Credit Portfolio Selection with Decaying Contagion Intensities

Lijun Bo * Agostino Capponi † Peng-Chu Chen ‡

Abstract

We develop a fixed income portfolio framework capturing the exponential decay of contagious intensities between successive default events. We show that the value function of the control problem is the classical solution to a recursive system of second-order uniformly parabolic Hamilton-Jacobi-Bellman (HJB) partial differential equations (PDEs). We analyze the interplay between risk premia, decay of default intensities, and their volatilities. Our comparative statics analysis finds that the investor chooses to go long only if he is capturing enough risk premia. If the default intensities deteriorate faster, the investor increases the size of his position if he goes short, or reduces the size of his position if he goes long.

AMS 2000 subject classifications: Primary: 91G40; secondary: 91G10, 93E20

Keywords: fixed income investment, decay of default intensities, dynamic programming, parabolic PDEs

1 Introduction

The fixed income market, including public debt securities as well as corporate bonds issued by institutions (financial and otherwise), reached approximately $93 trillion according to the report "Mapping Global Capital Markets 2011" released by the McKinsey Global Institute. This corresponds to almost twice the capitalization of the equity market. Fixed-income securities are attractive for many investors because they generate a steady flow of income if kept until maturity, and coupon payments are usually known in advance. This makes the optimal fixed income selection problem at least as significant as the equity selection problem. Nevertheless, most studies have been devoted to equity portfolio problems (e.g., the seminal work of Merton (1969) and the subsequent developments).

When investing in fixed income securities, the risk of default strongly impacts the distribution of portfolio returns and consequently affects the optimal investment decisions. Although the investor optimizes his expected utility under the probability measure describing the actual distribution of risk factors, he must take into account that prices of fixed income securities are observed under the pricing measure. Consequently, it becomes crucial to specify a dynamic model governing the relation between actual and risk-neutral default intensities. Such a relation has been analyzed, for instance, by Giesecke et al. (2014) who solve the static selection problem of a credit swaps portfolio taking into account solvency and trading constraints.

The vast majority of the literature studying fixed income portfolio selection problems has focused on the self-exciting feature, i.e., the impact that other defaults in the market have on the optimal investment in securities referencing the surviving firms. Those studies include Bo and Capponi (2016), who develop a dynamic portfolio optimization framework for credit derivatives with interacting default intensities, and Bo and Capponi (2017) who generalize it to account for robustness against misspecification of the parameters of the default intensity model. In their studies, the authors assume that the default intensities are piecewise constant, and only jump at the occurrence of a default event. In the context of optimal investment-consumption, self-excitation has been considered by Ait-Sahalia and Hurd (2016). They

*Email: lijunbo@ustc.edu.cn, School of Mathematical Sciences, University of Science and Technology of China, Hefei, Anhui Province, 230026, China, and Wu Wen Tsun Key Laboratory of Mathematics, Chinese Academy of Science, Hefei, Anhui Province 230026, China.

†E-mail: ac3827@columbia.edu, Department of Industrial Engineering and Operations Research, Columbia University, New York, 10027, NY, USA.

‡E-mail: pcchen@hku.hk, Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Pokfulam, Hong Kong.

restrict their attention to default-free stocks and use Hawkes processes to model self-excitation in the jump component of the price process.[1]

We consider a portfolio model in which the intensities between defaults exponentially decay toward their long run mean levels. To the best of our knowledge, ours is the first study to analyze the impact that the deteriorations of intensities between defaults have on the optimal investment decisions. The time decay of default intensities is supported by empirical research. Azizpour, Giesecke, and Schwenkler (2016) use historical corporate default data and show that the impact of default events on the default probabilities of other firms is statistically significant. Their analysis predicts that such an impact fades away with time, and the default intensity mean reverts to its long run average. In our default model, firms are exposed to common sources of risk representing fluctuations in the macro-economic environment modeled as diffusion processes. Moreover, a default event may lead to a jump of the surviving firms' default intensities because of direct contagion effects. For example, the creditors of the defaulted firm may only be able to recover a fraction of their interbanking claims, or their business activities may be negatively affected if they belong to the same industrial sector. We use mean-reverting diffusion processes enhanced with self-excitation to capture all the effects discussed above.

We study the dynamic investment problem of a power investor who distributes his wealth among risky bonds and the money market account. Using the dynamic programming principle, we show that the value function of the control problem solves a recursive system of second-order uniformly parabolic HJB-PDEs, each corresponding to a default state of the economy. Concretely, the optimal expected utility achievable by an investor in a default state depends on all optimal expected utilities that he can achieve if a default event were to occur. The latter would have a contagious effect on his portfolio of bond securities, and also decrease the set of available investment opportunities. We investigate the existence of classical solutions to this system of PDEs, defined on the positive real line, using the Bellman approximation technique in the policy space; see for example Davis and Lieo (2013) and Fleming and Rishel (1975) for its application in the stochastic control literature. We prove a verification theorem, showing the equivalence between the value function and the solution to the recursive system of HJB-PDEs, after establishing the uniform integrability of the related truncated family of wealth processes. In the degenerate case, i.e., when firms' default intensities are not driven by diffusive risk factors but only jump at other firms' defaults, we recover a closed-form representation for the optimal feedback strategies and the solution of the (degenerate) recursive system of HJB-PDEs. In this case, the optimal bond investment strategy is given by the product of two terms, (i) the inverse of a matrix measuring the bond price depreciations experienced at the default events, and (ii) a vector capturing the relation between the optimal expected utility (value function) of the investor in the current default state and in future states reached when a new firm defaults.

We conduct a numerical study to analyze the dependence of the investor's strategies on default risk premia, speed of decay of default intensities, and their volatilities. We consider a minimal market model consisting of two risky bonds to highlight the primary economic forces driving investment decisions. We find that the investor trades-off the benefit from holding the bond security and capturing the default risk premium with the cost of a negative return in case the bond security defaults. Of particular interest is the impact that the speed of the default intensity's decay between successive default events has on the investment decisions. Depending on whether the investor goes long or short in the bond security, the speed of decay of the default intensity toward its long run mean may have a different effect. More specifically, it leads the investor to increase the size of his position if he goes short, because he does not capture enough premium for holding the risky bond. Under these circumstances, the bond appreciates in value (lower default risk) if the default intensity decays faster, and thus the investor would be able to sell it at a higher price. In contrast, if the investor goes long because he captures high compensation for holding the bond security, he would reduce his holdings given that the bond becomes more expensive if the default intensity deteriorates faster. The investor benefits from a higher volatility of the firm "1"'s default intensity if he goes short. This not only directly increases the probability that firm "1" defaults, but also indirectly increases the probability that firm "2" defaults because it increases the likelihood that the default intensity of firm "2" ramps up as a result of the default of firm "1". On the other hand, an investor who is long in the bond security would reduce his position if the default intensity process becomes more volatile because of the higher probability of default.

[1]Other studies have considered credit portfolio selection without accounting for self-excitation. Kraft and Steffensen (2008) consider an investor who can allocate his wealth across multiple defaultable bonds, assuming constant default intensity. Wise and Bhansali (2002) analyze optimal allocation of capital to corporate bonds under a structural default model.

The rest of the paper is organized as follows. Section 2 develops the model. Section 3 formulates the optimal portfolio problem and studies the optimal bond investment strategy. Section 4 analyzes the recursive system of HJB equations and proves a verification theorem. Section 5 develops a numerical analysis. Section 6 concludes the paper. Technical proofs are delegated to Appendix A. The details of the finite difference method used in the numerical implementation are reported in Appendix B.

Notations and definitions. We give notations and definitions used throughout the paper. Let $N \geq 2$ be an integer and $\mathcal{S} := \{0,1\}^N$. We also adopt the following shorthand notation for the following domain: $\mathbb{R}_+ := (0,\infty)$. The vector $z = (z_1,\ldots,z_N) \in \mathcal{S}$ is used to denote the default state of the portfolio, with $z_i = 0$ if the firm i is alive and $z_i = 1$ if it has defaulted. For each $z \in \mathcal{S}$ such that $z_j = 0$, we use

$$z^j := (z_1,\ldots,z_{j-1},1-z_j,z_{j+1},\ldots,z_N), \qquad j=1,\ldots,N \tag{1}$$

to denote the vector obtained from z by setting its j-th component to one. Let $m \in \{1,\ldots,N\}$ and $j_1,\ldots,j_m \in \{1,\ldots,N\}$ be m distinct integers. Given $z \in \mathcal{S}$ such that $z_{j_1} = \cdots = z_{j_m} = 0$, we use the shorthand notation $z^{j_1,\ldots,j_m} := \big((z^{j_1})^{\cdots}\big)^{j_m}$ for the vector obtained from z by setting its components $j_1, j_2, \ldots j_m$ to one. In other words, $z^{j_1,\ldots,j_m}$ denotes a default state where the firms $j_1, j_2, \ldots, j_m$ have defaulted. We set $z^{j_1,\ldots,j_m} = z$ if $m = 0$. Consider a function $f(t,x,z)$ with $(t,x,z) \in [0,T] \times \mathbb{R}_+^N \times \mathcal{S}$. We introduce the shorthand notation $f_{j_1,\ldots,j_m}(t,x) := f(t,x,0^{j_1,\ldots,j_m})$, where 0 denotes the N-dimensional row zero vector. Moreover, if $j \notin \{j_1,\ldots,j_m\}$, we set $g_{j;j_1,\ldots,j_m}(t,x) := g_j(t,x,0^{j_1,\ldots,j_m})$ for a given function $g_j(t,x,z)$ depending on the default state and on the index j of the firm. We use $\mathbb{E}$ to denote the expectation under the risk-neutral measure $\mathbb{Q}$, and $\mathbb{E}^{\mathbb{P}}$ for the expectation under the actual probability measure $\mathbb{P}$.

2 The Model

We introduce the default model with decaying contagion intensities in Section 2.1, and describe the risky bond securities in Section 2.2.

2.1 The Default Model

We consider a model in which default intensities decay exponentially between successive default events. A default event induces a sudden increase in the default intensities of all surviving firms. Moreover, the impact of such a contagion effect decreases with the passage of time, as empirically observed by Azizpour, Giesecke, and Schwenkler (2016) via a statistical analysis of corporate default timing data from 1/1/1970 to 12/31/2012. A firm's default intensity is driven by a source of risk factors common to all firms. These are macroeconomic factors, such as the Treasury term structure level, its slope, and trailing returns of stock price indices reflecting a broad class of the different industries in the economy. Duffie, Saita, and Wang (2007) consider historical data of US-listed industrial firms, and find that the estimated term structures of default intensities depend significantly, in level and shape, on the current macroeconomic state.

The default intensity dynamics of each firm $j = 1,\ldots,N$ is governed by the following stochastic differential equation (SDE)

$$dX_j(t) = \big(\kappa_j - \nu_j X_j(t)\big)dt + \sum_{k=1}^{K} \sigma_{jk}(X_j(t))dW_k(t) + dL_j(t), \tag{2}$$

where $W(t) = (W_k(t))_{k=1,\ldots,K}^{\top}$ with $t \geq 0$ is a K-dimensional Brownian motion, $\nu_j > 0$ is the mean reversion speed of the default intensity of firm j, and $\frac{\kappa_j}{\nu_j}$ is its mean reversion level. The volatility matrix of the default intensity of firm j is given by $\sigma_{jk}(x_j) = \sigma_{jk}\sqrt{x_j}$ for $x_j > 0$ and $\sigma_{jk} > 0$. We assume that $2\kappa_j \geq \sum_{k=1}^{K} \sigma_{jk}^2$ for $j = 1,\ldots,N$ so that the default intensity cannot reach zero. The impact of past defaults on the default intensity of firm j is captured by the following process

$$L_j(t) := \sum_{i=1}^{N} w_{ij} Z_i(t), \qquad t \geq 0. \tag{3}$$

The j-th entry of the weight vector $w_j = (w_{ij})_{i=1,\ldots,N} \in \mathbb{R}_+^N$ measures the extent to which the default of firm i impacts the default intensity of firm j. If the weight w_{ij} is high, the default of firm i leads to a substantial increase of the default intensity of firm j. In this case, the probability that firm j defaults soon after firm i is quite high. In practice, we expect that if firm j is a direct counterparty of firm i or it belongs to the same industrial sector, then it will be affected more by the default of firm i relative to another firm k which is in a different business sector and does not have direct contractual relationships with it.[2] The process $Z_j(t) := \mathbf{1}_{\tau_j \le t}$ is the default indicator process of firm j. Let $Z(t) = (Z_1(t), \ldots, Z_N(t))$, for $t \ge 0$, be the N-dimensional default indicator process. We treat (X, Z) as a joint Markov process with state space $\mathbb{R}_+ \times \mathcal{S}$. Further, for $j = 1, \ldots, N$ and $t > 0$, $Z(t)$ transits to a neighbouring state $Z^j(t) := (Z_1(t), \ldots, Z_{j-1}(t), 1 - Z_j(t), Z_{j+1}(t), \ldots, Z_N(t))$ at the state-dependent rate $\mathbf{1}_{\{Z_j(t)=0\}} X_j(t)$. Let $\mathcal{F}_t := \sigma(X(u), Z(u);\ u \le t)$ for $t \ge 0$. The global market information is given by the right continuous filtration $\mathbb{G} = (\mathcal{G}_t)_{t\ge 0}$ where $\mathcal{G}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$. Using the Dykin's formula (see e.g., Eq. (10.13) in Rogers and Williams (2000), pp. 254), it follows that

$$M_j(t) := Z_j(t) - \int_0^{t\wedge\tau_j} X_j(u)du, \qquad t \ge 0 \tag{4}$$

is a $\mathbb{Q}$-martingale w.r.t. the total market information $\mathbb{G}$.

It has been empirically shown by Yu (2005) that if the common factor is properly calibrated, then the default intensity model in Eq. (2) is able to reproduce the levels of default correlations observed historically. Moreover, in a single firm context, the square root diffusion model has been successfully calibrated to the term structure of credit default swaps, and also used to accurately price caps for the interest rate market and options on CDSs; see Brigo and Alfonsi (2004) for additional discussions. The square root dynamics enhanced with feedbacks from defaults has been used by Errais, Giesecke, and Goldberg (2010) (see Example 4.2 therein) in the pricing context. Therein, they consider a top down model for the portfolio default intensity. By contrast, in this paper we consider a bottom-up multivariate default intensity model. The construction of the default intensity model (X, Z) is nontrivial because of the coupling between default intensities and default states. The following lemma proves the existence of such a credit model.

Lemma 2.1. *There exists a joint Markov process $(X(t), Z(t))_{t\ge0}$ satisfying (2), (3) and (4).*

The proof is based on a recursive procedure. We report the details of such a construction in Appendix A. Throughout the paper let $\mu(x) = (\kappa_j - \nu_j x_j)_{j=1,\ldots,N}^\top$ and $\sigma(x) = (\sigma_{jk}\sqrt{x_j})_{j=1,\ldots,N,k=1,\ldots,K}$ for $x \in \mathbb{R}_+^N$. We assume that $\det(\sigma\sigma^\top(x)) \ne 0$ for $x \in \mathbb{R}_+^N$. We refer to the model with default intensity given by Eq. (2) as the non-degenerate default model (see also Lemma 3 in Heath and Schweizer (2001)).

Remark 2.2. *When the volatility in Eq. (2) is zero, the non-degeneracy condition $\det(\sigma\sigma^\top(x)) \ne 0$ fails to hold. The dynamics of the default intensity reduces to*

$$dX_j(t) = \big(\kappa_j - \nu_j X_j(t)\big)dt + dL_j(t), \tag{5}$$

for all $j = 1, \ldots, N$. We refer to it as the degenerate default intensity model. For $s \ge t$, define

$$s \to \mathcal{X}_s^{(t,x_j,j)} := \frac{\kappa_j}{\nu_j} + \left(x_j - \frac{\kappa_j}{\nu_j}\right) e^{-\nu_j(s-t)}. \tag{6}$$

Then Eq. (5) admits the closed-form solution given by, for $j = 1, \ldots, N$,

$$\begin{aligned} X_j(t) &= e^{-\nu_j t} x_j + \frac{\kappa_j}{\nu_j}\big(1 - e^{-\nu_j t}\big) + \sum_{i=1}^N w_{ij} \int_0^t e^{-\nu_j(t-s)} dZ_i(s) \\ &= \underbrace{\mathcal{X}_t^{(0,x_j,j)}}_{\text{mean reverting}} + \underbrace{\sum_{i=1}^N w_{ij} e^{-\nu_j(t-\tau_i)} Z_i(t)}_{\text{default contagion}}. \end{aligned} \tag{7}$$

[2] Davis and Lo (2001) introduce a stylized credit contagion model. They use independent and identically distributed Bernoulli random variables to model both direct and contagious defaults. Kraft and Steffensen (2007) develop a Markov chain model to capture the contagious effects of defaults on the intensities of the surviving firms, and apply their results to price credit contingent claims. Yu (2007) consider an interacting intensity model with cyclical dependence among defaults, where firms hold the other firms' debt and are thus affected by their defaults. He develops a simulation method for generating the correlated default times under different structures of interaction among defaults.

From Eq. (7), it can be seen that the default intensity tends to mean revert to its long-run level given by $\frac{\kappa_j}{\nu_j} > 0$ between two consecutive default events. This captures the time decaying effect of default intensities. When a firm i defaults, the default intensity of firm j instantaneously jumps upward. As the time elapsed since the default of firm i increases, the contagion effect is reduced by an exponential factor.

2.2 The Portfolio Securities

The investor can allocate his wealth among the following securities:

- **Money market account**. The investor borrows and lends at constant risk-free rate $r > 0$. Hence the time-t price of one share of his account, denoted by $B(t)$, is given by $B(t) = e^{rt}$ for $t \geq 0$.

- **Risky bonds**. We consider N coupon paying bonds with unit notional, underwritten by risky firms. The bond underwritten by the i-th firm has maturity $T_i > 0$ and generates the following dividend process, for $i = 1, \ldots, N$,

$$D_i(t) := \underbrace{\int_0^t C_i(1 - Z_i(u))du}_{\text{coupon payments}} + \underbrace{\int_0^t R_i(Z(u-))dZ_i(u)}_{\text{recovery amount}} + \underbrace{(1 - Z_i(T_i))\mathbf{1}_{t \geq T_i}}_{\text{terminal payoff}}. \tag{8}$$

where $C_i \geq 0$ is the coupon payment of the i-th bond so that $\int_0^t C_i(1 - Z_i(u))du$ is the continuous stream of coupon payments received by the bond holder until the earliest of time t and the default of firm i. Moreover, $R_i(z) \in [0, 1)$ is the recovery rate paid at the default time τ_i. This depends on the default state $z \in \mathcal{S}$ of the portfolio, in line with empirical evidence suggesting that the fraction of defaulting firms in the economy and average recovery rates are correlated (see e.g., Altman et al. (2015) and Acharya, Bharath, and Srinivasan (2007) who provide support for a negative correlation). The quantity $(1 - Z_i(T_i))\mathbf{1}_{t \geq T_i}$ is the unit notional payment received by the bond holder at the maturity T_i if the firm i has not defaulted. Assuming that the bond market is free of arbitrage opportunities, the bond price is equal to the expected discounted payoff under the risk-neutral probability measure $\mathbb{Q}$. Hence the ex-dividend price of the i-th bond at time $t \leq T_i$ is given by

$$P_i(t) = \mathbb{E}_t\left[\int_t^{T_i} e^{-r(u-t)} dD_i(u)\right]. \tag{9}$$

The conditional expectation $\mathbb{E}_t[\cdot]$ is taken under the risk-neutral measure $\mathbb{Q}$ and conditional on the information set available by time t. From the price formula (9), we deduce that

$$\begin{aligned} P_i(t) = \mathbb{E}_t\left[\int_t^{T_i} C_i(1 - Z_i(u))e^{-r(u-t)}du\right] + \mathbb{E}_t\left[\int_t^{T_i} e^{-r(u-t)} R_i(Z(u-))dZ_i(u)\right] \\ + \mathbb{E}_t\left[(1 - Z_i(T_i))e^{-r(T_i - t)}\right]. \end{aligned} \tag{10}$$

Despite the complex default dependence structure across firms, we can characterize the bond prices in terms of classical solutions to a recursive system of Feynman-Kac's PDEs in the non-degenerate default case (see Remark A.1). Further, the price dynamics of the risky bonds are given in the following proposition.

Proposition 2.3. *Let $i = 1, \ldots, N$ and $t \in [0, T_i)$. Then the dynamics of the i-th risky bond price process is given by*

$$\begin{aligned} dP_i(t) = \underbrace{\left[rP_i(t) - (1 - Z_i(t))\left(C_i + R_i(Z(t))X_i(t)\right)\right]dt}_{\text{idiosyncratic price dynamics component}} \\ + \underbrace{P_i(t)\sum_{k=1}^{K}\left(\sum_{j=1}^{N} H_{(i,j)}(t, X(t), Z(t))\sigma_{jk}(X(t))\right) dW_k(t)}_{\text{systematic component}} \end{aligned}$$

$$+ P_i(t-) \underbrace{\sum_{j=1}^{N} \bar{H}_{(i,j)}(t, X(t-), Z(t-)) dM_j(t)}_{\text{contagion influence due to default of firm j}}$$

with terminal payoff $P_i(T_i) = 1 - Z_i(T_i)$. *For* $i, j = 1, \ldots, N$, $k = 1, \ldots, K$, *and* $(t, x, z) \in [0, T_i] \times \mathbb{R}_+^N \times \mathcal{S}$,

$$H_{(i,j)}(t, x, z) := \frac{\frac{\partial F_i(t,x,z)}{\partial x_j}}{F_i(t, x, z)}, \qquad \bar{H}_{(i,j)}(t, x, z) := \frac{F_i(t, x + w_j, z^j)}{F_i(t, x, z)} - 1, \tag{11}$$

where the price function $F_i(t, x, z)$ *is given by* (A.5) *in Appendix A.*

The quantity $\sum_{k=1}^{K} (\sum_{j=1}^{N} H_{(i,j)}(t, X(t), Z(t)) \sigma_{jk}(X(t))) dW_k(t)$ captures the influence of past defaults on the volatility of the i-th bond price process. The third term on the r.h.s. of $dP_i(t)$, instead, reflects the instantaneous impact of other firms' defaults on the i-th bond price.

3 The Optimal Investment Problem

We formulate the utility maximization problem in Section 3.1. We give the dynamic programming formulation in Section 3.2. We analyze the optimal bond investment strategy in Section 3.3.

3.1 The Optimization Problem

The investor dynamically allocates his wealth across risky bonds, and the risk-free money market account. Denote by $\psi_i(t)$ the number of shares that the investor holds in the i-th bond at time t ($\psi_i(t) > 0$ if he is long, and $\psi_i(t) < 0$ if he is short). Further, we use $\psi_B(t)$ to denote the number of shares held in the money market account at time t. Let $T \in (0, \min_{i=1,\ldots,N} T_i)$ be a finite horizon. Such an assumption comes without any loss of generality because we can always remove from the portfolio those bonds which have already matured by the time the investment is made. The wealth of the investor at time $t \in [0, T]$ is given by

$$V^{\psi}(t) = \sum_{i=1}^{N} \psi_i(t) P_i(t) + \psi_B(t) B(t). \tag{12}$$

For future purposes, we set $\psi(t) := (\psi_i(t))_{i \in \{1,\ldots,N,B\}}$ for $t \in [0, T]$. The wealth is obtained multiplying the holdings of the investor in each security by its corresponding price. As usual, we require the portfolio process ψ to be $\mathbb{G}$-predictable. A $\mathbb{G}$-predictable portfolio process $\psi = (\psi(t))_{t \in [0,T]}$ is said to be *self-financing* if $V^{\psi}(t) = V^{\psi}(0) + \Upsilon^{\psi}(t)$, where the gains process $\Upsilon^{\psi}(t)$, $t \in [0, T]$, is given by

$$\Upsilon^{\psi}(t) = \sum_{i=1}^{N} \int_0^t \psi_i(u) d\big(P_i(u) + D_i(u)\big) + \int_0^t \psi_B(u) dB(u). \tag{13}$$

Moreover, for $t \in [0, T]$, we define $\tilde{\pi}_i(t) := \frac{\psi_i(t) P_i(t-)}{V^{\psi}(t-)}$, $i = 1, \ldots, N$, and $\tilde{\pi}_B(t) := \frac{\psi_B(t) B(t)}{V^{\psi}(t-)} = 1 - \sum_{i=1}^{N} \tilde{\pi}_i(t)$ to denote the proportion of wealth invested in the i-th risky bond, and in the money market account respectively.

We consider a risk-averse investor who wants to find the optimal admissible strategy, i.e the one maximizing his expected utility from terminal wealth under the probability measure $\mathbb{P}$ which describes the actual distribution of risk factors. In other words, he wants to optimize the criterion

$$\mathbb{E}^{\mathbb{P}} \left[U \left(V^{\tilde{\pi}}(T) \right) \right] \tag{14}$$

over all admissible strategies $\tilde{\pi} \in \tilde{\mathcal{U}}_0$ which will be introduced in the following subsection. We choose the utility function $U : \mathbb{R}_+ \to \mathbb{R}$ given by $\frac{1}{\gamma} v^{\gamma}$ for $v > 0$ and $\gamma \in (0, 1)$. Notice the key role played by the two probability measures in the optimization criterion (14). Although the bond prices are given under the risk-neutral measure (see Eq. (9)), the investor wishes to optimize his expected utility under the measure describing the actual distribution of risk factors. Consequently, it becomes important to specify a dynamic model governing the relation between actual and risk-neutral default intensities. For $k = 1, \ldots, K$, and $i = 1, \ldots, N$, let $\phi_k(t, x, z)$ and $h_i(t, x, z)$ be sufficiently regular functions in $(t, x) \in [0, T] \times \mathbb{R}_+^N$ for each default state $z \in \mathcal{S}$, taking values respectively on $\mathbb{R}$ and $(-1, \infty)$. Under the actual probability measure $\mathbb{P}$, it holds that for each $j = 1, \ldots, N$, and $k = 1, \ldots, K$,

$$
\begin{aligned}
W_k^{\mathbb{P}}(t) &:= W_k(t) - \int_0^t \phi_k(u, X(u), Z(u))du, \qquad t \in [0, T], \quad \text{and} \\
M_j^{\mathbb{P}}(t) &:= M_j(t) - \int_0^{t\wedge\tau_j} X_j(u)h_j(u, X(u), Z(u))du, \qquad t \in [0, T]
\end{aligned} \tag{15}
$$

are $\mathbb{P}$-martingales. Hence, the coefficients ϕ_k and h_j represent, respectively, the market price of (diffusion) risk and the default risk premium. The formal statement on the relation between risk-neutral probability measure $\mathbb{Q}$ and the actual probability measure $\mathbb{P}$ is given in Lemma A.2 of Appendix A.

3.2 Dynamic Programming Formulation

The objective of this section is to derive the HJB equation associated with the control problem in (14). Let $\tilde{\pi} \in \tilde{\mathcal{U}}_0$ and $t \in [0, T]$. Using (13), the wealth process admits the dynamics

$$
\frac{dV^{\tilde{\pi}}(t)}{V^{\tilde{\pi}}(t-)} = \sum_{i=1}^N \tilde{\pi}_i(t) \frac{d\big(P_i(t) + D_i(t)\big)}{P_i(t-)} + \tilde{\pi}_B(t)\frac{dB(t)}{B(t)}, \tag{16}
$$

i.e., the relative change in wealth is given by the gains from the bond investment, and from interest rates proceeds. Using (16) and Lemma A.3 in Appendix A, we obtain the following lemma which characterizes the dynamics of the wealth process.

Lemma 3.1. *Let $\tilde{\pi} \in \tilde{\mathcal{U}}_0$ and $t \in [0, T]$. Under the probability measure $\mathbb{P}$ describing the actual distribution of risk factors, the dynamics of the wealth process (16) admits the following representation*

$$
\begin{aligned}
\frac{dV^{\tilde{\pi}}(t)}{V^{\tilde{\pi}}(t-)} = &\Bigg\{ r + \sum_{i=1}^N \tilde{\pi}_i(t) \left(\sum_{k=1}^K \left(\sum_{j=1}^N H_{(i,j)}(t, X(t), Z(t))\sigma_{jk}(X(t)) \right) \phi_k(t, X(t), Z(t)) \right) \\
&+ \sum_{j=1}^N \left(\sum_{i=1}^N \tilde{\pi}_i(t) G_{(i,j)}(t, X(t), Z(t)) \right) (1 - Z_j(t))X_j(t)h_j(t, X(t), Z(t)) \Bigg\} dt \\
&+ \sum_{k=1}^K \left\{ \sum_{i=1}^N \tilde{\pi}_i(t) \left(\sum_{j=1}^N H_{(i,j)}(t, X(t), Z(t))\sigma_{jk}(X(t)) \right) \right\} dW_k^{\mathbb{P}}(t) \\
&+ \sum_{j=1}^N \left(\sum_{i=1}^N \tilde{\pi}_i(t) G_{(i,j)}(t, X(t-), Z(t-)) \right) dM_j^{\mathbb{P}}(t).
\end{aligned} \tag{17}
$$

We recall that $H_{(i,j)}(t, x, z)$ is defined by (11) and $G_{(i,j)}(t, x, z)$ is defined by (A.19) in Appendix A.

Definition 3.1. *Let $t \in [0, T]$. The t-admissible control set $\tilde{\mathcal{U}}_t = \tilde{\mathcal{U}}_t(v, x, z)$, $(v, x, z) \in \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{S}$, is a class of $\mathbb{G}$-predictable feedback trading strategies of the form*

$$
\tilde{\pi}(u) = (\tilde{\pi}_i(u))_{i=1,\ldots,N} = \left(\pi_i\left(u, V^{\tilde{\pi}}(u-), X(u-), Z(u-)\right)\right)_{i=1,\ldots,N}, \quad u \in [t, T],
$$

where $\pi_i(\cdot)$ is the i-th deterministic locally bounded feedback control function taking values on $\mathcal{J}$ for $[t, T] \times \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{S}$. The admissible set is given by

$$
\mathcal{J} := \left\{ \pi \in \mathbb{R}^N;\ 1 + \sum_{i=1}^N \pi_i \theta_{ij} \geq 0,\ j = 1, \ldots, N,\ \forall\ \theta_{ij} \in [m_{ij}, M_{ij}] \right\}.
$$

For $i, j = 1, \ldots, N$, $G_{(i,j)}(t, x, z) \in [m_{ij}, M_{ij}]$ for some finite constants $m_{ij} < 0 < M_{ij}$ in Remark A.1.

Because of the wealth dynamics (17), the constraints imposed on the $\mathcal{J}$-valued feedback control function $\pi_i(\cdot)$ guarantee that the wealth process stays nonnegative when jumps due to defaults occur. In the sequel, we also use $\mathcal{U}_t$ to denote the set of t-admissible feedback control functions $\pi = (\pi_i)_{i=1,\dots,N}^\top$. In the remainder of the paper, we write $V^\pi(t)$, where $\pi \in \mathcal{U}_0$, to emphasize the dependence of the wealth on the strategy. We consider the following dynamic optimization problem. For $(t, v, x, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{S}$, define the value function by

$$\eta(t, v, x, z) := \sup_{\pi \in \mathcal{U}_t} \mathbb{E}^{\mathbb{P}}\Big[U\big(V^\pi(T)\big)\big|V^\pi(t) = v, X(t) = x, Z(t) = z\Big] \tag{18}$$

representing the optimal expected utility under the actual probability measure $\mathbb{P}$ and conditioned on the information set available at time t.

Remark 3.2. *Differently from the optimization problem considered by Ait-Sahalia and Hurd (2016) where self-excitation is in the jump component of the default-free stock price process, in our case the occurrence of defaults affects the default intensities of the surviving firms and reduces the universe of investment securities. As a consequence, $(V^\pi(t), X(t), Z(t))$ is Markovian but $(V^\pi(t), X(t))$ is not. This implies that our HJB equation takes a recursive form and its solution depends on the solutions of the HJB-PDEs associated with the control problem in any possible default state $Z(t)$ of the portfolio (see Eq. (22) and the expression of the recursive Hamiltonian given in Eq. (23)). By contrast, the absence of default risk invalidates the recursivity property of the HJB equation in Ait-Sahalia and Hurd (2016).*

Assuming that the above defined value function is $C^{1,2,2}$ in (t, v, x) for each default state $z \in \mathcal{S}$ (we will prove rigorously that this is the case in the verification theorem), it follows from standard arguments that it satisfies the following HJB equation: for $(t, v, x) \in [0, T) \times \mathbb{R}_+ \times \mathbb{R}_+^N$,

$$\sup_{\pi \in \mathcal{U}} \left(\frac{\partial}{\partial t} + \mathcal{L}_c^\pi + \mathcal{L}_J^\pi + \hat{\mathcal{A}}\right) \eta(t, v, x, z) = 0 \tag{19}$$

with terminal condition $\eta(T, v, x, z) = U(v)$ for all $(v, x, z) \in \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{S}$. Above, $\mathcal{L}_c^\pi$ and $\mathcal{L}_J^\pi$ are operators depending on the strategy $\pi = (\pi_i)_{i=1,\dots,N}^\top \in \mathbb{R}^N$, and acting on a function $f(t, v, x, z)$, $(t, v, x, z) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{S}$, which is $C^{1,2,2}$ in (t, v, x), as follows

$$\begin{aligned}
\mathcal{L}_c^\pi f(t, v, x, z) &:= v\frac{\partial f(t, v, x, z)}{\partial v}\Bigg[r + \sum_{i=1}^N \pi_i(1 - z_i) \sum_{k=1}^K \phi_k(t, x, z)\left(\sum_{j=1}^N H_{(i,j)}(t, x, z)\sigma_{jk}(x)\right) \\
&\quad - \sum_{j=1}^N \left(\sum_{i=1}^N \pi_i(1 - z_i)G_{(i,j)}(t, x, z)\right)(1 - z_j)x_j\Bigg] \\
&\quad + \frac{v^2}{2}\frac{\partial^2 f(t, v, x, z)}{\partial v^2} \sum_{k=1}^K \left(\sum_{i=1}^N \pi_i(1 - z_i)\left(\sum_{j=1}^N H_{(i,j)}(t, x, z)\sigma_{jk}(x_j)\right)\right)^2 \\
&\quad + \sum_{j=1}^N v\frac{\partial^2 f(t, v, x, z)}{\partial v \partial x_j} \sum_{k=1}^K \sigma_{jk}(x_j) \sum_{i=1}^N \pi_i(1 - z_i)\left(\sum_{j=1}^N H_{(i,j)}(t, x, z)\sigma_{jk}(x_j)\right), \\
\mathcal{L}_J^\pi f(t, v, x, z) &:= \sum_{j=1}^N \Bigg[f\Bigg(t, v + v\bigg(\sum_{i=1}^N \pi_i(1 - z_i)G_{(i,j)}(t, x, z)\bigg), x + w_j, z^j\Bigg) \\
&\quad - f(t, v, x, z)\Bigg](1 - z_j)x_j\big(1 + h_j(t, x, z)\big).
\end{aligned} \tag{20}$$

From the above equations, we can easily deduce that the operator $\mathcal{L}_c^\pi$ is related to the continuous part, while $\mathcal{L}_J^\pi$ to the jump part of the dynamic programming problem. The operator $\hat{\mathcal{A}}$ is a second-order differential operator acting on any function $f(t, x, z)$ which is continuously differentiable in t and is continuously twice differentiable in x for each $z \in \mathcal{S}$, and given by

$$\hat{\mathcal{A}} f(t, x, z) := \big(\mu(x)^\top + \sigma(x)\phi(t, x, z)\big)D_x f(t, x, z) + \frac{1}{2}\mathrm{Tr}[(\sigma\sigma^\top)(x)D_x^2 f(t, x, z)]. \tag{21}$$

Above Tr denotes the trace operator.

3.3 Optimal Feedback Strategies

We analyze the optimal feedback control functions. These are denoted by $\pi^*(t,x,z) = (\pi_i^*(t,x,z))_{i=1,\dots,N}^\top$ for $(t,x,z) \in [0,T] \times \mathbb{R}_+^N \times \mathcal{S}$. For $(\pi,t,v,x,z) \in \mathbb{R}^N \times [0,T] \times \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{S}$, $\big(\mathcal{L}_c^\pi + \mathcal{L}_J^\pi + \hat{\mathcal{A}}\big)\eta(t,v,x,z)$ denotes the corresponding Hamiltonian. We postulate (and later prove) that the value function admits the decomposition given by $\eta(t,v,x,z) = v^\gamma Q(t,x,z)$, where $Q(\cdot,z)$ is a positive $C^{1,2}$-function on $[0,T] \times \mathbb{R}_+^N$ for each $z \in \mathcal{S}$. The HJB equation (19) may be equivalently rewritten as follows: we solve for the function $Q(\cdot,z)$, $z \in \mathcal{S}$, satisfying

$$\frac{\partial Q(t,x,z)}{\partial t} + \hat{\mathcal{A}}Q(t,x,z) + \sup_{\pi \in \mathcal{U}_t} \mathcal{H}(t,x,z,\pi) = 0, \qquad \text{in } [0,T) \times \mathbb{R}_+^N, \tag{22}$$

with terminal condition $Q(T,x,z) = \gamma^{-1}$ for all $(x,z) \in \mathbb{R}_+^N \times \mathcal{S}$. For $(t,x,z,\pi) \in [0,T] \times \mathbb{R}_+^N \times \mathcal{S} \times \mathbb{R}^N$, the function

$$\begin{aligned}
\mathcal{H}(t,x,z,\pi) = \gamma Q(t,x,z) \Bigg[r &+ \sum_{i=1}^N \pi_i (1-z_i) \sum_{k=1}^K \phi_k(t,x,z) \left(\sum_{j=1}^N H_{(i,j)}(t,x,z) \sigma_{jk}(x_j) \right) \\
&- \sum_{i=1}^N \pi_i (1-z_i) \left(\sum_{j=1}^N (1-z_j) x_j G_{(i,j)}(t,x,z) \right) \Bigg] - Q(t,x,z) \sum_{j=1}^N (1-z_j) x_j \big(1 + h_j(t,x,z)\big) \\
&+ \frac{\gamma(\gamma-1)}{2} Q(t,x,z) \sum_{k=1}^K \left(\sum_{i=1}^N \pi_i (1-z_i) \left(\sum_{j=1}^N H_{(i,j)}(t,x,z) \sigma_{jk}(x_j) \right) \right)^2 \\
&+ \gamma \sum_{j=1}^N \frac{\partial Q(t,x,z)}{\partial x_j} \left[\sum_{i=1}^N \pi_i (1-z_i) \left(\sum_{k=1}^K \sigma_{jk}(x_j) \left(\sum_{j=1}^N H_{(i,j)}(t,x,z) \sigma_{jk}(x_j) \right) \right) \right] \\
&+ \sum_{j=1}^N \left(1 + \sum_{i=1}^N \pi_i (1-z_i) G_{(i,j)}(t,x,z) \right)^\gamma Q(t, x + w_j, z^j)(1-z_j) x_j (1 + h_j(t,x,z)).
\end{aligned} \tag{23}$$

From the above expression of $\mathcal{H}(t,x,z,\pi)$, it is clear that the optimal feedback control function $\pi^*(t,x,z)$ depends on the default state $z \in \mathcal{S}$. Recall the notation $z = 0^{j_1,\dots,j_m}$ introduced at the end of Section 1. Notice also that no investment can be made in a defaulted bond, i.e., $\pi_i^*(t,x,z) = 0$ if $z_i = 1$.

Next, we analyze the feedback control function $\pi_i^*(t,x,z)$ when $i \notin \{j_1,\dots,j_m\}$. The main result is stated in the following proposition.

Proposition 3.3. *Let $z = 0^{j_1,\dots,j_m}$ for $m = 0,1,\dots,N-1$. There exists a unique (measurable) optimal feedback control function $\pi^*(t,x) = (\pi_i^*(t,x);\ i \notin \{j_1,\dots,j_m\})^\top$ such that*

$$\pi^*(t,x) \in \underset{\pi \in \mathcal{J}^{(N-m)}}{\arg\max}\ \mathcal{H}(t,x,0^{j_1,\dots,j_m},\pi)$$

for almost every $(t,x) \in [0,T] \times \mathbb{R}_+^N$. The admissible set in the default state $z = 0^{j_1,\dots,j_m}$ is given by

$$\mathcal{J}^{(N-m)} := \left\{ \pi \in \mathbb{R}^{N-m};\ 1 + \sum_{i \notin \{j_1,\dots,j_m\}} \pi_i \theta_{ij} \geq 0,\ j \notin \{j_1,\dots,j_m\},\ \forall\ \theta_{ij} \in [m_{ij}, M_{ij}] \right\}. \tag{24}$$

In general, closed form expressions for optimal feedback control function are not available. In the degenerate model (see Remark 2.2), however, the absence of diffusion risk in the default intensity processes makes it possible to obtain more explicit representations for the strategies. These also carry clear economic interpretations, and hence allow to understand the influence of default contagion on optimal portfolio allocations. Indeed, if the volatility term in the default intensities vanish, the Hamiltonian given by (23) reduces to

$$\mathcal{H}(t,x,z,\pi) = \gamma Q(t,x,z) \left[r - \sum_{i=1}^N \pi_i (1-z_i) \left(\sum_{j=1}^N (1-z_j) x_j G_{(i,j)}(t,x,z) \right) \right]$$

$$-Q(t,x,z)\sum_{j=1}^{N}(1-z_j)x_j(1+h_j(t,x,z)) \tag{25}$$
$$+\sum_{j=1}^{N}\left(1+\sum_{i=1}^{N}\pi_i(1-z_i)G_{(i,j)}(t,x,z)\right)^{\gamma}Q(t,x+w_j,z^j)(1-z_j)x_j(1+h_j(t,x,z)),$$

for $(t,x,z)\in[0,T]\times\mathbb{R}_+^N\times\mathcal{S}$. Above, $Q(t,x,z)$ denotes the positive solution to the recursive HJB equation in the degenerate case. Introduce the following shorthand notation: $Q_j(t,x) := Q(t,x,0^{j_1,\dots,j_m,j})$, $h_j(t,x) := h_j(t,x,0^{j_1,\dots,j_m})$, and $Q(t,x) := Q(t,x,0^{j_1,\dots,j_m})$, i.e., we omit the subscript $(j_1,\dots,j_m)$ for notational convenience. The first-order condition in the default state $z=0^{j_1,\dots,j_m}$, and $i\notin\{j_1,\dots,j_m\}$, is given by

$$0=-Q(t,x)\sum_{j\notin\{j_1,\dots,j_m\}}x_jG_{(i,j)}(t,x)$$
$$+\sum_{j\notin\{j_1,\dots,j_m\}}\left(1+\sum_{l\notin\{j_1,\dots,j_m\}}\pi_lG_{(l,j)}(t,x)\right)^{\gamma-1}G_{(i,j)}(t,x)Q_j(t,x+w_j)x_j(1+h_j(t,x)).$$

The above equation can be solved explicitly and yield the optimal feedback control function $\pi^*(t,x)$ given by

$$\pi^*(t,x)=\left(G^\top(t,x)\right)^{-1}\Theta(t,x). \tag{26}$$

Here $G(t,x) := (G_{(i,j)}(t,x))_{(i,j)\notin\{j_1,\dots,j_m\}^2}$ and the $(N-m)$-dimensional column vector $\Theta(t,x)$ is defined by

$$\Theta(t,x) := \left[\left(\frac{Q(t,x)}{Q_j(t,x+w_j)(1+h_j(t,x))}\right)^{\frac{1}{\gamma-1}}-1\right]^\top_{j\notin\{j_1,\dots,j_m\}}.$$

It can be seen that $1+\sum_{l\notin\{j_1,\dots,j_m\}}\pi_l^*(t,x)G_{(l,j)}(t,x)>0$ for all $(t,x)\in[0,T]\times\mathbb{R}_+^N$. From the wealth dynamics given by (3.1) by setting $\sigma(\cdot)=0$ therein, it follows that the above optimal strategy $\pi^*(t,x)$ in the degenerate case is admissible.

Using the closed form representation in Eq. (26), we can perform a decomposition which allows isolating idiosyncratic from contagious influences. We omit the variable $(t,x)\in[0,T]\times\mathbb{R}_+^N$ in the following matrix functions. Assume to be in a state where the firms $j_1,j_2,\dots,j_m$ have defaulted, and let Π be the $(N-m)$-dimensional diagonal matrix whose entries are $\Pi_{j,j}=G_{(j,j)}$ for $j\in\{j_{m+1},\dots,j_N\}$. We may rewrite

$$(G^\top)^{-1}=\Pi^{-1}\left(G^\top\Pi^{-1}\right)^{-1}=\Pi^{-1}\sum_{k=0}^{\infty}\left(I-G^\top\Pi^{-1}\right)^k, \tag{27}$$

provided that the eigenvalues of $I-G^\top\Pi^{-1}$ are strictly smaller than one. Above, I denotes the $(N-m)$-dimensional identity matrix. The term $k=0$ gives the idiosyncratic contribution of default risk on the optimal strategies. Indeed, when $k=0$, we have that $(G^\top)^{-1}=\Pi^{-1}$, and the zeroth-order approximation for the strategy is given by

$$\pi_j^{0,*}=\frac{1}{G_{(j,j)}}\left(\frac{Q}{Q_j(1+h_j)}\right)^{\frac{1}{\gamma-1}}-1, \qquad j=j_{m+1},\dots,j_N.$$

Clearly, the above quantity only depends on the default risk of firm $j\in\{j_{m+1},\dots,j_N\}$. When $k=1$ and the first term in the expansion (27) is included, we have the first round of contagious influences on the optimal bond allocation strategies. More specifically, we have

$$I-G^\top\Pi^{-1}=\begin{bmatrix} 0 & -\frac{G_{(j_{m+2},j_{m+1})}}{G_{(j_{m+1},j_{m+1})}} & -\frac{G_{(j_{m+3},j_{m+1})}}{G_{(j_{m+1},j_{m+1})}} & \cdots & -\frac{G_{(j_N,j_{m+1})}}{G_{(j_{m+1},j_{m+1})}} \\ -\frac{G_{(j_{m+1},j_{m+2})}}{G_{(j_{m+2},j_{m+2})}} & 0 & -\frac{G_{(j_{m+3},j_{m+2})}}{G_{(j_{m+2},j_{m+2})}} & \cdots & -\frac{G_{(j_N,j_{m+2})}}{G_{(j_{m+2},j_{m+2})}} \\ \vdots & & \ddots & \vdots & \\ -\frac{G_{(j_{m+1},j_N)}}{G_{(j_N,j_N)}} & -\frac{G_{(j_{m+2},j_N)}}{G_{(j_N,j_N)}} & -\frac{G_{(j_{m+3},j_N)}}{G_{(j_N,j_N)}} & \cdots & 0 \end{bmatrix}.$$

The above matrix characterizes the first-order influence of default contagion on the optimal bond allocation strategies. Indeed, we can write the first-order approximation of the optimal strategy as

$$\pi^{1,*} = \pi^{0,*} + \underbrace{\Pi^{-1}\left(I - G^\top \Pi^{-1}\right)\Theta}_{\text{first-order contagion correction}} .$$

The propagation of contagion effects continues in subsequent rounds, but dampens as $k \to \infty$ because the series converges.

4 HJB Equations and Verification Theorem

Section 4.1 analyzes the recursive system of HJB equations (22) associated with the control problem. Section 4.2 proves a verification theorem.

4.1 Recursive Classical Solutions of HJB Equations

This section studies existence and uniqueness of a classical solution to the recursive system of HJB equations in the non-degenerate case. Let $z = 0^{j_1,\dots,j_m}$, where $m = 0, 1, \dots, N$. If $m = N$, i.e., all firms have defaulted, the investor can only deposit his wealth in the bank account and accrue interest. Indeed, Eq. (22) reduces to

$$\frac{\partial Q(t,x,1)}{\partial t} + \hat{\mathcal{A}} Q(t,x,1) + \gamma r Q(t,x,1) = 0, \qquad \text{in } [0,T) \times \mathbb{R}_+^N \tag{28}$$

with terminal condition $Q(T,x,1) = \gamma^{-1}$ for all $x \in \mathbb{R}_+^N$. Eq. (28) is a linear PDE which admits a unique classical solution given by

$$Q(t,x,1) = \gamma^{-1} e^{\gamma r (T-t)}, \qquad \text{for } (t,x) \in [0,T] \times \mathbb{R}_+^N. \tag{29}$$

We next study the existence and uniqueness of a classical solution to the HJB equation (22), when $m \le N-1$. We proceed inductively and assume that the HJB equation (22) has a unique classical solution when $z = 0^{j_1,\dots,j_m,j}$ for $j \notin \{j_1,\dots,j_m\}$. We denote such a solution by $Q_j(t,x) := Q_{j_1,\dots,j_m,j}(t,x)$. Using Eq. (22), the function $u(t,x) := Q_{j_1,\dots,j_m}(t,x)$ solves the following HJB-PDE

$$\frac{\partial u}{\partial t} + \hat{\mathcal{A}} u + \bar{\mathcal{H}}(t,x,u,D_x u) = 0, \qquad \text{in } [0,T) \times \mathbb{R}_+^N \tag{30}$$

with terminal condition $u(T,x) = \gamma^{-1}$ for all $x \in \mathbb{R}_+^N$. Above, for $(t,x,u,p) \in [0,T] \times \mathbb{R}_+^N \times \mathbb{R}_+ \times \mathbb{R}^N$, the Hamiltonian

$$\bar{\mathcal{H}}(t,x,u,p) := \sup_{\pi \in \mathcal{U}_t} \left\{ g(t,x,\pi) u + f(t,x,\pi) p + l(t,x,\pi) \right\}, \tag{31}$$

where for $(t,x,\pi) \in [0,T] \times \mathbb{R}_+^N \times \mathbb{R}^{N-m}$, the functions

$$\begin{aligned}
f(t,x,\pi) &:= \gamma \pi^\top H(t,x) \sigma(x) \sigma(x)^\top, \\
g(t,x,\pi) &:= \gamma r + \gamma \pi^\top \left(H(t,x) \sigma(x) \phi(t,x) - G(t,x) (x_j)^\top_{j \notin \{j_1,\dots,j_m\}} \right) \\
&\qquad - \sum_{j \notin \{j_1,\dots,j_m\}} x_j (1 + h_j(t,x)) + \frac{\gamma(\gamma-1)}{2} \left| \sigma(x)^\top H(t,x)^\top \pi \right|^2, \\
l(t,x,\pi) &:= \sum_{j \notin \{j_1,\dots,j_m\}} \left(1 + \sum_{i \notin \{j_1,\dots,j_m\}} \pi_i G_{(i,j)}(t,x) \right)^\gamma Q_j(t, x + w_j) x_j (1 + h_j(t,x)).
\end{aligned} \tag{32}$$

Here we have used the following representations

$$H(t,x) := (H_{(i,j)}(t,x,0^{j_1,\dots,j_m}))_{(i,j) \in \{j_{m+1},\dots,j_N\} \times \{1,\dots,N\}}, \; \sigma(x) := (\sigma_{jk}(x))_{(j,k) \in \{1,\dots,N\} \times \{1,\dots,K\}},$$
$$\phi(t,x) := (\phi_k(t,x,0^{j_1,\dots,j_m}))^\top_{k=1,\dots,K}, \qquad x = (x_j)_{j \notin \{j_1,\dots,j_m\}}.$$

Also notice that $[H(t,x)\sigma(x)]_{(i,k)} = \sum_{j=1}^N H_{(i,j)}(t,x)\sigma_{jk}(x) = \sum_{j\notin\{j_1,\dots,j_m\}} H_{(i,j)}(t,x)\sigma_{jk}(x)$ by Remark A.4 in Appendix A.

We notice that the solution of the HJB equation in each state z is recursively linked to the solutions of the HJB equations in the states z^j reached upon the default of an additional firm j. That is, the optimal expected utility in the state z depends on all optimal expected utilities in the state z^j through the terms $Q_{j_1,\dots,j_m,j}$. When firm j defaults, the default intensity of each surviving firm i increases by an amount w_{ji} due to contagion. The following theorem establishes existence of classical solutions for the recursive system of HJB equations (22). The proof of this theorem relies on an argument based on the approximation of the policy space, and on the approximation of the unbounded domain by bounded domains.

Theorem 4.1. *Suppose the HJB equation (22) admits a positive classical solution $Q(t,x,0^{j_1,\dots,j_m,j})$ when the default state is $z=0^{j_1,\dots,j_m,j}$, $j\notin\{j_1,\dots,j_m\}$. Then there exists a positive classical solution $Q(t,x,0^{j_1,\dots,j_m})$ to the HJB equation (22) when the default state is $z=0^{j_1,\dots,j_m}$ (such an equation takes the form given in (30)).*

We know that, in the state where all firms have defaulted, the HJB equation (22) admits the positive classical solution $Q(t,x,1)=\gamma^{-1}e^{\gamma r(T-t)}$ for all $(t,x)\in[0,T]\times\mathbb{R}_+^N$. Using Theorem 4.1, this implies the existence of a classical solution to the HJB equation (22) via a backward recursion on the default states.

Remark 4.2. *For $(t,x,z)\in[0,T]\times\mathbb{R}_+^N\times\mathcal{S}$, define*

$$\bar{L}(t,x,z) := H(t,x,z)\sigma(x)\phi(t,x,z) - D(z)G(t,x,z)x^\top,$$
$$\bar{h}_j(t,x,z) := x_j(1+h_j(t,x,z)), \qquad j=1,\dots,N,$$

where we set $D(z) := diag(1-z_j;\ j=1,\dots,N)$ and recall that ϕ and h are the vector of market and default risk premium coefficients. Assume the following condition

(C1) *The functions $\bar{L}(\cdot,z)$ and $\bar{h}_j(\cdot,z)$ are bounded for each $z\in\mathcal{S}$.*

Then the classical solution given in Theorem 4.1 can be bounded (see also the proof of Theorem 4.1 in Appendix A). In other words, if the system of HJB equations admits a unique bounded positive classical solution when $z=0^{j_1,\dots,j_m,j}$, $j\notin\{j_1,\dots,j_m\}$, then there exists a unique positive bounded classical solution to the system (30) when the default state is $z=0^{j_1,\dots,j_m}$. In this case, the coefficient $g^*(t,x)$ given in Eq. (A.32) of Appendix A is bounded from above given that $\gamma\in(0,1)$, and the function $l^*(t,x)$ given in (A.32) is positive and bounded. This implies that the classical solution is bounded from above using the representation of solutions (A.33) in Appendix A.

The recursive HJB equations admit an explicit solution if diffusion risk is absent. In this case, the HJB equation (22) reduces to

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial t}+\hat{\mathcal{A}}_0\right)Q(t,x,z) + \mathcal{C}\big(t,x,z,\pi^*(t,x,z)\big)Q(t,x,z) \\ &\quad + \sum_{j=1}^N (1-z_j)x_j\big(1+h_j(t,x,z)\big)\left(1+\sum_{l=1}^N \pi_l^*(t,x,z)G_{(l,j)}(t,x,z)\right)^\gamma Q(t,x+w_j,z^j) \end{aligned} \tag{33}$$

with terminal condition $Q(T,x,z)=\gamma^{-1}$ for all $(x,z)\in\mathbb{R}_+^N\times\mathcal{S}$. The operator $\hat{\mathcal{A}}_0$ is defined as (21) by setting $\sigma(\cdot)=0$. Moreover, the coefficient

$$\mathcal{C}(t,x,z,\pi) := \gamma r - \gamma\sum_{j=1}^N\left(\sum_{l=1}^N \pi_l G_{(l,j)}(t,x,z)\right)(1-z_j)x_j - \sum_{j=1}^N (1-z_j)x_j\big(1+h_j(t,x,z)\big). \tag{34}$$

Let $z=0^{j_1,\dots,j_m}$, where $j_1,\dots,j_m$ are the defaulted firms. If $z=1$, then Eq. (33) reduces to

$$\left(\frac{\partial}{\partial t}+\hat{\mathcal{A}}_0\right)Q(t,x,1)+\gamma r Q(t,x,1)=0, \qquad \text{in } (t,x)\in[0,T)\times\mathbb{R}_+^N$$

with terminal condition $Q(T,x,1)=\gamma^{-1}$ for all $x\in\mathbb{R}_+^N$. The solution is given by $Q(t,x,1)=\gamma^{-1}e^{\gamma r(T-t)}$ for all $(t,x)\in[0,T]\times\mathbb{R}_+^N$, as in non-degenerate case (see Eq. (29)). Next, consider $0\le m\le N-1$. We first define the following quantities by

$$\mathcal{C}^a(t,x):=\gamma r+\gamma\sum_{j\notin\{j_1,\dots,j_m\}}x_j-\sum_{j\notin\{j_1,\dots,j_m\}}x_j\big(1+h_j(t,x)\big),$$
$$\mathcal{C}^b(t,x):=(1-\gamma)\sum_{j\notin\{j_1,\dots,j_m\}}x_j\Big\{Q_j(t,x+w_j)\big(1+h_j(t,x)\big)\Big\}^{\frac{1}{1-\gamma}}. \tag{35}$$

Above, we recall that $Q_j(t,x):=Q(t,x,0^{j_1,\dots,j_m,j})$, $j\notin\{j_1,\dots,j_m\}$, is the solution to the PDE (33) when the default state is augmented with the default of firm j, i.e., it becomes $0^{j_1,\dots,j_m,j}$. Here $h_j(t,x):=h_j(t,x,0^{j_1,\dots,j_m})$ and $Q(t,x):=Q(t,x,0^{j_1,\dots,j_m})$, i.e., we omit the subscript $(j_1,\dots,j_m)$ for the notational convenience. When this happens, the default intensities of all surviving firms increase by an amount w_j due to contagion, hence explaining why $Q_{j_1,\dots,j_m,j}$ in Eq. (35) is evaluated at $x+w_j$. Then, Eq. (33) reduces to the following *Bernoulli*'s type PDEs given by

$$\left(\frac{\partial}{\partial t}+\hat{\mathcal{A}}_0\right)Q(t,x)+\mathcal{C}^a(t,x)Q(t,x)+\mathcal{C}^b(t,x)Q^{\frac{\gamma}{\gamma-1}}(t,x)=0. \tag{36}$$

In the above equations, the nonlinearity of the PDE is only manifested through the term $Q^{\frac{\gamma}{\gamma-1}}(t,x)$ and not through the coefficients as in the original equivalent PDE (33). The quantity $\frac{\gamma}{\gamma-1}$ in the above nonlinear terms is usually called the Legendre-Frechel exponent of the risk-aversion parameter γ. It is easy to see that $\frac{\gamma}{\gamma-1}<0$ if $\gamma\in(0,1)$. We can guarantee the existence of a unique classical solution to the degenerate HJB equation.

Proposition 4.3. *Assume the existence of a unique positive, continuously differentiable solution $Q_j(t,x):=Q(t,x,0^{j_1,\dots,j_m,j})$, $j\notin\{j_1,\dots,j_m\}$, to Eq. (33) in the default state $z=0^{j_1,\dots,j_m,j}$. Then there exists a unique solution to Eq. (33) in the default state $z=0^{j_1,\dots,j_m}$. More specifically, for $(t,x)\in[0,T]\times\mathbb{R}_+^N$, we have*

$$Q(t,x)=\hat{Q}^{1-\gamma}(t,x), \tag{37}$$

where the positive function $\hat{Q}(t,x)$ is given by

$$\begin{aligned}\hat{Q}(t,x)=\gamma^{-\frac{1}{1-\gamma}}\exp\left\{\frac{1}{1-\gamma}\int_t^T\mathcal{C}^a\big(u,(\mathcal{X}_u^{(t,x_j,j)};\ j=1,\dots,N)\big)du\right\}\\+\frac{1}{1-\gamma}\int_t^T\mathcal{C}^b\big(s,(\mathcal{X}_s^{(t,x_j,j)};\ j=1,\dots,N)\big)\\\times\exp\left\{\frac{1}{1-\gamma}\int_t^s\mathcal{C}^a\big(u,(\mathcal{X}_u^{(t,x_j,j)};\ j=1,\dots,N)\big)du\right\}ds.\end{aligned} \tag{38}$$

Here $\mathcal{X}_u^{(t,x_j,j)}$ is given by Eq. (6) in Section 2.1.

4.2 Verification Theorem

We provide a verification theorem which

(i) Guarantees that the optimal feedback strategy at time t is obtained by evaluating the deterministic feedback control function at the current default state and default intensity vector.

(ii) Establishes the correspondence between the value function (18) of the control problem and the solution to the HJB equation (30).

Theorem 4.4. *Let $m=0,1,\dots,N$ and $Q(t,x,0^{j_1,\dots,j_m})$ be the classical solution to the HJB equation (30) with terminal condition $Q(T,x,0^{j_1,\dots,j_m})=\gamma^{-1}$ for all $x\in\mathbb{R}_+^N$ in the default state $z=0^{j_1,\dots,j_m}$. Assume that the condition* **(C1)** *in Remark 4.2 holds in the non-degenerate case. Then the value function (18) admits an explicit expression given by*

$$\eta(t,v,x,z)=v^\gamma Q(t,x,z),\qquad (t,v,x,z)\in[0,T]\times\mathbb{R}_+\times\mathbb{R}_+^N\times\mathcal{S}. \tag{39}$$

Moreover, the predictable optimal feedback strategy is given by the Markov control $\pi^*(t)=\pi^*(t,X(t-),Z(t-))$ with $t\in[0,T]$, where for $m=0,1,\dots,N$, $\pi^*(t,x,0^{j_1,\dots,j_m})$, $(t,x)\in[0,T]\times\mathbb{R}_+^N$, is given by (A.21).

5 Numerical Analysis

We consider a special case of our framework consisting of two firms, i.e., we set $N = 2$. This enables us to highlight the extent to which the various model parameters affect the optimal allocation decisions.

Benchmark Parameters. Throughout the section, we use the following benchmark parameters. We choose the interest rate $r = 0.05$. We consider two common risk factors, i.e., $K = 2$. We choose $w_{12} = w_{21} = 0.2$ to quantify the instantaneous jump of the default intensity of firm "2" (firm "1") at the default of firm "1" (firm "2"). Such a setting is expected to hold for a scenario in which the two firms belong to the same industrial sector, and have similar risk exposures to each other. We assume zero default risk premia prior to the occurrence of any default event, $h_1(t, x, (0, 0)) = h_2(t, x, (0, 0)) = 0$. We choose $h_1(t, x_1, (0, 1)) = h_2(t, x_2, (1, 0)) = -0.8$, where $x = (x_1, x_2)$, i.e., the default intensity of both firms is lower under the actual probability measure than under the market measure in the case when one of the two firms has defaulted. This specification is designed to capture empirical evidence suggesting that lower-rated corporate bonds have a higher risk premium than that of higher-rated corporate bonds. We choose constant market risk premia: $\phi_1(t, x_1, (0, 1)) = \phi_2(t, x_2, (0, 1)) = \phi_1(t, x_2, (1, 0)) = \phi_2(t, x_2, (1, 0)) = 0.4$, and $\phi_1(t, x, (0, 0)) = \phi_2(t, x, (0, 0)) = 0.1$. The contractual bond parameters are set to $C_1 = C_2 = 0.7$ for the coupon payments, and to $R_1(0, 1) = R_1(0, 0) = R_2(1, 0) = R_2(0, 0) = 0.2$ for the recovery payments. The maturity of the bonds is $T_1 = T_2 = 4$, and the investment horizon is $T = 2$. Unless otherwise specified, we set the investment time $t = 0$ and the initial values of the intensities to $x_1 = x_2 = 1.5$, that is the feedback functions are plotted for $t = 0, x_1 = x_2 = 1.5$.

Bond Prices. The price of a bond with longer maturity is higher than that with a shorter maturity, if the default risk of the underlying firm is low. This is because of the higher value of the stream of coupon payments received. However, if the default intensity becomes higher, then the stream of coupon payments terminate earlier, the bond price jumps to the recovery payment, and the nominal value is not paid. Under these circumstances, we expect the prices of a long maturity and of a short maturity bond to be nearly the same. These statements are visually confirmed by the plots in Figure 1.

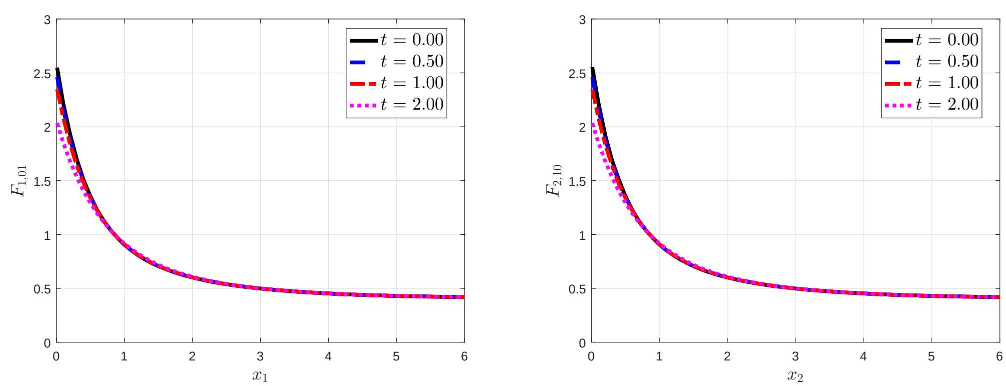

Figure 1: Left: Price of the bond underwritten by firm "1" after default of firm "2". Right: Price of the bond underwritten by firm "2" after default of firm "1". The default intensity parameters are set to $\nu_1 = \nu_2 = 0.1$, $\kappa_1 = \kappa_2 = 0.1$ for mean reversion speeds and levels, $\sigma_{11} = \sigma_{12} = \sigma_{21} = \sigma_{22} = 0.01$ for volatilities.

5.1 Analysis of Investment Strategies

We perform a comparative statics analysis to examine how the default risk premia, the deterioration of the intensity toward its long run mean, and the volatility of the default intensity process affect the optimal allocation decisions of the investor. A common pattern in Figures 2, 3 and 4 is that the investor increases the size of his investment position as his risk aversion decreases.

Default Risk Premia. Consider first the case that both firms are alive. Under our benchmark parameters specification, there is zero default risk premia if no security has defaulted. If the default intensity of the firm is low, the investor shorts the bond security. This is because the price of the bond is higher and consequently the jump of the bond price to the recovery payment at default is larger, allowing the investor to make a larger gain at the default event. In contrast, if the default intensity is high, the downward jump in the bond price at the default event of the firm is low; on the other hand, a positive default risk premium is captured by the investor, if the other firm were to default (recall that under our benchmark setup, when one firm defaults the market attributes higher default risk to the surviving firm than the history, which means that the investor finds it cheaper to buy and hold the bond). If the cost of the downward jump outweighs the gain from the default risk premium, the investor goes short, otherwise he goes long. These intuitive arguments are supported by the left graph in Figure 2. After the default of a firm, the investor always captures a positive default risk premium for holding the bond security, and always goes long in the bond security; see the right graph of Figure 2.

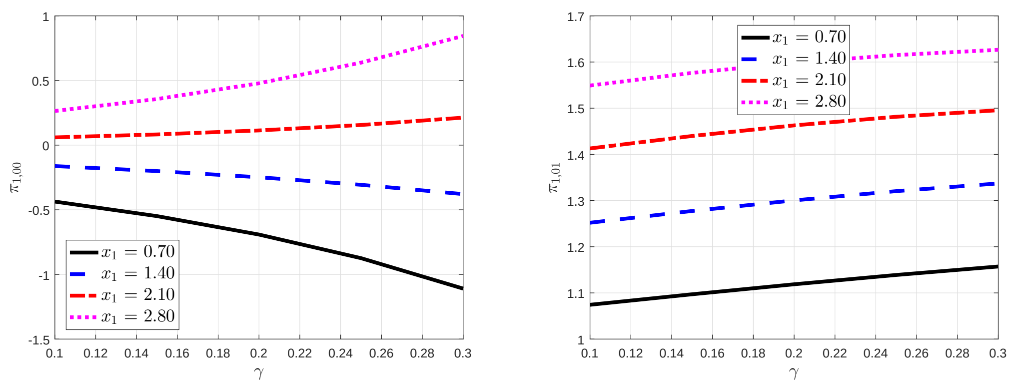

Figure 2: We set the default intensity parameters to $\nu_1 = \nu_2 = 0.1$, $\kappa_1 = \kappa_2 = 0.1$ for mean reversion speeds and levels, $\sigma_{11} = \sigma_{12} = \sigma_{21} = \sigma_{22} = 0.01$ for volatilities.

Deterioration of Default Intensities. In the numerical example of Figure 3, we fix the mean reversion level of the firm "1"'s default intensity to $\frac{\kappa_1}{\nu_1} = x_1 - 0.3$, and vary the mean reversion speed ν_1. Increasing ν_1 makes the default intensity deteriorate faster toward its long run mean, which is lower than the initial value x_1. Because of the lower default risk, the bond of firm "1" can be sold at a higher price. As a consequence, the investor prefers to short a higher number of shares of bond "1" to profit of the larger downward jump in price when bond "1" defaults. The investor uses the proceeds from the short sale of bond "1" to buy back shares of the bond "2", i.e., to reduce his short position in bond "2". After the default of a firm, the investor captures a default risk premium if he goes long in the bond security. However, as ν_1 increases he allocates a lower fraction of his wealth to shares of bond "1". This is because the bond becomes more expensive if the default intensity mean-reverts faster toward a lower level.

Volatility of Default Intensities. Consider first the situation in which both firms are alive. The top left graph of Figure 4 indicates that, as the firm "1"'s default intensity becomes more volatile, i.e., σ_{11} increases, the investor increases the size of his short position in the bond security. Under the benchmark parameter configuration, the investor shorts the bond because he wants to profit of the gain realized at the default event. Clearly, higher values of σ_{11} imply a larger volatility for the default intensity, and thus a higher probability that firm "1" defaults. Interestingly, even if the default intensity of firm "2" does not depend directly on σ_{11}, the investor's holdings in bond "2" are still affected. As σ_{11} increases, the probability that the default intensity of firm "2" jumps upward as a result of firm "1"'s default increases. Hence, the investor would be able to profit of his short position in the bond "2" with a higher probability. This line of reasoning is confirmed by the top right graph of Figure 4. After the default of firm "2", the investor finds it profitable to go long in the bond "1". In this case, a higher volatility would hurt the investor because it increases the probability of a default event, in which case he realizes a negative return from holding a long position in the bond (see bottom left graph of Figure 4). Moreover, after firm "1" defaults, no contagion effect is present, i.e., the higher volatility of firm 1's default intensity has no impact on the default intensity of firm "2". As a consequence, the investor's allocation decisions on the bond "2" are insensitive to the value of σ_{11} (see bottom right graph of Figure 4).

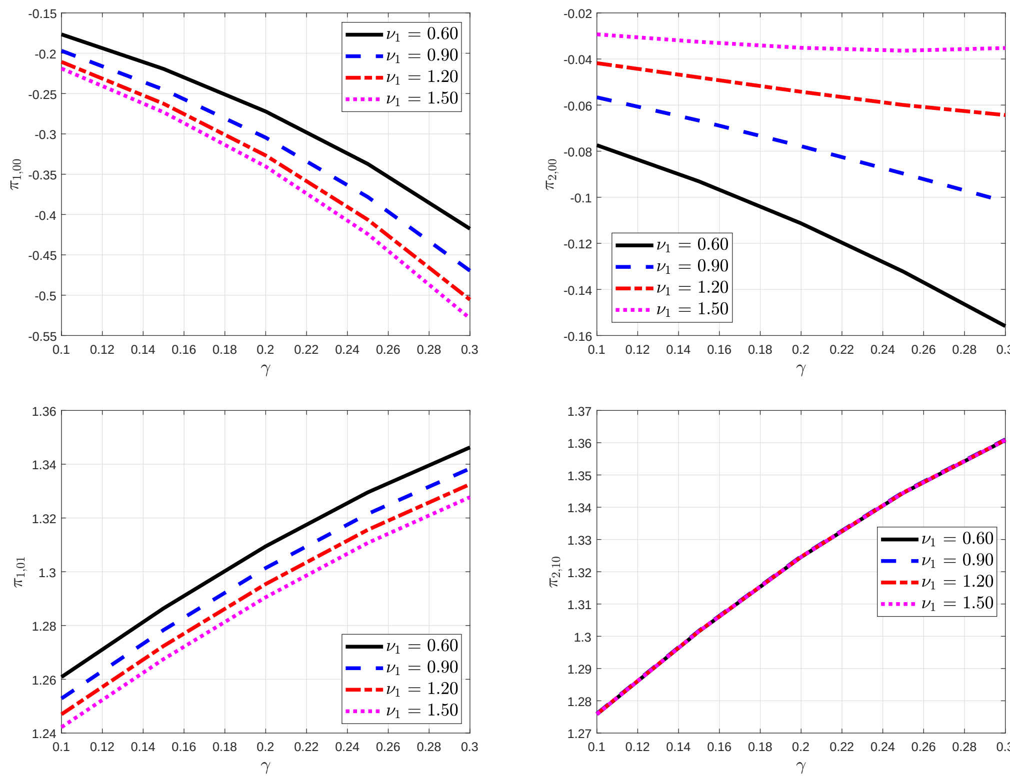

Figure 3: We set $\nu_2 = 0.1$, $\kappa_2 = 0.1$, and $\kappa_1 = \nu_1 \times (x_1 - 0.3)$ where $x_1 = 1.5$. We set the volatility parameters $\sigma_{11} = \sigma_{12} = \sigma_{21} = \sigma_{22} = 0.01$.

6 Conclusion

We have developed a fixed-income portfolio optimization framework featuring a default model in which the jump-diffusive default intensities of the firms ramp up at the occurrence of a default event, and then decay exponentially fast toward their long-run mean levels. The default of a firm can bring a loss to the investor if he is long in the bond security, and additionally, it permanently decreases the set of portfolio securities at his disposal. Mathematically, the dependence structure between the default states of the economy is encoded by a recursive system of uniformly second-order parabolic HJB-PDEs. We have studied the existence of classical solutions to the recursive system, and obtained explicit representations for them in the degenerate case of vanishing volatility. We have analyzed how default contagion impacts the optimal allocation decisions, both analytically and numerically. In particular, we have decomposed the optimal portfolio into an idiosyncratic component and higher-order contagious influences manifested when firms default. Our comparative statics analysis indicates that the decay speed of the default intensities toward their long run levels, default risk premia, and volatility of the default intensity processes have a significant impact on the optimal allocation decisions of the investor. They determine whether the investor should go long or short in the bond securities, and also affect the size of the optimal allocation decisions.

Acknowledgments

The authors gratefully acknowledge the constructive comments provided by two anonymous reviewers and the associate editor, which led to improving the quality of the manuscript. This research of L. Bo was supported in part by the NSF of China under grant 11471254, The Key Research Program of Frontier Sciences, CAS under grant QYZDB-SSW-SYS009 and Fundamental Research Funds for Central Universities under grant WK3470000008. The research of A. Capponi was partially supported

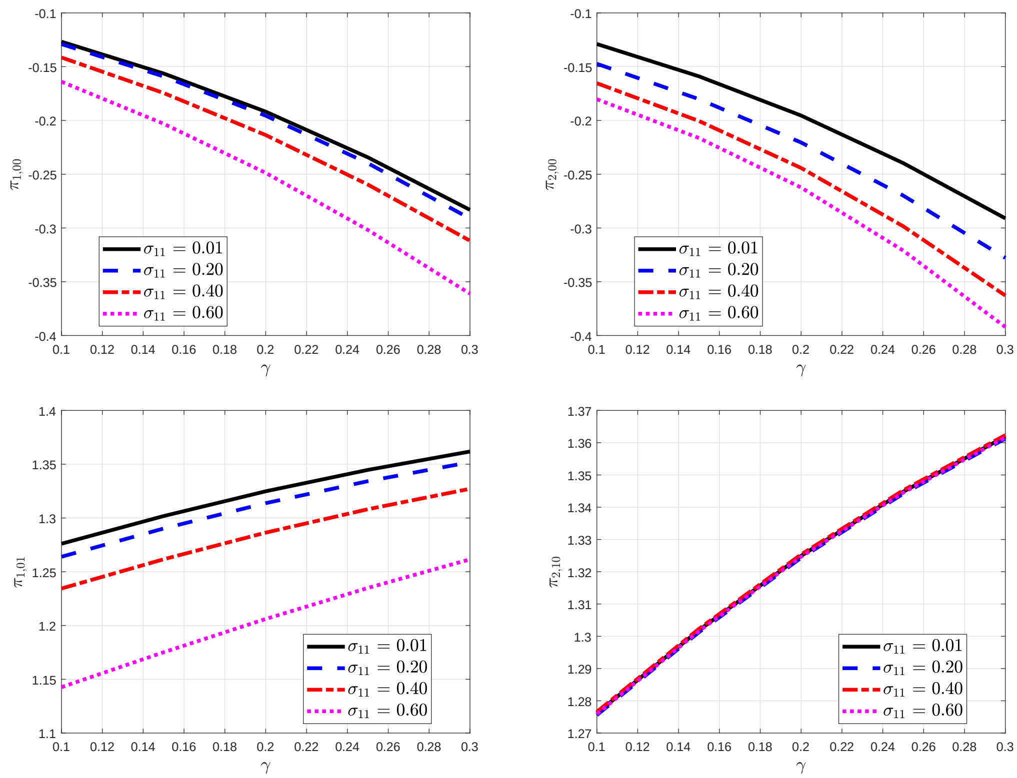

Figure 4: We set the parameters $\nu_1 = \nu_2 = 0.1$, $\kappa_1 = \kappa_2 = 0.1$. We set the volatility parameters $\sigma_{12} = \sigma_{21} = \sigma_{22} = 0.01$.

by the National Science Foundation under grant DMS 1716145. The research of P.C. Chen was partially supported by the HKU Seed Fund for Basic Research under grant 201705159004.

References

Acharya, V.-V., Bharath, S.-T., & Srinivasan, A. (2007). Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. *Journal of Financial Economics, 85*, 787-821.

Ait-Sahalia, Y., & Hurd, T.-R. (2016). Portfolio choice in markets with contagion. *Journal of Financial Econometrics, 14*, 1-28.

Altman, E. I., Brady, B., Resti, A., & Sironi, A. (2005). The link between default and recovery rates: theory, empirical evidence, and implications. *Journal of Business, 78*, 2203-2228.

Azizpour, S., Giesecke, K., & Schwenkler, G. (2017). Exploring the sources of default clustering. *Journal of Financial Economics*, Forthcoming.

Becherer, D., & Schweizer, M. (2005). Classical solutions to reaction-diffusion systems for hedging problems with interacting Itô and point processes. *Annals of Applied Probability, 15*, 1111-1144.

Bo, L., & Capponi, A. (2016). Optimal investment in credit derivatives portfolio under contagion risk. *Mathematical Finance, 26*, 785-834.

Bo, L., & Capponi, A. (2017). Robust optimization of credit portfolios. *Mathematics of Operations Research, 42*, 30-56.

Brigo, D., & Alfonsi, A. (2004). Credit default swap calibration and derivatives pricing under the SSRD stochastic intensity model. *Finance and Stochastics, 9*, 29-42.

Chow, Y., & Teicher, H. (1978). *Probability Theory.* New York: Springer.

Davis, M., & Lo, V. (2001). Infectious defaults. *Quantitative Finance, 1*, 305-308.

Davis, M., & Lieo, S. (2013). Jump-diffusion risk-sensitive asset management II: jump-diffusion factor model. *SIAM Journal on Control and Optimization, 51*, 1441-1480.

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics, 83*, 635-655.

Errais, E., Giesecke, K., & Goldberg, L. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics, 1*, 642-655.

Fleming, H., & Rishel, R. (1975). *Deterministic and Stochastic Optimal Control.* Berlin: Springer.

Giesecke, K., Kim, B., Kim, J., & Tsoukalas, G. (2014). Optimal credit swap portfolios. *Management Science, 60*, 2291-2307.

Heath, D., & Schweizer, M. (2001). Martingales versus PDEs in finance: An equivalence result with examples. *Journal of Applied Probability, 37*, 947-957.

Karatzas, I., & Shreve, S. (1991). *Brownian Motion and Stochastic Calculus* (2nd ed.) New York: Springer.

Kraft, H., & Steffensen, M. (2008). How to invest optimally in corporate bonds. *Journal of Economic Dynamics and Control, 32*, 348-385.

Kraft, H., & Steffensen, M. (2007). Bankruptcy, counterparty risk, and contagion. *Review of Finance, 11*, 209-252.

Ladyzenskaja, O. A., Solonnikov, V. A. & Uralceva, N. N. (1968). *Linear and Quasilinear Equations of Parabolic Type* (Translations of Mathematical Monographs 23). Providence, RI: American Mathematical Society.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research, 2*, 99-120.

Merton, R. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics, 51*, 247-257.

Rogers, L.-C.-G., & Williams, D. (2000). *Diffusions, Markov Processes and Martingales* (Vol. 1, 2nd ed.). Cambridge: Cambridge University Press.

Yu, F. (2007). Correlated defaults in intensity-based models. *Mathematical Finance, 17*, 155-173.

Yu, F. (2005). Default correlation in reduced-form models. *Journal of Investment Managment, 3*, 33-42.

Wise, M.-B., & Bhansali, V. (2002). Portfolio allocation to corporate bonds with correlated defaults. *Journal of Risk, 5*, 39-58.

A Technical Proofs

Proof of Lemma 2.1. We construct $(X(t), Z(t))$ for $t \geq 0$ based on an iterative procedure as in Section 4 of Lando (1998). More precisely, for $j = 1, \ldots, N$, we first consider the following SDE given by

$$dX_j^{(0)}(t) = \left(\kappa_j - \nu_j X_j^{(0)}(t)\right)dt + \sum_{k=1}^{K} \sigma_{jk}(X_j^{(0)}(t))dW_k(t) \tag{A.1}$$

with initial condition $X_j^{(0)}(0) = X_j(0) \in \mathbb{R}_+$. Observe that for each $j = 1, \ldots, N$, Eq. (A.1) is a SDE with a linear drift and a square root volatility coefficient. Then Proposition 2.13 in Chapter 5 of Karatzas and Shreve (1991) implies that there is a unique (strong) solution to Eq. (A.1) for each $j = 1, \ldots, N$.

Let $(\Theta_{ij};\ i,j=1,\dots,N)$ be independent standard exponentially distributed random variables, independent of the K-dimensional Brownian motion. Initially, no firm has defaulted, i.e., $Z(0)=0$. Then the first default time $\hat{\tau}_1$ is

$$\begin{aligned}\hat{\tau}_1 &= \min_{j=1,\dots,N} \tau_{1j},\\ \tau_{1j} &:= \inf\left\{t\geq 0;\ \int_0^t X_j^{(0)}(u)du \geq \Theta_{1j}\right\},\qquad j=1,\dots,N.\end{aligned}$$

For $j=1,\dots,N$, we set $X_j(u)=X_j^{(0)}(u)$ and $Z(u)=Z(0)=0$ when $u\in[0,\hat{\tau}_1)$. Further define $j_1 := \underset{j=1,\dots,N}{\arg\min}\,\tau_{1j}$. We next consider the following system of SDEs, on $t\geq\hat{\tau}_1$: for $j\in\{1,\dots,N\}\setminus\{j_1\}$, and $t\geq\hat{\tau}_1$,

$$X_j^{(1)}(t) = X_j^{(0)}(\hat{\tau}_1) + \int_{\hat{\tau}_1}^t \left(\kappa_j - \nu_j X_j^{(1)}(u)\right)du + \sum_{k=1}^K \int_{\hat{\tau}_1}^t \sigma_{jk}(X_j^{(1)}(u))dW_k^{(1)}(u) + w_{j_1 j}. \tag{A.2}$$

Above, $W_k^{(1)}(t) := W_k(t+\hat{\tau}_1) - W_k(\hat{\tau}_1)$ for $t\geq 0$. Obviously Eq. (A.2) admits a unique (strong) solution $X_j^{(1)}(t)$ on $t\geq\hat{\tau}_1$. Further, define the second default time by

$$\begin{aligned}\hat{\tau}_2 &= \min_{j\in\{1,\dots,N\}\setminus\{j_1\}} \tau_{2j},\\ \tau_{2j} &:= \inf\left\{t\geq \hat{\tau}_1;\ \int_{\hat{\tau}_1}^t X_j^{(1)}(u)du \geq \Theta_{2j}\right\},\qquad j\in\{1,\dots,N\}\setminus\{j_1\}.\end{aligned}$$

Similarly to the above construction of $(X(t),Z(t))$ on $t\in[0,\hat{\tau}_1)$, for $u\in[\hat{\tau}_1,\hat{\tau}_2)$, we set $X_j(u)=X_j^{(1)}(u)$ for all $j\in\{1,\dots,N\}\setminus\{j_1\}$, and $Z(u)=Z(\hat{\tau}_1)=0^{j_1}$. Moreover, denote by $j_2 := \underset{j\in\{1,\dots,N\}\setminus\{j_1\}}{\arg\min}\tau_{2j}$. More generally, for $n=3,\dots,N$, we consider the following n-th default time among the N firms

$$\begin{aligned}\hat{\tau}_n &= \min_{j\in\{1,\dots,N\}\setminus\{j_1,\dots,j_{n-1}\}} \tau_{nj},\\ \tau_{nj} &:= \inf\left\{t\geq \hat{\tau}_{n-1};\ \int_{\hat{\tau}_{n-1}}^t X_j^{(n-1)}(u)du \geq \Theta_{nj}\right\},\quad j\in\{1,\dots,N\}\setminus\{j_1,\dots,j_{n-1}\}.\end{aligned}$$

The indices $j_1,\dots,j_{n-1}$ are recursively defined in a similar way to j_1 and j_2. For $j\in\{1,\dots,N\}\setminus\{j_1,\dots,j_{n-1}\}$, and $t\geq\hat{\tau}_{n-1}$,

$$\begin{aligned}X_j^{(n-1)}(t) &= X_j^{(n-2)}(\hat{\tau}_{n-1}) + \int_{\hat{\tau}_{n-1}}^t \left(\kappa_j - \nu_j X_j^{(n-1)}(u)\right)du\\ &\quad + \sum_{k=1}^K \int_{\hat{\tau}_{n-1}}^t \sigma_{jk}(X_j^{(n-1)}(u))dW_k^{(n-1)}(u) + \sum_{i\in\{j_1,\dots,j_{n-1}\}} w_{ij},\end{aligned} \tag{A.3}$$

where $W_k^{(n-1)}(t) := W_k(t+\hat{\tau}_{n-1}) - W_k(\hat{\tau}_{n-1})$ for $t\geq 0$. Obviously, Eq. (A.3) admits a unique (strong) solution. We can repeat the above recursive procedure for $n=1,2$ and construct $(X(t),Z(t))$ on $t\in[\hat{\tau}_{n-1},\hat{\tau}_n)$ until $n=N$. If $t\geq\hat{\tau}_N$, then all firms in the pool have defaulted. This completes the proof of the lemma. □

Proof of Proposition 2.3. Combining the bond price (9) with the price representation (10), we have

$$P_i(t) = F_i(t, X(t), Z(t)),\qquad i=1,\dots,N. \tag{A.4}$$

Here the price function F_i is defined as

$$F_i(t,x,z) := F_i^a(t,x,z) + F_i^b(t,x,z) + F_i^c(t,x,z), \tag{A.5}$$

for $(t,x,z)\in[0,T]\times\mathbb{R}_+^N\times\mathcal{S}$. On the event $\tau_i>t$, the price subfunctions are given by

$$F_i^a(t,x,z) := \mathbb{E}\left[\int_t^{T_i\wedge\tau_i} R_i(Z(u))X_i(u)e^{-r(u-t)}du \Big| X(t)=x, Z(t)=z\right],$$

$$F_i^b(t,x,z) := \mathbb{E}\left[\int_t^{T_i\wedge\tau_i} C_i e^{-r(u-t)}du \middle| X(t)=x, Z(t)=z\right], \tag{A.6}$$
$$F_i^c(t,x,z) := \mathbb{E}\left[(1-Z_i(T_i))e^{-r(T_i-t)}\middle| X(t)=x, Z(t)=z\right].$$

Let the difference-differential operator $\mathcal{A}$ acting on the smooth function $f(\cdot,z)$ for each $z\in\mathcal{S}$ be defined as

$$\mathcal{A}f(x,z) := \mu(x)^\top D_x f(x,z) + \frac{1}{2}\mathrm{Tr}[(\sigma\sigma^\top)(x)D_x^2 f(x,z)]$$
$$+\sum_{j=1}^N \left[f(x+w_j,z^j)-f(x,z)\right](1-z_j)x_j. \tag{A.7}$$

Using the Feynman-Kac's formula, we deduce that F_i^a, F_i^b and F_i^c defined by (A.6) satisfy PDEs

$$\begin{aligned}
&\left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i^a(t,x,z)+R_i(z)(1-z_i)x_i = rF_i^a(t,x,z), \qquad F_i^a(T_i,x,z)=0,\\
&\left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i^b(t,x,z)+C_i(1-z_i) = rF_i^b(t,x,z), \qquad F_i^b(T_i,x,z)=0,\\
&\left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i^c(t,x,z) = rF_i^c(t,x,z), \qquad F_i^c(T_i,x,z)=1-z_i.
\end{aligned} \tag{A.8}$$

Then the price function F_i given by (A.5) satisfies

$$\begin{aligned}
\left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i(t,x,z) &= \left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i^a(t,x,z)+\left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i^b(t,x,z)+\left(\frac{\partial}{\partial t}+\mathcal{A}\right)F_i^c(t,x,z)\\
&= rF_i^a(t,x,z)-R_i(z)(1-z_i)x_i+rF_i^b(t,x,z)-C_i(1-z_i)+rF_i^c(t,x,z)\\
&= rF_i(t,x,z)-R_i(z)(1-z_i)x_i-C_i(1-z_i).
\end{aligned}$$

Thus from Itô's formula, it follows that

$$\begin{aligned}
&F_i(t,X(t),Z(t)) = F_i(0,X(0),Z(0))\\
&\quad+\int_0^t \{rF_i(u,X(u),Z(u))-R_i(Z(u))(1-Z_i(u))X_i(u)-C_i(1-Z_i(u))\}\,du\\
&\quad+\int_0^t D_xF_i(u,X(u),Z(u))^\top\sigma(X(u))dW(u)\\
&\quad+\sum_{j=1}^N\int_0^t \left[F_i(u,X(u-)+w_j,Z^j(u-))-F_i(u,X(u-),Z(u-))\right]dM_j(u),
\end{aligned}$$

Then the dynamics of price $P_i(t)$ in Proposition 2.3 follows from (A.4). □

Remark A.1. *Using the operator $\mathcal{A}$ defined by (A.7), we have that the price function F_i indeed solves*

$$\left(\frac{\partial}{\partial t}+\bar{\mathcal{A}}\right)F_i(t,x,z)-\left(r+\sum_{j=1}^N(1-z_j)x_j\right)F_i(t,x,z)+R_i(z)(1-z_i)x_i+C_i(1-z_i)$$
$$+\sum_{j=1}^N F_i(t,x+w_j,z^j)(1-z_j)x_j=0 \tag{A.9}$$

with terminal condition $F_i(T,x,z)=1-z_i$. The operator $\bar{\mathcal{A}}$ is defined as

$$\bar{\mathcal{A}}f(t,x,z) := \mu(x)^\top D_x f(t,x,z)+\frac{1}{2}\mathrm{Tr}[(\sigma\sigma^\top)(x)D_x^2 f(t,x,z)]. \tag{A.10}$$

Recall the notation introduced at the end of Section 1, and let the $z=0^{j_1,\ldots,j_m}$ for $m=0,1,\ldots,N$. We can solve the above PDE recursively. When $m=N$, Eq. (A.9) reduces to

$$\left(\frac{\partial}{\partial t}+\bar{\mathcal{A}}\right)F_i(t,x,1)-rF_i(t,x,1)=0, \qquad F_i(T_i,x,1)=0.$$

Then $F_i(t,x,1)=0$. Consider the case $m=N-1$, Eq. (A.9) is given by

$$0=\left(\frac{\partial}{\partial t}+\bar{\mathcal{A}}\right)F_{i;j_1,\dots,j_{N-1}}(t,x)-(r+x_{j_N})\,F_{i;j_1,\dots,j_{N-1}}(t,x)$$
$$+R_{i;j_1,\dots,j_{N-1}}\mathbf{1}_{i=j_N}x_{j_N}+C_{j_N}\mathbf{1}_{i=j_N} \tag{A.11}$$

with terminal condition $F_i(T,x,z)=1-\mathbf{1}_{i\neq j_N}$. Hence, it holds that $F_{i;j_1,\dots,j_{N-1}}(t,x)=0$ for $i\neq j_N$. Using Proposition 3.2 in Becherer and Schweizer (2005) along with the fact that $\det(\sigma\sigma^\top(x))\neq 0$, Eq. (A.11) admits a unique classical solution given by

$$F_{j_N;j_1,\dots,j_{N-1}}(t,x)$$
$$=\mathbb{E}_{t,x}\left[e^{-\int_t^{T_i}(r+\bar{X}_{j_N}(s))ds}+\int_t^{T_i}\left(R_{i;j_1,\dots,j_{N-1}}\bar{X}_{j_N}(u)+C_{j_N}\right)e^{-\int_t^u(r+\bar{X}_{j_N}(s))ds}\right], \tag{A.12}$$

where the underlying state process is the multiple-dimensional CIR process given by

$$d\bar{X}_j(t)=(\kappa_j-\nu_j\bar{X}_j(t))dt+\sum_{k=1}^K\sigma_{jk}\sqrt{\bar{X}_j(t)}dW_k(t). \tag{A.13}$$

It can also be seen that $F_{j_N;j_1,\dots,j_{N-1}}(t,x)$ only depends on (t,x_{j_N}). For the case $0\le m\le N-2$, we have

$$\left(\frac{\partial}{\partial t}+\bar{\mathcal{A}}\right)F_{i;j_1,\dots,j_m}(t,x)-\left(r+\sum_{j\notin\{j_1,\dots,j_m\}}x_j\right)F_{i;j_1,\dots,j_m}(t,x)+R_{i;j_1,\dots,j_m}\mathbf{1}_{i\notin\{j_1,\dots,j_m\}}x_i$$
$$+C_{i;j_1,\dots,j_m}\mathbf{1}_{i\notin\{j_1,\dots,j_m\}}+\sum_{j\notin\{j_1,\dots,j_m\}}F_{i;j_1,\dots,j_m,j}(t,x+w_j)x_j=0 \tag{A.14}$$

with terminal condition $F_i(T,x,z)=1-\mathbf{1}_{i\in\{j_1,\dots,j_m\}}$. Notice that $F_{i;j_1,\dots,j_m,j}(t,x)$ is the classical solution to Eq. (A.9) in the default state $z=0^{j_1,\dots,j_m,j}$. Then, for all $i\in\{j_1,\dots,j_m\}$, we have that $F_{i;j_1,\dots,j_m}(t,x)=0$. Proposition 3.2 in Becherer and Schweizer (2005) along with the fact that $\det(\sigma\sigma^\top(x))\neq 0$ imply that Eq. (A.14) admits a classical solution given by

$$F_{i;j_1,\dots,j_m}(t,x) \tag{A.15}$$
$$=\mathbb{E}_{t,x}\left[e^{-\int_t^{T_i}\left(r+\sum_{j\notin\{j_1,\dots,j_m\}}\bar{X}_j(s)\right)ds}+\int_t^{T_i}\left(R_{i;j_1,\dots,j_m}\bar{X}_i(u)+C_i\right)e^{-\int_t^u\left(r+\sum_{j\notin\{j_1,\dots,j_m\}}\bar{X}_j(s)\right)ds}\right],$$

for all $i\notin\{j_1,\dots,j_m\}$. Moreover, $F_{i;j_1,\dots,j_m}(t,x)$ only depends on $(t,x_{j_{m+1}},\dots,x_{j_N})$. From the expression (10), it can be easily seen that $F_i(\cdot,z)$ admits a strictly positive lower and upper bound. By the results in Chapter VI of Ladyzenskaja, Solonnikov, and Uralceva (1968) (see also the form (3.1) therein), it can be verified that $H_{(i,j)}(\cdot,z)\in C^{1,2}$ and that $G_{(i,j)}(\cdot,z)\in C^{1,2}$ is bounded for any $z\in\mathcal{S}$.

Lemma A.2. *For $k=1,\dots,K$ and $i=1,\dots,N$, let $\phi_k(t,x,z)$ and $h_i(t,x,z)$ be sufficiently regular functions in $(t,x)\in[0,T]\times\mathbb{R}_+^N$ for each default state $z\in\mathcal{S}$, taking values on $\mathbb{R}$ and $(-1,\infty)$, respectively. Assume that the positive process $\xi_{\phi,h}=(\xi_{\phi,h}(t))_{t\ge0}$, satisfies the following SDE*

$$\frac{d\xi_{\phi,h}(t)}{\xi_{\phi,h}(t-)}=\sum_{k=1}^K\phi_k(t,X(t),Z(t))dW_k(t)+\sum_{j=1}^N h_j(t,X(t-),Z(t-))dM_j(t),\quad \xi_{\phi,h}(0)=1. \tag{A.16}$$

Let $T>0$. Define a new probability measure $\mathbb{P}\ll\mathbb{Q}$ on $\mathcal{G}_T$ by $d\mathbb{P}=\xi_{\phi,h}(T)d\mathbb{Q}$. Then, for each $j=1,\dots,N$,

$$W_k^{\mathbb{P}}(t):=W_k(t)-\int_0^t\phi_k(u,X(u),Z(u))du,\qquad t\in[0,T],\quad\text{and}$$
$$M_j^{\mathbb{P}}(t):=M_j(t)-\int_0^t(1-Z_j(u))X_j(u)h_j(u,X(u),Z(u))du,\qquad t\in[0,T] \tag{A.17}$$

are $\mathbb{P}$-martingales.

Lemma A.3. *For $i = 1, \dots, N$, under the actual probability measure $\mathbb{P}$, it holds that for $t \in [0, T_i)$,*

$$\frac{d\big(P_i(t) + D_i(t)\big)}{P_i(t-)} = rdt \tag{A.18}$$
$$+ \sum_{k=1}^{K} \left(\sum_{j=1}^{N} H_{(i,j)}(t, X(t), Z(t))\sigma_{jk}(X(t)) \right) d\left\{ W_k^{\mathbb{P}}(t) + \int_0^t \phi_k(u, X(u), Z(u))du \right\}$$
$$+ \sum_{j=1}^{N} G_{(i,j)}(t, X(t-), Z(t-))d\left\{ M_j^{\mathbb{P}}(t) + \int_0^t (1 - Z_j(u))X_j(u)h_j(u, X(u), Z(u))du \right\}.$$

Above, for $(t, x, z) \in [0, T_i] \times \mathbb{R}_+^N \times \mathcal{S}$, we have defined the functions

$$G_{(i,j)}(t, x, z) := \bar{H}_{(i,j)}(t, x, z), \ \text{if } i \neq j, \quad \text{and} \quad G_{(i,i)}(t, x, z) := \frac{R_i(z)}{F_i(t, x, z)} - 1. \tag{A.19}$$

For $i, j = 1, \dots, N$, the functions $H_{(i,j)}(t, x, z)$ and $\bar{H}_{(i,j)}(t, x, z)$ have been defined in (11).

Proof of Lemma A.3. Using Proposition 2.3, it follows that for $t \in [0, T_i)$,

$$P_i(t) = P_i(0) + \int_0^t \left\{ rP_i(u) - (1 - Z_i(u))[C_i + R_i(Z(u))X_i(u)] \right\} du$$
$$+ \int_0^t P_i(u) \sum_{k=1}^{K} \left(\sum_{j=1}^{N} H_{(i,j)}(u, X(u), Z(u))\sigma_{jk}(X(u)) \right) dW_k(u)$$
$$+ \int_0^t P_i(u-) \sum_{j=1}^{N} \bar{H}_{(i,j)}(u, X(u-), Z(u-))dM_j(u).$$

Taking the dividend given by (8) into account, we obtain from (A.4) that, for $t \in [0, T_i)$,

$$P_i(t) + D_i(t) = P_i(0) + D_i(0) + r\int_0^t P_i(u)du + \int_0^t P_i(u-)\frac{R_i(Z(u-))}{F_i(u, X(u-), Z(u-))}dM_i(u)$$
$$+ \int_0^t P_i(u) \sum_{k=1}^{K} \left(\sum_{j=1}^{N} H_{(i,j)}(u, X(u), Z(u))\sigma_{jk}(X(u)) \right) dW_k(u)$$
$$+ \int_0^t P_i(u-) \sum_{j=1}^{N} \bar{H}_{(i,j)}(t, X(u-), Z(u-))dM_j(u).$$

This yields the dynamics given by (A.18) using (A.17). □

Remark A.4. *Let $(t, x) \in [0, T] \times \mathbb{R}_+^N$. By Remark A.1 and (11), for all $(i, j) \in \{j_{m+1}, \dots, j_N\} \times \{j_1, \dots, j_m\}$,*

$$H_{(i,j)}(t, x, 0^{j_1,\dots,j_m}) = \frac{\frac{\partial F_i(t, x_{j_{m+1}}, \dots, x_{j_N}, 0^{j_1,\dots,j_m})}{\partial x_j}}{F_i(t, j_{m+1}, \dots, x_{j_N}, 0^{j_1,\dots,j_m})} = 0,$$

and for all $(i, j) \in \{j_{m+1}, \dots, j_N\}^2$, it holds that

$$H_{(i,j)}(t, x, 0^{j_1,\dots,j_m}) = \frac{\frac{\partial F_i(t, x_{j_{m+1}}, \dots, x_{j_N}, 0^{j_1,\dots,j_m})}{\partial x_j}}{F_i(t, j_{m+1}, \dots, x_{j_N}, 0^{j_1,\dots,j_m})} =: H_{(i,j)}(t, j_{m+1}, \dots, x_{j_N}, 0^{j_1,\dots,j_m}),$$

i.e., $H_{(i,j)}(t, x, 0^{j_1,\dots,j_m})$ only depends on the arguments $(t, j_{m+1}, \dots, x_{j_N})$. Similarly, we have that for all $(i, j) \in \{j_{m+1}, \dots, j_N\}^2$, $G_{(i,j)}(t, x, 0^{j_1,\dots,j_m})$ only depends on the arguments $(t, j_{m+1}, \dots, x_{j_N})$.

Proof of Proposition 3.3. Fix $(t, x) \in [0, T] \times \mathbb{R}_+^N$, and omit the argument $(j_1, \dots, j_m)$ to lighten notation in the remaining part of the section. Using (23), from Remark A.4, it follows that the objective function (23) in the default state $z = 0^{j_1,\dots,j_m}$ is

$$\mathcal{H}(t, x, \pi) = Q(t, x) \left[\gamma r - \sum_{j \notin \{j_1,\dots,j_m\}} x_j(1 + h_j(t, x)) \right]$$

$$+\gamma Q(t,x)\sum_{i\notin\{j_1,\dots,j_m\}}\pi_i\left[\sum_{k=1}^K\phi_k(t,x)\left(\sum_{j\notin\{j_1,\dots,j_m\}}H_{(i,j)}(t,x)\sigma_{jk}(x_j)\right)-\sum_{j\notin\{j_1,\dots,j_m\}}x_jG_{(i,j)}(t,x)\right]$$

$$+\gamma\sum_{i\notin\{j_1,\dots,j_m\}}\pi_i\left[\sum_{k=1}^K\left(\sum_{j=1}^N\frac{\partial Q(t,x)}{\partial x_j}\sigma_{jk}(x_j)\right)\left(\sum_{l\notin\{j_1,\dots,j_m\}}H_{(i,l)}(t,x)\sigma_{lk}(x_l)\right)\right]$$

$$+\frac{\gamma(\gamma-1)}{2}Q(t,x)\sum_{k=1}^K\left(\sum_{i\notin\{j_1,\dots,j_m\}}\pi_i\left(\sum_{j\notin\{j_1,\dots,j_m\}}H_{(i,j)}(t,x)\sigma_{jk}(x_j)\right)\right)^2$$

$$+\sum_{j\notin\{j_1,\dots,j_m\}}\left(1+\sum_{i\notin\{j_1,\dots,j_m\}}\pi_iG_{(i,j)}(t,x)\right)^\gamma Q_j(t,x)x_j(1+h_j(t,x)). \tag{A.20}$$

It holds that $\lim_{|\pi|\to\infty}\mathcal{H}(t,x,\pi)=-\infty$ for almost every $(t,x)\in[0,T]\times\mathbb{R}_+^N$ because $\gamma\in(0,1)$ and $Q(t,x):=Q(t,x,0^{j_1,\dots,j_m})>0$. Fix some point $\pi_0\in\mathcal{U}$, we have that for almost every $(t,x)\in[0,T]\times\mathbb{R}_+^N$, there exists a positive constant $\zeta(t,x)>0$ such that for all $|\pi|>\zeta(t,x)$, $\mathcal{H}(t,x,\pi_0)>\mathcal{H}(t,x,\pi)$. This implies that the set of solutions to the maximization problem of $\mathcal{H}(t,x,\pi)$ over $\pi\in\mathcal{J}^{(N-m)}$ coincides with the set of solutions to the maximization problem of $\mathcal{H}(t,x,\pi)$ over the compact set of $\mathbb{R}^{N-m}$ given by $\pi\in\{\pi\in\mathbb{R}^{N-m};\ |\pi|\le\zeta(t,x)\}\cap\mathcal{J}^{(N-m)}$. Here $\mathcal{J}^{(N-m)}$ is the admissible set in the default state $z=0^{j_1,\dots,j_m}$, which is given by (24). By continuity of the function $\mathcal{H}(t,x,\pi)$ in π, we deduce the existence of

$$\pi^*(t,x)\in\underset{\pi\in\{\pi\in\mathbb{R}^{N-m};\ |\pi|\le\zeta(t,x)\}\cap\mathcal{J}^{(N-m)}}{\arg\max}\mathcal{H}(t,x,\pi)=\underset{\pi\in\mathcal{J}^{(N-m)}}{\arg\max}\ \mathcal{H}(t,x,\pi) \tag{A.21}$$

for almost every $(t,x)\in[0,T]\times\mathbb{R}_+^N$. Moreover, one can choose a Borel measurable version by a classical measurable selection theorem (see, e.g., Appendix B in Fleming and Rishel (1975)). Because $\mathcal{H}(t,x,\pi)$ is strictly concave, and hence $\pi^*(t,x)$ is unique. □

Proof of Theorem 4.1. Let $Q_j(t,x):=Q(t,x,0^{j_1,\dots,j_m,j})$ be the classical solution of the HJB equation in the state $z=0^{j_1,\dots,j_m,j}$ for $j\notin\{j_1,\dots,j_m\}$. Then $Q_j(t,x)$ belongs to $C^{1,2}$. We next introduce the following space of functions. For a set $\mathcal{O}\subset[0,T]\times\mathbb{R}^N$ and $p\in(1,+\infty)$, let $L^p(\mathcal{O})$ be the space of p-th order integrable functions on $\mathcal{O}$ with the norm $\|\cdot\|_{L^p(\mathcal{O})}$ in $L^p(\mathcal{O})$. Let $H^p(\mathcal{O})$ be the Sobolev space of functions f such that f together with all generalized partial derivatives of first and second order are in $L^p(\mathcal{O})$. Denote by $\|\cdot\|_{H^p(\mathcal{O})}$ the Sobolev norm of the space $H^p(\mathcal{O})$. We also introduce the norms of Hölder type. For any $\gamma\in(0,1]$, and $\mathcal{O}=[T_0,T_1]\times K\subset[0,T]\times\mathbb{R}^m$, define $\|f\|_{\mathcal{O}}:=\sup_{(t,y)\in\mathcal{O}}|f(t,y)|$, and

$$\|f\|_{\gamma,\mathcal{O}}:=\|f\|_{\mathcal{O}}+\sup_{t\in[T_0,T_1],\ (x,y)\in K,\ x\neq y}\frac{|f(t,x)-f(t,y)|}{|x-y|^\gamma}+\sup_{(s,t)\in[T_0,T_1],\ s\neq t,\ x\in K}\frac{|f(t,x)-f(s,x)|}{|t-s|^\gamma}.$$

We also define $\|f\|^{(1)}_{\gamma,\mathcal{O}}:=\|f\|_{\gamma,\mathcal{O}}+\sum_{i=1}^N\|D_{x_i}f\|_{\gamma,\mathcal{O}}$ and $\|f\|^{(2)}_{\gamma,\mathcal{O}}:=\|f\|^{(1)}_{\gamma,\mathcal{O}}+\|D_tf\|^{(1)}_{\gamma,\mathcal{O}}+\sum_{i,j=1}^N\|D^2_{x_ix_j}f\|^{(1)}_{\gamma,\mathcal{O}}$.

Consider the bounded domain $B_\nu:=\{x\in(0,\infty)^N;\ |x|<\nu\}$ for any $\nu>0$. For $(t,x)\in E_\nu:=(0,T)\times B_\nu$, introduce the following problem

$$\frac{\partial u^\nu}{\partial t}+\hat{\mathcal{A}}u^\nu+\bar{\mathcal{H}}(t,x,u^\nu,D_xu^\nu)=0 \tag{A.22}$$

subject to boundary conditions

$$u^\nu(t,x)=\varphi(t,x)\qquad\text{for }(t,x)\in\partial E_\nu:=((0,T)\times\partial E_\nu)\cup(\{T\}\times E_\nu). \tag{A.23}$$

The function $\varphi\in C^{1,2}(\bar{E}_\nu)$ and satisfy $\varphi(T,x)=\gamma^{-1}$ for all $x\in B_\nu$. We next define sequences of functions $(u_k^\nu=u_k^\nu(t,x))_{k\ge1}$ on E_ν and of bounded feedback control functions $(\pi_k=\pi_k(t,x))_{k\ge0}$ as follows. Here, π_0 is an arbitrary control. Define u_{k+1}^ν to be the (Sobolev) solution of the following problem

$$\begin{cases}0=\frac{\partial u_{k+1}^\nu}{\partial t}+\hat{\mathcal{A}}u_{k+1}^\nu+g(t,x,\pi_k^\nu(t,x))u_{k+1}^\nu+f(t,x,\pi_k^\nu(t,x))D_xu_{k+1}^\nu+l(t,x,\pi_k^\nu(t,x))\\ \text{and subject to terminal condition } u_{k+1}^\nu(t,x)=\varphi(t,x),\ \text{for }(t,x)\in\partial E_\nu,\\ \pi_k^\nu(t,x):=\underset{\pi\in\mathcal{J}^{(N-m)}}{\arg\max}\ \{g(t,x,\pi)u_k^\nu(t,x)+f(t,x,\pi)D_xu_k^\nu(t,x)+l(t,x,\pi)\}\\ \text{for almost all }E_\nu.\end{cases} \tag{A.24}$$

We can rewrite Eq. (A.24) in the following semi-linear form given by

$$0 = \frac{\partial u_{k+1}^\nu}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 u_{k+1}^\nu] + f_k^\nu(t,x)D_x u_{k+1}^\nu + g_k^\nu(t,x)u_{k+1}^\nu + l_k^\nu(t,x), \tag{A.25}$$

where for $(t,x)\in E_\nu$, the coefficients

$$\begin{aligned} &f_k^\nu(t,x) := \mu(x)^\top + \sigma(x)\phi(t,x) + f(t,x,\pi_k^\nu(t,x)),\\ &g_k^\nu(t,x) := g(t,x,\pi_k^\nu(t,x)), \quad l_k^\nu(t,x) := l(t,x,\pi_k^\nu(t,x)). \end{aligned} \tag{A.26}$$

Obviously $\sigma\sigma^\top(x)$ is bounded and Lipschitz continuous on $\bar{E}_\nu$. As $\det(\sigma\sigma^\top(x)) \neq 0$, it follows that $\sigma\sigma^\top(x)$ is uniformly elliptic on $\mathbb{R}_+^N$ for $x\in B_\nu$ using Lemma 3 in Heath and Schweizer (2001). Because $H(\cdot)\in C^{1,1}$ and $G(\cdot), Q_j(\cdot)\in C^{1,2}$, it holds that $f_k^\nu(t,x)$ and $g_k^\nu(t,x)$ are bounded on E_ν. Also, notice that $l_k^\nu(t,x)$ is bounded on E_ν, and hence $\|l_k^\nu\|_{L^p(E_\nu)} < +\infty$ for $1<p<\infty$. Using the existence and uniqueness of Sobolev space-valued solutions to the linear PDE in Fleming and Rishel (1975), pag. 207, Eq. (A.25) admits a unique solution $u_{k+1}^\nu \in H^p(E_\nu)$ for any $1<p<\infty$; moreover the following estimate holds

$$\left\|u_{k+1}^\nu\right\|_{H^p(E_\nu)} \le C_\nu\big(\|l_k^\nu\|_{L^p(E_\nu)} + \|\varphi\|_{H^p(\partial E_\nu)}\big). \tag{A.27}$$

Using the estimate (E.9) in Fleming and Rishel (1975), pag. 207, it follows that

$$\left\|u_{k+1}^\nu\right\|_{\rho,E_\nu}^{(1)} \le C_{\nu,p}\left\|u_{k+1}^\nu\right\|_{H^p(E_\nu)}, \tag{A.28}$$

for $\rho = 1-\frac{N+2}{p}$ provided $p>N+2$.

We next develop the related estimates (A.27) and (A.28) for the solution u_{k+1}^ν. First, we establish the limit of u_k^ν as $k\to\infty$ in some appropriate spaces. This limit is verified to be the classical solution to Eq. (A.22) subject to boundary conditions (A.23). For $k\ge 1$, using (A.24), it follows that

$$\begin{aligned} &\frac{\partial u_k^\nu}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 u_k^\nu] + f_k^\nu(t,x)D_x u_k^\nu + g_k^\nu(t,x)u_k^\nu + l_k^\nu(t,x)\\ &\quad\ge \frac{\partial u_k^\nu}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 u_k^\nu] + f_{k-1}^\nu(t,x)D_x u_k^\nu + g_{k-1}^\nu(t,x)u_k^\nu + l_{k-1}^\nu(t,x) = 0, \text{ in } E_\nu, \end{aligned}$$

and hence setting $v_k^\nu := u_{k+1}^\nu - u_k^\nu$, it holds that

$$\frac{\partial v_k^\nu}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 v_k^\nu] + f_k^\nu(t,x)D_x v_k^\nu + g_k^\nu(t,x)v_k^\nu + l_k^\nu(t,x) \le 0, \quad \text{in } E_\nu$$

subject to the boundary condition $v_k(T,x)=0$ on ∂E_ν. The maximum principle implies that $v_k^\nu \ge 0$ on E_ν for $k\ge 1$. Thus, the constructed sequence of functions $(u_k^\nu)_{k\ge1}$ determined by (A.24) is non-decreasing. Using the estimates (A.27) and (A.28) established above, we have for $p>N+2$,

$$\left\|u_{k+1}^\nu\right\|_{\rho,E_\nu}^{(1)} \le C_{\nu,p}\left\|u_{k+1}^\nu\right\|_{H^p(E_\nu)} \le C_\nu C_{\nu,p}\big(\|l_k^\nu\|_{L^p(E_\nu)} + \|\varphi\|_{H^p(\partial E_\nu)}\big). \tag{A.29}$$

Then u_{k+1}^ν is bounded and $u_{k+1}^\nu, Du_{k+1}^\nu$ are continuous on E_ν. Further, as $k\to\infty$ we have that u_{k+1}^ν, Du_{k+1}^ν converges to the limit (denote by u^ν) and the gradient (denote by Du^ν) of u_{k+1}^ν uniformly on $\bar{E}_\nu$. The functions $\frac{\partial u_{k+1}^\nu}{\partial t}$ and Du_{k+1}^ν weakly converge, respectively, to $\frac{\partial u^\nu}{\partial t}$ and D^2u^ν in $L^p(E_\nu)$. For any admissible feedback control $\pi(t,x)$, we have

$$\begin{aligned} &\frac{\partial u_k^\nu}{\partial t} + \hat{\mathcal{A}}u_k^\nu + g(t,x,\pi(t,x))u_k^\nu + f(t,x,\pi(t,x))D_x u_k^\nu + l(t,x,\pi(t,x))\\ &\quad\le \frac{\partial u_k^\nu}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 u_k^\nu] + f_k^\nu(t,x)D_x u_k^\nu + g_k^\nu(t,x)u_k^\nu + l_k^\nu(t,x)\\ &\quad= -\frac{\partial v_k^\nu}{\partial t} - \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 v_k^\nu] - f_k^\nu(t,x)D_x v_k^\nu - g_k^\nu(t,x)v_k^\nu, \text{ in } E_\nu. \end{aligned}$$

As $k\to\infty$, the left hand side of the above display weakly converges to $\frac{\partial u^\nu}{\partial t} + \hat{\mathcal{A}}u^\nu + g(t,x,\pi(t,x))u^\nu + f(t,x,\pi(t,x))D_x u^\nu + l(t,x,\pi(t,x))$, while the right hand side converges to zero. This implies that, for almost all $(t,x)\in E_\nu$,

$$\frac{\partial u^\nu}{\partial t} + \hat{\mathcal{A}}u^\nu + g(t,x,\pi(t,x))u^\nu + f(t,x,\pi(t,x))D_x u^\nu + l(t,x,\pi(t,x)) \le 0. \tag{A.30}$$

For almost all $(t,x) \in E_\nu$, define

$$\pi^{\nu,*}(t,x) \in \underset{\pi \in \mathcal{J}^{(N-m)}}{\arg\max}\left\{g(t,x,\pi)u^\nu(t,x) + f(t,x,\pi)D_x u^\nu(t,x) + l(t,x,\pi)\right\}$$

which can be Borel measurable using Lemma VI.6.1 of Fleming and Rishel (1975). Thus, we have that for almost all $(t,x) \in E_\nu$,

$$\begin{aligned}
&\frac{\partial u^\nu}{\partial t} + \hat{A}u^\nu + g(t,x,\pi^{\nu,*}(t,x))u^\nu + f(t,x,\pi^{\nu,*}(t,x))D_x u^\nu + l(t,x,\pi^{\nu,*}(t,x)) \\
&\quad \geq \frac{\partial u^\nu}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2 u^\nu] + f_k^\nu(t,x)D_x u^\nu + g_k^\nu(t,x)u_k^\nu + l_k^\nu(t,x) \\
&\quad = \frac{\partial(u^\nu - u_{k+1}^\nu)}{\partial t} + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)D_x^2(u^\nu - u_{k+1}^\nu)] + f_k^\nu(t,x)D_x(u^\nu - u_{k+1}^\nu) \\
&\qquad + g_k^\nu(t,x)(u^\nu - u_{k+1}^\nu).
\end{aligned}$$

Letting $k \to \infty$ along with the inequality (A.30), we deduce that Eq. (A.22) holds for almost all $(t,x) \in E_\nu$. Further, we have that $u^\nu \in H^p(E_\nu)$. By the locally Lipschitz continuity of $\bar{\mathcal{H}}$, and using the estimate (A.29) with $p > N+2$, we deduce that $u^\nu \in C^{1,2}(E_\nu)$ from the estimate (E.10) in Fleming and Rishel (1975), pp. 208.

We next show the existence of a classical solution to Eq. (30) in the unbounded domain $[0,T] \times \mathbb{R}_+^N$, using a localization argument as in Davis and Lieo (2013). For $\nu \in \mathbb{N}$, let $\chi_\nu(x)$ be a nonnegative C^∞-function satisfying $|D\chi_\nu|_{\mathbb{R}_+^N} \leq 2$. The function $\chi_\nu(x) = 1$ if $x \in B_\nu$ and $\chi_\nu(x) = 0$ if $x \in \mathbb{R}_+^N \setminus B_{\nu+1}$. Let w^ν be the solution of the equation

$$\frac{\partial w^\nu}{\partial t} + \hat{A}w^\nu + \chi_\nu(x)\bar{\mathcal{H}}(t,x,w^\nu,D_x w^\nu) = 0 \tag{A.31}$$

with terminal condition $w^\nu(T,x) = \gamma^{-1}\chi_\nu(x)$ for all $x \in \mathbb{R}_+^N$. This equation takes the same form as Eq. (30); in particular, the functions $f(t,x,\pi)$, $g(t,x,\pi)$ and $l(t,x,\pi)$ are replaced, respectively, by $\chi_\nu(x)f(t,x,\pi)$, $\chi_\nu(x)g(t,x,\pi)$ and $\chi_\nu(x)l(t,x,\pi)$. Using the local estimate (A.27) above, it follows that $\|w^\nu\|_{H^p(E)}$ is bounded for any bounded subset $E \subset [0,T] \times \mathbb{R}_+^N$ for $1 < p < \infty$. Hence, Dw^ν satisfies a uniform Hölder condition on each bounded set E from (A.28). It can also be seen that, for any fixed $\nu_0 > 0$, w^ν is a solution of Eq. (30) in E_{ν_0} with $w^\nu(T,x) = \gamma^{-1}$ for $x \in B_{\nu_0}$ if $\nu \geq \nu_0$. Because $\mathcal{H}$ is locally Lipschitz continuous, both $\frac{\partial w^\nu}{\partial t}$ and $D^2 w^\nu$ satisfy a uniform Hölder condition on any compact subset $E \subset [0,T] \times \mathbb{R}_+^N$. Using Arzela-Ascoli's theorem, there exists a subsequence (ν_k) of $\mathbb{N}$ such that w^{ν_k} goes to a limit w uniformly on any compact subset of $[0,T] \times \mathbb{R}_+^N$, and $\frac{\partial w^{\nu_k}}{\partial t}, D_x w^{\nu_k}, D_x^2 w^{\nu_k}$ converge respectively to $\frac{\partial w}{\partial t}, D_x w, D_x^2 w$ uniformly on any compact set of $[0,T] \times \mathbb{R}_+^N$, as $k \to \infty$. From Eq. (A.31), we can conclude that w is a classical solution of Eq. (30) with terminal condition $w(T,x) = \gamma^{-1}$ for all $x \in \mathbb{R}_+^N$. To show the positivity of the solution, because $w \in C^{1,2}$, define for almost all $(t,x) \in [0,T] \times \mathbb{R}_+^N$,

$$\pi^*(t,x) := \underset{\pi \in \mathcal{J}^{(N-m)}}{\arg\max}\left\{g(t,x,\pi)w(t,x) + f(t,x,\pi)D_x w(t,x) + l(t,x,\pi)\right\}$$

which is Borel measurable using Lemma VI.6.1 of Fleming and Rishel (1975). Then in $(t,x) \in [0,T) \times \mathbb{R}_+^N$,

$$\frac{\partial w}{\partial t} + \hat{A}w + g^*(t,x)w + f^*(t,x)D_x w + l^*(t,x) = 0 \tag{A.32}$$

with terminal condition $w(T,x) = \gamma^{-1}$ for all $x \in D$. Above, $f^*(t,x) := f(t,x,\pi^*(t,x))$, $g^*(t,x) := g(t,x,\pi^*(t,x))$ and $l^*(t,x) := l(t,x,\pi^*(t,x))$ for $(t,x) \in [0,T] \times \mathbb{R}_+^N$. By virtue of (32), we have that $l^*(t,x)$ is strictly positive for all $(t,x) \in [0,T] \times \mathbb{R}_+^N$, and hence

$$\begin{aligned}
w(t,x) &= \mathbb{E}\left[\gamma^{-1}e^{\int_t^T g^*(s,\zeta(s))ds} + \int_t^T l^*(s,\zeta(s))e^{\int_t^s g^*(u,\zeta(u))du}ds \middle| \zeta(t) = x\right] \\
&> \mathbb{E}\left[\gamma^{-1}e^{\int_t^T g^*(s,\zeta(s))ds} \middle| \zeta(t) = x\right] > 0,
\end{aligned} \tag{A.33}$$

where the process $\zeta = (\zeta(t))_{t \in [0,T]}$ satisfies the SDE given by

$$d\zeta(t) = \left\{\mu(\zeta(t)) + f^*(t,\zeta(t))\right\}dt + \sigma(\zeta(t))dW(t). \tag{A.34}$$

This shows that the solution has a positive lower bound function given by $\gamma^{-1}e^{\gamma r(T-t)}$ for $t\in[0,T]$. $\square$

Proof of Theorem 4.4. We only prove the non-degenerate case. For any admissible feedback control $\pi=(\pi_i(t);\ t\in[0,T])_{i=1,\ldots,N}^{\top}\in\mathcal{U}$, we rewrite the wealth process given by (3.1) as under the physical probability measure $\mathbb{P}$,

$$\begin{aligned}\frac{dV^{\pi}(t)}{V^{\pi}(t-)}&=rdt+\pi(t)^{\top}\bar{L}(t,X(t),Z(t))dt+\pi(t)^{\top}H(t,X(t),Z(t))\sigma(X(t))dW^{\mathbb{P}}(t)\\&\quad+\pi(t)^{\top}G(t,X(t-),Z(t-))dZ(t)^{\top},\end{aligned}\tag{A.35}$$

where the coefficient $\bar{L}(\cdot)$ is defined in Remark 4.2. The default intensity processes may be rewritten as

$$dX(t)=\left\{\mu(X(t))+\sigma(X(t))\phi(t,X(t),Z(t))\right\}dt+\sigma(X(t))dW^{\mathbb{P}}(t)+w^{\top}dZ(t)^{\top}.$$

For $(t,v,x,z)\in[0,T]\times\mathbb{R}_+\times\mathbb{R}_+^N\times\mathcal{S}$, define the function $\xi(t,v,x,z):=v^{\gamma}Q(t,x,z)$. Because $Q(t,x,0^{j_1,\ldots,j_m})$ is a classical solution to the HJB equation (30) at the default state $z=0^{j_1,\ldots,j_m}$ for any $m=0,1,\ldots,N$, we can apply Itô's formula and obtain

$$d\xi(t,V^{\pi}(t),X(t),Z(t))=(V^{\pi}(t))^{\gamma}\left\{\frac{\partial}{\partial t}+\mathcal{L}_c^{\pi}+\mathcal{L}_J^{\pi}+\hat{\mathcal{A}}\right\}Q(t,X(t),Z(t))dt+dY^{\pi}(t),\tag{A.36}$$

where the operators $\mathcal{L}_c^{\pi}$, $\mathcal{L}_J^{\pi}$, $\hat{\mathcal{A}}$ are defined in (20) and (21) respectively and $Y^{\pi}=(Y^{\pi}(t))_{t\in[0,T]}$ is a $\mathbb{P}$-(local) martingale given by

$$\begin{aligned}Y^{\pi}(t):=&\int_0^t(V^{\pi}(s))^{\gamma}\left\{\gamma Q(s,X(s),Z(s))\pi(s)^{\top}H(s,X(s),Z(s))\right.\\&\qquad\left.+D_xQ(s,X(s),Z(s))^{\top}\right\}\sigma(X(s))dW(s)\\&+\int_0^t(V^{\pi}(s-))^{\gamma}\sum_{j=1}^N\left[\left(1+\sum_{i=1}^N\pi_i(t)G_{(i,j)}(t,X(s-),Z(s-))\right)^{\gamma}\right.\\&\qquad\left.\times Q(s,X(s-)+w_j,Z^j(s-))-Q(s,X(s-),Z(s-))\right]dM_j^{\mathbb{P}}(s).\end{aligned}\tag{A.37}$$

Recall that the optimal strategy $\pi^*\in\mathcal{J}^{(N-m)}$ is given by (A.21). Then using (A.36), for any $u\in[t,T]$,

$$\xi(u,V^{\pi^*}(u),X(u),Z(u))=\xi(t,V^{\pi^*}(t),X(t),Z(t))+Y^{\pi^*}(T)-Y^{\pi^*}(t).\tag{A.38}$$

We next fix $(t,v,x,z)\in[0,T]\times\mathbb{R}_+\times\mathbb{R}_+^N\times\mathcal{S}$. Denote by $\mathbb{E}_{t,v,x,z}^{\mathbb{P}}[\cdot]$ the conditional expectation with time-t quantities $V^{\pi^*}(t)=v$, $X(t)=x$ and $Z(t)=z$. Moreover, we define the stopping time $\tau_{ab}:=\inf\{s\geq t;\ V^{\pi^*}(s)\geq b^{-1}$ or $V^{\pi^*}(s)\leq a$ or $|X(s)|\geq b^{-1}\}$ where $0<a<v<b^{-1}<+\infty$. Notice that for any $z\in\mathcal{S}$, $H(\cdot,z),G(\cdot,z),Q(\cdot,z)\in C^{1,2}$, both $H(\cdot,z)$ and $G(\cdot,z)$ are bounded and $\sigma(\cdot)$ is sufficiently smooth. Moreover, $\pi^*(t,x,z)\in\mathcal{J}$ and is bounded. Then, for all $t\in[0,T]$,

$$\mathbb{E}_{t,v,x,z}^{\mathbb{P}}\left[Y^{\pi^*}(T\wedge\tau_{ab})-Y^{\pi^*}(t)\right]=0.$$

Hence, by choosing $u=T\wedge\tau_{ab}$ in Eq. (A.38), we obtain

$$\mathbb{E}_{t,v,x,z}^{\mathbb{P}}\left[\xi(T\wedge\tau_{ab},V^{\pi^*}(T\wedge\tau_{ab}),X(T\wedge\tau_{ab}),Z(T\wedge\tau_{ab}))\right]=\xi(t,v,x,z).\tag{A.39}$$

We next want to prove that

$$\begin{aligned}&\lim_{a,b\to0}\mathbb{E}_{t,v,x,z}^{\mathbb{P}}\left[\xi(T\wedge\tau_{ab},V^{\pi^*}(T\wedge\tau_{ab}),X(T\wedge\tau_{ab}),Z(T\wedge\tau_{ab}))\right]\\&\quad=\mathbb{E}_{t,v,x,z}^{\mathbb{P}}\left[\xi(T,V^{\pi^*}(T),X(T),Z(T))\right].\end{aligned}\tag{A.40}$$

By virtue of Corollary 7.1.5 in Chow and Teicher (1978) and Remark 4.2 above, in order to prove (A.40), it suffices to prove that there exists a constant $C>0$ so that for some $p>1$,

$$\sup_{a,b}\mathbb{E}_{t,v,x,z}^{\mathbb{P}}\left[(V^{\pi^*}(T\wedge\tau_{ab}))^{\gamma p}\right]\leq C_{t,v,x,z},\tag{A.41}$$

where $C_{t,v,x,z} > 0$ is a positive constant depending on (t, v, x, z) but independent of (a, b). Here, we take $p \in (1, \gamma^{-1})$ because $\gamma \in (0, 1)$ and hence $\gamma p \in (0, 1)$. Notice that $G(\cdot, z)$ and the default indicator process Z are bounded. Then we have for $s \in [t, T]$,

$$V^{\pi^*}(s) \le V^{\pi^*}(t) + \int_t^s V^{\pi^*}(u)\big\{r + \pi(u)^\top \bar{L}(u, X(u), Z(u))\big\}du$$
$$+ \int_t^s V^{\pi^*}(u)\pi(u)^\top H(u, X(u), Z(u))\sigma(X(u))dW(u) + C_N,$$

for some $C_N > 0$. Then there exists some positive constant $C_L > 0$ such that

$$\mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[V^{\pi^*}(T \wedge \tau_{ab})\right] \le v + (r + C_L)\mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[\int_t^T V^{\pi^*}(u \wedge \tau_{ab})du\right] + C_N.$$

By virtue of the Gronwall's lemma, it holds that $\mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[V^{\pi^*}(T \wedge \tau_{ab})\right] \le (v + C_N)e^{(r+C_L)T}$. Using the Jensen's inequality with $\gamma p \in (0, 1)$, we have

$$\sup_{a,b} \mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[(V^{\pi^*}(T \wedge \tau_{ab}))^{\gamma p}\right] \le \sup_{a,b}\left\{\mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[V^{\pi^*}(T \wedge \tau_{ab})\right]\right\}^{\gamma p} \le (v + C_N)^{\gamma p} e^{\gamma p(r+C_L)T},$$

which is independent of (a, b). This yields the above estimate (A.41). From (A.39) and (A.40), it follows that

$$\mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[\xi(T, V^{\pi^*}(T), X(T), Z(T))\right] = \xi(t, v, x, z) = v^\gamma Q(t, x, z), \tag{A.42}$$

while using the terminal condition $Q(T, x, z) = \gamma^{-1}$ for all $(x, z) \in D \times \mathcal{S}$, it follows that

$$\mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[\xi(T, V^{\pi^*}(T), X(T), Z(T))\right] = \mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[\frac{1}{\gamma}\big(V^{\pi^*}(T)\big)^\gamma\right] = \mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[U\big(V^{\pi^*}(T)\big)\right].$$

Using the equality (A.42), we then obtain that

$$\xi(t, v, x, z) = v^\gamma Q(t, x, z) = \mathbb{E}^{\mathbb{P}}_{t,v,x,z}\left[U\big(V^{\pi^*}(T)\big)\right] = \eta(t, v, x, z).$$

Hence, the proof of the verification theorem is completed. □

Proof of Proposition 4.3. In the default state $z = 0^{j_1,\dots,j_m}$, it follows from (25) that the Hamiltonian evaluated at the optimum feedback function is given by

$$\mathcal{H}(t, x, \pi^*(t, x)) = Q(t, x)\left\{\gamma r + \gamma \sum_{j \notin \{j_1,\dots,j_m\}} x_j - \sum_{j \notin \{j_1,\dots,j_m\}} x_j\big(1 + h_j(t, x)\big)\right\}$$
$$- \gamma Q(t, x) \sum_{j \notin \{j_1,\dots,j_m\}} x_j \left\{\frac{Q(t, x)}{Q_j(t, x + w_j)\big(1 + h_j(t, x)\big)}\right\}^{\frac{1}{\gamma-1}}$$
$$+ \sum_{j \notin \{j_1,\dots,j_m\}} x_j\big(1 + h_j(t, x)\big)\left\{\frac{Q(t, x)}{Q_j(t, x + w_j)\big(1 + h_j(t, x)\big)}\right\}^{\frac{\gamma}{\gamma-1}} Q_j(t, x + w_j).$$

Using the expressions for the coefficients $\mathcal{C}^a(\cdot)$ and $\mathcal{C}^b(\cdot)$ defined in (35), we can rewrite the above Hamiltonian as

$$\mathcal{H}(t, x, \pi^*(t, x)) = \mathcal{C}^a(t, x)Q(t, x) + \mathcal{C}^b(t, x)Q^{\frac{\gamma}{\gamma-1}}(t, x). \tag{A.43}$$

We may then rewrite the HJB equation as in Eq. (36). Recall that by the inductive hypothesis, for $j \notin \{j_1, \dots, j_m\}$, $Q_j(t, x)$, $(t, x) \in [0, T] \times \mathbb{R}^N_+$, is the positive continuously differentiable solution of Eq. (33) when the default state at time t is $z = 0^{j_1,\dots,j_m,j}$. Then the coefficients $\mathcal{C}^a(\cdot)$ and $\mathcal{C}^b(\cdot)$ defined in Eq. (35) are all continuously differentiable. Moreover, the coefficient $\mathcal{C}^b(\cdot) > 0$ because $\gamma < 1$ and $h_j(\cdot) > -1$. Solving Eq. (36) yields the solutions given by (37) using the power transformation of the solutions for the Bernoulli's type equation. This completes the proof of the theorem. □

B The Crank-Nicolson Method

This section describes the Crank-Nicolson scheme used to solve the HJB equation (30) for the case $N = 2$. In the default state $z = (0, 1)$, the HJB equation takes the form

$$\frac{\partial Q(t, x_1, (0,1))}{\partial t} + \hat{\mathcal{A}}_{01} Q(t, x_1, (0,1)) + \mathcal{H}_{01}(t, x_1, Q(t, x_1, (0,1)), D_{x_1} Q(t, x_1, (0,1))) = 0 \tag{A.44}$$

with terminal condition $Q(T, x_1, (0,1)) = \gamma^{-1}$ for all $x_1 \in \mathbb{R}_+$. Above, the operator

$$\hat{\mathcal{A}}_{01} f(t, x_1) := \left(\kappa_1 - \nu_1 x_1 + \sum_{k=1}^{2} \sigma_{1k} \sqrt{x_1} \phi_k(t, x_1) \right) \frac{\partial f}{\partial x_1} + \frac{1}{2} \left(\sum_{k=1}^{2} \sigma_{1k}^2 \right) x_1 \frac{\partial^2 f}{\partial x_1^2},$$

and the function

$$\mathcal{H}_{01}(t, x_1, u, p) := g^{(0,1)}(t, x_1) u + f^{(0,1)}(t, x_1) p + l^{(0,1)}(t, x_1).$$

We first discretize the time and space axes by the grid points $\{(t^m, x_1^i)\}_{m \in \{1,\dots,M\}, i \in \{1,\dots,I\}}$ with $M, I \in \mathbb{N}$, $\Delta t = t^{m+1} - t^m$ and $\Delta x = x_1^{i+1} - x_1^i$. Throughout the section, we abbreviate $\kappa_1 - \nu_1 x_1 + \sum_{k=1}^2 \sigma_{1k} \sqrt{x_1} \phi_k(t, x_1)$ with $a(t, x_1)$, and $\frac{1}{2}(\sum_{k=1}^2 \sigma_{1k}^2) x_1$ with $b(x_1)$. Then, the discretized version of Eq. (A.44) is given by

$$\begin{aligned}
&\frac{Q_i^{m+1} - Q_i^m}{\Delta t} + \left[a\left(t^m + \frac{\Delta t}{2}, x_1^i\right) + f^{(0,1)}\left(t^m + \frac{\Delta t}{2}, x_1^i\right) \right] \frac{1}{2} \left(\frac{Q_{i+1}^{m+1} - Q_{i-1}^{m+1}}{2\Delta x} + \frac{Q_{i+1}^m - Q_{i-1}^m}{2\Delta x} \right) \\
&\quad + b(x_1^i) \times \frac{1}{2} \left(\frac{Q_{i+1}^{m+1} - 2Q_i^{m+1} + Q_{i-1}^{m+1}}{(\Delta x)^2} + \frac{Q_{i+1}^m - 2Q_i^m + Q_{i-1}^m}{(\Delta x)^2} \right) \\
&\quad + g^{(0,1)}\left(t^m + \frac{\Delta t}{2}, x_1^i\right) \frac{Q_i^{m+1} + Q_i^m}{2} + l^{(0,1)}\left(t^m + \frac{\Delta t}{2}, x_1^i\right) = 0.
\end{aligned}$$

for $m = 1, \dots, M-1$, $i = 2, \dots, I-1$ and $Q_i^M = \gamma^{-1}$ for $i = 1, \dots, I$, where Q_i^m is the value at the space point i and time step m of the function Q, i.e., $Q_i^m := Q(t^m, x_1^i, (0,1))$. Rearranging the above equation and omitting the argument $\left(t^m + \frac{\Delta t}{2}, x_1^i\right)$ and x_1^i, we obtain

$$\begin{aligned}
&\left(\frac{a + f^{(0,1)}}{4\Delta x} - \frac{b}{2(\Delta x)^2} \right) Q_{i-1}^m + \left(\frac{1}{\Delta t} - \frac{g^{(0,1)}}{2} + \frac{b}{(\Delta x)^2} \right) Q_i^m + \left(-\frac{a + f^{(0,1)}}{4\Delta x} - \frac{b}{2(\Delta x)^2} \right) Q_{i+1}^m \\
&= \left(-\frac{a + f^{(0,1)}}{4\Delta x} + \frac{b}{2(\Delta x)^2} \right) Q_{i-1}^{m+1} + \left(\frac{1}{\Delta t} + \frac{g^{(0,1)}}{2} - \frac{b}{(\Delta x)^2} \right) Q_i^{m+1} \\
&\quad + \left(\frac{a + f^{(0,1)}}{4\Delta x} + \frac{b}{2(\Delta x)^2} \right) Q_{i+1}^{m+1} + l^{(0,1)}.
\end{aligned} \tag{A.45}$$

Assume that Q_1^m and Q_I^m satisfy the following quadratic extrapolation given by

$$(Q_4^m - Q_3^m) - (Q_3^m - Q_2^m) = (Q_3^m - Q_2^m) - (Q_2^m - Q_1^m) \tag{A.46}$$

$$(Q_I^m - Q_{I-1}^m) - (Q_{I-1}^m - Q_{I-2}^m) = (Q_{I-1}^m - Q_{I-2}^m) - (Q_{I-2}^m - Q_{I-3}^m). \tag{A.47}$$

Equations (A.45)-(A.47) yield

$$\begin{bmatrix}
1 & -3 & 3 & -1 & 0 & \cdots & \cdots & 0 \\
c_{2,1} & c_{2,2} & c_{2,3} & 0 & 0 & \cdots & \cdots & 0 \\
0 & c_{3,2} & c_{3,3} & c_{3,4} & 0 & \cdots & \cdots & 0 \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & & & 0 & c_{I-1,I-2} & c_{I-1,I-1} & c_{I-1,I} \\
0 & \cdots & & 0 & 1 & -3 & 3 & -1
\end{bmatrix}
\begin{bmatrix} Q_1^m \\ Q_2^m \\ Q_3^m \\ \vdots \\ Q_{I-1}^m \\ Q_I^m \end{bmatrix}
=
\begin{bmatrix} 0 \\ d_2 \\ d_3 \\ \vdots \\ d_{I-1} \\ 0 \end{bmatrix},$$

where for $i = 2, \dots, I-1$, $c_{i,i-1}, c_{i,i}, c_{i,i+1}$ denote, respectively, the coefficients of Q_{i-1}^m, Q_i^m and Q_i^{m+1} in the l.h.s. of Eq. (A.45), and d_i denotes the r.h.s. of the same equation. Solving the above system of equations sequentially from $m = M-1$ to $m = 1$, we obtain Q_i^m for $i = 1, \dots, I, m = 1, \dots, M-1$. The

HJB equation in the state $z = (1, 0)$ has the same form as Eq. (A.44), and can be solved by applying the same procedure described above.

In the state $z = (0, 0)$, the HJB equation takes the form

$$\frac{\partial Q(t, x, (0,0))}{\partial t} + \hat{\mathcal{A}}_{00} Q(t, x, (0,0)) + \mathcal{H}_{00}(t, x, Q(t, x, (0,0)), DQ(t, x, (0,0))) = 0 \tag{A.48}$$

with terminal condition $Q(T, x, (0,0)) = \gamma^{-1}$ for all $x \in \mathbb{R}^2_+$. Above, the operator

$$\hat{\mathcal{A}}_{00} f(t,x) := \sum_{j=1}^{2} \left(\kappa_j - \nu_j x_j + \sum_{k=1}^{2} \sigma_{jk} \sqrt{x_j} \phi_k(t, x, (0,0)) \right) \frac{\partial f}{\partial x_j} + \frac{1}{2} \sum_{i,j=1}^{2} \left(\sum_{k=1}^{2} \sigma_{ik} \sigma_{jk} \right) \sqrt{x_i} \sqrt{x_j} \frac{\partial^2 f}{\partial x_i \partial x_j},$$

and the function

$$\mathcal{H}_{00}(t, x, u, p) := g^{(0,0)}(t,x) u + \sum_{j=1}^{2} f_j^{(0,0)}(t,x) p_j + l^{(0,0)}(t,x).$$

To numerically solve the above equation, we discretize the time and space axes by the grid points

$$\{(t^m, x_1^i, x_2^j)\}_{m \in \{1,\dots,M\}, i \in \{1,\dots,I\}, j \in \{1,\dots,J\}}$$

with $M, I, J \in \mathbb{N}$, $\Delta t = t^{m+1} - t^m$, $\Delta x = x_1^{i+1} - x_1^i = x_2^{j+1} - x_2^j$. For the ease of notation, let

$$a_j(t,x) := \kappa_j - \nu_j x_j + \sum_{k=1}^{2} \sigma_{jk} \sqrt{x_j} \phi_k(t, x, (0,0)), \quad b_{ij}(x) := \frac{1}{2} \sum_{i,j=1}^{2} \left(\sum_{k=1}^{2} \sigma_{ik} \sigma_{jk} \right) \sqrt{x_i} \sqrt{x_j}.$$

Then, the discretized version of Eq. (A.48) has the form

$$\begin{aligned}
&\frac{Q_{i,j}^{m+1} - Q_{i,j}^m}{\Delta t} \\
&+ \left[a_1 \left(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j \right) + f_1^{(0,0)} \left(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j \right) \right] \frac{1}{2} \left(\frac{Q_{i+1,j}^{m+1} - Q_{i-1,j}^{m+1}}{2\Delta x} + \frac{Q_{i+1,j}^m - Q_{i-1,j}^m}{2\Delta x} \right) \\
&+ \left[a_2 \left(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j \right) + f_2^{(0,0)} \left(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j \right) \right] \frac{1}{2} \left(\frac{Q_{i,j+1}^{m+1} - Q_{i,j-1}^{m+1}}{2\Delta x} + \frac{Q_{i,j+1}^m - Q_{i,j-1}^m}{2\Delta x} \right) \\
&+ b_{11}(x_1^i, x_2^j) \frac{1}{2} \left(\frac{Q_{i+1,j}^{m+1} - 2Q_{i,j}^{m+1} + Q_{i-1,j}^{m+1}}{(\Delta x)^2} + \frac{Q_{i+1,j}^m - 2Q_{i,j}^m + Q_{i-1,j}^m}{(\Delta x)^2} \right) \\
&+ b_{22}(x_1^i, x_2^j) \frac{1}{2} \left(\frac{Q_{i,j+1}^{m+1} - 2Q_{i,j}^{m+1} + Q_{i,j-1}^{m+1}}{(\Delta x)^2} + \frac{Q_{i,j+1}^m - 2Q_{i,j}^m + Q_{i,j-1}^m}{(\Delta x)^2} \right) \\
&+ \left[b_{12}(x_1^i, x_2^j) + b_{21}(x_1^i, x_2^j) \right] \frac{1}{2} \left(\frac{Q_{i+1,j+1}^{m+1} - Q_{i+1,j-1}^{m+1} - Q_{i-1,j+1}^{m+1} + Q_{i-1,j-1}^{m+1}}{4(\Delta x)^2} \right. \\
&\qquad\qquad \left. + \frac{Q_{i+1,j+1}^m - Q_{i+1,j-1}^m - Q_{i-1,j+1}^m + Q_{i-1,j-1}^m}{4(\Delta x)^2} \right) \\
&+ g^{(0,0)} \left(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j \right) \frac{Q_{i,j}^{m+1} + Q_{i,j}^m}{2} + l^{(0,0)} \left(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j \right) = 0.
\end{aligned}$$

for $m = 1, \dots, M-1$, $i = 2, \dots, I-1$, $j = 2, \dots, J-1$ and $Q_{ij}^M = \gamma^{-1}$ for $i = 1, \dots, I$ and $j = 1, \dots, J$, where $Q_{i,j}^m$ is the value of Q at the space point (i,j) and time step m, i.e., $Q_{i,j}^m := Q(t^m, (x_1^i, x_2^j), (0,0))$. Omitting the argument $(t^m + \frac{\Delta t}{2}, x_1^i, x_2^j)$ and (x_1^i, x_2^j), the above equation may be rewritten as

$$\begin{aligned}
&\left(\frac{-a_1 - f_1^{(0,0)}}{4\Delta x} - \frac{b_{11}}{2(\Delta x)^2} \right) Q_{i+1,j}^m + \left(\frac{1}{\Delta t} + \frac{b_{11} + b_{22}}{(\Delta x)^2} - \frac{g^{(0,0)}}{2} \right) Q_{ij}^m + \left(\frac{a_1 + f_1^{(0,0)}}{4\Delta x} - \frac{b_{11}}{2(\Delta x)^2} \right) Q_{i-1,j}^m \\
&+ \left(\frac{-a_2 - f_2^{(0,0)}}{4\Delta x} - \frac{b_{22}}{2(\Delta x)^2} \right) Q_{i,j+1}^m + \left(\frac{a_2 + f_2^{(0,0)}}{4\Delta x} - \frac{b_{22}}{2(\Delta x)^2} \right) Q_{i,j-1}^m
\end{aligned}$$

$$+\left(\frac{-b_{12}-b_{21}}{8(\Delta x)^2}\right)Q^m_{i+1,j+1}+\left(\frac{b_{12}+b_{21}}{8(\Delta x)^2}\right)Q^m_{i+1,j-1}+\left(\frac{b_{12}+b_{21}}{8(\Delta x)^2}\right)Q^m_{i-1,j+1}+\left(\frac{-b_{12}-b_{21}}{8(\Delta x)^2}\right)Q^m_{i-1,j-1}$$

$$=\left(\frac{a_1+f_1^{(0,0)}}{4\Delta x}+\frac{b_{11}}{2(\Delta x)^2}\right)Q^{m+1}_{i+1,j}+\left(\frac{1}{\Delta t}-\frac{b_{11}+b_{22}}{(\Delta x)^2}+\frac{g^{(0,0)}}{2}\right)Q^{m+1}_{ij}+\left(\frac{-a_1-f_1^{(0,0)}}{4\Delta x}+\frac{b_{11}}{2(\Delta x)^2}\right)Q^{m+1}_{i-1,j}$$

$$+\left(\frac{a_2+f_2^{(0,0)}}{4\Delta x}+\frac{b_{22}}{2(\Delta x)^2}\right)Q^{m+1}_{i,j+1}+\left(\frac{-a_2-f_2^{(0,0)}}{4\Delta x}+\frac{b_{22}}{2(\Delta x)^2}\right)Q^{m+1}_{i,j-1}$$

$$+\left(\frac{b_{12}+b_{21}}{8(\Delta x)^2}\right)Q^{m+1}_{i+1,j+1}+\left(\frac{-b_{12}-b_{21}}{8(\Delta x)^2}\right)Q^{m+1}_{i+1,j-1}+\left(\frac{-b_{12}-b_{21}}{8(\Delta x)^2}\right)Q^{m+1}_{i-1,j+1}+\left(\frac{b_{12}+b_{21}}{8(\Delta x)^2}\right)Q^{m+1}_{i-1,j-1}$$

$$+l^{(0,0)}. \tag{A.49}$$

Assume that $Q^m_{1,j}, Q^m_{I,j}, Q^m_{i,1}, Q^m_{i,J}$ for $m=1,\dots,M-1$, $i=1,\dots,I$, and $j=1,\dots,J$ satisfy

$$Q^m_{1,j}=Q^{m+1}_{1,j},\quad Q^m_{I,j}=Q^{m+1}_{I,j},\quad Q^m_{i,1}=Q^{m+1}_{i,1},\quad Q^m_{i,J}=Q^{m+1}_{i,J}. \tag{A.50}$$

Sequentially solving the system of Eqs. (A.49)-(A.50) from $m=M-1$ to $m=1$, we obtain $Q^m_{i,j}$ for $i=1,\dots,I$, $j=1,\dots,J$, and $m=1,\dots,M-1$.