Robust Asset-Liability Management

Tjeerd de Vries* Alexis Akira Toda†

April 15, 2025

Abstract

How should financial institutions hedge their balance sheets against interest rate risk when they have long-term assets and liabilities? Using the perspective of functional and numerical analysis, we propose a model-free bond portfolio selection method that generalizes classical immunization and accommodates arbitrary liability structure, portfolio constraints, and perturbations in interest rates. We prove the generic existence of an immunizing portfolio that maximizes the worst-case equity with a tight error estimate and provide a solution algorithm. Numerical evaluations using empirical and simulated yield curves from a no-arbitrage term structure model support the feasibility and accuracy of our approach relative to existing methods.

Keywords: immunization, interest rate risk, maxmin, model-free, robustness.

JEL codes: C65, G11, G12, G22.

1 Introduction

Many financial institutions have long-term commitments. For instance, insurance companies promise annuities or life insurance payments to customers; (defined-benefit) pension plans promise predetermined pension payments to retirees; or commercial banks may make long-term loans at fixed interest rates and thus commit to receiving certain future cash flows in exchange of funding the projects with short-term deposits. In such circumstances, it becomes crucial for financial

*Department of Finance, HEC Paris. Email: de-vries@hec.fr.

†Department of Economics, Emory University. Email: alexis.akira.toda@emory.edu.

institutions to effectively manage their assets and liabilities to hedge against interest rate risk. The recent gilt market crisis in the UK showcases the importance of liability-driven investing strategies and the risk associated with interest rate changes, which eventually led to an £65 billion emergency intervention by the Bank of England.[1] Even more recently, Silicon Valley Bank and First Republic Bank collapsed as a result of increased interest rates and the subsequent decline in value of long-term bonds and mortgages.[2,3]

If zero-coupon bonds of all maturities were to exist, any deterministic future cash flow can be replicated by these bonds (which is called a "dedication" strategy), and the problem becomes trivial, at least theoretically. However, in practice dedication is infeasible due to market incompleteness: there are fewer bonds available for trade than the number of payment dates of the liability, or the long-term liability could have a longer maturity than the government bond with longest maturity. Thus, in general, one can only hope to hedge against interest rate risk approximately. The question of fundamental practical importance is how to achieve this goal given the set of bonds available for trade.

In this article, we propose a new method to construct a hedging portfolio that maximizes equity (asset minus liability) under the most adversarial interest rate shock. This so-called *maxmin* problem originates in the work of Fisher and Weil (1971), who show that a portfolio that matches value and duration (weighted average time to payment) is maxmin against parallel shocks to the yield curve. In that and subsequent works, the liability is assumed to be a zero-coupon bond and no-shortsale constraints are imposed (or implicitly assumed not to bind). These restrictions are undesirable in practice because most liabilities pay out over time and shortsales are essential when liabilities have very long maturities (like pensions). Furthermore, there is no systematic analysis of the existence, uniqueness, and optimality of the solution method as well as explicit or tight error estimates.

Our approach overcomes these shortcomings using techniques from functional and numerical analysis. First we argue that the most general formulation of the maxmin problem is intractable because the objective function is not convex and the space has infinite dimension. To make the problem manageable, we approximate the objective function using the Gateaux differential with respect to basis functions that approximate yield curve shifts. This allows us to recast the maxmin

[1] https://www.bankofengland.co.uk/speech/2022/november/sarah-breeden-speech-at-isda-aimi-boe-on-nbfi-and-leverage

[2] https://www.ft.com/content/f9a3adce-1559-4f66-b172-cd45a9fa09d6

[3] https://www.economist.com/finance-and-economics/2023/05/03/what-the-first-republic-deal-means-for-americas-banks

problem as a saddle point (minmax) problem where the inner maximization is a large linear programming problem and the outer minimization is a small convex programming problem, which is computationally tractable. We prove that a robust immunizing portfolio generically exists (Proposition 3.1) and its solution achieves the smallest error order and maximizes the worst-case equity (Theorem 1). This maxmin result is significantly different from the existing literature because both the liability structure and bond portfolio constraint are arbitrary and the guaranteed equity bound is tight. When the majority of yield curve changes are captured by a small number of principal components such as the level of the overall interest rate, we improve this guaranteed equity bound by incorporating moment matching (e.g., duration matching) in the portfolio constraint (Theorem 4). We also propose particular basis functions (transformation of Chebyshev polynomials) that are motivated by approximation theory.

An alternative approach to asset-liability management, referred to as classical immunization (see, e.g., Redington (1952)), involves matching the interest rate sensitivity of asset and liability. A common measure of interest rate sensitivity is duration, and matching the duration of asset and liability makes equity insensitive to small interest rate changes. Although classical immunization is intuitive and elegant, by assumption it only allows for small parallel shifts in the yield curve. Furthermore, when there are multiple bonds, it is not obvious how to construct the portfolio because there are infinitely many linear combinations that achieve the same duration. Extensions such as high-order duration matching (which are designed to allow non-infinitesimal or non-parallel shifts in the yield curve) result in unstable portfolio weights and extreme leverage, leading to poor performance (Mantilla-Garcia et al., 2022). Our approach contains classical immunization and its extensions as a special case by choosing a monomial basis and imposing only a value matching constraint. In simulation, we show that our preferred robust immunization method that combines moment matching and a Chebyshev polynomial basis does not suffer from extreme leverage and significantly outperforms existing methods.

The simulation exercise uses historical yield curve data to evaluate the change in equity resulting from instantaneous yield curve shocks. A hedging method's success is measured by its ability to minimize these equity changes. Indeed, we find that robust immunization generates approximation errors that are an order of magnitude smaller than the existing approaches and has lower downside risk, in line with our maxmin result. This numerical experiment has a static flavor, since we only consider one-time perturbations. In a separate simulation based on a no-arbitrage term structure model, we consider the dynamic properties of robust immunization, allowing for portfolio rebalancing every three months. Over a 10-year period of rebalancing, robust immunization achieves an approximation error at least 24% lower in the 1% worst-case scenario compared to existing methods. Because our approach is model-free, we expect our proposed method to be useful for practitioners in asset-liability management.[4]

1.1 Related literature

When inputs to a problem such as beliefs, information, or shocks are complicated, it is common to optimize against the worst case scenario, i.e., solve the *maxmin* problem (Gilboa and Schmeidler, 1989; Bergemann and Morris, 2005; Du, 2018; Brooks and Du, 2021). In the context of asset-liability management, Redington (1952, p. 290) considers the Taylor expansion of assets minus liabilities in response to a small change in the (constant) interest rate and anticipates the importance of convexity to guarantee the portfolio value. Fisher and Weil (1971) formalize this idea and show that if the liability is a zero-coupon bond and a bond portfolio matches the value and duration, then the portfolio value can never fall below liabilities under any parallel shift to the yield curve. Bierwag and Khang (1979) show that when the investor has a fixed budget to invest in bonds, then classical immunization (duration matching) is maxmin in the sense that it maximizes the worst possible rate of return under any parallel shift to the yield curve. Fong and Vasicek (1984) consider any perturbation to the forward curve such that the slope of the forward curve is bounded by some constant and derive a lower bound on the portfolio return over the investment horizon that is proportional to it. The constant of proportionality is a measure of interest rate risk and is called "M-squared". Minimization of M-squared renders a portfolio that minimizes the likelihood of a deviation from liabilities. Bowden (1997) proposes measuring sensitivity to interest rate risk using the Fréchet derivative, which Balbás and Ibáñez (1998) use to prove a maxmin result under the assumption of no shortsales. Zheng (2007) considers perturbations to the forward rate that are Lipschitz continuous, derives the maximum deviation of the bond value, and applies it to a portfolio choice problem.

Several classical books and papers such as Macaulay (1938), Hicks (1939,

[4] This statement is similar to the fact that the Black and Scholes (1973) option pricing model has been hugely successful precisely because the model requires only a few assumptions, namely the absence of arbitrage and the stock price following a geometric Brownian motion, and no assumptions on investor preferences are required.

pp. 184-188), and Samuelson (1945) discovered that the average time to payment ("duration") of a bond captures the interest rate sensitivity of the bond with respect to parallel shifts in the yield curve. Redington (1952) suggested matching the duration of the asset and liability ("immunization") to hedge against interest rate risk. Chambers et al. (1988), Nawalkha and Lacey (1988), and Prisman and Shores (1988) use polynomials to approximate the yield curve and discuss immunization using high-order duration measures. Ho (1992) introduced the concept of "key rate duration", which is the bond price sensitivity with respect to local shifts in the yield curve at certain key rates (e.g., 10-year yield). In the language of Vayanos and Vila (2021), hedging key rate risk is relevant for preferred-habitat investors with specific maturity preferences. Litterman and Scheinkman (1991) use principal component analysis (PCA) to identify common factors that affect bond returns and find that the three factors called *level*, *slope*, and *curvature* explain a large fraction of the variations in returns. Using these factors, Willner (1996) defines level, slope, and curvature durations and shows how they can be used for asset-liability management. See Sydyak (2016) for a review of this literature. In a recent paper, Onatski and Wang (2021) argue that PCA based on the yield curve is prone to spurious analysis since bond yields are highly persistent. As a result, Crump and Gospodinov (2022) show that PCA tends to favor a much lower dimension of the factor space than the true dimension, which can lead to large costs in bond portfolio management. We further discuss our contribution relative to the literature in Section 3.2.

2 Problem statement

2.1 Classical immunization

We start the discussion with a brief review of classical immunization. Consider cash flows (liabilities) $f_1, \dots, f_N > 0$ paid out at time $t_1 < \dots < t_N$. Assuming a constant (continuously-compounded) interest rate r, the present value of liabilities is

$$P := \sum_{n=1}^{N} \mathrm{e}^{-rt_n} f_n. \tag{2.1}$$

Because each term in the sum (2.1) is decreasing in r, so is the present value P. Therefore we may define the interest rate sensitivity of liability by

$$D := -\frac{\partial \log P}{\partial r} = -\frac{1}{P}\frac{\partial P}{\partial r} = \frac{1}{P}\sum_{n=1}^{N} t_n \mathrm{e}^{-rt_n} f_n > 0. \tag{2.2}$$

Intuitively, D is the percentage change in present value with respect to a one percentage point change in the interest rate. The quantity D in (2.2) is called the *duration* of the cash flow $(f_1, \dots, f_N)$ because it can be interpreted as the average time to payment. To see why, define the weight $w_n = \mathrm{e}^{-rt_n} f_n / P$, which by (2.1) is the fraction of present value of time t_n payment. Since $w_n > 0$ and $\sum_{n=1}^{N} w_n = 1$, the definition of duration (2.2) implies $D = \sum_{n=1}^{N} w_n t_n$ is indeed a weighted average of time to payment.

In general, the interest rate need not be constant across different maturities. If $y(t)$ denotes the pure yield for maturity t (so by definition the price of a zero-coupon bond with face value 1 and maturity t is $\mathrm{e}^{-y(t)t}$), then the present value and duration of the cash flows become

$$\begin{aligned}
P(y) &= \sum_{n=1}^{N} \mathrm{e}^{-y(t_n)t_n} f_n, \\
D(y) &= -\frac{1}{P(y)} \lim_{\Delta \to 0} \frac{P(y+\Delta) - P(y)}{\Delta} = \frac{1}{P(y)} \sum_{n=1}^{N} t_n \mathrm{e}^{-y(t_n)t_n} f_n,
\end{aligned}$$

respectively. Here the duration is the sensitivity of the present value with respect to an infinitesimal parallel shift in the yield curve. The idea of classical *immunization* is to match the duration of asset and liability so that equity (asset minus liability) is insensitive to yield curve shifts (Macaulay, 1938; Samuelson, 1945; Redington, 1952). Below, we generalize classical immunization to overcome the limitations discussed in the introduction.

2.2 Robust immunization problem

We now turn to the description of the robust immunization problem. Time is continuous and denoted by $t \in [0, T]$, where $T > 0$ is the planning horizon. There are finitely many bonds available for trade indexed by $j = 1, \dots, J$, where $J \geq 2$. The cumulative payout of bond j is denoted by the (weakly) increasing function $F_j : [0, T] \to \mathbb{R}_+$. For instance, if bond j is a zero-coupon bond with face value normalized to 1 and maturity t_j, then

$$F_j(t) = \begin{cases} 0 & \text{if } 0 \le t < t_j, \\ 1 & \text{if } t_j \le t \le T. \end{cases} \tag{2.3}$$

Similarly, if bond j continuously pays out coupons at rate $c_j > 0$ and has zero face value, then $F_j(t) = c_j t$ for $0 \le t \le T$.

The fund manager seeks to immunize future cash flows against interest rate risk by forming a portfolio of bonds $j = 1, \dots, J$. Let $F : [0, T] \to \mathbb{R}_+$ be the cumulative cash flow to be immunized and $y : [0, T] \to \mathbb{R}$ be the yield curve, which the fund manager takes as given. The present value of cash flows is given by the Riemann-Stieltjes integral

$$\int_0^T \mathrm{e}^{-ty(t)}\,\mathrm{d}F(t). \tag{2.4}$$

Because the expression $ty(t)$ appears frequently, it is convenient to introduce the notation $x(t) := ty(t)$. Note that by the definition of the instantaneous forward rate $f(t)$ at term t, we have

$$x(t) = \int_0^t f(u)\,\mathrm{d}u. \tag{2.5}$$

Because x is the integral of forward rates, we refer to it as the *cumulative discount rate*. Using x, we can rewrite the present value of cash flows (2.4) as

$$P(x) := \int_0^T \mathrm{e}^{-x(t)}\,\mathrm{d}F(t), \tag{2.6}$$

which is a functional of x. We can define the price $P_j(x)$ of bond j analogously. The fund manager's problem is to approximate $P(x)$ using a linear combination of bonds $\{P_j(x)\}_{j=1}^J$ in a way such that the approximation is robust against perturbations to the yield curve y (and hence the cumulative discount rate x).

To define this portfolio choice problem, let $\mathcal{Z} \subset \mathbb{R}^J$ and $\mathcal{H}$ be the sets of admissible portfolios and perturbations to the cumulative discount rate, to be specified later. For portfolio $z \in \mathcal{Z}$ and perturbation $h \in \mathcal{H}$, define the "asset" by

$$V(z, x+h) := \sum_{j=1}^J z_j P_j(x+h)$$

and the "equity" by asset minus liability

$$
\begin{aligned}
E(z, x+h) &\coloneqq V(z, x+h) - P(x+h) \\
&= \sum_{j=1}^{J} z_j P_j(x+h) - P(x+h).
\end{aligned} \tag{2.7}
$$

We assume the fund manager has a maxmin preference and seeks to solve

$$
\sup_{z \in \mathcal{Z}} \inf_{h \in \mathcal{H}} E(z, x+h). \tag{2.8}
$$

The interpretation of the maxmin problem (2.8) is as follows. Given the portfolio $z \in \mathcal{Z}$, nature chooses the most adversarial perturbation $h \in \mathcal{H}$ to minimize equity. The fund manager chooses the portfolio z that guarantees the highest equity under the worst possible perturbation. We refer to this problem as the *robust immunization problem*. Maxmin preferences are common in the literature on bond portfolio choice; see, e.g., the structural models of Gagliardini et al. (2008) and Vayanos and Vila (2021, Appendix B). The objective function in (2.8) can be microfounded using the ambiguity-averse preferences of Gilboa and Schmeidler (1989), assuming the fund manager is risk-neutral.

2.3 Assumptions

The maxmin problem (2.8) is intractable because we have not yet specified the admissible sets $\mathcal{Z}, \mathcal{H}$ and the objective function is nonlinear (not even convex) in h. We thus impose several assumptions to make progress.

Assumption 1 (Discrete payouts)**.** *The bonds and liability pay out on finitely many dates, whose union is denoted by* $\{t_n\}_{n=1}^N \subset (0, T]$.

Assumption 1 always holds in practice. Under this assumption, each F_j is a step function with discontinuities at points contained in $\{t_n\}_{n=1}^N$, and integrals of the form (2.6) reduce to summations.

Assumption 2 (Portfolio constraint)**.** *The set of admissible portfolios* $\mathcal{Z} \subset \mathbb{R}^J$ *is nonempty and closed. Furthermore, all* $z \in \mathcal{Z}$ *satisfy value matching:*

$$
P(x) = \sum_{j=1}^{J} z_j P_j(x). \tag{2.9}
$$

Value matching (2.9) is merely a normalization to make the initial equity (asset minus liability) equal to 0. This assumption is standard in the immunization literature (see, for example, Bierwag and Khang (1979)). Note that Assumption 2 allows for short-sale constraints either on the entire portfolio or on individual bonds. The latter constraint may arise if the portfolio manager is concerned about short-selling illiquid long-term bonds.

We now specify the space of cumulative discount rates and their perturbations. Let the yields be in $C[0,T]$; the vector space of continuous functions on $[0,T]$ endowed with the supremum norm denoted by $\|\cdot\|_\infty$.[5] By the definition of the cumulative discount rate, we obtain that $x : [0,T] \to \mathbb{R}$ defined by $x(t) = ty(t)$ is continuous with $x(0) = 0$. We thus define the space of cumulative discount rates by

$$\mathcal{X} = \{x \in C[0,T] : x(0) = 0\}. \tag{2.10}$$

Lemma A.1 shows that $\mathcal{X}$ is a Banach space. The next assumption allows us to approximate any element $x \in \mathcal{X}$, which is important for reducing the problem to a finite dimension.

Assumption 3. *There exists a countable basis* $\{h_i\}_{i=1}^\infty$ *of* $\mathcal{X}$ *such that for each* $I \in \{1, \dots, N\}$*, the* $I \times N$ *matrices* $H := (h_i(t_n))$ *and* $G := (h_i(t_n)/t_n)$ *have full row rank.*

We refer to each h_i as a *basis function*. The matrices H and G will be used to approximate the discount rate and yield curve at the payout dates.[6] Assumption 3 says that the basis functions are linearly independent when evaluated on the payout dates. We impose this assumption to avoid portfolio indeterminacy. In practice, we can always ensure that H and G have full row rank by removing redundant basis functions if necessary. A typical example satisfying Assumption 3 is to let h_i be a polynomial of degree i with $h_i(0) = 0$ (Lemma A.2).

Finally, we specify the set of admissible perturbations to the cumulative discount rate. For any $\Delta > 0$, define

$$\mathcal{H}_I(\Delta) := \left\{ h \in \operatorname{span}\{h_i\}_{i=1}^I : (\forall n) \, |h(t_n)/t_n| \le \Delta \right\}. \tag{2.11}$$

Because h is a perturbation to the cumulative discount rate, choosing $h \in \mathcal{H}_I(\Delta)$ amounts to allowing the yields to change by at most $\pm\Delta$ within the span of the

[5] As we use several different norms in this paper, we use subscripts to distinguish them. An example is the ℓ^p norm on $\mathbb{R}^J$ for $p = 1, 2$, which we denote by $\|\cdot\|_p$.

[6] Recall that the yield curve satisfies $y(t) = x(t)/t$.

first I basis functions. With this choice of $\mathcal{H}$, the robust immunization problem (2.8) becomes

$$\sup_{z\in\mathcal{Z}} \inf_{h\in\mathcal{H}_I(\Delta)} E(z, x+h). \tag{2.12}$$

Example 1 (Classical Immunization)**.** The setting in classical immunization corresponds to $I = 1$ and $h_1(t) = t$ (hence $h_1(t)/t = 1$), which implies that the perturbations to the yield curve are restricted to parallel shifts.

Example 2 (Vasicek Model)**.** In the Vasicek (1977) model, the change in the yields from time s to $s + S$ is given by

$$y_{s+S}(t) - y_s(t) = [r_{s+S} - r_s] \frac{1 - e^{-at}}{at} =: [r_{s+S} - r_s] \, h(t)/t, \tag{2.13}$$

where r_s is the spot rate that solves the stochastic differential equation

$$\mathrm{d}r_s = a(b - r_s)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s.$$

A similar, though more complicated expression for the yield changes holds in the equilibrium model of Vayanos and Vila (2021, Appendix B). Equation (2.13) implies that the model does not allow for parallel shifts of the yield curve. As we shall see, classical immunization is therefore never a maxmin strategy.

Example 3 (Principal Components)**.** Similar to Diebold and Li (2006), we can specify a 3-factor model for the yield curve at time s by:

$$y_s(t) = \beta_{1s}h_1(t)/t + \beta_{2s}h_2(t)/t + \beta_{3s}h_3(t)/t,$$

where $h_1(t)/t$ represents the loading on the level factor, $h_2(t)/t$ is the loading on the slope factor, and $h_3(t)/t$ is the loading on the curvature factor (see Figure 1). For example, an increase in β_{2s} will increase short yields more relative to long yields, thereby changing the slope of the yield curve. Similarly, an increase in β_{3s} primarily increases medium yields around the two-year maturity, while short and long yields remain unaffected. This leads to an increase in the curvature of the yield curve. In this setting, perturbations to the yield curve are restricted to changes in level, slope, and curvature. Litterman and Scheinkman (1991) find that these components are the main drivers of changes in the yield curve, although recent work of Crump and Gospodinov (2022) suggests that the factor dimension may be larger.

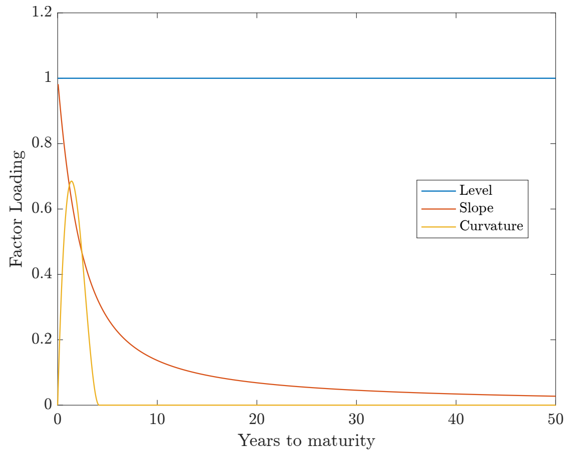


Figure 1: Factor loadings on the level, slope and curvature factors.

3 Solving robust immunization problem

In this section we solve the maxmin problem (2.12) in the limit as $\Delta \downarrow 0$.

3.1 Robust immunization

As the set of cumulative discount rates $\mathcal{X}$ forms an infinite-dimensional vector space, we employ tools from functional analysis to analyze how prices change in response to perturbations in the discount rate $h \in \mathcal{H}_I(\Delta)$. We assess the price change following an arbitrary shift in the cumulative discount rate by using the Gateaux differential of $P(x)$:

$$\delta P(x; h) := \lim_{\alpha \to 0} \frac{1}{\alpha}(P(x + \alpha h) - P(x)) = -\int_0^T \mathrm{e}^{-x(t)} h(t) \, \mathrm{d}F(t). \tag{3.1}$$

Remark 1. The operator $h \mapsto \delta P(x; h)$ defined by (3.1) is a bounded linear operator from $\mathcal{X}$ to $\mathbb{R}$ (Lemma A.3), which is called the Fréchet derivative and denoted by $P'(x)$. Thus by definition $P'(x)h = \delta P(x; h)$. In broad terms, $P'(x)h$ quantifies the first-order impact on price change when the cumulative discount rate curve is perturbed by h.

We construct a solution to the maxmin problem (2.12) by assessing the sensitivity of asset and liability to perturbations in specific directions h. Specifically, given the basis functions $\{h_i\}_{i=1}^I$ and bonds $j = 1, \dots, J$, we define the *sensitivity vector* $b = (b_i) \in \mathbb{R}^I$ of liabilities by

$$b_i := -\frac{P'(x)h_i}{P(x)} = -\frac{\delta P(x; h_i)}{P(x)} = \frac{1}{P(x)} \int_0^T \mathrm{e}^{-x(t)} h_i(t) \, \mathrm{d}F(t). \tag{3.2}$$

Note that the duration (2.2) corresponds to the special case of $h_i(t) = t$, and therefore b_i is a generalization. Intuitively, each entry b_i represents the sensitivity of liability to a perturbation evaluated at $h = h_i$. Similarly, we define the *sensitivity matrix* $A = (a_{ij}) \in \mathbb{R}^{I \times J}$ by

$$a_{ij} := -\frac{P_j'(x)h_i}{P(x)} = -\frac{\delta P_j(x; h_i)}{P(x)} = \frac{1}{P(x)} \int_0^T \mathrm{e}^{-x(t)} h_i(t) \, \mathrm{d}F_j(t). \tag{3.3}$$

Division by $P(x)$ is merely a normalization to make a_{ij} unit-free. Again, each entry a_{ij} represents the sensitivity of bond j (with $F = F_j$) to a perturbation evaluated at $h = h_i$.

If $h \in \mathcal{H}_I(\Delta)$ in (2.11), so $h = \Delta \sum_{i=1}^I w_i h_i$ for some $w \in \mathbb{R}^I$ (the coefficient Δ is to make w scale-free), then using the definition of equity (2.7) and noting that $E(z, x) = 0$ by Assumption 2, we obtain the sensitivity of equity

$$\lim_{\Delta \to 0} \frac{1}{\Delta P(x)} E(z, x + h) = -\langle w, Az - b \rangle, \tag{3.4}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. Hence, the change in equity following an infinitesimal perturbation in the discount rate is governed by the Fréchet derivatives of the asset and liability. For $h = \Delta \sum_{i=1}^I w_i h_i \in \mathcal{H}_I(\Delta)$, the coefficients (w_i) need to satisfy particular restrictions. Using the definition of the matrix G in Assumption 3 and (2.11), it is straightforward to show $h = \Delta \sum_{i=1}^I w_i h_i \in \mathcal{H}_I(\Delta)$ if and only if $G'w \in [-1, 1]^N$. This observation as well as (2.12) and (3.4) motivate us to define the set

$$\mathcal{W} := \left\{ w \in \mathbb{R}^I : G'w \in [-1, 1]^N \right\} \tag{3.5}$$

and the minmax problem

$$V_I(\mathcal{Z}) := \inf_{z \in \mathcal{Z}} \sup_{w \in \mathcal{W}} \langle w, Az - b \rangle. \tag{3.6}$$

The next proposition establishes the existence of a solution to the minmax problem (3.6). Before stating this result, it is convenient to introduce notation for the value matching constraint, which is always assumed to hold (Assumption 2). Specifically, set $h_0 \equiv 1$ and define a_{0j} using (3.3). Define the $1 \times J$ vector $a_0 := (a_{0j})$ and the $(I + 1) \times J$ matrix and $(I + 1) \times 1$ vector

$$A_+ := \begin{bmatrix} a_0 \\ A \end{bmatrix} \quad \text{and} \quad b_+ := \begin{bmatrix} 1 \\ b \end{bmatrix}. \tag{3.7}$$

In what follows, longer proofs are deferred to Appendix B.

Proposition 3.1 (Minmax)**.** *Suppose Assumptions 1–3 hold, $I \geq J-1$, and A_+ in (3.7) has full column rank. Then the following statements are true.*

(i) There exists $(z^*, w^*) \in \mathcal{Z} \times \mathcal{W}$ that achieves the minmax value (3.6).

(ii) $V_I(\mathcal{Z}) \geq 0$, and $z \in \mathcal{Z}$ achieves $V_I(\mathcal{Z}) = 0$ if and only if $A_+ z = b_+$.

The solution z to the minmax problem (3.6) depends on the basis functions $\{h_i\}_{i=1}^I$ only through its span and it is immaterial how we parameterize these functions.

Proposition 3.2 (Basis invariance)**.** *Let everything be as in Proposition 3.1 and $\mathcal{Z}^*$ be the set of solutions $z^* \in \mathcal{Z}$ to the minmax problem (3.6). Then $V_I(\mathcal{Z})$ and $\mathcal{Z}^*$ depend on the basis functions $\{h_i\}_{i=1}^I$ only through its span.*

Proposition 3.1 assumes that A_+ in (3.7) has full column rank, which holds under weak conditions. If the cumulative payouts of bonds $\{F_j\}$ and the basis functions $\{h_i\}$ are linearly independent, the matrix A_+ generically has full column rank and therefore a solution $(z, w) \in \mathcal{Z} \times \mathcal{W}$ to the minmax problem (3.6) generically exists. In Appendix B.1 we make this statement more precise.

Before presenting our main result, we introduce one last piece of notation. For any bond portfolio $z \in \mathcal{Z}$, define the portfolio share $\theta = (\theta_j) \in \mathbb{R}^J$ by

$$\theta_j := z_j P_j(x)/P(x). \tag{3.8}$$

Assuming value matching (Assumption 2), the portfolio share θ satisfies $\sum_{j=1}^J \theta_j = 1$. Therefore the ℓ^1 norm $\|\theta\|_1 = \sum_{j=1}^J |\theta_j|$ satisfies $\|\theta\|_1 = 1$ if and only if $\theta_j \geq 0$ for all j, and $\|\theta\|_1 > 1$ is equivalent to $\theta_j < 0$ for some j. Thus $\|\theta\|_1$ can be interpreted as a measure of leverage, which we refer to as the *gross leverage*.

Theorem 1 (Robust immunization)**.** *Let everything be as in Proposition 3.1 and $\mathcal{H}_I(\Delta)$ be as in (2.11). Then the following statements are true.*

(i) The guaranteed equity satisfies

$$\lim_{\Delta \downarrow 0} \frac{1}{\Delta} \sup_{z \in \mathcal{Z}} \inf_{h \in \mathcal{H}_I(\Delta)} E(z, x+h) = -P(x) V_I(\mathcal{Z}). \tag{3.9}$$

(ii) *Letting $z^* \in \mathcal{Z}$ be the solution to the minmax problem* (3.6) *and $\theta = (\theta_j) \in \mathbb{R}^J$ be the corresponding portfolio share defined by* (3.8)*, then*

$$\sup_{h \in \mathcal{H}_I(\Delta)} |E(z^*, x+h)| \le \Delta P(x) \left(V_I(\mathcal{Z}) + \frac{1}{4}\Delta T^2 \mathrm{e}^{\Delta T}(1 + \|\theta\|_1) \right). \tag{3.10}$$

Theorem 1 has several implications. First, (3.9) shows that, to the first order, the guaranteed equity is exactly $-\Delta P(x) V_I(\mathcal{Z})$ when yields are perturbed by at most $\pm\Delta$ within the span of the basis functions. The minmax value $V_I(\mathcal{Z})$ has a natural interpretation and is the answer to the following question: "if yields change by at most one percentage point, what is the largest percentage point decline in the portfolio value?" The maxmin formula (3.9) provides an exact characterization of the worst-case outcome, and the number $V_I(\mathcal{Z})$ can be solved as the minmax value (3.6). Second, the error estimate (3.10) shows that the solution $z^* \in \mathcal{Z}$ to the minmax problem (3.6) achieves the lower bound in (3.9), to the first order. In this sense z^* is an optimal portfolio, which we refer to as the *robust immunizing portfolio*. Clearly, this immunizing portfolio is independent of $\Delta > 0$ as the minmax problem (3.6) does not involve Δ. Third, because the second order term in (3.10) is proportional to $1 + \|\theta\|_1$, leverage can negatively affect the immunization performance.

3.2 Relation to existing literature

In this section we discuss in some detail how Theorem 1 is related to the existing literature. The following corollary shows that when $I = J - 1$ and there is no portfolio constraint beyond value matching, the immunizing portfolio can be solved explicitly.

Corollary 2 (Robust immunization with $I = J - 1$)**.** *Let everything be as in Proposition 3.1 and suppose that the only portfolio constraint is value matching* (2.9)*, so the set of admissible portfolios is*

$$\mathcal{Z}_0 := \left\{ z \in \mathbb{R}^J : P(x) = \sum_{j=1}^J z_j P_j(x) \right\}. \tag{3.11}$$

If $I = J - 1$ and the square matrix A_+ in (3.7) *is invertible, then the unique solution to* (3.6) *is $z^* = A_+^{-1} b_+$, with $V_I(\mathcal{Z}) = 0$.*

Proof. Immediate from the proof of Proposition 3.1. ∎

Remark 2. The special case of Corollary 2 with $I = J - 1 = 1$ and $h_1(t) = t$ reduces to classical immunization that matches the bond value and duration. To see this, recall that by the definition (2.2), the duration of the cash flow F equals the weighted average time to payment

$$D = \frac{\int_0^T t\mathrm{e}^{-ty(t)}\,\mathrm{d}F(t)}{\int_0^T \mathrm{e}^{-ty(t)}\,\mathrm{d}F(t)}.$$

Using the definition $x(t) = ty(t)$ and (3.1), the duration can be rewritten as

$$D = \frac{\int_0^T t\mathrm{e}^{-x(t)}\,\mathrm{d}F(t)}{\int_0^T \mathrm{e}^{-x(t)}\,\mathrm{d}F(t)} = -\frac{P'(x)h_1}{P(x)} = b_1,$$

where $h_1(t) = t$ and we have used (3.2). A similar calculation implies that the duration of the immunizing portfolio is

$$-\frac{\sum_{j=1}^J z_j P_j'(x)h_1}{\sum_{j=1}^J z_j P_j(x)} = -\frac{\sum_{j=1}^J z_j P_j'(x)h_1}{P(x)} = \sum_{j=1}^J a_{1j}z_j$$

using value matching (2.9) and (3.3). Therefore if $z = A_+^{-1}b_+$, so $A_+z = b_+$, the duration is matched. By the same argument, setting $I = J - 1$ and $h_i(t) = t^i$ reduces to high-order duration matching ($I = J - 1 = 2$ is convexity matching). If, instead, we use the basis functions corresponding to the factor loadings on the yield curve's level, slope, and curvature as described in Example 3, then $A_+z = b_+$ amounts to factor duration matching.

In addition to the setting in Corollary 2, if the liability pays out on a single date and the immunizing portfolio does not involve shortsales, we can obtain the following global result.

Proposition 3.3 (Guaranteed funding)**.** *Let everything be as in Corollary 2 and suppose that the liability pays out on a single date. If* $z^* = A_+^{-1}b_+ \geq 0$*, then for all* $h \in \operatorname{span}\{h_i\}_{i=1}^I$ *we have* $E(z^*, x + h) \geq 0$*.*

Remark 3. Our maxmin result (Theorem 1) is quite different from the existing literature such as Fisher and Weil (1971) and Bierwag and Khang (1979). To the best of our knowledge, in this literature it is always assumed that the liability pays out on a single date and the portfolio does not involve shortsales ($z \geq 0$), yet this constraint is implicitly assumed not to bind. Under these assumptions, Proposition 3.3 shows that the immunizing portfolio always funds the liability, which generalizes the result of Fisher and Weil (1971) (who proved Proposition

3.3 for $I = J - 1 = 1$ and $h_1(t) = t$). However, this result is quite restrictive because liabilities are typically paid out over time and shortsales are essential when the maturity of the liability is very long (such as pensions). Our maxmin result (3.9) accommodates arbitrary liability structures and portfolio constraints. Furthermore, we allow the number of basis functions to exceed the number of bonds ($I \gg J - 1$), which can significantly improve hedging performance by making the portfolio more robust to perturbations in the yield curve.

3.3 Robust immunization with ℓ^p perturbations

The portfolio solution in (3.6) can be obtained using linear programming techniques. However, computing the solution can become slow when the number of basis functions is large, due to the high number of linear constraints. In addition, there is no closed-form expression for the portfolio, which makes it difficult to analyze how changes – such as adding basis functions – affect the solution.

In order to overcome these problems, we generalize the set of admissible perturbations which leads to a portfolio solution that can be obtained in closed form. So far we defined the set of admissible perturbations as $\mathcal{H}_I(\Delta)$ in (2.11). More generally, for measuring the magnitude of perturbations, for $p \in [1, \infty]$ we may use the ℓ^p norm on the payment dates $t = (t_1, \dots, t_N)$ defined by

$$\|h(t)/t\|_{p,t} := \begin{cases} \left(\sum_{n=1}^N |h(t_n)/t_n|^p\right)^{1/p}, & (p < \infty) \\ \max_n |h(t_n)/t_n|, & (p = \infty) \end{cases} \tag{3.12}$$

and define the set of admissible perturbations by

$$\mathcal{H}_I^p(\Delta) := \left\{ h \in \operatorname{span}\{h_i\}_{i=1}^I : \|h(t)/t\|_{p,t} \leq \Delta \right\}. \tag{3.13}$$

The case $p = \infty$ corresponds to what we treated in Section 3.1. With this generalization, we need to redefine the minmax problem (3.6) using

$$\mathcal{W}^p := \left\{ w \in \mathbb{R}^I : \|G'w\|_p \leq 1 \right\}, \tag{3.14}$$

$$V_I^p(\mathcal{Z}) := \inf_{z \in \mathcal{Z}} \sup_{w \in \mathcal{W}^p} \langle w, Az - b \rangle, \tag{3.15}$$

where $\|\cdot\|_p$ denotes the usual ℓ^p norm for vectors. Noting the equivalence of norms for finite-dimensional spaces, it is straightforward to generalize Theorem 1 in this setting, though the expression for the high-order term in the error estimate in

(3.10) needs to be modified appropriately.

The special case of $p = 2$ (Euclidean norm for perturbations) with linear constraints is particularly analytically tractable, as the following theorem shows.[7]

Theorem 3 (Robust immunization with ℓ^2 perturbations)**.** *Suppose Assumptions 1–3 hold and the portfolio constraint is given by* $\mathcal{Z} = \left\{z \in \mathbb{R}^J : Rz = r\right\}$*, where* $R \in \mathbb{R}^{M\times J}$ *has full row rank and* $r \in \mathbb{R}^M$*. If* $p = 2$*,* A *has full column rank, and* $\tilde{A} := (GG')^{-1}A$*, then the unique solution to the minmax problem (3.15) is*

$$z = (\tilde{A}'A)^{-1}\tilde{A}'b + (\tilde{A}'A)^{-1}R'[R(\tilde{A}'A)^{-1}R']^{-1}(r - R(\tilde{A}'A)^{-1}\tilde{A}'b), \tag{3.16}$$

where the inverses all exist.

The portfolio solution can also be viewed as a projection of the liability sensitivity vector b on the bond sensitivity matrix A. Specifically, consider the model

$$b = Az + \varepsilon \qquad \varepsilon|A \sim \mathsf{N}\left(0, GG'\right),$$

where $\mathsf{N}\left(\mu, \Sigma\right)$ denotes the multivariate normal distribution with mean μ and variance-covariance matrix Σ. Using maximum likelihood estimation for the constrained problem renders the identical solution as (3.16). Hence, the maxmin portfolio (3.16) also corresponds to a constrained generalized least squares (GLS) solution.

3.4 Robust immunization as a regularized HD portfolio

As is well known, high-order duration matching ($I = J-1$, no portfolio constraint, and $h_i(t) = t^i$) does not necessarily have a good performance due to extreme leverage (Mantilla-Garcia et al., 2022). In the portfolio literature, a useful remedy to extreme portfolio weights is to apply shrinkage methods (Ledoit and Wolf, 2003; Kozak et al., 2020). Here we show that overidentification of the portfolio ($I > J - 1$) can also be interpreted as a form of regularization, leading to less extreme portfolio weights and providing a theoretical explanation for why the robust immunization portfolio may outperform high-order duration matching.

It turns out that there is a close connection between high-order duration matching (Corollary 2) and robust immunization with ℓ^2 perturbations (Theorem 3). To spell this out in more detail we introduce some terminology. For any matrix

[7] In Section 3.5, we show why it can sometimes be beneficial to impose multiple linear constraints in the portfolio solution.

$A \in \mathbb{R}^{I\times J}$ with $\operatorname{rank}(A) = J$, define $A(s) \in \mathbb{R}^{J\times J}$ to be the matrix A with row index in $s \subset \{1, \ldots, I\}$ and $|s| = J$. Similarly, let $b(s) \in \mathbb{R}^J$ denote the vector b with row index in s. We call $z(s) = A(s)^{-1}b(s) \in \mathbb{R}^J$ an elemental estimate. Jacobi (1841) showed that the ordinary least squares estimator can be expressed as the expected value of the elemental estimates, with the probability measure given by $\Pr(S = s) = \det[A_+(s)]^2 / \sum_s \det[A_+(s)]^2$ (see also Knight (2019)). In Appendix B.6 we generalize these results to the constrained regression case. Applied to our context, the result is as follows.

Proposition 3.4. *Let $A_{HD} \in \mathbb{R}^{(J-1)\times J}$ be the bond sensitivity matrix corresponding to high-order duration matching. Let $A \in \mathbb{R}^{I\times J}$ ($I > J-1$) be the bond sensitivity matrix of the robust immunization portfolio in Theorem 3, which can be partitioned as $A' = [A'_{HD},\ A(\{J, \ldots, I\})']$, where $A(\{J, \ldots, I\}) \in \mathbb{R}^{(I-J+1)\times J}$. Define $\tilde{A} := (GG')^{-1}A$, and let*

$$A_+ = \begin{bmatrix} a_0 \\ A \end{bmatrix},\ \tilde{A}_+ = \begin{bmatrix} a_0 \\ \tilde{A} \end{bmatrix},\ b_+ = \begin{bmatrix} 1 \\ b \end{bmatrix}.$$

Then the solution to the minmax problem in (3.16) with a value matching constraint can be expressed as

$$z = \frac{\sum_{1\subset s} \det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right] z(s)}{\sum_{1\subset s} \det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right]} =: \sum_{[M]\subset s} \lambda(s) z(s), \tag{3.17}$$

where $z(s) = A_+^{-1}(s) b_+(s)$, $\sum_{1\subset s}$ denotes the sum over all subsets s of cardinality J that contain 1, and

$$\lambda(s) = \frac{\det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right]}{\sum_{1\subset s} \det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right]} \quad \textit{with} \sum_{1\subset s} \lambda(s) = 1.$$

One of the summands in (3.17) includes the set $s = \{1, \ldots, J\}$. In that case, $z(s)$ is given by the `HD` solution in Corollary 2. The weight given to the `HD` solution is proportional to $\det(A_+(s))$. Hence, if the sensitivity matrix of the `HD` solution is close to singular, it receives little weight in the robust immunization portfolio. In this sense robust immunization with $I > J-1$ helps to regularize the portfolio solution, since extreme portfolio weights or leverage induced by a close to singular sensitivity matrix have less bearing on the portfolio solution in (3.17).

Furthermore, the solution in (3.17) is a linear combination of different maxmin solutions, since by Proposition 3.1 and Theorem 1, $z(s)$ solves a minmax problem for every s with $V_{|s|}(\mathcal{Z}) = 0$, where different basis functions correspond to the rows s. This decomposition sheds light on the choice of basis functions. Consider the robust immunization portfolio with $I = J$, and compare this to the HD solution. If the perturbation to the yield curve is solely due to the Ith basis function, it follows from (3.17) that the robust immunization portfolio always does better than HD, provided the weight given to the HD portfolio is between zero and one. On the other hand, if the perturbation is a weighted average of several basis functions, it can be that the HD portfolio outperforms robust immunization. The choice of basis functions thus hinges on how close they are to spanning the shock and on how much each basis function contributes to the shock. In Figure 3 below, we empirically find that the first 10 basis functions strike a good balance.

3.5 Robust immunization with principal components

So far we have put no structure on the basis functions $\{h_i\}_{i=1}^I$ beyond Assumption 3. The set of admissible perturbations (2.11) depends only on span $\{h_i\}_{i=1}^I$ (Proposition 3.2) and the particular order or parameterization does not matter. However, in practice there could be some factor structure in the yield curve. For instance, a typical shift to the yield curve might be decomposed into the sum of a parallel shift and a nonparallel shift of a smaller size. In fact, according to Figure 3 below, the first few basis functions explain a large fraction of variations in the yield curve changes. Therefore it could be important to account for explanatory power of different basis functions in constructing the robust immunizing portfolio.

We formalize this idea and extend Theorem 1 to a setting in which the perturbation in a particular direction (principal component) could be larger. For any $\Delta_1, \Delta_2 > 0$, consider the following admissible set of perturbations:

$$\begin{aligned}&\mathcal{H}_I(\Delta_1, \Delta_2)\\ &:= \left\{ h \in \operatorname{span} \{h_i\}_{i=1}^I : (\exists \alpha)(\forall n) \, |\alpha h_1(t_n)/t_n| \le \Delta_1, |h(t_n)/t_n - \alpha h_1(t_n)/t_n| \le \Delta_2 \right\}.\end{aligned} \tag{3.18}$$

Choosing $h \in \mathcal{H}_I(\Delta_1, \Delta_2)$ amounts to perturbing the yield curve in the direction spanned by the first component $(h_1(t)/t)$ by a magnitude at most Δ_1, and then perturbing in an arbitrary direction spanned by the first I basis functions by a magnitude at most Δ_2. Thus setting $\Delta_1 \gg \Delta_2$ captures the idea that h_1 is the first principal component. In this setting, we can generalize Theorem 1 as follows.

Theorem 4 (Robust immunization with principal components)**.** *Let everything be as in Proposition 3.1 and suppose the set*

$$\mathcal{Z}_1 := \left\{ z \in \mathcal{Z} : \sum_{j=1}^{J} a_{1j} z_j = b_1 \right\} \tag{3.19}$$

is nonempty, where a_{1j} and b_1 are defined by (3.3) and (3.2) with $i = 1$. Let $\mathcal{H}_I(\Delta_1, \Delta_2)$ be as in (3.18). Then the following statements are true.

(i) The guaranteed equity satisfies

$$\lim \frac{1}{\Delta_2} \sup_{z \in \mathcal{Z}_1} \inf_{h \in \mathcal{H}_I(\Delta_1,\Delta_2)} E(z, x+h) = -P(x) V_I(\mathcal{Z}_1), \tag{3.20}$$

where the limit is taken over $\Delta_1, \Delta_2 \to 0$, $\Delta_1/\Delta_2 \to \infty$, and $\Delta_1^2/\Delta_2 \to 0$.

(ii) Letting $z^* \in \mathcal{Z}_1$ be the solution to the minmax problem (3.6) with portfolio constraint $\mathcal{Z}_1$, we have

$$\sup_{h \in \mathcal{H}_I(\Delta_1,\Delta_2)} |E(z^*, x+h)| \le \Delta_2 P(x) \left(V_I(\mathcal{Z}_1) + O(\Delta_2 + \Delta_1^2/\Delta_2) \right). \tag{3.21}$$

Imposing the portfolio constraint $\mathcal{Z}_1 \subset \mathcal{Z}$ may improve or worsen the performance. To explain why, we first present the following simple result.

Proposition 3.5 (Monotonicity of minmax value)**.** *Let everything be as in Proposition 3.1. If $I < I'$ and $\mathcal{Z} \subset \mathcal{Z}'$, then $V_I(\mathcal{Z}) \le V_{I'}(\mathcal{Z})$ and $V_I(\mathcal{Z}) \ge V_I(\mathcal{Z}')$.*

The claim $V_I(\mathcal{Z}) \le V_{I'}(\mathcal{Z})$ is obvious because the more basis functions we include, the more freedom nature has to select adversarial perturbations. The claim $V_I(\mathcal{Z}) \ge V_I(\mathcal{Z}')$ is also obvious because the larger the set of admissible portfolios is, the more freedom the fund manager has to select portfolios.

Comparing to (3.18) to (2.11) and applying the triangle inequality

$$|h(t)/t| \le \underbrace{|\alpha h_1(t)/t|}_{\le \Delta_1} + \underbrace{|h(t)/t - \alpha h_1(t)/t|}_{\le \Delta_2},$$

we obtain $\mathcal{H}_I(\Delta_1, \Delta_2) \subset \mathcal{H}_I(\Delta_1 + \Delta_2)$. Therefore to the first order, the maximum portfolio return loss can be bounded as

$$\underbrace{\Delta_2 V_I(\mathcal{Z}_1)}_{\text{Theorem 4}} \le (\Delta_1 + \Delta_2) V_I(\mathcal{Z}_1) \ge \underbrace{(\Delta_1 + \Delta_2) V_I(\mathcal{Z})}_{\text{Theorem 1}},$$

where the right inequality follows from Proposition 3.5. Thus if $\Delta_1 \gg \Delta_2$ in typical situations (which we document in Figure 3), then imposing the constraint $\mathcal{Z}_1$ in (3.19) improves the performance because the loss in the minmax value $V_I(\mathcal{Z}_1) \geq V_I(\mathcal{Z})$ from imposing the constraint is compensated by the gain in the coefficient $\Delta_2 \ll \Delta_1 + \Delta_2$. However, if it so happens that $\Delta_1 \sim \Delta_2$, then imposing the constraint $\mathcal{Z}_1$ worsens the performance.

This result sheds further light on the sometimes poor performance of high-order duration matching ($I = J - 1$, no portfolio constraint, and $h_i(t) = t^i$). First, as we increase I while setting $I = J - 1$, both the span of basis functions and the set of admissible portfolios $\mathcal{Z}$ expand. Because increasing I makes $V_I(\mathcal{Z})$ larger but expanding $\mathcal{Z}$ makes it smaller, the combined effect could go either way. This observation explains the poor performance of high-order duration matching. Second, in the setting of Theorem 4, if $\Delta_1 \sim \Delta_2$, then imposing the constraint $\mathcal{Z}_1$ worsens the performance. To see why, by Proposition 3.5 we have $V_I(\mathcal{Z}_1) \geq V_I(\mathcal{Z})$, so if $\Delta_1 \sim \Delta_2$, then

$$\underbrace{\Delta_2 V_I(\mathcal{Z}_1)}_{\text{Theorem 4}} > \underbrace{(\Delta_1 + \Delta_2) V_I(\mathcal{Z})}_{\text{Theorem 1}}.$$

Remark 4. Theorem 4 can be further generalized if we allow larger perturbations spanned by the first few basis functions. For instance, if we use the first two basis functions, we can define $\mathcal{H}_I(\Delta_1, \Delta_2, \Delta_3)$ analogously to (3.18) by incorporating the constraints $|\alpha_i h_i(t_n)/t_n| \leq \Delta_i$ for $i = 1, 2$ and

$$|h(t_n)/t_n - \alpha_1 h_1(t_n)/t_n - \alpha_2 h_2(t_n)/t_n| \leq \Delta_3.$$

The portfolio constraint (3.19) then becomes

$$\mathcal{Z}_2 := \left\{ z \in \mathcal{Z} : \sum_{j=1}^{J} a_{ij} z_j = b_i \text{ for } i = 1, 2 \right\}, \tag{3.22}$$

and the maxmin formula (3.20) involves $V_I(\mathcal{Z}_2)$.

By a similar argument as above, imposing the constraint $\mathcal{Z}_2$ improves the performance relative to $\mathcal{Z}_1$ if $\Delta_2 \gg \Delta_3$, whereas it worsens the performance if $\Delta_2 \sim \Delta_3$.

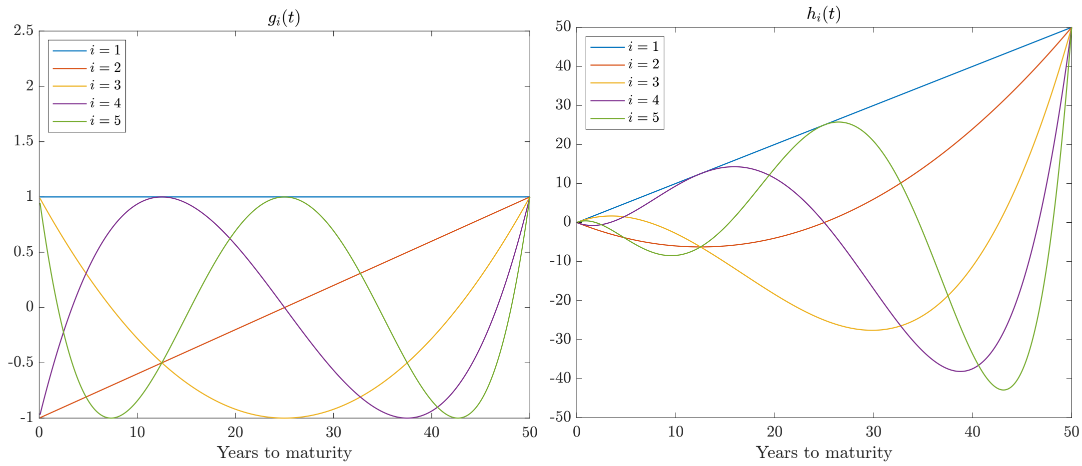
3.6 Implementation

To implement robust immunization, we need to choose the basis functions $\{h_i\}_{i=1}^{I}$. Although the conclusion of Theorem 1 holds regardless of the choice of the basis functions, here we propose a particular choice.

For each i, it is natural to choose h_i such that h_i is a polynomial of degree i with $h_i(0) = 0$, for Assumption 3 then holds (Lemma A.2). By basis invariance (Proposition 3.2), any choice of such a basis will result in the same immunizing portfolio. However, we suggest using Chebyshev polynomials because they enjoy good numerical properties (Trefethen, 2019, Ch. 2–4). To be more specific, let $T_n : [-1, 1] \to \mathbb{R}$ be the n-degree Chebyshev polynomial defined by $T_n(\cos\theta) = \cos n\theta$ and setting $x = \cos\theta$. We map $[0, T]$ to $[-1, 1]$ using the affine transformation $t \mapsto x = 2t/T - 1$, and define $g_i : [0, T] \to \mathbb{R}$ by

$$g_i(t) = T_{i-1}(2t/T - 1) \tag{3.23}$$

so that we can allow any (continuous) perturbation to the yield curve for $t \in [0, T]$. Then we can define the basis functions for perturbing the cumulative discount rate by $h_i(t) = tg_i(t)$. Figure 2a shows the graphs of g_i in (3.23) for a maturity $T = 50$ years, which are the rows of the matrix G in Proposition 3.1. Figure 2b shows the graphs of the basis functions $h_i(t)$.



(a) Yield curve. (b) Cumulative discount rate.

Figure 2: Basis functions of robust immunization.

Because Theorem 1 takes the number of basis functions I as given, a natural question is how to select it. Choosing a small I restricts the space of admissible perturbations and may lead to non-robustness against model misspecification. To address this concern, we evaluate the goodness-of-fit of approximating discount rate changes by basis functions. For this purpose, we use the daily yield curve data described in Section 4. Let I be the number of basis functions to include, d the number of days ahead, and $\{t_n\}_{n=1}^N$ the set of terms (in years) to evaluate the cumulative discount rates, where we set $t_n = n/12$ and $N = 360$ so payouts correspond to a 30-year horizon at a monthly interval. Let $y_s(t)$ be the yield curve on day s at maturity t. We use the following procedure.

(i) For each day s and term t_n, calculate the d-day ahead change in the yield curve $y_{s+d}(t_n) - y_s(t_n)$.

(ii) For each s and d, estimate

$$y_{s+d}(t_n) - y_s(t_n) = \sum_{i=1}^{I} \gamma_{isd} g_i(t_n) + \epsilon_{sd}(t_n), \quad n = 1, \ldots, N \tag{3.24}$$

by ordinary least squares (OLS), where g_i is as in (3.23).

(iii) For each basis function i, decompose the R^2 of the regression (3.24) using the Shapley value (see Huettner and Sunder (2012)):[8]

$$R_i^2 = \frac{1}{k} \sum_{S \setminus \{i\}} \binom{I-1}{|S|}^{-1} \left(R^2(S \cup \{i\}) - R^2(S) \right).$$

The sum is taken over all subsets S of $\{1, \ldots, I\}$ that do not include basis function g_i. To calculate $R^2(S)$ for a model that includes basis functions in S, we use 0 as the benchmark instead of the sample mean, since $g_1 \equiv 1$ is already a constant function. As a result, R_i^2 measures the explanatory power of basis function i relative to the other basis functions.

(iv) Let $\{\widehat{\gamma}_{isd}\}_{i=1}^{I}$ be the OLS estimator, calculate the overall goodness-of-fit measure

$$R_d^2 := \frac{\sum_{s=1}^{S} \sum_{n=1}^{N} \left(\sum_{i=1}^{I} \widehat{\gamma}_{isd} g_i(t_n) \right)^2}{\sum_{s=1}^{S} \sum_{n=1}^{N} (y_{s+d}(t_n) - y_s(t_n))^2}. \tag{3.25}$$

The left panel of Figure 3 shows the Shapley decomposition of the R^2 with $I = 6$ basis functions. The Shapley values are averaged across all dates in our sample. The explanatory power of each basis function is roughly constant across different horizons d. The first basis function (constant) explains around 60% of variations in the yield curve changes. In more than 91% of time periods, the R^2 is above 95%. Furthermore, most of the explanatory power is contributed by the first, second and third basis functions, while the other basis functions generally

[8]The Shapley value of the R^2 has some desirable properties, such as efficiency and monotonicty (Huettner and Sunder, 2012). Furthermore, the order of the regressors is irrelevant.

contribute less than 5%. The right panel shows the unexplained component $1-R_d^2$ as we include more basis functions. We can see that setting $I = 10$ captures about 99.88% ($1 - R_d^2 \sim 10^{-3}$) of variations in the yield curve changes.

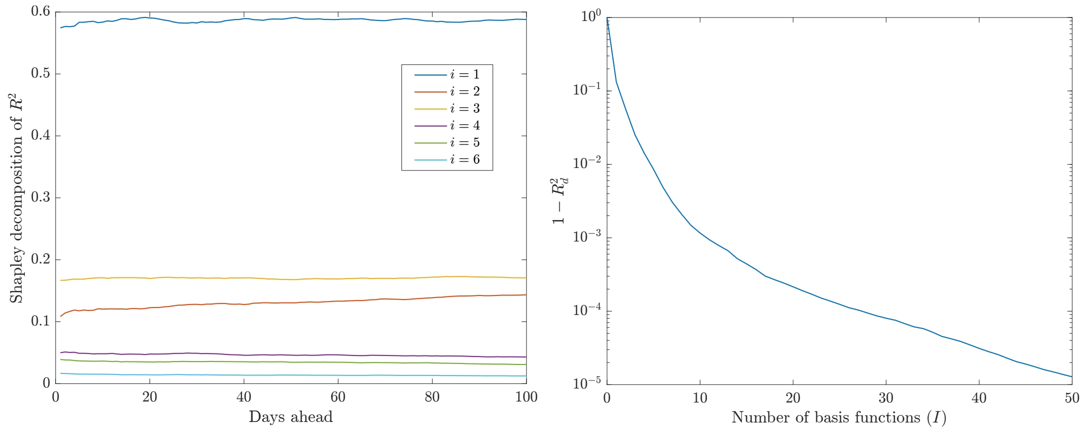


Figure 3: Goodness-of-fit of yield curve change approximation.

Note: The left panel shows the decomposed R^2 using the Shapley value corresponding to regression (3.24) with $I = 6$ basis functions. The Shapley values are averaged across all dates in the sample. The right panel shows the combined $1 - R_d^2$ as we increase the number of basis functions I. See Section 4 for data description.

We now describe the algorithm to implement robust immunization in practice. Although the underlying theory (which heavily relies on functional and numerical analysis) may not be familiar to practitioners, the implementation requires little more than basic linear algebra and linear programming.

Robust Immunization.

(i) Let $\mathbf{t} = (t_1, \ldots, t_N)$ be the $1 \times N$ vector of asset/liability payout dates and $T = t_N$ be the planning horizon. Let $\mathbf{y} = (y_1, \ldots, y_N)$ be the $1 \times N$ vector of yields, $\mathbf{f} = (f_1, \ldots, f_N)$ the $1 \times N$ vector of liabilities, and $\mathbf{F} = (f_{jn})$ the $J \times N$ matrix of bond payouts.

(ii) Let $I \geq J-1$, define the basis functions by (3.23), evaluate at each t_n, and construct the $I \times N$ matrix of basis functions $\mathbf{H} = (t_n g_i(t_n)) = (h_i(t_n))$ and $\mathbf{G} = (g_i(t_n))$. Define the $1 \times N$ vector of zero-coupon bond prices $\mathbf{p} = \exp(-\mathbf{y} \odot \mathbf{t})$, where $\odot$ denotes entry-wise multiplication (Hadamard product).

(iii) Define the $I \times J$ matrix A, $I \times 1$ vector b, and $1 \times J$ vector a_0 by

$$A := (\mathbf{H}\,\mathrm{diag}(\mathbf{p})\mathbf{F}')/(\mathbf{p}\mathbf{f}'), \quad b := \mathbf{H}\,\mathrm{diag}(\mathbf{p})\mathbf{f}'/(\mathbf{p}\mathbf{f}'), \quad a_0 := \mathbf{p}\mathbf{F}'/(\mathbf{p}\mathbf{f}'),$$

where $\mathrm{diag}(\mathbf{p})$ denotes the diagonal matrix with diagonal entries given by $\mathbf{p}$. Define the $(I+1) \times J$ matrix A_+ and $(I+1) \times 1$ vector b_+ by

$$A_+ := \begin{bmatrix} a_0 \\ A \end{bmatrix} \quad \text{and} \quad b_+ := \begin{bmatrix} 1 \\ b \end{bmatrix}.$$

(iv) If $I = J - 1$ and there are no portfolio constraints, calculate the immunizing portfolio as $z^* = A_+^{-1} b_+$. Otherwise, numerically solve the minmax problem (3.6) (if using the ℓ^∞ norm) or use (3.16) (if using the ℓ^2 norm and constraints are linear). Nonlinear constraints, such as short-sale restrictions, can be easily incorporated. The resulting optimization problem can be solved using quadratic programming.

Note that the inner maximization in (3.6) is a linear programming problem with I variables and $2N$ inequality constraints, which is straightforward to solve numerically even when N is large (a few hundred in typical applications). The outer minimization is a convex minimization problem with J variables, which is also straightforward to solve numerically.[9]

4 Evaluation: static and dynamic hedging

In this section we evaluate the performance of robust immunization and other existing methods using a numerical experiment in static and dynamic settings.

4.1 Data and yield curve model

We obtain daily U.S. Treasury nominal yield curve data from November 25, 1985 to December 2023 from Liu and Wu (2021).[10] These yield curves are estimated using a non-parametric method that accommodates general yield curve perturbations. This is important because more complex perturbations require a greater number

[9]In simulations, however, repeatedly solving the linear programming problem can be computationally expensive.

[10]https://sites.google.com/view/jingcynthiawu/yield-data

of basis functions, which our approach can handle, unlike high-order duration matching. We index the dates by $s = 1, \ldots, S$, where $S = 9{,}526$ is the sample length.

Remark 5. The estimated yields of Liu and Wu (2021) go back all the way to 1961, but we only use their data beyond 11/25/1985 when bonds with a maturity of 30 years were introduced in the market. Gürkaynak et al. (2007) caution against extrapolation of the yield curve beyond the maximum available bond maturity. Anticipating our empirical application, we need to obtain yields with maturity up to 50 years. Since extrapolation is still necessary in this case, we extrapolate the forward rate by a constant beyond the 30-year maturity. This approach is motivated by no-arbitrage arguments which stipulate that the long term forward rate is constant (Dybvig et al., 1996). In Appendix D.1, we show how the constant forward rate assumption affects our estimate of the yield curve.

The observed yield curve data is only one sample and thus inadequate for evaluating the performance in a dynamic hedging experiment. Therefore in addition to the observed yield curve data, we also use simulated yield curves generated from a no-arbitrage term structure model. Specifically, we apply the Ang et al. (2008) 3-factor regime switching model. By simulating yields from this stationary regime-switching model, we can evaluate the performance of various immunization methods under a wide variety of yield curves. See Appendix C for details.

4.2 Cash flow and immunization methods

We consider several cash flow schemes for the liability, assuming a fixed time horizon of $T = 50$ years. The specifications are as follows: (i) a constant cash flow of 1 each month throughout the horizon (`fullHorizon`); (ii) a constant cash flow of 1 after year 20 (`longRun`); (iii) a constant cash flow of 1 between years 15 and 35 (`medium`); (iv) a cash flow of 1 between years 1 and 15, and again between years 35 and 50 (`shortAndLong`). In all cases, we normalize the cash flows so that their cumulative sum equals 1. The bonds available for trade are zero-coupon bonds with face value 1 and years to maturity being $\{1, 2, 5, 10, 30\}$. We intentionally choose a long maturity of 50 years for the cash flows because it is of interest to study how the yield curve at the long end affects the performance of the immunization methods.

We consider three immunization methods. The first method is high-order duration matching (`HD`) explained in Remark 2, which is a special case of robust immunization by setting $I = J - 1$ and $h_i(t) = t^i$. By basis invariance (Proposition

3.2), we can choose any polynomial basis, so we use the Chebyshev polynomial basis in (3.23) with $T = 50$. The second method is key rate duration matching (KRD) proposed by Ho (1992) and explained in Appendix D.2. In short, this method is designed to match the liability and portfolio sensitivity to interest rate changes at pre-specified maturities. The third method is our proposed robust immunization method (RI) with the Chebyshev polynomial basis for the yield curve in (3.23). Motivated by the right panel of Figure 3, we set the number of basis functions to $I = 10$. For the portfolio constraint, motivated by Theorem 4 and the left panel of Figure 3, we consider value matching only ($\mathcal{Z}_0$ in (3.11)), value- and duration matching ($\mathcal{Z}_1$ in (3.19)), and value-, duration- and convexity matching. We denote these methods by RI(0), RI(1), RI(2). Throughout the simulation, we report results only for the robust immunization portfolio based on the ℓ^2 norm (see (3.16)), as it is computationally much faster than the portfolio based on the ℓ^∞ norm (see (3.6)). Unreported simulations using the ℓ^∞ norm yield similar results.

4.3 Static hedging

Suppose that on date s, the fund manager immunizes future cash flows with a bond portfolio $z_s = (z_{sj})$ constructed by the HD, KRD, and RI methods. Letting x_s be the cumulative discount rate on date s and h be a perturbation, we evaluate the performance of each method using the *funding ratio* defined by

$$\varphi_s(h) := \frac{1}{P(x_s + h)} \sum_{j=1}^{J} z_{sj} P_j(x_s + h). \tag{4.1}$$

We suppose that the fund manager is worried about underfunding, so we further define the "underfunding ratio" by

$$1 - \min \{\varphi_s(h), 1\}. \tag{4.2}$$

A higher underfunding ratio makes it less likely that a fund will cover its liabilities. As we are interested in realistic yield changes and portfolio holding periods, we let the perturbation h to be the change in the cumulative discount rates from date s to $s + d$ for $d = 1, \ldots, 100$ days.

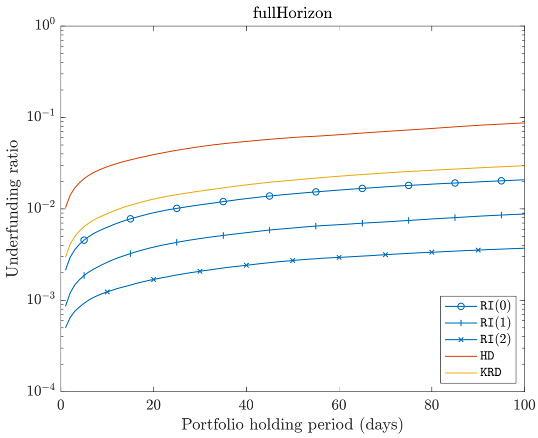
Figure 4 shows the underfunding ratio (4.2) averaged over the sample period. The performance worsens with longer portfolio holding periods (d) for all different liabilities and immunization methods because of greater yield curve fluctuations.

In all cases, RI(2) outperforms the other methods, followed by RI(1) and RI(0). This suggests that the robust immunization methodology provides a better hedge than existing approaches, and that adding sensitivity constraints – such as duration or convexity matching – tends to improve performance. In three out of the four cases shown in Figure 4, the classical HD method performs notably worse than the competing methods, partly due to the excessive leverage in the resulting portfolios. Leverage tends to exacerbate hedging errors (Mantilla-Garcia et al., 2022), and Table 2 shows that leverage under classical HD matching is orders of magnitude higher than in the other methods. We also observe that leverage tends to increase when additional constraints are imposed on the robust immunization portfolio. Section 3.4 explains why: RI portfolios are constructed as averages over different HD portfolios, which naturally reduces leverage. Imposing more constraints in the RI portfolio implies that the average is taken over fewer HD portfolios, which tends to increase leverage. The fact that, in most cases, all portfolios require leverage highlights the importance of allowing short sales in Theorems 1 and 3.

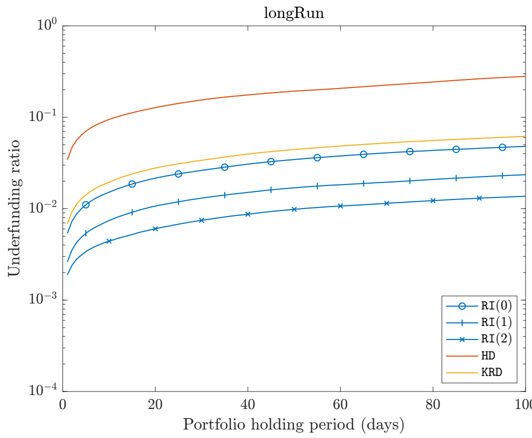
Figure 5 shows the portfolio weights over time for the RI(2) and HD portfolios corresponding to the fullHorizon liability. The difference in leverage is striking, especially on the three dates when both portfolios exhibit their highest leverage (indicated in black). On each of these dates, which correspond to peaks in the 2008 financial crisis and the COVID-19 crisis, yields fell and the yield curve flattened amid a "flight to safety".[11] During such periods of market distress, the classical HD portfolio shows extreme leverage, which compromises its performance. In contrast, the RI(2) portfolio remains only mildly levered. The extreme leverage of the HD portfolio can be explained using numerical linear algebra: when yields are all close to zero – as is typical in a flight to safety – the HD sensitivity matrix approaches a Vandermonde matrix, which is known to be ill-conditioned (Beckermann, 2000).

Figure 4 reports only average underfunding ratios. To assess each method's performance under adverse conditions, Table 1 presents the 90th, 95th, and 99th percentiles of the underfunding ratio (4.2) over a 30-day holding period. According to the table, the HD method performs comparably to the other approaches only when the liability pays out over the medium time horizon. The KRD method performs slightly better, but is still outperformed by the robust immunization portfolios. As before, RI(2) delivers the best performance, regardless of liability type.

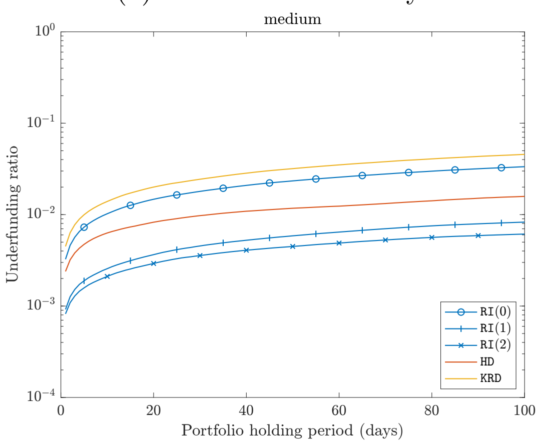
[11] https://www.chicagobooth.edu/review/how-treasury-yield-curve-reflects-worry



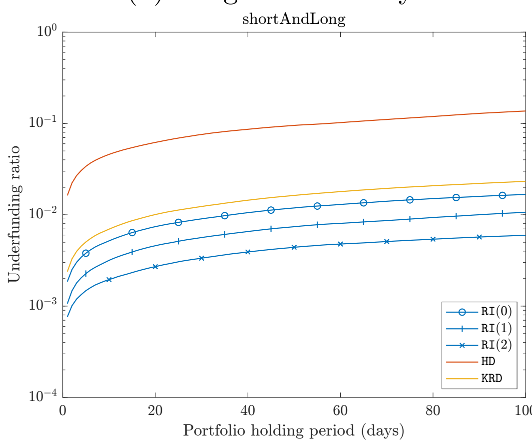
(a) Full horizon liability.



(b) Long term liability.



(c) Medium term liability.



(d) Short and long term liability.

Figure 4: Underfunding ratio for different holding periods.

Note: The figure shows the underfunding ratio (4.2) over various holding periods, averaged over the entire sample period. `RI`(0): robust immunization with a value matching; `RI`(1): robust immunization with value and duration matching; `RI`(2): robust immunization with value, duration, and convexity matching; `HD`: high-order duration matching; `KRD`: key rate duration matching. Each panel corresponds to different number of bonds J used to construct the immunizing portfolio.

4.4 Dynamic hedging

Although the static hedging experiment in Section 4.3 may be informative, it only addresses the performance of various immunization methods under a one-shot instantaneous change in the yield curve. In practice, the fund manager will rebalance the portfolio over time, in which case the yield curve as well as the bond maturities change. In this section, to evaluate the performance of various immunization methods under practical situations, we conduct a dynamic hedging experiment using simulated yield curves.

Let $\{s_n\}_{n=0}^N$ be the portfolio rebalancing dates (with the normalization $s_0 = 0$)

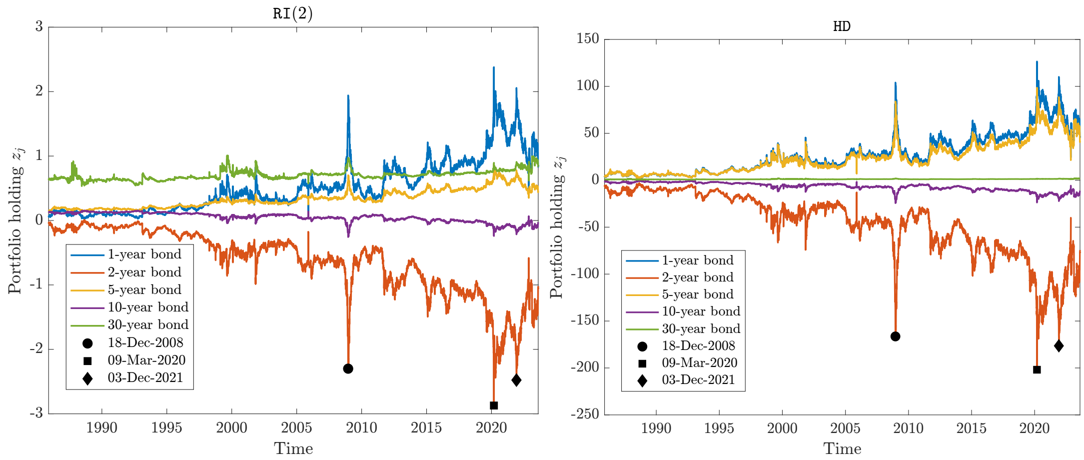


Figure 5: Portfolio weights over time.

Note: The figure shows portfolio weights for the `RI`(2) and `HD` portfolios in the static hedging experiment with `fullHorizon` liability. The black symbols indicate the dates on which leverage was most pronounced.

Table 1: Underfunding ratio (%) for 30-day holding period.

Method:	`RI`(0)	`RI`(1)	`RI`(2)	`HD`	`KRD`
90$^{\text{th}}$ percentile					
`fullHorizon`	3.20	1.31	0.54	13.34	4.50
`longRun`	7.72	3.66	2.03	43.21	10.17
`medium`	5.21	1.42	1.08	3.09	7.10
`shortAndLong`	2.53	1.52	0.89	21.22	3.47
95$^{\text{th}}$ percentile					
`fullHorizon`	4.42	2.03	0.98	22.28	6.04
`longRun`	10.60	6.07	3.35	69.95	13.62
`medium`	6.83	1.89	1.45	4.40	9.41
`shortAndLong`	3.80	2.53	1.62	35.53	4.95
99$^{\text{th}}$ percentile					
`fullHorizon`	7.97	4.71	2.94	55.35	10.74
`longRun`	18.20	13.85	10.75	150.25	22.88
`medium`	11.34	3.15	2.68	7.17	15.81
`shortAndLong`	8.57	6.64	4.33	89.64	9.39

Note: This table shows the quantiles of the underfunding ratio (4.2). See Figure 4 caption for explanation of methods.

and assume that the coupon payment dates of the liability are contained in this set. For simplicity let $s_n = n\Delta$ with $\Delta > 0$ so the dates are evenly spaced, although this is inessential. The liability pays $f_s \geq 0$ at time $s > 0$. The fund manager can use J zero-coupon bonds with face value 1 and maturities $\{t_j\}_{j=1}^{J}$ to

Table 2: ℓ^1 norm of investment shares.

Liability Type	RI(0)	RI(1)	RI(2)	HD	KRD
Median					
fullHorizon	1.21	3.07	3.65	213.47	1.00
longRun	6.57	4.52	14.99	709.70	1.00
medium	3.55	3.06	2.51	42.38	1.00
shortAndLong	1.01	3.08	4.35	331.46	1.00
95^{th} percentile					
fullHorizon	1.77	4.03	6.71	426.72	1.00
longRun	7.43	4.93	17.89	985.75	1.00
medium	4.03	3.63	2.63	53.62	1.00
shortAndLong	1.02	4.86	9.32	698.41	1.00
99^{th} percentile					
fullHorizon	1.88	4.25	7.61	489.13	1.00
longRun	7.68	5.50	19.23	1119.07	1.00
medium	4.10	3.71	2.74	56.00	1.00
shortAndLong	1.02	5.30	10.73	800.39	1.00

Note: This table shows the ℓ^1 norm of the investment shares, $\|\theta\|_1$. See Figure 4 caption for explanation of methods.

hedge the liability. We introduce the following notations:

$$\begin{aligned}
x_s(t) &= \text{cumulative discount rate for term } t \text{ at time } s, \\
P_s &= \text{present value of liability at time } s, \\
V_s &= \text{net asset value (NAV) of fund at time } s, \\
z_s = (z_{sj}) &= \text{immunizing portfolio at time } s, \\
C_s &= \text{cash position at time } s, \\
R_s &= \text{gross short rate at time } s.
\end{aligned}$$

We now describe how to calculate these quantities recursively. At time s, the present value of the liability (after coupon payment) is

$$P_s := \sum_{n:s_n>s} \mathrm{e}^{-x_s(s_n-s)} f_{s_n}.$$

Note that at time s, the remaining term of the n-th payment is $s_n - s$ and we only retain future payments in the sum. Let $s^- = s - \Delta$ denote the previous rebalancing period. The NAV of the fund consists of the present value of the bond and cash positions carried over from the previous period minus the current liability payment, which is

$$V_s := \underbrace{R_{s^-} C_{s^-}}_{\text{cash}} + \underbrace{\sum_{j=1}^{J} z_{s^- j} \mathrm{e}^{-x_s (t_j - \Delta)}}_{\text{bond}} - \underbrace{f_s}_{\text{liability}}.$$

Here, note that the cash position earns a (predetermined) gross return R_{s^-}, and the zero-coupon bonds have shorter maturities $t_j - \Delta$ because time has passed. The equity (asset minus liability) is therefore

$$\begin{aligned} E_s &:= V_s - P_s \\ &= R_{s^-} C_{s^-} + \sum_{j=1}^{J} z_{s^- j} \mathrm{e}^{-(x+h)(t_j - \Delta)} - f_s - \sum_{n: s_n > s} \mathrm{e}^{-(x+h)(s_n - s)} f_{s_n} \\ &= R_{s^-} C_{s^-} - f_s + \sum_{j=1}^{J} z_{s^- j} \mathrm{e}^{-(x+h)(t_j - \Delta)} - \sum_{n: s_n - \Delta - s^- > 0} \mathrm{e}^{-(x+h)(s_n - \Delta - s^-)} f_{s_n}, \end{aligned} \tag{4.3}$$

where $x = x_{s^-}$ denotes the cumulative discount rate at s^- and $h = x_s - x_{s^-}$ denotes the perturbation in the cumulative discount rate. As an illustration, consider the robust immunization method introduced in Section 3. The fund manager's problem at time s^- is to maximize the worst case equity, where the equity is defined by E_s in (4.3). Shifting s^- to s, the time s objective function is then

$$E_{s+\Delta}(z, x+h) := R_s C_s - f_{s+\Delta} + \sum_{j=1}^{J} z_{sj} \mathrm{e}^{-(x+h)(t_j - \Delta)} - \sum_{n: s_n - \Delta - s > 0} \mathrm{e}^{-(x+h)(s_n - \Delta - s)} f_{s_n},$$

where $x = x_s$ is the current cumulative discount rate. Because $f_{s+\Delta}$ is predetermined and C_s is determined by the budget constraint and hence independent of the perturbation h, the dynamic hedging problem reduces to the static hedging problem discussed in Section 3 except that *all payments need to be treated as if their maturities are reduced by* Δ. This modification takes into account the passage of time and hence the reduction in bond maturities by the next rebalancing date. For example, if the time to rebalancing is one quarter, a 1-year zero coupon bond is treated as if it is a 9-month bond.

Given the current cumulative discount rate x_s, it is straightforward to apply various immunizing methods to bonds and liability with maturities reduced by Δ.

Suppose the new (time s) immunizing portfolio $z_s = (z_{sj})$ is chosen. Then the cash position is the difference between the NAV and portfolio value, which is

$$C_s = V_s - \sum_{j=1}^{J} z_{sj} \mathrm{e}^{-x_s(t_j)}.$$

Note that although we reduce the maturities by Δ to form the portfolio, we use the actual maturities to evaluate the portfolio value and define the cash position. Initializing at $V_0 = P_0$ (100% funding), we can implement dynamic hedging by repeating this procedure for $s = \Delta, 2\Delta, \dots$. We evaluate the quality of the hedge at time s using the absolute return error

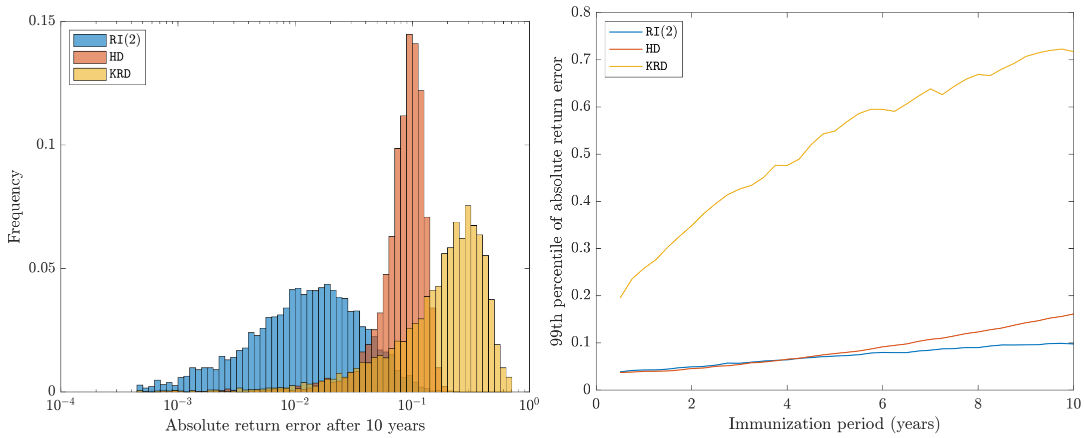
$$\frac{1}{P_{s^-}} |V_s - P_s| . \tag{4.4}$$

We implement the dynamic hedging approach using the `fullHorizon` liability and the same zero-coupon bonds as in the static problem. Among the robust portfolio methods, we focus solely on `RI`(2), as it performed best in the static case. Since the yield curve model of Ang et al. (2008) is estimated using quarterly data, we assume the immunizing portfolio is rebalanced every quarter. We evaluate performance over a 10-year horizon, repeating the simulation 5,000 times.

The results are summarized in Figure 6. The left panel shows the histogram of absolute return errors at the end of the 10-year period across all simulations.[12] Overall, it is clear that `RI`(2) is the superior method, since it has more mass in the left tail where the absolute return error is small. Also, the MSE is 11 times smaller compared to `HD`, which comes second best. The worst performing method is `KRD`, which has a MSE that is almost 100 times higher than `RI`(2), and is much more left skewed.

The right panel of Figure 6 sheds light on the maxmin property by showing the 99th percentile of the absolute return error for each method throughout the 10-year period across all simulations. Interestingly, `RI`(2) and `HD` are quite comparable according to this metric, especially during earlier years. Only after 6 years do we observe a noticeable difference between the two methods. Due to increased uncertainty, the percentiles are naturally increasing over time. `KRD` compares poorly to the other methods because of outliers in the right tail, especially at the end of the immunization period.

[12] I.e. the absolute return error in (4.4) evaluated at $s = 40$.



(a) Histogram of return error.

(b) 99th percentile of return error.

Figure 6: Distribution of absolute return error

The left panel shows the histogram of absolute return errors calculated at the end of the 10-year immunization period. The right panel shows the 99th percentile of the absolute return error throughout the 10-year immunization period, calculated across all 5,000 simulations.

5 Conclusion

This paper uses techniques from functional and numerical analysis to study the classical portfolio immunization problem. The goal is to construct a portfolio that protects a financial institution against interest rate risk. We use the concept of Fréchet derivatives to find a portfolio that hedges against general perturbations to the cumulative discount rate. Subsequently, we present a maxmin result that proves existence of an immunizing portfolio which maximizes the worst-case equity loss and we provide a solution algorithm. This maxmin portfolio, which we refer to as robust immunization, contains duration and convexity matching as a special case. In our empirical applications, we show that a judicious choice of basis functions for the discount rate leads to a robust immunization method that outperforms existing approaches in the static and dynamic case.

A Space of cumulative discount rates

Lemma A.1. *Let* $\mathcal{X} = \{x \in C[0,T] : x(0) = 0\}$ *be the vector space of continuous functions on* $[0,T]$ *with* $x(0) = 0$*. For* $x \in \mathcal{X}$*, define* $\|x\|_{\mathcal{X}} = \sup_{t\in[0,T]} |x(t)|$*. Then* $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ *is a Banach space.*

Proof. It is well known that the space $C[0,T]$ endowed with the supremum norm is a Banach space (e.g. Folland (1999, Chapter 5)). Let $L_0 x = x(0)$ be the evaluation functional at 0. Since L_0 is a continuous map and $\{0\}$ is closed, we get that

$L_0^{-1}\{0\} = \mathcal{X}$ is a closed set. Since a closed subspace of a Banach space is a Banach space, the proof is complete. $\square$

Lemma A.2 (Polynomial basis). *Suppose Assumption 1 holds and h_i is a polynomial of degree i with $h_i(0) = 0$. Then Assumption 3 holds.*

Proof. Since h_i is a polynomial of degree i with $h_i(0) = 0$, without loss of generality we may assume $h_i(t) = t^i$. By the Stone-Weierstrass theorem (Folland, 1999, p. 139), $\operatorname{span}\{h_i\}_{i=1}^{\infty}$ is dense in $\mathcal{X}$ since it separates the points and contains a non-zero constant function. By Assumption 1, we can choose I distinct points $\{t_{n_j}\}_{j=1}^{I}$. Consider the $I \times I$ submatrix of H defined by $\tilde{H} = (h_i(t_{n_j})) = (t_{n_j}^i)$. Dividing the j-th column by $t_{n_j} > 0$, $\tilde{H}$ reduces to a Vandermonde matrix, which is invertible. Therefore H has full row rank. The same argument applies to G. $\square$

Lemma A.3. *The Fréchet derivative $P'(x)$ defined by $h \mapsto \delta P(x;h)$ in (3.1) is a bounded linear operator.*

Proof. Clearly $P'(x)$ is a linear operator. If $h \in \mathcal{X}$, then

$$|P'(x)h| \le \int_0^T \mathrm{e}^{-x(t)} |h(t)| \,\mathrm{d}F(t) \le \|h\|_{\mathcal{X}} \int_0^T \mathrm{e}^{-x(t)} \,\mathrm{d}F(t),$$

so $P'(x)$ is a bounded linear operator with operator norm less than or equal to $\int_0^T \mathrm{e}^{-x(t)} \,\mathrm{d}F(t)$. $\square$

B Proofs

B.1 Proof of Proposition 3.1

Let us first show that $\mathcal{W}$ in (3.5) is compact, convex, and contains 0 in the interior. Clearly $0 \in \mathcal{W}$. Since $w \mapsto G'w$ is linear (hence continuous) and $G'0 = 0$ is an interior point of $[-1,1]^N$, 0 is an interior point of $\mathcal{W}$. Since $\mathcal{W}$ is defined by weak linear inequalities, it is closed and convex. Let us show compactness. By Assumption 3, H has full row rank, and so does G. Take $n_1, \dots, n_I$ such that the $I \times I$ matrix $\tilde{G} := (g_{i,n_j})$ is invertible. Define

$$\tilde{\mathcal{W}} := \left\{ w \in \mathbb{R}^I : \tilde{G}'w \in [-1,1]^I \right\} = (\tilde{G}')^{-1}[-1,1]^I.$$

Since $\tilde{\mathcal{W}}$ is defined by a subset of inequalities that define $\mathcal{W}$, clearly we have $\mathcal{W} \subset \tilde{\mathcal{W}}$. Furthermore, $\tilde{\mathcal{W}}$ is compact because it is the image of the compact set

$[-1,1]^I$ under the linear (hence continuous) map $(\tilde{G}')^{-1} : \mathbb{R}^I \to \mathbb{R}^I$. Therefore $\mathcal{W} \subset \tilde{\mathcal{W}}$ is compact.

Next, let us show that the minmax problem (3.6) has a solution $(z^*, w^*) \in \mathcal{Z} \times \mathcal{W}$. Since $\mathcal{W}$ is nonempty and compact and $w \mapsto \langle w, Az - b\rangle$ is linear (hence continuous),

$$M(z) := \max_{w\in\mathcal{W}} \langle w, Az - b\rangle \tag{B.1}$$

exists. The maximum theorem (Berge, 1963, p. 116) implies that M is continuous. Furthermore, since $0 \in \mathcal{W}$, we have $M(z) \geq 0$ and hence $V_I(\mathcal{Z}) = \inf_{z\in\mathcal{Z}} M(z) \geq 0$. Let $\|\cdot\|_2$ denote the ℓ^2 (Euclidean) norm. Since 0 is an interior point of $\mathcal{W}$, there exists $\epsilon > 0$ such that $w \in \mathcal{W}$ whenever $\|w\|_2 \leq \epsilon$. If $Az \neq b$, setting $w = \epsilon \frac{Az-b}{\|Az-b\|_2}$, we obtain

$$M(z) \geq \left\langle \epsilon \frac{Az - b}{\|Az - b\|_2}, Az - b \right\rangle = \epsilon \|Az - b\|_2 . \tag{B.2}$$

Note that the lower bound (B.2) is valid even if $Az = b$.

To bound (B.2) from below, let us show that

$$\|Az - b\|_2 = \|A_+ z - b_+\|_2 \tag{B.3}$$

when $z \in \mathcal{Z}$. Using the definition (3.7), it suffices to show that $a_0 z - 1 = 0$ if $z \in \mathcal{Z}$. But since by Assumption 2 value matching holds, dividing (2.9) by $P(x)$ and using (3.3) for $i = 0$ (hence $h_0 \equiv 1$), we obtain

$$1 = \frac{1}{P(x)} \sum_{j=1}^{J} z_j P_j(x) = \sum_{j=1}^{J} a_{0j} z_j = a_0 z,$$

which implies (B.3). Define $m := \min_{\|z\|_2=1} \|A_+ z\|_2$, which is achieved because $\|z\|_2 = 1$ is a nonempty compact set and $z \mapsto \|A_+ z\|_2$ is continuous. Since by assumption A_+ has full column rank, we have $A_+ z = 0$ only if $z = 0$, so $m > 0$. Therefore it follows from (B.2) and (B.3) that for any $z \in \mathcal{Z}$,

$$M(z) \geq \epsilon \|Az - b\|_2 = \epsilon \|A_+ z - b_+\|_2 \geq \epsilon(m \|z\|_2 - \|b_+\|_2) \to \infty \tag{B.4}$$

as $\|z\|_2 \to \infty$, so we may restrict the minimization of $M(z)$ to a compact subset of $\mathcal{Z}$. Since $M(z)$ is continuous, the minmax value $V_I(\mathcal{Z})$ is achieved.

Finally, let us show that $z \in \mathcal{Z}$ achieves $V_I(\mathcal{Z}) = 0$ if and only if $A_+ z = b_+$. If $A_+ z = b_+$, then $Az = b$ so clearly $M(z) = 0$ and $V_I(\mathcal{Z}) = 0$. If $V_I(\mathcal{Z}) = 0$, then for any $z \in \mathcal{Z}$ with $M(z) = V_I(\mathcal{Z}) = 0$, (B.2) and (B.3) imply $\|A_+ z - b_+\|_2 = 0$ and therefore $A_+ z = b_+$. $\square$

B.2 Proof of Proposition 3.2

Suppose that $\text{span}\{\tilde{h}_i\}_{i=1}^I = \text{span}\,\{h_i\}_{i=1}^I$. Since $\{h_i\}_{i=1}^I$ span $\{\tilde{h}_i\}_{i=1}^I$, there exists an $I \times I$ matrix $C = (c_{ij})$ such that $\tilde{h}_i = \sum_{j=1}^I c_{ij} h_j$. Since $\{h_i\}_{i=1}^I$ are linearly independent, C is unique. Since $\{\tilde{h}_i\}_{i=1}^I$ also span $\{h_i\}_{i=1}^I$, C must be invertible. Then $\tilde{H} = CH$, $\tilde{A} = CA$, $\tilde{b} = Cb$, $\tilde{G} = CG$, so setting $w = C'\tilde{w}$, we obtain

$$\tilde{M}(z) \coloneqq \sup_{\tilde{w}:\tilde{G}'\tilde{w}\in[-1,1]^N} \left\langle \tilde{w}, \tilde{A}z - \tilde{b} \right\rangle = \sup_{w:G'w\in[-1,1]^N} \langle w, Az - b \rangle =: M(z).$$

Therefore the minimizers of M and $\tilde{M}$ agree and the conclusion holds. □

B.3 Proposition B.1

Proposition B.1 (Generic full column rank of A_+)**.** *Let $I \geq J-1$, $\{h_i\}_{i=1}^I$ be the basis functions, and set $h_0 \equiv 1$. Suppose that there exist $\{i_\ell\}_{\ell=1}^J \subset \{0, 1, \dots, I\}$ with $i_1 = 0$ and $\{\tau_j\}_{j=1}^J \subset (0, T]$ such that (i) at date τ_j, bond j makes a lump-sum payout $f_j \coloneqq F_j(\tau_j) - F_j(\tau_j-) > 0$, and (ii) the $J \times J$ matrix $\tilde{H} = (h_{i_\ell}(\tau_j))$ is invertible. Then there exists a closed set $S \subset \mathbb{R}^J$ with Lebesgue measure 0 such that the matrix A_+ in (3.7) has full column rank whenever $(f_1, \dots, f_J) \notin S$.*

If in addition all bonds are zero-coupon bonds, then A_+ has full column rank.

The fact that the set of $(f_1, \dots, f_J)$ for which A_+ has rank deficiency is contained in a closed set with Lebesgue measure 0 implies that the set of rank deficiency is nowhere dense (has empty interior). In this sense the rank deficiency of A_+ is "rare". To prove the result, we need the following lemma.

Lemma B.1. *Let A, B be $N \times N$ matrices and define $\phi : \mathbb{R}^N \to \mathbb{R}$ by $\phi(x) = \det(A \operatorname{diag}(x) + B)$, where $\operatorname{diag}(x)$ denotes the diagonal matrix with diagonal entries $x_1, \dots, x_N$. If $\det A \neq 0$, then for any $c \in \mathbb{R}$ the set*

$$\phi^{-1}(c) \coloneqq \left\{x \in \mathbb{R}^N : \phi(x) = c\right\}$$

is closed and has Lebesgue measure 0.

Proof. Since

$$\begin{aligned} \det(A \operatorname{diag}(x) + B) &= \det(A(\operatorname{diag}(x) + A^{-1}B)) \\ &= \det(A) \times \det(\operatorname{diag}(x) + A^{-1}B), \end{aligned}$$

without loss of generality we may assume that A is the identity matrix. Let $B = (b_{mn})$. That $\phi^{-1}(c)$ is closed is obvious because f is continuous.

Let us show by induction on the dimension N that $\phi^{-1}(c)$ is a null set. If $N = 1$, then $\phi(x) = x_1 + b_{11}$, so $\phi^{-1}(c) = \{c - b_{11}\}$ is a singleton, which is a null set. Suppose the claim holds when $N = n - 1$ and consider n. Let B_n be the $n \times n$ matrix obtained from the first n rows and columns of B, and let

$$\phi_n(x_1, \dots, x_n) = \det(\operatorname{diag}(x_1, \dots, x_n) + B_n).$$

Clearly ϕ_n is affine in each variable $x_1, \dots, x_n$. Using the Laplace expansion along the n-th column, it follows that

$$\phi_n(x_1, \dots, x_n) = (x_n + b_{nn})\phi_{n-1}(x_1, \dots, x_{n-1}) + \psi_{n-1}(x_1, \dots, x_{n-1})$$

for some function ψ_{n-1} that is affine in each variable $x_1, \dots, x_{n-1}$.

Define the sets $\phi_{n-1}^{-1}(0) \subset \mathbb{R}^{n-1}$ and $G \subset \mathbb{R}^n$ by

$$\begin{aligned}
\phi_{n-1}^{-1}(0) &\coloneqq \{(x_1, \dots, x_{n-1}) : \phi_{n-1}(x_1, \dots, x_{n-1}) = 0\}, \\
G &\coloneqq \left\{(x_1, \dots, x_n) : (x_1, \dots, x_{n-1}) \notin \phi_{n-1}^{-1}(0), x_n = (c - \psi_{n-1})/\phi_{n-1} - b_{nn}\right\}.
\end{aligned}$$

Then $\phi_n^{-1}(c) \subset (\phi_{n-1}^{-1}(0) \times \mathbb{R}) \cup G$. By the induction hypothesis, $\phi_{n-1}^{-1}(0)$ has measure 0 in $\mathbb{R}^{n-1}$. Since G is the graph of a Borel measurable function, by Fubini's theorem it has measure 0. Therefore $\phi_n^{-1}(c)$ is a null set. $\square$

Proof of Proposition B.1. Define $\mathbf{h} : [0, T] \to \mathbb{R}^I$ by $\mathbf{h}(t) = (h_0(t), h_1(t), \dots, h_I(t))'$. Let the j-th column vector of A_+ be $\mathbf{a}_j = (a_{0j}, \dots, a_{Ij})'$. By assumption, bond j pays $f_j > 0$ at $\tau_j \in (0, T]$, so it follows from (3.3) that

$$\mathbf{a}_j = \frac{1}{P(x)} \int_{[0,T] \setminus \{\tau_j\}} \mathrm{e}^{-x(t)} \mathbf{h}(t) \, \mathrm{d}F_j(t) + \frac{1}{P(x)} \mathrm{e}^{-x(\tau_j)} f_j \mathbf{h}(\tau_j) \eqqcolon \mathbf{p}_j f_j + \mathbf{q}_j. \tag{B.5}$$

Collecting (B.5) into a matrix, we can write $A_+ = P \operatorname{diag}(f) + Q$, where P, Q are matrices with j-th column vectors $\mathbf{p}_j, \mathbf{q}_j$ and $f = (f_1, \dots, f_J)$. To show that A_+ generically has full column rank, let $\tilde{A}_+$ be the $J \times J$ matrix obtained by taking its i_ℓ-th row for $\ell = 1, \dots, J$. Define $\tilde{P}, \tilde{Q}$ similarly. Then $\tilde{A}_+ = \tilde{P} \operatorname{diag}(f) + \tilde{Q}$. Since $\mathbf{p}_j = \mathrm{e}^{-x(\tau_j)} \mathbf{h}(\tau_j)/P(x)$, we obtain

$$\det \tilde{P} = P(x)^{-J} \left(\prod_{j=1}^{J} \mathrm{e}^{-x(\tau_j)} \right) \det \tilde{H} \neq 0.$$

Therefore by Lemma B.1, $\tilde{A}_+$ is generically invertible, so A_+ has generically full column rank.

If in addition all bonds are zero-coupon bonds, then (B.5) reduces to $\mathbf{a}_j = \mathrm{e}^{-x(\tau_j)} f_j \mathbf{h}(\tau_j)/P(x)$, where τ_j is the maturity. Then $A_+ = P \operatorname{diag}(f)$, which has full column rank because $\det \tilde{P} \neq 0$ and $f_j > 0$ for all j. $\square$

The following example shows that the zero-coupon bond assumption in Proposition B.1 is essential.

Example 4. Suppose $I = J - 1 = 1$ and the basis function is $h_1(t) = t$. Bond 1 is a zero-coupon bond with face value $f_1 > 0$ and maturity t_1. Bond 2 pays $f_n > 0$ at time t_n, where $n = 2, 3$. To simplify notation, write $x(t_1) = x_1$ etc. The determinant of the matrix A_+ is

$$\begin{aligned}\det A_+ &= P(x)^{-2} \det \begin{bmatrix} f_1 \mathrm{e}^{-x_1} & f_2 \mathrm{e}^{-x_2} + f_3 \mathrm{e}^{-x_3} \\ f_1 \mathrm{e}^{-x_1} t_1 & f_2 \mathrm{e}^{-x_2} t_2 + f_3 \mathrm{e}^{-x_3} t_3 \end{bmatrix} \\ &= P(x)^{-2} f_1 \mathrm{e}^{-x_1} \left(f_2 \mathrm{e}^{-x_2}(t_2 - t_1) + f_3 \mathrm{e}^{-x_3}(t_3 - t_1) \right).\end{aligned}$$

Therefore for any $t_2 < t_1 < t_3$ and $f_3 > 0$, we have $\det A_+ = 0$ if and only if

$$(f_1, f_2) \in \left\{ (f_1, f_2) \in \mathbb{R}^2_{++} : f_2 = f_3 \mathrm{e}^{x_2 - x_3} \frac{t_3 - t_1}{t_1 - t_2} \right\}. \tag{B.6}$$

The closure of the rank deficiency set (B.6) is a ray in $\mathbb{R}^2$ and has measure 0.

B.4 Proof of Theorem 1

To prove Theorem 1, we recall Taylor's theorem with the integral form for the remainder term.

Lemma B.2 (Taylor's theorem). *Let $f \in C^{n+1}[0,1]$, so $f : [0,1] \to \mathbb{R}$ is $n+1$ times continuously differentiable. Then*

$$f(1) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} + \int_0^1 f^{(n+1)}(s) \frac{(1-s)^n}{n!} \, \mathrm{d}s. \tag{B.7}$$

Proof of Theorem 1. For any $x, h \in \mathbb{R}$, define $f : [0,1] \to \mathbb{R}$ by $f(s) = \mathrm{e}^{-x-sh}$. Applying Lemma B.2 for $n = 1$, we obtain

$$\mathrm{e}^{-x-h} = \mathrm{e}^{-x} - \mathrm{e}^{-x} h + \int_0^1 (1-s) \mathrm{e}^{-x-sh} h^2 \, \mathrm{d}s.$$

Setting $x = x(t)$ and $h = h(t)$ for $x, h \in \mathcal{X}$ and integrating both sides on $[0, T]$ with respect to F, we obtain

$$\int_0^T \mathrm{e}^{-x(t)-h(t)}\,\mathrm{d}F(t) = \int_0^T \mathrm{e}^{-x(t)}\,\mathrm{d}F(t) - \int_0^T \mathrm{e}^{-x(t)}h(t)\,\mathrm{d}F(t) \\ + \int_0^T \int_0^1 (1-s)\mathrm{e}^{-x(t)-sh(t)}h(t)^2\,\mathrm{d}s\,\mathrm{d}F(t).$$

Using the definition of P and P', we obtain

$$P(x+h) = P(x) + P'(x)h + \int_0^T \int_0^1 (1-s)\mathrm{e}^{-x(t)-sh(t)}h(t)^2\,\mathrm{d}s\,\mathrm{d}F(t). \tag{B.8}$$

A similar equation holds for each P_j. Hence for any $z = (z_j) \in \mathbb{R}^J$ we have

$$E(z, x+h) = \sum_{j=1}^J z_j P_j(x+h) - P(x+h) = E_0 + E_1 + E_2, \tag{B.9}$$

where

$$E_0 \coloneqq \sum_{j=1}^J z_j P_j(x) - P(x), \tag{B.10a}$$

$$E_1 \coloneqq \left(\sum_{j=1}^J z_j P_j'(x) - P'(x)\right) h, \tag{B.10b}$$

$$E_2 \coloneqq \int_0^T \int_0^1 (1-s)\mathrm{e}^{-x(t)-sh(t)}h(t)^2\,\mathrm{d}s\,\mathrm{d}\left(\sum_{j=1}^J z_j F_j(t) - F(t)\right). \tag{B.10c}$$

Since $\mathcal{Z}$ satisfies value matching by Assumption 2, we have $E_0 = 0$ by (B.10a). Inspection of Assumption 3, (2.11), and (3.5) reveals that any $h \in \mathcal{H}_I(\Delta)$ can be expressed as $h = \Delta \sum_{i=1}^I w_i h_i$ for some $w \in \mathcal{W}$. Using (B.10b), (3.3), and (3.2), we obtain

$$E_1 = \left(\sum_{j=1}^J z_j P_j'(x) - P'(x)\right) h = -\Delta P(x) \langle w, Az - b \rangle. \tag{B.11}$$

To bound E_2, note that the last integral in (B.8) is nonnegative because $1 - s \geq 0$ on $s \in [0, 1]$ and F is increasing. Furthermore, it can be bounded above by

$$\int_0^T \int_0^1 (1-s)\mathrm{e}^{-x(t)+\|h\|_\infty} \|h\|_\infty^2\,\mathrm{d}s\,\mathrm{d}F(t) = \frac{1}{2}\|h\|_\infty^2\,\mathrm{e}^{\|h\|_\infty} P(x),$$

where $\|h\|_\infty = \max_n |h(t_n)|$.[13] Therefore E_2 in (B.10c) can be bounded as

$$\frac{1}{2}\|h\|_\infty^2 \mathrm{e}^{\|h\|_\infty}\left(\sum_{z_j<0} z_j P_j(x) - P(x)\right) \le E_2 \le \frac{1}{2}\|h\|_\infty^2 \mathrm{e}^{\|h\|_\infty}\sum_{z_j\ge 0} z_j P_j(x). \tag{B.12}$$

Using (2.9) and (3.8), we obtain

$$\begin{aligned} P(x) - \sum_{z_j<0} z_j P_j(x) &= \sum_{z_j\ge 0} z_j P_j(x) = \frac{1}{2}\left(P(x) + \sum_{j=1}^J |z_j| P_j(x)\right) \\ &= \frac{1}{2}P(x)\left(1 + \sum_{j=1}^J |\theta_j|\right) = \frac{1}{2}P(x)(1+\|\theta\|_1). \end{aligned} \tag{B.13}$$

Noting that $\|h\|_\infty \le \Delta T$ for $h \in \mathcal{H}_I(\Delta)$, it follows from (B.12) and (B.13) that

$$|E_2| \le \frac{1}{4}\Delta^2 T^2 \mathrm{e}^{\Delta T} P(x)(1+\|\theta\|_1). \tag{B.14}$$

Combining (B.9), $E_0 = 0$, (B.11), and (B.14), we obtain

$$\begin{aligned} &-\langle w, Az - b\rangle - \frac{1}{4}\Delta T^2 \mathrm{e}^{\Delta T}(1+\|\theta\|_1) \\ &\quad \le \frac{1}{\Delta P(x)} E(z, x+h) \le -\langle w, Az - b\rangle + \frac{1}{4}\Delta T^2 \mathrm{e}^{\Delta T}(1+\|\theta\|_1). \end{aligned} \tag{B.15}$$

Since by (3.8) θ_j is proportional to z_j, there exists some constant $c(x) > 0$ that depends only on x such that $\|\theta\|_1 \le c(x)\|z\|_2$. Therefore minimizing (B.15) over $w \in \mathcal{W}$, it follows from the definition of $M(z)$ in (B.1) that

$$\begin{aligned} &-M(z) - \frac{1}{4}\Delta T^2 \mathrm{e}^{\Delta T}(1 + c(x)\|z\|_2) \\ &\quad \le \frac{1}{\Delta P(x)} \inf_{h\in\mathcal{H}_I(\Delta)} E(z, x+h) \le -M(z) + \frac{1}{4}\Delta T^2 \mathrm{e}^{\Delta T}(1 + c(x)\|z\|_2). \end{aligned} \tag{B.16}$$

Let $m, \epsilon > 0$ be as in the proof of Proposition 3.1 and take $\bar{\Delta} > 0$ such that $\epsilon m = \frac{1}{4}\bar{\Delta} T^2 \mathrm{e}^{\bar{\Delta} T} c(x)$. Then if $0 < \Delta < \bar{\Delta}$, by (B.4) both sides of (B.16) tend to $-\infty$ as $\|z\|_2 \to \infty$. Therefore when we take the supremum of (B.16) with respect to $z \in \mathcal{Z}$, we may restrict it to some compact subset $\mathcal{Z}' \subset \mathcal{Z}$. Therefore there

[13] Recall that we only put restriction on h at the payout dates.

exists a constant $c' > 0$ such that

$$-M(z) - c'\Delta \leq \frac{1}{\Delta P(x)} \inf_{h \in \mathcal{H}_I(\Delta)} E(z, x+h) \leq -M(z) + c'\Delta$$

for all $z \in \mathcal{Z}'$ and $\Delta \in (0, \bar{\Delta})$. Taking the supremum over $z \in \mathcal{Z}$ (which is achieved in $\mathcal{Z}'$) and letting $\Delta \to 0$, by the definition of $V_I(\mathcal{Z})$ in (3.6), we obtain (3.9).

To show the error estimate (3.10), let $z^* \in \mathcal{Z}$ be a solution to the minmax problem (3.6). It follows from (B.15) that

$$\frac{1}{\Delta P(x)} |E(z^*, x+h)| \leq |\langle w, Az^* - b\rangle| + \frac{1}{4}\Delta T^2 \mathrm{e}^{\Delta T}(1 + \|\theta\|_1).$$

Taking the supremum over $w \in \mathcal{W}$ and noting that $\mathcal{W}$ is symmetric ($w \in \mathcal{W}$ implies $-w \in \mathcal{W}$), it follows from the definition of $V_I(\mathcal{Z})$ in (3.6) that (3.10) holds. $\square$

B.5 Proof of Theorem 3

To solve the inner maximization of (3.15), fix z and let $c = Az - b$. Then the optimization problem reduces to

$$\begin{aligned} &\text{maximize} && \langle c, w\rangle \\ &\text{subject to} && \frac{1}{2}w'GG'w \leq \frac{1}{2}. \end{aligned}$$

If $c = 0$, the maximum value is 0, and any w is optimal. Therefore assume $c \neq 0$ and define the Lagrangian by

$$L(w, \lambda) := \langle c, w\rangle + \frac{\lambda}{2}(1 - w'GG'w),$$

where $\lambda \geq 0$ is the Lagrange multiplier. Since by Assumption 3 the matrix G has full row rank, GG' is positive definite. Then the Slater constraint qualification is satisfied and we can apply the Karush-Kuhn-Tucker theorem. The first-order condition is

$$0 = \nabla_w L = c - \lambda GG'w \iff w = \frac{1}{\lambda}(GG')^{-1}c.$$

The complementary slackness condition is

$$1 = w'GG'w = \frac{1}{\lambda^2}c'(GG')^{-1}c \iff \lambda = \sqrt{c'(GG')^{-1}c}.$$

Therefore the maximum value is

$$\langle c, w \rangle = \frac{c'(GG')^{-1}c}{\lambda} = \sqrt{c'(GG')^{-1}c},$$

which is also valid when $c = 0$.

Noting that $c = Az - b$, the outer minimization reduces to

$$\begin{aligned} &\text{minimize} && \frac{1}{2}\left\langle Az - b, (GG')^{-1}(Az - b)\right\rangle \\ &\text{subject to} && Rz = r. \end{aligned}$$

Let $\mu \in \mathbb{R}^M$ be the Lagrange multiplier for the equality constraint and define the Lagrangian by

$$L(z, \mu) := \frac{1}{2}\left\langle Az - b, (GG')^{-1}(Az - b)\right\rangle + \mu'(r - Rz).$$

Letting $\tilde{A} := (GG')^{-1}A$, the first-order condition is

$$\tilde{A}'(Az - b) - R'\mu = 0 \iff z = (\tilde{A}'A)^{-1}(\tilde{A}'b + R'\mu),$$

where we used the fact that $\tilde{A}'A = A'(GG')^{-1}A$ is invertible because A has full column rank. The complementary slackness condition implies

$$r = Rz = R(\tilde{A}'A)^{-1}(\tilde{A}'b + R'\mu) \iff \mu = [R(\tilde{A}'A)^{-1}R']^{-1}(r - R(\tilde{A}'A)^{-1}\tilde{A}'b),$$

where we used the fact that $R(\tilde{A}'A)^{-1}R'$ is invertible because R has full row rank. Therefore the solution to the minmax problem (3.15) for $p = 2$ is

$$z = (\tilde{A}'A)^{-1}\tilde{A}'b + (\tilde{A}'A)^{-1}R'[R(\tilde{A}'A)^{-1}R']^{-1}(r - R(\tilde{A}'A)^{-1}\tilde{A}'b). \qquad \square$$

B.6 Proof of Proposition 3.4

We prove Proposition 3.4 by deriving a general result that expresses a constrained GLS estimator in terms of elemental estimates. We first derive the analogous result for the constrained OLS estimator, which might be of independent interest. In this section we always assume $A \in \mathbb{R}^{I\times J}$, $b \in \mathbb{R}^I$ with $I \geq J$, and $\operatorname{rank}(A) = J$. Furthermore, $[M]$ denotes the set of integers $\{1, \dots, M\}$, and s denotes any subset of size J from the set $\{1, \dots, I\}$.

Proposition B.2. *Let $z_{cls} \in \mathbb{R}^J$ be the solution to the constrained least squares* *problem* $\min_z \|b - Az\|_2^2$ *subject to the linear constraints* $Rz = r$*, where* $R \in \mathbb{R}^{M\times J}$ *has full row rank and* $r \in \mathbb{R}^M$*. Define the augmented matrices*

$$A_+ = \begin{bmatrix} R \\ A \end{bmatrix} \in \mathbb{R}^{(M+I)\times J} \qquad b_+ = \begin{bmatrix} r \\ b \end{bmatrix} \in \mathbb{R}^{M+I},$$

and let $z(s) = A_+(s)^{-1} b_+(s) \in \mathbb{R}^J$ *denote the elemental estimate based on the rows of* A_+ *and* b_+ *that are in* s*. Then,*

$$\begin{aligned} z_{cls} = \mathrm{E}(z(S)|[M] \subset S) &= \sum_{[M]\subset s} \frac{\det[A_+(s)]^2}{\sum_{[M]\subset s} \det[A_+(s)]^2} z(s) \qquad \text{(B.17)} \\ &=: \sum_{[M]\subset s} \lambda(s) z(s), \end{aligned}$$

where

$$\lambda(s) = \frac{\det[A_+(s)]^2}{\sum_{[M]\subset s} \det[A_+(s)]^2} \geq 0 \quad \textit{and} \quad \sum_{[M]\subset s} \lambda(s) = 1.$$

In the proof we use the generalized Cauchy-Binet formula (Chapman and Miyake, 2018, Appendix B).

Lemma B.3 (Generalized Cauchy-Binet formula)**.** *Let* $C \in \mathbb{R}^{m\times n}$ *and* $D \in \mathbb{R}^{n\times m}$*, and let* T *denote the set* $\{n-j+1,\dots,n\}$*. Let* s *denote any subset of* $\{1,\dots,n\}$ *of cardinality* m*. Denote* $C(s) \in \mathbb{R}^{m\times|s|}$ *by the matrix* C *with column index in* s *and* $D(s) \in \mathbb{R}^{|s|\times m}$ *by the matrix* D *with row index in* s*. Then,*

$$\sum_{T\subset s} \det\left[C(s)\right] \det\left[D(s)\right] = (-1)^j \det \begin{bmatrix} 0_{j\times j} & D(s) \\ C(s) & C([n-j])D([n-j]) \end{bmatrix}.$$

Proof of Proposition B.2. By construction the expression for z_{cls} in (B.17) satisfies $Rz_{\text{cls}} = r$ since the sum is taken only over rows that also contain the constraints. Using Lemma B.3 we can write

$$\begin{aligned} &\sum_{[M]\subset s} \det[A_+(s)]^2 = \sum_{[M]\subset s} \det[A_+(s)'] \det[A_+(s)] \\ &= (-1)^M \det \begin{bmatrix} A_+([M])' & A_+(\{M+1,\dots,I\})' A_+(\{M+1,\dots,I\}) \\ 0_{M\times M} & A_+([M]) \end{bmatrix} \\ &= (-1)^M \det \begin{bmatrix} R' & A'A \\ 0_{M\times M} & R \end{bmatrix} \\ &= (-1)^M \det[A'A] \det\left[R(A'A)^{-1}R'\right]. \end{aligned}$$

The final equality follows from the determinant formula for block matrices. Now we define $z^j(s)$ to be the jth element of $z(s)$. Using Cramer's rule we have $z^j(s) = \det\left[A_+^j(s)\right] / \det\left[A_+(s)\right]$, where $A_+^j(s)$ is $A_+(s)$ except that the jth column is replaced by $b_+(s)$. Similarly R^j is defined such that the jth column of R is replaced by r. It then follows that

$$\begin{aligned}
&\sum_{[M]\subset s} \det[A_+(s)]^2 z^j(s) = \sum_{[M]\subset s} \det\left[A_+(s)'\right] \det\left[A_+^j(s)\right] \\
&= (-1)^M \det \begin{bmatrix} A_+([M])' & A_+\left(\{M+1,\dots,I\}\right)' A_+^j\left(\{M+1,\dots,I\}\right) \\ 0_{M\times M} & A_+^j\left([M]\right) \end{bmatrix} \\
&= (-1)^M \det \begin{bmatrix} R' & A'A^j \\ 0_{M\times M} & R^j \end{bmatrix} \\
&= (-1)^M \det\left[A'A^j\right] \det\left[R^j (A'A^j)^{-1} R'\right].
\end{aligned}$$

In conclusion we have shown that the jth element of z_{cls} in (B.17) can be expressed as

$$\begin{aligned}
z_{\text{cls}}^j &= \frac{\det\left[A'A^j\right] \det\left[R^j (A'A^j)^{-1} R'\right]}{\det\left[A'A\right] \det\left[R (A'A)^{-1} R'\right]} \\
&= z_{\text{ols}}^j \frac{\det\left[R^j (A'A^j)^{-1} R'\right]}{\det\left[R (A'A)^{-1} R'\right]},
\end{aligned} \tag{B.18}$$

where z_{ols}^j is the jth element of the unconstrained OLS solution $\min_z \|Az - b\|_2^2$. The latter observation follows immediately from Cramer's rule applied to the first order condition $A'Az_{\text{ols}} = A'b$.

It remains to verify that the solution in (B.18) concurs with the constrained least squares solution at the top of Proposition B.2. Let $\mu \in \mathbb{R}^M$ be the Lagrange multiplier for the equality constraint and consider the associated Lagrangian

$$L(z,\mu) = \frac{1}{2}\|Az - b\|_2^2 + \mu'(Rz - r).$$

The first order conditions for this problem imply

$$\begin{bmatrix} A'A & R' \\ R & 0 \end{bmatrix} \begin{bmatrix} z_{\text{cls}} \\ \mu \end{bmatrix} = \begin{bmatrix} A'b \\ r \end{bmatrix}.$$

Using Cramer's rule to get the jth element of z_{cls} we get

$$z_{\text{cls}}^{j} = \frac{\det \begin{bmatrix} A'A^j & R' \\ R^j & 0 \end{bmatrix}}{\det \begin{bmatrix} A'A & R' \\ R & 0 \end{bmatrix}} = \frac{\det[A'A^j] \det[R^j (A'A^j)^{-1} R']}{\det[A'A] \det[R(A'A)^{-1} R']}.$$

$\square$

Proposition B.3. *Consider the constrained GLS solution*

$$z_{cgls} := \arg\min_{z} \left\langle Az - b, \Omega^{-1}(Az - b) \right\rangle \quad s.t.\ \ Rz = r. \tag{B.19}$$

Define $\tilde{A} := \Omega^{-1}A$ *and let* $\tilde{A}_{+}' = [R, \tilde{A}]'$. *Let* A_+ *and* $z(s)$ *be the same as in Proposition B.2. Then the solution in (B.19) can be expressed as*

$$z_{cgls} = \frac{\sum_{[M]\subset s} \det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right] z(s)}{\sum_{[M]\subset s} \det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right]} =: \sum_{[M]\subset s} \lambda(s) z(s),$$

where

$$\lambda(s) = \frac{\det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right]}{\sum_{[M]\subset s} \det\left[\tilde{A}_+(s)\right] \det\left[A_+(s)\right]} \quad \textit{with} \ \sum_{[M]\subset s} \lambda(s) = 1.$$

Remark 6. In this case $\lambda(\cdot)$ does not define a conditional probability measure since it can take on negative values.

Proof. The steps are similar to the proof of Proposition B.2. The first order conditions for the optimization problem in (B.19) imply

$$\begin{bmatrix} A'\Omega^{-1}A & R' \\ R & 0 \end{bmatrix} \begin{bmatrix} z_{cgls} \\ \mu \end{bmatrix} = \begin{bmatrix} A'\Omega^{-1}b \\ r \end{bmatrix}.$$

Applying Cramer's rule yields

$$
\begin{aligned}
z^j_{cgls} &= \frac{\det \begin{bmatrix} A'\Omega^{-1}A^j & R' \\ R^j & 0 \end{bmatrix}}{\det \begin{bmatrix} A'\Omega^{-1}A & R' \\ R & 0 \end{bmatrix}} \\
&= \frac{\det\left[A'\Omega^{-1}A^j\right]\det\left[R^j(A'\Omega^{-1}A^j)^{-1}R'\right]}{\det\left[A'\Omega^{-1}A\right]\det\left[R(A'\Omega^{-1}A)^{-1}R'\right]} \\
&= \frac{\det\left[\tilde{A}'A\right]\det\left[R^j(\tilde{A}'A^j)^{-1}R'\right]}{\det\left[\tilde{A}'A\right]\det\left[R(\tilde{A}'A)^{-1}R'\right]}.
\end{aligned}
$$

The rest of the proof is identical to Proposition B.2. $\square$

Finally we prove Proposition 3.4 as a special case of Proposition B.3.

Proof of Proposition 3.4. Due to the basis invariance of the robust immunization portfolio in (3.16), we can use the polynomial basis $h_i(t) = t^i$. Therefore, the first $J-1$ rows of A are equal to A_{HD}. The result now follows from Proposition B.3 with $\Omega = GG'$, $R = a_0$, and $r = 1$. $\square$

B.7 Proof of Theorem 4

Because the proof is similar to that of Theorem 1, we only provide a sketch.

By assumption, $\mathcal{Z}_1$ in (3.19) is nonempty, and it is clearly closed. Hence by Proposition 3.1 the minmax value $V_I(\mathcal{Z}_1)$ defined by (3.6) is achieved by some $z^* \in \mathcal{Z}_1$. Inspection of Assumption 3, (3.18), and (3.5) reveals that any $h \in \mathcal{H}_I(\Delta_1, \Delta_2)$ can be expressed as $h = \Delta_1 v h_1 + \Delta_2 \sum_{i=1}^I w_i h_i$ for some $w \in \mathcal{W}$ and $v \in \mathbb{R}$ with $|v| \le \min_n 1/\left|h(t_n)/t_n\right| =: \bar{v} \in (0,\infty)$. Applying a similar argument to the derivation of (B.15), we obtain

$$
\frac{1}{P(x)}E(z, x+h) = -\Delta_1 v(Az-b)_1 - \Delta_2 \langle w, Az-b\rangle + O(\Delta_1^2 + \Delta_2^2),
$$

where $(Az-b)_1$ denotes the first entry of the vector $Az-b$. Minimizing both sides over $h \in \mathcal{H}_I(\Delta, \Delta_1)$, we obtain

$$
\inf_{h\in\mathcal{H}_I(\Delta,\Delta_1)} \frac{1}{P(x)}E(z, x+h) = -\Delta_1 \bar{v}\left|(Az-b)_1\right| - \Delta_2 M(z) + O(\Delta_1^2 + \Delta_2^2),
$$

where $M(z)$ is defined by (B.1). Dividing both sides by $\Delta_2 > 0$ and letting

$\Delta_2 \to 0$, $\Delta_1/\Delta_2 \to \infty$, and $\Delta_1^2/\Delta_2 \to 0$, the objective function remains finite only if $(Az-b)_1 = 0$, which is equivalent to $z \in \mathcal{Z}_1$. Under this condition, we have

$$\frac{1}{\Delta} \inf_{h \in \mathcal{H}_I(\Delta_1,\Delta_2)} \frac{1}{P(x)} E(z, x+h) = -M(z) + O(\Delta_2 + \Delta_1^2/\Delta_2).$$

Maximizing over $z \in \mathcal{Z}_1$ and letting $\Delta_2 \to 0$, we obtain (3.20). The proof of (3.21) is similar. $\square$

B.8 Proof of Proposition 3.5

For each I, let $M_I(z) = \sup_{w \in \mathcal{W}_I} \langle w, A_I z - b_I \rangle$, where A_I, b_I denote the sensitivity matrix and vector A, b defined by (3.3), (3.2) and $\mathcal{W}_I$ denotes the set $\mathcal{W}$ defined by (3.5). Suppose $I < I'$. Letting 0_N denote the zero vector of $\mathbb{R}^N$, we have $\mathcal{W}_I \times \{0_{I'-I}\} \subset \mathcal{W}_{I'}$, so

$$\begin{aligned}(z) &= \sup_{w \in \mathcal{W}_I} \langle w, A_I z - b_I \rangle = \sup_{w \in \mathcal{W}_I \times \{0_{I'-I}\}} \langle w, A_{I'} z - b_{I'} \rangle \\ &\le \sup_{w \in \mathcal{W}_{I'}} \langle w, A_{I'} z - b_{I'} \rangle = M_{I'}(z).\end{aligned}$$

Taking the infimum over $z \in \mathcal{Z}$, we obtain $V_I(\mathcal{Z}) \le V_{I'}(\mathcal{Z})$. Similarly,

$$V_I(\mathcal{Z}) = \inf_{z \in \mathcal{Z}} M_I(z) \ge \inf_{z \in \mathcal{Z}'} M_I(z) = V_I(\mathcal{Z}'). \qquad \square$$

B.9 Proof of Proposition 3.3

Suppose that the liability has maturity s with face value 1. Then the value of the liability is

$$P(x) = \int_0^T \mathrm{e}^{-x(t)} \, \mathrm{d}F(t) = \mathrm{e}^{-x(s)}.$$

Let $z^* = A_+^{-1} b_+$ be the immunizing portfolio and assume $z^* \ge 0$. Take any perturbation $h \in \operatorname{span}\{h_i\}_{i=1}^I$ and write $h = \sum_{i=1}^I w_i h_i$. Then the funding ratio is

$$\phi(w) := \frac{\sum_{j=1}^J z_j^* P_j(x+h)}{P(x+h)} = \sum_{j=1}^J z_j^* \int_0^T \mathrm{e}^{-x(t)+x(s)-h(t)+h(s)} \, \mathrm{d}F_j(t).$$

Since $z^* \ge 0$ and the exponential function is convex, $\phi(w)$ is convex in $w \in \mathbb{R}^I$.

Let us show that $\nabla\phi(0) = 0$. To this end we compute

$$
\begin{aligned}
\frac{\partial\phi}{\partial w_i}(0) &= \sum_{j=1}^{J} z_j^* \int_0^T \mathrm{e}^{-x(t)+x(s)}(-h_i(t) + h_i(s))\,\mathrm{d}F_j(t) \\
&= \mathrm{e}^{x(s)} \sum_{j=1}^{J} z_j^* \left(-\int_0^T \mathrm{e}^{-x(t)} h_i(t)\,\mathrm{d}F_j(t) + h_i(s)\int_0^T \mathrm{e}^{-x(t)}\,\mathrm{d}F_j(t)\right) \\
&= \mathrm{e}^{x(s)} \left(-P(x)\sum_{j=1}^{J} a_{ij} z_j^* + h_i(s)\sum_{j=1}^{J} z_j^* P_j(x)\right),
\end{aligned}
\tag{B.20}
$$

where the last line uses (3.3) and (2.6) for each bond j. Using value matching (2.9) and the fact that the liability is a zero-coupon bond, we obtain

$$
h_i(s)\sum_{j=1}^{J} z_j^* P_j(x) = h_i(s)P(x) = \mathrm{e}^{-x(s)}h_i(s) = \int_0^T \mathrm{e}^{-x(t)}h_i(t)\,\mathrm{d}F(t) = P(x)b_i,
\tag{B.21}
$$

where the last equality uses (3.2). Combining (B.20) and (B.21), we obtain

$$
\nabla\phi(0) = b - Az^* = 0. \tag{B.22}
$$

Since ϕ is convex, it follows that $\phi(w) \geq \phi(0) = 1$ for all w, which implies $E(z^*, x + h) \geq 0$. □

References

Alvarez, F. and U. J. Jermann (2005). "Using asset prices to measure the persistence of the marginal utility of wealth". *Econometrica* 73.6, 1977–2016. DOI: 10.1111/j.1468-0262.2005.00643.x.

Ang, A., G. Bekaert, and M. Wei (2008). "The term structure of real rates and expected inflation". *Journal of Finance* 63.2, 797–849. DOI: 10.1111/j.1540-6261.2008.01332.x.

Balbás, A. and A. Ibáñez (1998). "When can you immunize a bond portfolio?" *Journal of Banking & Finance* 22.12, 1571–1595. DOI: 10.1016/S0378-4266(98)00070-3.

Beckermann, B. (2000). "The condition number of real Vandermonde, Krylov and positive definite Hankel matrices". *Numerische Mathematik* 85.4, 553–577. DOI: 10.1007/PL00005392.

Berge, C. (1963). *Topological Spaces*. Edinburgh: Oliver & Boyd.

Bergemann, D. and S. Morris (2005). “Robust mechanism design”. *Econometrica* 73.6, 1771–1813. DOI: 10.1111/j.1468-0262.2005.00638.x.

Bierwag, G. O. and C. Khang (1979). “An immunization strategy is a minimax strategy”. *Journal of Finance* 34.2, 389–399. DOI: 10.2307/2326978.

Black, F. and M. Scholes (1973). “The pricing of options and corporate liabilities”. *Journal of Political Economy* 81.3, 637–654. DOI: 10.1086/260062.

Bowden, R. J. (1997). “Generalising interest rate duration with directional derivatives: direction x and applications”. *Management Science* 43.5, 586–595. DOI: 10.1287/mnsc.43.5.586.

Brooks, B. and S. Du (2021). “Optimal auction design with common values: An informationally robust approach”. *Econometrica* 89.3, 1313–1360. DOI: 10.3982/ECTA16297.

Chambers, D. R., W. T. Carleton, and R. W. McEnally (1988). “Immunizing default-free bond portfolios with a duration vector”. *Journal of Financial and Quantitative Analysis* 23.1, 89–104. DOI: 10.2307/2331026.

Chapman, A. and A. Miyake (2018). “Classical simulation of quantum circuits by dynamical localization: analytic results for pauli-observable scrambling in time-dependent disorder”. *Physical Review A* 98 (1), 012309. DOI: 10.1103/PhysRevA.98.012309.

Crump, R. K. and N. Gospodinov (2022). “On the factor structure of bond returns”. *Econometrica* 90.1, 295–314. DOI: 10.3982/ECTA17943.

Diebold, F. X. and C. Li (2006). “Forecasting the term structure of government bond yields”. *Journal of Econometrics* 130.2, 337–364. DOI: 10.1016/j.jeconom.2005.03.005.

Du, S. (2018). “Robust mechanisms under common valuation”. *Econometrica* 86.5, 1569–1588. DOI: 10.3982/ECTA14993.

Dybvig, P. H., J. E. Ingersoll, and S. A. Ross (1996). “Long forward and zero-coupon rates can never fall”. *Journal of Business* 69.1, 1–25. DOI: 10.1086/209677.

Fisher, L. and R. L. Weil (1971). “Coping with the risk of interest-rate fluctuations: Returns to bondholders from naïve and optimal strategies”. *Journal of Business* 44.4, 408–431. DOI: 10.1086/295402.

Folland, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications*. 2nd ed. Hoboken, NJ: John Wiley & Sons.

Fong, G. H. and O. A. Vasicek (1984). “A risk minimizing strategy for portfolio immunization”. *Journal of Finance* 39.5, 1541–1546. DOI: 10.1111/j.1540-6261.1984.tb04923.x.

Gagliardini, P., P. Porchia, and F. Trojani (2008). “Ambiguity aversion and the term structure of interest rates”. *Review of Financial Studies* 22.10, 4157–4188. DOI: 10.1093/rfs/hhn092.

Gilboa, I. and D. Schmeidler (1989). “Maxmin expected utility with non-unique prior”. *Journal of Mathematical Economics* 18.2, 141–153. DOI: 10.1016/0304-4068(89)90018-9.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007). “The U.S. Treasury yield curve: 1961 to the present”. *Journal of Monetary Economics* 54.8, 2291–2304. DOI: 10.1016/j.jmoneco.2007.06.029.

Hicks, J. R. (1939). *Value and Capital*. Oxford, UK: Oxford University Press.

Ho, T. S. Y. (1992). “Key rate durations: measures of interest rate risks”. *Journal of Fixed Income* 2.2, 29–44. DOI: 10.3905/jfi.1992.408049.

Huettner, F. and M. Sunder (2012). “Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values”. *Electronic Journal of Statistics* 6, 1239–1250. DOI: 10.1214/12-EJS710.

Jacobi, C. G. J. (1841). “De formatione et proprietatibus determinatium.” *Journal für die reine und angewandte Mathematik* 22, 285–318.

Knight, K. (2019). “Elemental estimates, influence, and algorithmic leveraging”. In: *Nonparametric Statistics: 3rd ISNPS Avignon*, 219–231.

Kozak, S., S. Nagel, and S. Santosh (2020). “Shrinking the cross-section”. *Journal of Financial Economics* 135.2, 271–292. DOI: 10.1016/j.jfineco.2019.06.008.

Ledoit, O. and M. Wolf (2003). “Improved estimation of the covariance matrix of stock returns with an application to portfolio selection”. *Journal of Empirical Finance* 10.5, 603–621. DOI: 10.1016/S0927-5398(03)00007-0.

Litterman, R. B. and J. Scheinkman (1991). “Common factors affecting bond returns”. *Journal of Fixed Income* 1.1, 54–61. DOI: 10.3905/jfi.1991.692347.

Liu, Y. and J. C. Wu (2021). “Reconstructing the yield curve”. *Journal of Financial Economics* 142.3, 1395–1425. DOI: 10.1016/j.jfineco.2021.05.059.

Macaulay, F. R. (1938). *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1856*. National Bureau of Economic Research.

Mantilla-Garcia, D., L. Martellini, V. Milhau, and H. E. Ramirez-Garrido (2022). “Improving interest rate risk hedging strategies through regularization”. *Financial Analysts Journal* 78.4, 18–36. DOI: 10.1080/0015198X.2022.2095193.

Nawalkha, S. K. and N. J. Lacey (1988). “Closed-form solutions of higher-order duration measures”. *Financial Analysts Journal* 44.6, 82–84. DOI: 10.2469/faj.v44.n6.82.

Onatski, A. and C. Wang (2021). “Spurious factor analysis”. *Econometrica* 89.2, 591–614. DOI: 10.3982/ECTA16703.

Prisman, E. Z. and M. R. Shores (1988). “Duration measures for specific term structure estimations and applications to bond portfolio immunization”. *Journal of Banking and Finance* 12.3, 493–504. DOI: 10.1016/0378-4266(88)90011-8.

Redington, F. M. (1952). “Review of the principles of life-office valuations”. *Journal of the Institute of Actuaries* 78.3, 286–340. DOI: 10.1017/S0020268100052811.

Samuelson, P. A. (1945). “The effect of interest rate increases on the banking system”. *American Economic Review* 35.1, 16–27.

Sydyak, O. (2016). “Interest rate risk management and asset liability management”. In: *Handbook of Fixed-Income Securities*. Ed. by P. Veronesi. John Wiley & Sons. Chap. 7, 119–146. DOI: 10.1002/9781118709207.ch7.

Trefethen, L. N. (2019). *Approximation Theory and Approximation Practice*. Extended. Philadelphia, PA: Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9781611975949.

Vasicek, O. (1977). “An equilibrium characterization of the term structure”. *Journal of Financial Economics* 5.2, 177–188. DOI: https://doi.org/10.1016/0304-405X(77)90016-2.

Vayanos, D. and J.-L. Vila (2021). “A preferred-habitat model of the term structure of interest rates”. *Econometrica* 89.1, 77–112. DOI: 10.3982/ECTA17440.

Willner, R. (1996). “A new tool for portfolio managers: level, slope, and curvature durations”. *Journal of Fixed Income* 6.1, 48–59. DOI: 10.3905/jfi.1996.408171.

Zheng, H. (2007). “Macaulay durations for nonparallel shifts”. *Annals of Operations Research* 151, 179–191. DOI: 10.1007/s10479-006-0115-7.

Online Appendix

C No-arbitrage term structure model

The no-arbitrage term structure model of Ang et al. (2008) features multiple factors, regime switching, and closed-form solutions for bond prices, which is convenient for simulating yield curves. This appendix summarizes their model and presents parameter estimates based on our yield curve data.

C.1 Model and bond price formula

The equation numbers follow that of Ang et al. (2008). The model has three factors denoted by $X_t = (q_t, f_t, \pi_t)'$. The dynamics of factors follows the regime-dependent VAR process

$$X_{t+1} = \mu(s_{t+1}) + \Phi X_t + \Sigma(s_{t+1})\varepsilon_{t+1}, \tag{2}$$

where

$$\mu(s_t) = \begin{bmatrix} \mu_q \\ \mu_f(s_t) \\ \mu_\pi(s_t) \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_{qq} & 0 & 0 \\ \Phi_{fq} & \Phi_{ff} & 0 \\ \Phi_{\pi q} & \Phi_{\pi f} & \Phi_{\pi\pi} \end{bmatrix}, \quad \Sigma(s_t) = \begin{bmatrix} \sigma_q & 0 & 0 \\ 0 & \sigma_f(s_t) & 0 \\ 0 & 0 & \sigma_\pi(s_t) \end{bmatrix}, \tag{3}$$

and ε is IID $N(0, I_3)$. The regime s_t is a finite-state Markov chain taking values denoted by $k = 1, \dots, K$ with transition probability matrix $\Pi = (p_{kk'})$. The real short rate is given by

$$\widehat{r}_t = \delta_0 + \delta_1' X_t. \tag{4}$$

The regime-dependent price of risk is denoted by $\lambda(s_t) = (\lambda_f(s_t), \lambda_\pi(s_t))'$. Furthermore, define

$$\gamma_t = \gamma_0 + \gamma_1 q_t = \gamma_0 + \gamma_1 e_1' X_t, \tag{6}$$

where e_n denotes the n-th unit vector.

With this notation, the price of zero-coupon bonds can be obtained in closed form (Ang et al., 2008, Proposition B). For each maturity n, the nominal zero-coupon bond price in regime i and factor X is given by

$$P_n(i, X) = \exp(A_n(i) + B_n X), \tag{B1}$$

where the scalar $A_n(i)$ and the $M \times 1$ vector B_n can be computed as follows.

Let $M = 3$ be the number of factors and $M_1 = 2$ be the number of non-q factors. Partition B_n as $B_n = [B_{nq}; B_{nx}]$, where B_{nq} is a scalar and B_{nx} is 2×1. Similarly, let $\Sigma_x(i)$ be the lower 2×2 block of $\Sigma(i)$.

First, define $A_0(i) = 0$ and $B_0 = 0$. Then define $\{(A_n, B_n)\}_{n=1}^{\infty}$ recursively by

$$\begin{aligned} A_{n+1}(i) = -\delta_0 - B_{nq}\sigma_q\gamma_0 + \log \sum_j p_{ij} \exp\Big(A_n(j) + (B_n - e_M)'\mu(j) \\ - (B_{nx} - e_{M_1})'\Sigma_x(j)\lambda(j) + \frac{1}{2}(B_n - e_M)'\Sigma(j)\Sigma(j)'(B_n - e_M)\Big), \end{aligned} \tag{B2.a}$$

$$B_{n+1} = -\delta_1 + \Phi'(B_n - e_M) - B_{nq}\sigma_q\gamma_1 e_1. \tag{B2.b}$$

C.2 Data

We use end of the quarter yield data from Liu and Wu (2021) for the period of 1985:Q4 to 2022:Q4; a total of 149 quarterly observations.[14] The authors use a nonparametric approach to estimate the yield curve up to the 30-year maturity, which allows us to infer the long end of the yield curve consistently over time. The inflation data for the same period are obtained from the Bureau of Labor Statistics, from the *CPI for All Urban Consumers series* (seasonally adjusted).

We estimate the term structure model ourselves instead of using the parameters reported in Ang et al. (2008, Table III) to better reflect the evolution in yields over the last decade. In our dynamic hedging experiment, we need to infer the yields up to a maturity of 50 years. Estimating the term structure model, we sometimes find counterfactual steep declines in the yield curve for long maturities, depending on the maturities used for estimation. To mitigate this issue, we incorporate 50-year yields in the estimation, treating them as equivalent to the observed 30-year yields. This inclusion proves essential for generating yield curves that remain relatively "flat" over long horizons, thereby preventing the possibility of counterfactual steep declines at the long end of the yield curve. Additionally, we incorporate 1-year yields in the estimation to capture short-run dynamics.[15]

C.3 Parameter estimates

We consider the benchmark model IV$^{\text{C}}$ of Ang et al. (2008, §I.B.4). This model has four regimes. There are two state variables denoted by s^f, s^π, which both take

[14] https://sites.google.com/view/jingcynthiawu/yield-data

[15] Unlike Ang et al. (2008), we do not use additional yield data as overidentifying restrictions.

values in $\{1, 2\}$. The combined state s thus takes four values

$$\begin{aligned}
s = 1 &:= (s^f = 1, s^\pi = 1), \\
s = 2 &:= (s^f = 1, s^\pi = 2), \\
s = 3 &:= (s^f = 2, s^\pi = 1), \\
s = 4 &:= (s^f = 2, s^\pi = 2).
\end{aligned}$$

We also impose the following restrictions consistent with Ang et al. (2008):

$$\begin{aligned}
\delta_0 &= 0.0077, && \text{(mean of nominal short rate)} \\
\delta_1 &= (1, 1, \delta_\pi)', \\
\Phi_{fq} &= 0, \\
\mu_q &= 0, \\
\gamma_0 &= 0, \\
\lambda_\pi(s_t) &= 0.
\end{aligned}$$

We estimate the model using maximum likelihood using the parameters from Ang et al. (2008, Table III) as starting values. Table 3 below summarizes the resulting parameter estimates.

Table 3: Parameter estimates

Real short rate		δ_1		
	δ_0	q	f	π
	0.008	1.000	1.000	-0.199
Companion Form Φ		q	f	π
	q	0.962	0.000	0.000
	f	0.000	0.969	0.000
	π	-0.139	0.246	0.178
Moments of X_t				
			Regime 1	Regime 2
		$\mu_q \times 100$	0.000	0.000
		$\mu_f(s_t^f) \times 100$	-0.621	-0.020
		$\mu_\pi(s_t^\pi) \times 100$	-0.789	0.726
		$\sigma_q \times 100$	0.054	0.054
		$\sigma_f(s_t^f) \times 100$	0.400	0.108
		$\sigma_\pi(s_t^\pi) \times 100$	0.048	0.624
Prices of Risk			$\lambda_f(s_t^\pi)$	
		γ_1	Regime 1	Regime 2
		-84.137	-19.734	0.051
Transition Probabilities Π				
	$s_{t+1} = 1$	$s_{t+1} = 2$	$s_{t+1} = 3$	$s_{t+1} = 4$
$s_t = 1$	0.744	0.174	0.037	0.045
$s_t = 2$	0.685	0.216	0.052	0.047
$s_t = 3$	0.001	0.001	0.354	0.645
$s_t = 4$	0.000	0.000	0.020	0.980

D Miscellaneous results

D.1 Bias in the estimated yield curve

In our empirical application in Section 4, we assume that the forward rate is constant beyond the 30-year maturity, $f(t) = f(30)$ for all $t \geq 30$. As a result, the inferred date s yield curve with term $t \geq 30$ satisfies[16]

$$
\begin{aligned}
\widehat{y}_s(t) &:= \frac{1}{t}\int_0^t f_s(u)\,\mathrm{d}u = \frac{1}{t}\int_0^{30} f_s(u)\,\mathrm{d}u + \frac{1}{t}\int_{30}^t f_s(30)\,\mathrm{d}u \\
&= \frac{1}{t}\int_0^{30} f_s(u)\,\mathrm{d}u + f_s(30) - \frac{30}{t} f_s(30) \\
&= f_s(30) + O\left(\frac{1}{t}\right).
\end{aligned}
$$

Taking unconditional expectations and comparing to the true (unobserved) yield, we obtain

$$
\begin{aligned}
\mathbb{E}\left[\widehat{y}_s(t) - y_s(t)\right] &= \mathbb{E}\left[f_s(30) - y_s(t)\right] + O\left(\frac{1}{t}\right) \\
&= \mathbb{E}\left[f_s(30) - f_s(t)\right] + \mathbb{E}\left[f_s(t) - y_s(t)\right] + O\left(\frac{1}{t}\right). \qquad \text{(D.1)}
\end{aligned}
$$

Under integrability conditions on $y_s(t)$ and a mild stationarity assumption on bond returns, Alvarez and Jermann (2005, Proposition 5) show that

$$
\mathbb{E}\left[\lim_{t\to\infty} f_s(t)\right] = \mathbb{E}\left[\lim_{t\to\infty} y_s(t)\right]. \qquad \text{(D.2)}
$$

Using the dominated convergence theorem and (D.2) in (D.1), we get

$$
\mathbb{E}\left[\widehat{y}_s(t)\right] = \mathbb{E}\left[y_s(t)\right] + \mathbb{E}\left[f_s(30) - f_s(t)\right] + o(1).
$$

Hence, on average we estimate the correct yield plus a bias term that reflects the average gap between the 30-year forward rate and long forward rate.

D.2 Key rate duration matching

This appendix explains the key rate duration matching method of Ho (1992). The key rate duration of a bond with yield curve y and yield change Δ at time to

[16] Throughout we ignore the approximation error coming from misspecification of the forward rate model.

maturity t is defined by

$$\mathrm{KRD}(y, t, \Delta) := \frac{P(y_-) - P(y_+)}{2\Delta P(y)},$$

where $y_\pm$ denotes the yield curve after changing $y(t)$ to $y(t) \pm \Delta$ at a specific term t and linearly interpolating between the adjacent terms. Following the literature, we set the shift to $\Delta = 0.01$ (100 basis points). Figure 7 illustrates the procedure for a set of key rates on December 2, 2016.[17]

In our simulation, we consider five zero-coupon bonds and aim to match the liability's key rate exposures at the maturities of these bonds. In addition, we impose a value-matching constraint to ensure consistency with other immunization methods. This leads to six restrictions and five unknowns. To determine the optimal portfolio, we minimize the mean squared distance between the key rate exposures of the bond portfolio and those of the liability, subject to the value-matching constraint.

[17]The key rate duration of a zero-coupon bond with maturity t is equal to t and zero otherwise. Since we use linear interpolation after a key rate perturbation to keep the yield curve continuous, the key rate for a zero-coupon bond with maturity t is not exactly equal to t in our application.

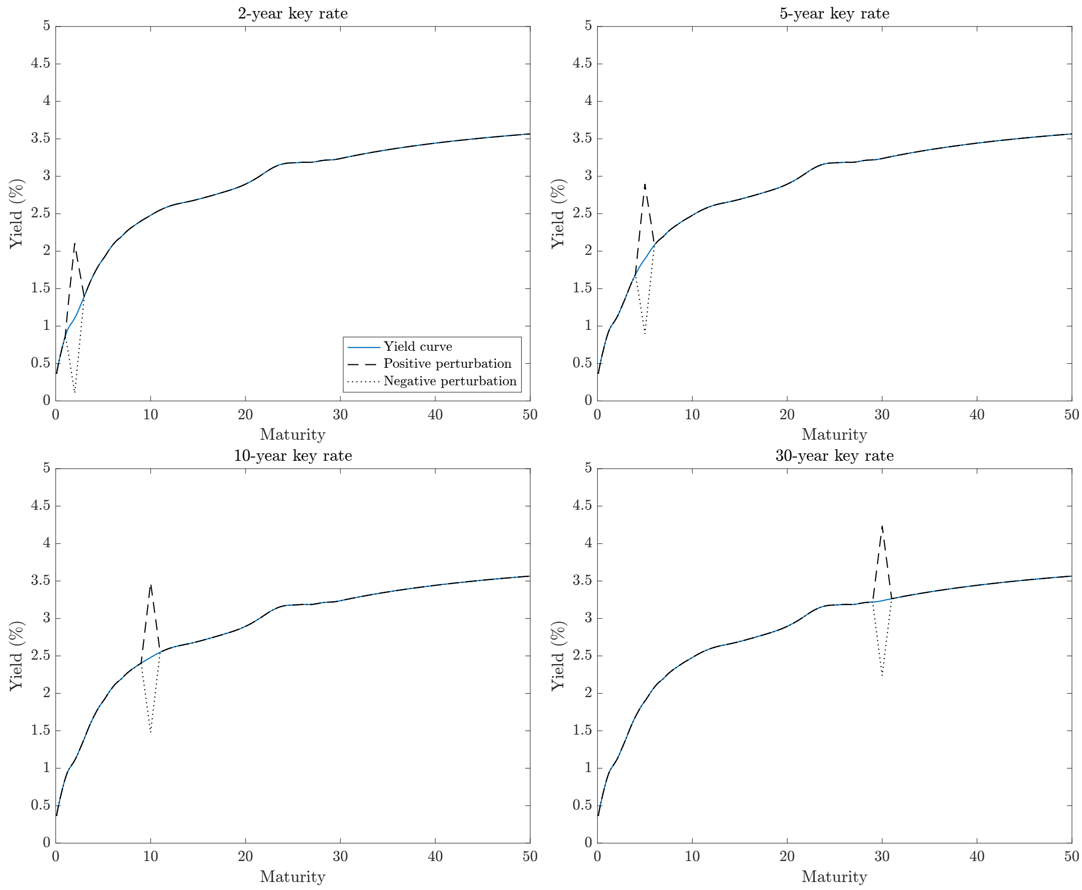


Figure 7: Key rate perturbations.

Note: The figures show positive and negative perturbations to the yield curve due to a 1% change in the respective key rate. We linearly interpolate the yields after a change in the key rate to ensure that the yield curve remains continuous. The true yield curve (in blue) is calculated on December 2, 2016.