Journal of Banking and Finance*, 21(2):127–141, 1997. doi: 10.1016/S0378-4266(96)00031-3.

A. Carvalhal and S. P. Daumas. A Minimizing Risk Strategy for Brazilian Fixed-Income Portfolios. *Latin American Business Review*, 11(1):1–18, 2010. doi: 10.1080/10978521003698471.

D. R. Chambers, W. T. Carleton, and R. W. McEnally. Immunizing Default-Free Bond Portfolios with a Duration Vector. *The Journal of Financial and Quantitative Analysis*, 23(1):89, 1988. doi: 10.2307/2331026.

I. A. Cooper. Asset Values, Interest-Rate Changes, and Duration. *The Journal of Financial and Quantitative Analysis*, 12(5):701, 1977. doi: 10.2307/2330252.

J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross. Duration and the Measurement of Basis Risk. *The Journal of Business*, 52(1):51, 1979. doi: 10.1086/296033.

R. E. Dattatreya and F. J. Fabozzi. The Risk-Point Method for Measuring and Controlling Yield Curve Risk. *Financial Analysts Journal*, 51(4):45–54, 1995. doi: 10.2469/faj.v51.n4.1920.

R. L. D'ecclesia and S. A. Zenios. Risk Factor Analysis and Portfolio Immunization in the Italian Bond Market. *The Journal of Fixed Income*, 4(2):51–58, 1994. doi: 10.3905/jfi.1994.408113.

A. Diaz, M. González, and E. Navarro Arribas. Bond Portfolio Immunization, Immunization Risk and Idiosyncratic Risk. *Revista de economía financiera*, 16:8–25, 2008.

A. Díaz, M. d. l. O. González, E. Navarro, and F. S. Skinner. An evaluation of contingent immunization. *Journal of Banking and Finance*, 33(10):1874–1883, 2009. doi: 10.1016/j.jbankfin.2009.04.008.

E. J. Elton, M. J. Gruber, and R. Michaely. The Structure of Spot Rates and Immunization. *The Journal of Finance*, 45(2):629–642, 1990. doi: 10.1111/j.1540-6261.1990.tb03708.x.

L. Fisher and R. L. Weil. Coping with the Risk of Interest-Rate Fluctuations: Returns to Bondholders From Naïve and Optimal Strategies. *Journal of Business*, 44(4):408–431, 1971.

H. G. Fong and O. A. Vasicek. A Risk Minimizing Strategy for Portfolio Immunization. *The Journal of Finance*, 39(5):1541–1546, 1984. doi: 10.1111/j.1540-6261.1984.tb04923.x.

I. Fooladi and G. S. Roberts. Bond portfolio immunization: Canadian tests. *Journal of Economics and Business*, 44(1):3–17, 1992. doi: 10.1016/0148-6195(92)90003-S.

N. B. Guletkin and R. J. Rogalski. Alternative Duration Specifications and the Measurement of Basis Risk: Empirical Tests. *The Journal of Business*, 57(2):241–264, 1984.

C. F. Hill and S. Vaysman. An approach to scenario hedging: Using principal components analysis. *Journal of Portfolio Management*, 24(2):83–92, 1998. doi: 10.3905/jpm.24.2.83.

T. S. Ho. Key Rate Durations. *The Journal of Fixed Income*, 2(2):29–44, 1992. doi: 10.3905/jfi.1992.408049.

W. Hürlimann. On immunization, stop-loss order and the maximum Shiu measure. *Insurance: Mathematics and Economics*, 31(3):315–325, 2002. doi: 10.1016/s0167-6687(02)00158-0.

A. Ilmanen. How Well Does Duration Measure Interest Rate Risk? *The Journal of Fixed Income*, 1(4):43–51, 1992. doi: 10.3905/jfi.1992.408032.

J. E. Ingersoll, J. Skelton, and R. L. Weil. Duration Forty Years Later. *The Journal of Financial and Quantitative Analysis*, 13(4):627–650, 1978. doi: 10.2307/2330468.

N. J. Lacey and S. K. Nawalkha. Convexity, Risk, and Returns. *Journal of Fixed Income*, 3(3):72–79, 1993. doi: 10.3905/jfi.1993.408086.

V. Lapshin. A nonparametric approach to bond portfolio immunization. *Mathematics*, 7(11):1121, 2019. doi: 10.3390/math7111121.

V. Lapshin. Immunizing a Marked-to-Model Obligation with Marked-to-Market Financial Instruments. Working Paper WP BRP 84/FE/2021, HSE University, Moscow, 2021. URL https://ssrn.com/abstract=3962205.

S. B. Lee and H. Y. Cho. A Rebalancing Discipline for an Immunization Strategy. *The Journal of Portfolio Management*, 18(4):56–62, 1992. doi: 10.3905/jpm.1992.409413.

R. B. Litterman and J. Scheinkman. Common Factors Affecting Bond Returns. *The Journal of Fixed Income*, 1(1):54–61, 1991. doi: 10.3905/jfi.1991.692347.

M. Á. M. Mato. Classic and modern measures of risk in fixed-income portfolio optimization. *Journal of Risk Finance*, 6(5):416–423, 2005. doi: 10.1108/15265940510633488.

S. K. Nawalkha and D. R. Chambers. An improved immunization strategy: M-absolute. *Financial Analysts Journal*, 52(5):69–76, 1996. doi: 10.2469/faj.v52.n5.2026.

S. K. Nawalkha and D. R. Chambers. The M-Vector Model. *The Journal of Portfolio Management*, 23(2):92–98, 1997. doi: 10.3905/jpm.23.2.92.

S. K. Nawalkha, G. M. Soto, and J. Zhang. Generalized M-vector models for hedging interest rate risk. *Journal of Banking and Finance*, 27(8):1581–1604, 2003. doi: 10.1016/S0378-4266(03)00089-X.

L. Oliveira, J. P. Vidal Nunes, and L. Malcato. The performance of deterministic and stochastic interest rate risk measures: Another Question of Dimensions? *Portuguese Economic Journal*, 13(3):141–165, 2014. doi: 10.1007/s10258-014-0104-8.

S. Ortobelli, S. Vitali, M. Cassader, and T. Tichý. Portfolio selection strategy for fixed income markets with immunization on average. *Annals of Operations Research*, 260(1-2):395–415, 2018. doi: 10.1007/s10479-016-2182-8.

E. Z. Prisman. Immunization as a maxmin strategy: A new look. *Journal of Banking and Finance*, 10(4):491–509, 1986. doi: 10.1016/S0378-4266(86)80002-4.

F. M. Redington. Review of the Principles of Life-Office Valuations. *Journal of the Institute of Actuaries*, 78(3):286–340, 1952. doi: 10.4324/9781315145976-9.

R. R. Reitano. Non–parallel yield curve shifts and durational leverage. *The Journal of Portfolio Management*, 16(4):62–67, 1990. doi: 10.3905/jpm.1990.409282.

R. R. Reitano. Non-Parallel Yield Curve Shifts and Immunization. *The Journal of Portfolio Management*, 18(3):36–43, 1992. doi: 10.3905/jpm.1992.409410.

E. S. Shiu. Immunization of multiple liabilities. *Insurance Mathematics and Economics*, 7(4):219–224, 1988. doi: 10.1016/0167-6687(88)90079-0.

G. M. Soto. Duration models and IRR management: A question of dimensions? *Journal of Banking and Finance*, 28(5):1089–1110, 2004. doi: 10.1016/S0378-4266(03)00044-X.

C. S. Tark. The Term Structure of Interest Rates, the Stochastic Bond Price, and Immunization. *The Korean Economic Review*, 6(2):41–60, 1990.

M. F. Theobald and P. J. Yallup. Liability-driven investment: Multiple liabilities and the question of the number of moments. *European Journal of Finance*, 16(5):413–435, 2010. doi: 10.1080/13518470903211681.

M. Uberti. Immunization of Portfolios with Liabilities. In M. Bonilla, T. Casasús, and R. Sala, editors, *Financial Modelling. Contributions to Management Science*, pages 391–400. Physica, Heidelberg, 2000. doi: 10.1007/978-3-642-57652-2_26.

R. Willner. A New Tool for Portfolio Managers: Level, Slope and Curvature Durations. *The Journal of Fixed Income*, 6(1):48–59, 1996. doi: 10.3905/jfi.1996.408171.

X. Wu. A new stochastic duration based on the vasicek and CIR term structure theories. *Journal of Business Finance and Accounting*, 27(7-8):911–932, 2000. doi: 10.1111/1468-5957.00339.

W. Zhu, C. H. Zhang, Q. Liu, and S. S. Zhu. Incorporating Convexity in Bond Portfolio Immunization Using Multifactor Model: A Semidefinite Programming Approach. *Journal of the Operations Research Society of China*, 6(1):3–23, 2018. doi: 10.1007/s40305-018-0196-4.