A Pricing Formula for Delayed Claims: Appreciating the Past to Value the Future

Enrico Biffis* Beniamin Goldys† Cecilia Prosdocimi‡
Margherita Zanella§

December 28, 2022

Abstract

We consider the valuation of contingent claims with delayed dynamics in a Samuelson complete market model. We find a pricing formula that can be decomposed into terms reflecting the current market values of the past and the future, showing how the valuation of prospective cashflows cannot abstract away from the contribution of the past. As a practical application, we provide an explicit expression for the market value of human capital in a setting with wage rigidity. The formula we derive has successfully been used to explicitly solve the infinite dimensional stochastic control problems addressed in [7], [6] and [16].

Key words: Stochastic functional differential equations, delay equations, no-arbitrage pricing, human capital, sticky wages.

AMS classification: 34K50, 91B25, 91G80.

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Contents

*Enrico Biffis (`e.biffis@imperial.ac.uk`), Imperial College Business School, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.

†Beniamin Goldys (`beniamin.goldys@sydney.edu.au`), School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia.

‡Cecilia Prosdocimi (`cecilia.prosdocimi@gmail.com`), LUISS "Guido Carli", Rome, Italy.

§Margherita Zanella (`margherita.zanella@polimi.it`), Politecnico di Milano, Dipartimento di Matematica ‘Francesco Brioschi’, Via Bonardi 13, 20133 Milano, Italy.

1 Introduction

It is a standard result in asset pricing theory that the absence of arbitrage opportunities is "essentially" equivalent to the existence of an equivalent probability measure under which the price of any contingent claim is a local martingale after deflation by the money market account; see [17, 24, 25]. In this paper, we preserve the standard formulation of arbitrage pricing in a complete market model with security prices evolving as geometric Brownian motions (GBM), in the spirit of Samuelson's contribution [36] (see [37] for an overview). The main novelty of our work is that we consider contingent claims that have dynamics described by a stochastic delay differential equation (SDDE).

It is perhaps surprising that using the no-arbitrage pricing machinery we are able to derive an explicit valuation formula for dynamics with memory, a case that is notoriously difficult to handle. However, we find a particularly appealing solution showing that the price can be decomposed into a term related to the 'current market value of the past' (in a sense to be made precise below), and the more traditional term reflecting the 'current market value of the future' (in the spirit of a discounted cashflow approach). In our setting the contribution of the past is represented by the portion of a contingent claim's past trajectory that shapes its dynamics going forward.[1] Our pricing formula demonstrates that the market consistent valuation of future cashflows cannot ignore the contribution of the past. This is important for a number of applications in which path-dependency is a key feature of the state variables, as we now discuss.

As a practical application of our results, we consider in detail contingent claims representing stochastic wages received by an agent over her lifetime (e.g., [20, 7]). It is well known that when labor income is spanned by the assets available for trade, the market value of human capital can be easily derived via risk-neutral valuation in a setting with labor income driven by a Stochastic Differential Equation (SDE); see [13] for an overview. However, the empirical literature on labor income dynamics widely relies on auto-regressive moving average (ARMA) processes (e.g., [31], [1], [26], [33]). The contributions by [35], [29], and [19] show how SDDEs can be understood as the weak limit of discrete time ARMA processes. Therefore specific classes of SDDEs can be used to model labor income, thereby providing the continuous time counterpart of ARMA models supported by the empirical literature. We therefore consider the introduction of delayed drift and volatility coefficients in a GBM labor income model to provide a tractable example of wage dynamics that adjusts slowly to financial market shocks. We obtain a closed form solution for human capital, which makes explicit the contributions of the market value of the past and the future. Our results

[1] The importance of the past in understanding the qualitative feature of a model with delay was also emphasized in Fabbri and Gozzi [21], although in a deterministic setting, when solving the endogenous growth model with vintage capital of Boucekkine *et al.* [10].

demonstrate that SDDEs are valuable modeling tools that can address the findings of a large body of empirical literature on wage rigidity (e.g., [28], [15], [3], [30]). Moreover, the results open the way to finding explicit solutions to interesting classes of lifecycle portfolio choice problems with state costraints (see [7, 6, 16]), as discussed in Section 3.

Although in this paper we discuss the application to human capital more extensively, the results can be used for other applications of interest. For instance, we provide some references to the literature on counterparty risk and derivatives valuation, in which analogous dynamics arise in the context of collateralization procedures entailing a delay in the marking-to-market procedure of over-the-counter derivatives (e.g., [11, 12]).

It should be noted that no-arbitrage pricing in the case of delayed price dynamics has been recently studied by several authors (see [2, 32], for example). Their focus, however, is on proving completeness of a market with security prices with delayed dynamics and hence very different from ours. On the other hand, their work suggests the opportunity to explore extensions of our results to broader settings in which market completeness is preserved, including those in which tradable assets have delayed drift and volatility terms.

The paper is organized as follows. In the next Section, we introduce the setup and state our main result. In Section 3 we discuss in detail an application to human capital valuation and briefly outline other areas of application. In Section 4, we prove the main result of the paper. Finally, section 5 offers concluding remarks. We relegate to an appendix the proofs of some more technical results.

2 Setup and statement of the main result

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and an $\mathbb{F}$-adapted vector valued process (S_0, S), representing the price evolution of a money market account, S_0, and n risky assets, $S = (S_1, ..., S_n)^\top$, whose dynamics is given by

$$\begin{cases} \mathrm{d}S_0(t) = S_0(t) r \mathrm{d}t, \\ \mathrm{d}S(t) = \mathrm{diag}(S(t)) \left(\mu \mathrm{d}t + \sigma \mathrm{d}Z(t)\right), \\ S_0(0) = 1, \\ S(0) \in \mathbb{R}^n_+, \end{cases} \tag{1}$$

where $\mu \in \mathbb{R}^n$, and the matrix $\sigma \in \mathbb{R}^{n\times n}$ is assumed to be invertible. Here Z is an n-dimensional Brownian motion and we assume that $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ is the filtration generated by Z, and enlarged with the $\mathbb{P}$-null sets. In the following, for given vectors $a, b \in \mathbb{R}^n$, we denote by $a \cdot b$ the scalar product in $\mathbb{R}^n$ and by $\|\cdot\|$ the corresponding norm.

We are interested in the valuation of a payment stream represented by an $\mathbb{F}$-adapted process $(y_t)_{t\geq 0}$. The payment stream can be thought of as capturing the mark-to-market process of a trading account, the flow of profits and losses generated from a trading strategy, the collateral flows arising from an over-the-counter derivative transaction, or the labor income received by an agent over time. The latter is the application we will consider in detail in the next section.

We assume that the payment stream y obeys the following SDDE with delay in both the drift and the diffusion terms:

$$\begin{cases} \mathrm{d}y(t) = & \left[y(t)\mu_y + \int_{-d}^0 y(t+s)\phi(\mathrm{d}s)\right] \mathrm{d}t \\ & + \left[y(t)\sigma_y + \begin{pmatrix} \int_{-d}^0 y(t+s)\varphi_1(\mathrm{d}s) \\ \vdots \\ \int_{-d}^0 y(t+s)\varphi_n(\mathrm{d}s) \end{pmatrix}\right] \cdot \mathrm{d}Z(t), \\ y(0) = & x_0, \\ y(s) = & x_1(s) \quad \text{for } s \in [-d, 0), \end{cases} \tag{2}$$

where $\mu_y \in \mathbb{R}$, $\sigma_y \in \mathbb{R}^n$, the signed measures ϕ, φ_i (for $i = 1, \dots, n$) have bounded variation on $[-d, 0]$ and $x_0 \in \mathbb{R}$, $x_1 \in L^2\big([-d, 0]; \mathbb{R}\big)$, where $L^2\big([-d, 0]; \mathbb{R}\big)$ denotes the space of Lebesgue square integrable (deterministic) functions on $[-d, 0]$. In the following, we denote by $|\phi|$ the total variation of a signed measure of bounded variation ϕ on $[-d, 0]$. Existence of a unique solution to SDDE (2) is ensured by the following result, which is proved in Appendix A.1.

Proposition 2.1. *For any given initial datum $(x_0, x_1) \in \mathbb{R} \times L^2(-d, 0; \mathbb{R})$ equation (2) admits a unique strong (in the probabilistic sense) solution $y \in L^2(\Omega \times [0, T])$, for all $T > 0$, with $\mathbb{P}$-a.s. continuous paths.*

Assuming the process y to represent an agent's income stream, we see that (2) provides a compelling example (which will be shown to be tractable) of wage dynamics adjusting slowly to financial market shocks. In particular, the measures $\phi, \varphi_1, ...\varphi_n$ can be thought as modulating the impact of past income's realizations on wages going forward. As discussed in [7], this setting can also be related to the literature on 'learning your income' (e.g., [22]), whereby an agent learns about her earning potential based on past wages. Here, moving averages can be used as a tractable substitute for fully fledged Bayesian filters and can be justified on the grounds of bounded rationality.

As the market is complete and the stream process y is spanned by the stock S, we can find uniquely the current value of the future uncertain income stream by considering the expected discounted value under the unique stochastic discount factor ξ (see [18]):

$$H(t_0) := \xi(t_0)^{-1} \mathbb{E}\left[\int_{t_0}^{+\infty} \xi(s) y(s) \mathrm{d}s \Big| \mathcal{F}_{t_0}\right], \tag{3}$$

where the process ξ is $\mathbb{F}$-adapted and evolves according to the following SDE,

$$\begin{cases} \mathrm{d}\xi(t) & = -\xi(t)\left(r\mathrm{d}t + \kappa \cdot \mathrm{d}Z(t)\right) \\ \xi(0) & = 1, \end{cases} \tag{4}$$

with the constant κ representing the market price of risk, which is given by

$$\kappa = (\sigma^\top)^{-1}(\mu - r\mathbf{1}), \tag{5}$$

with $\mathbf{1} = (1, \dots, 1)^\top$ denoting the unitary vector in $\mathbb{R}^n$.

Our main goal now is to provide an explicit formula for the quantity $H(t_0)$ given in (3). The challenging aspect of the problem lies in the fact that we consider a payment stream whose dynamics is path dependent, thus making the problem considerably different from and harder to prove than the cases discussed in the extant literature. Before stating the main result of the paper, we introduce the following conditions that will be assumed to hold throughout the paper.

Assumption 2.2. *The following conditions are assumed to apply:*

(i) The measure Φ on $[-d, 0]$ defined as

$$\Phi(\cdot) := \phi(\cdot) - (\varphi_1(\cdot), ..., \varphi_n(\cdot)) \cdot \kappa \tag{6}$$

is a signed measure of bounded variation.

(ii) The quantity

$$r - \mu_0 + \sigma_y \cdot \kappa - \int_{-d}^{0} e^{r\tau} |\Phi|(\mathrm{d}\tau) \tag{7}$$

is assumed to be strictly positive.

Under assumption 2.2, which will be discussed in details in Section 4.5, the following result holds.

Theorem 2.3. *Under Assumption 2.2, for any $t_0 \geq 0$, the quantity $H(t_0)$ defined in (3) admits the following explicit expression:*

$$H(t_0) = \frac{1}{K}\left(y(t_0) + \int_{-d}^{0} G(s)y(t_0+s)\,\mathrm{d}s\right), \quad \mathbb{P}-a.s., \tag{8}$$

where $y(t_0)$ denotes the solution at time t_0 of equation (2) and the constant K and function G are given by:

$$K := r - \mu_0 + \sigma_y \cdot \kappa - \int_{-d}^{0} e^{r\tau}\Phi(\mathrm{d}\tau),$$

$$G(s) := \int_{-d}^{s} e^{-r(s-\tau)}\Phi(\mathrm{d}\tau).$$

In expression (8), we can recognize an annuity factor, K^{-1}, multiplying a term representing the current market value of labor income, $y(t_0)$, and a term representing the current market value of the past trajectory of y over the time window $(t_0 - d, t_0)$. The 'market value of the past' trades off the returns on the payment stream against its exposure to financial risk, as can be seen from expression (6). When the delay terms in the drift and volatility coefficients vanish, the valuation of the payment stream reduces to $K^{-1}y(t_0)$, in line with [20], among others.

Remark 2.4. *The setup described above can be extended to the case of payments over a bounded horizon in some interesting situations. When the payment stream is received until an exogenous Poisson stopping time τ_δ (representing death or an irreversible unemployment shock when y represents labor income), expression (8) still applies, provided discounting is carried out at rate $r+\delta$ instead of r, where $\delta > 0$ represents the Poisson parameter; see [7], [6] and [16]. The case in which payments are received until a finite, deterministic time (capturing irreversible retirement when y represents labor income) can also be considered, at the price of additional technical work; see [5].*

3 Applications

3.1 Optimal portfolio problems with path dependent labor income

We now consider in detail the case in which the contingent claim y in (2) represents stochastic wages received by an agent over her lifetime. As discussed in the Introduction, SDDEs allow us to rely on continuous time labor income dynamics better matching some of the salient features documented in the empirical literature. The path-dependency of (2) captures the slow adjustment of labor income to financial market shocks and provides a counterpart to discrete time ARMA models used to model wage dynamics. See [7] for a comprehensive list of references and [5] for the complexities brought about by considering unspanned shocks in labor income dynamics.

In the context of lifecycle portfolio choice, obtaining an explicit expression for human capital is often crucial not only to reveal the structure of optimal solutions (e.g., [13]), but also to handle state constraints (e.g., [20],[7]). We will discuss these points by making reference to some applications in sections 3.1.1-3.1.3. Before proceeding, we note that assumption 2.2 is all we need to provide the explicit valuation result of Theorem 2.3, but the particular application to human capital requires labor income to be positive almost surely. A sufficient condition for this to be the case is as follows (see [7, Proposition 2.7] for a proof).

Remark 3.1. *When $\varphi_i = 0$ for all $i = 1...n$ (i.e., when the delay term in the volatility coefficient of (2) vanishes), the variation of constants formula yields the following explicit representation:*

$$y(t) = \mathcal{E}(t)\big(x_0 + \mathcal{I}(t)\big), \tag{9}$$

where

$$\mathcal{E}(t) := e^{(\mu_y - \frac{1}{2}|\sigma_y|^2)t + \sigma_y Z(t)}, \qquad \mathcal{I}(t) := \int_0^t \mathcal{E}^{-1}(u)\int_{-d}^{0} y(s+u)\phi(\mathrm{d}s)\,\mathrm{d}u.$$

One can then see that in this case $y(t) > 0$ $\mathbb{P}$-a.s. if $x_0 > 0$, $x_1 \geq 0$ a.s. and $\phi \geq 0$ a.s..

The results of Theorem 2.3 and the solution approach followed in this paper show how tools from infinite-dimensional analysis can be successfully used to address non-Markovian valuation problems, which are beyond the reach of conventional approaches. In the following, we illustrate some of those applications to show how the findings of Theorem 2.3 (or suitable generalization of it) can be successfully used to identify explicit solutions in some interesting situations.

3.1.1 Infinite horizon lifecycle portfolio choice

In [7] the authors solve an infinite horizon portfolio problem with borrowing constraints in which an agent receives labor income affected by financial market shocks in a path dependent way. The framework is the complete market model described by (1). An agent is endowed with initial wealth $w \geq 0$ and receives wages until death, which coincides with the first jump time τ_δ of a Poisson process with parameter $\delta > 0$. The financial wealth of the agent at time $t \geq 0$, denoted by $W(t)$, can be invested in the riskless and risky assets. The agent receives wages at rate $y(t)$ and consumes at rate $c(t) \geq 0$. The wealth amount allocated to the risky assets is $\theta(t) \in \mathbb{R}^n$ at each time $t \geq 0$. The agent sets a bequest target $B(\tau_\delta) \geq 0$ at death. In line with [20], the agent can purchase life insurance at the (actuarially fair) premium rate $\delta(B(t) - W(t))$, for $t < \tau_\delta$, to reach the bequest target and hence cover any shortfall $B(\tau_\delta) - W(\tau_\delta)$ at death. The agent's wealth (before death) is assumed to obey the standard dynamic budget constraint given by the SDE

$$\begin{cases} dW(t) = [W(t)r + \theta(t)\cdot(\mu - r\mathbf{1}) + y(t) - c(t) - \delta\,(B(t) - W(t))]\,\mathrm{d}t + \theta(t)\cdot\sigma\mathrm{d}Z(t), \\ W(0) = w. \end{cases} \tag{10}$$

In line with the empirical evidence on labour income dynamics recalled in the Introduction, labor income is modelled via the following SDDE:

$$\begin{cases} \mathrm{d}y(t) = \left[\mu_y y(t) + \int_{-d}^{0}\phi(s)y(t+s)\mathrm{d}s\right]\mathrm{d}t + y(t)\sigma_y\mathrm{d}Z(t), \\ y(0) = x_0, \quad y(s) = x_1(s) \text{ for } s \in [-d, 0), \end{cases} \tag{11}$$

where $\mu_y \in \mathbb{R}$, $\sigma_y \in \mathbb{R}^n$ and the functions $\phi(\cdot), x_1(\cdot)$ belong to $L^2(-d, 0; \mathbb{R})$ (thus making (11) a particular case of (2)).

Denoting by $k > 0$ the intensity of preference for leaving a bequest, by $\gamma \in (0,1) \cup (1, +\infty)$ the risk-aversion coefficient and by $\rho > 0$ the subjective discount rate, the objective is to maximize the expected utility from lifetime consumption and bequest, which on $\{\tau^\delta > t\}$ takes the form:

$$\mathbb{E}\left(\int_0^{+\infty} e^{-(\rho+\delta)t}\left(\frac{c(t)^{1-\gamma}}{1-\gamma} + \delta\frac{\left(kB(t)\right)^{1-\gamma}}{1-\gamma}\right)\mathrm{d}t\right). \tag{12}$$

Maximization is carried out over all triplets $(c, \theta, B) \in \Big\{\mathbb{F} - \text{predictable } c(\cdot), B(\cdot), \theta(\cdot) \colon c(\cdot), B(\cdot) \in L^1(\Omega \times [0, +\infty); \mathbb{R}_+), \theta(\cdot) \in L^2(\Omega \times \mathbb{R}; \mathbb{R}^n)\Big\}$ subject to the state constraint

$$W(t) + \xi^{-1}(t)\mathbb{E}\left(\int_t^{+\infty}\xi(u)y(u)\mathrm{d}u\middle|\mathcal{F}_t\right) \geq 0, \tag{13}$$

which is a 'no-borrowing-without-repayment' constraint [20]. As the second term appearing in (13) represents the agent's human capital at time t, constraint (13) captures the situation in which human capital can be pledged as collateral and represents the agent's maximum borrowinge capacity. In line with [7], we note that the triplets (c, θ, B) must be understood as representing the pre-death counterparts of the controls predictable relative to the reference filtration $\mathbb{F}$. Similarly, on the trace of $\mathbb{F}$ on $\{\tau^\delta > t\}$, the process ξ satisfies equation (4) with a drift of the form $-\xi(t)(r+\delta)$, as explained in Remark 2.4.

In [7], the authors find an explicit solution to the optimization problem under power utility. The proof of the result relies on the resolution of an infinite-dimensional Hamilton-Jacobi-Bellman (HJB) equation, which can be considered as an infinite-dimensional version of the classical Merton problem. From a technical point of view, the key idea is to extend the state space so as to include the past trajectory of y over $[-d, 0]$. In this way, the problem becomes infinite dimensional and Markovian (in the current wealth level and path of y over the time window $[-d, 0]$). In this infinite-dimensional reformulation of the problem, it becomes essential to rewrite the constraint (13) by exploiting the explicit expression given in Theorem 2.3. Importantly, the optimal risky asset allocation found in [7],

$$\theta^*(t) := (\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1})\frac{W^*(t) + H(t)}{\gamma} - \sigma^{-1}\sigma_y \frac{1}{K} y(t), \tag{14}$$

reveals that the decomposition of human capital into a past and future component is essential in understanding the agent's hedging demand, as we now discuss. We first note that the solution follows the logic of Merton's, in that the agent chooses constant fractions of total wealth (given by financial wealth and human capital). In line with [9, 20], the agent considers the capitalized value of future wages as if they were a traded asset. As the agent's labor income is instantaneously perfectly correlated with the risky assets, a negative income hedging demand arises (the term $\sigma^{-1}\sigma_y K^{-1} y(t)$), and the allocation to risky assets is reduced by a term proportional to the regression coefficient of labor income shocks on risky asset returns ([13]). The key difference with the benchmark model with no delay is that the hedging demand only depends on the 'future component' of human capital, and not on the capitalized market value of the labor income's past trajectory. The intuition is that when wages respond to market shocks with a delay, then human capital is more predictable and only the portion of human capital exposed to future market shocks drives the negative income hedging demand.

3.1.2 Extensions: robustness and finite horizon

The authors in [6] study a robust version of the lifecycle optimal portfolio choice problem presented in section 3.1.1. Again, the result in Theorem 2.3 is essential in obtaining an explicit solution of the problem considered.

The authors in [5] consider the same problem as in the previous section, but now with a finite time horizon, which can be interpreted as the agent's fixed retirement date. The authors extend the results of Theorem 2.3 to the finite horizon case, obtaining again a decomposition of human capital into two components pertaining to the past and the future evolution of labor income. They then proceed to solve the lifecycle portfolio choice model, which entails time-dependent state constraints, as horizon effects are now material during the working life of the agent. This appears to be a novel type of problem in the infinite dimensional stochastic control literature, which is again successfully solved by exploiting the structure of the explicit solution for human capital.

3.1.3 Mean-field games

Another generalization of the problem addressed in [7] is considered in [16]. Similarly to [7], the authors consider a lifecycle optimal portfolio choice problem faced by an agent receiving labor income and allocating her wealth to risky securities and a riskless asset subject to a borrowing constraint. However, in addition to assuming the dynamics of labor income to adjust slowly to financial market shocks, they also assume it to be benchmarked to the wages of a population of agents with comparable tasks and/or ranks. Specifically, each agent i's labor income y_i is benchmarked against the wages $y^n := (y_1, y_2, \ldots, y_n)$ of a population of n agents. As n grows larger, the problem falls into the family of optimal control of infinite dimensional McKean-Vlasov dynamics type. By adding a suitable new variable, the authors simplify the problem and are able to find explicitly the solution of the associated infinite-dimensional HJB equation and the optimal feedback controls. A necessary step to solve the problem is to provide a suitable reformulation of the no-borrowing without repayment constraint (13), where now labor income obeys an SDDE with a drift containing not only a path-dependent term but also a mean reverting term. This is carried out in [16, Section 3], where the authors provide a generalization of formula (8) by carefully adapting the techniques used in section 4 of this paper.

3.2 Counterparty risk and derivatives valuation

As a simple example of application of our setup to the context of over-the-counter derivatives, let us consider in equation (2) the case of $n = 1$, $\mu_0 = 0$, $\phi = 0$, $\sigma_0 = 0$, and $\varphi(s) = \delta_{-d}(s)$, where $\delta_a(s)$ indicates the delta-Dirac measure at a, so that equation (2) now reads

$$\mathrm{d}X_0(t) = X_0(t-d)\mathrm{d}Z(t). \tag{15}$$

Then, for $t \in [0, d)$ we have

$$X_0(t) = x_0 + \int_0^t X_0(s-d)\mathrm{d}Z(s) = x_0 + \int_{-d}^{t-d} x_1(\tau)\mathrm{d}Z(\tau + d). \tag{16}$$

In this case $X_0(t)$ is Gaussian, and dynamics (15) could be used to model, for example, the variation margin of an over-the-counter swap, when the collateralization procedure relies on a delayed mark-to-market value of the instrument (see [11], page 316, or [12] and [8] for some examples).

4 Proof of the result

Within this Section we consider the unique continuous $\mathbb{F}$-adapted solution y of (2) given in Proposition 2.1. The proof of Theorem 2.3 can be divided in the following steps:

- we incorporate the discount factor ξ in the equivalent risk-neural probability measure $\tilde{\mathbb{P}}$ and rewrite the dynamics of y under $\tilde{\mathbb{P}}$. Derive the deterministic delayed equation satisfies by the quantity $\tilde{\mathbb{E}}[y(t)|\mathcal{F}_{t_0}]$. (Subsection 4.1).
- We rewrite the delayed equation for $\tilde{\mathbb{E}}[y(t)|\mathcal{F}_{t_0}]$ as a deterministic evolution equation, which takes values in a suitable Hilbert space incorporatingthe past and the present in its structure. We will appeal to the so-called *product-space framework* for path-dependent equations (Subsection 4.2).
- We exploit suitable spectral properties of the operator that appears in the above mentioned infinite-dimensional formulation of the problem in order to obtain expression (8) for $H(t_0)$ (Subsections 4.3 and 4.4).
- We clarify the relationship between the spectral properties used and our Assumption 2.2 (Subsection 4.5).

The above first three steps will lead to Proposition 4.8, whereas the last step will be formalized in Lemmas 4.10 and 4.11. Theorem 2.3 will then follow as an immediate consequence of these results. For readability, we relegate to an appendix the proofs of some technical lemmas.

4.1 Equivalent probability measure

We find it more convenient to change perspective from a valuation formula relying on the use of the stochastic discount factor ξ under the measure $\mathbb{P}$ to the one discounting at the risk-free rate r the equivalent martingale measure $\tilde{\mathbb{P}}$, which is therefore called 'risk-neutral measure'. It is well know that the two approaches are equivalent under no arbitrage in our setting.

We start by considering the equivalent probability measure $\tilde{\mathbb{P}}$ on $\mathcal{F}_s$ such that[2]

[2]Recall that the state price density ξ characterizes the Radon-Nikodym derivative that defines the change of probability measure from the objective probability measure $\mathbb{P}$ to the risk-neutral measure $\tilde{\mathbb{P}}$ via the relationship $\xi(s) = e^{-rs}\rho(s) = e^{-rs}\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}(s)$.

$$\left.\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_s} = \exp\left(-\frac{1}{2}|\kappa|^2 s - \kappa \cdot Z(s)\right) = e^{rs}\xi(s). \tag{17}$$

By [27, Lemma 3.5.3] we can write

$$\mathbb{E}\left[\xi(s)y(s) \mid \mathcal{F}_{t_0}\right] = \xi(t_0)e^{-r(s-t_0)}\tilde{\mathbb{E}}_s\left[y(s) \mid \mathcal{F}_{t_0}\right],$$

and thus[3]

$$\int_{t_0}^{+\infty} \mathbb{E}\left[\xi(s)y(s) \mid \mathcal{F}_{t_0}\right] \mathrm{d}s = \xi(t_0)e^{rt_0} \int_{t_0}^{+\infty} e^{-rs}\tilde{\mathbb{E}}_s\left[y(s) \mid \mathcal{F}_{t_0}\right] \mathrm{d}s. \tag{18}$$

The idea is to now understand what kind of differential equation the quantity $\tilde{\mathbb{E}}\left[y(s) \mid \mathcal{F}_{t_0}\right] = \tilde{\mathbb{E}}_s\left[y(s) \mid \mathcal{F}_{t_0}\right]$ satisfies. Let $\tilde{\mathbb{P}}$ the measure such that $\left.\tilde{\mathbb{P}}\right|_{\mathcal{F}_s} = \tilde{\mathbb{P}}(s)$ for all $s \geq 0$. By the Girsanov Theorem the process $\tilde{Z}(t) = Z(t)+\kappa t$ is an n-dimensional Brownian motion under $\tilde{\mathbb{P}}$. The dynamics of y under $\tilde{\mathbb{P}}$ is then

$$dy(s) = \left[(\mu_y - \sigma_y \cdot \kappa)y(s) + \int_{-d}^0 y(s+\tau)\,\Phi(\mathrm{d}\tau)\right]\mathrm{d}s + \left[y(t)\sigma_y + \begin{pmatrix} \int_{-d}^0 y(s+\tau)\varphi_1(\mathrm{d}\tau) \\ \vdots \\ \int_{-d}^0 y(s+\tau)\varphi_n(\mathrm{d}\tau) \end{pmatrix}\right] \cdot \mathrm{d}\tilde{Z}(s), \tag{19}$$

where Φ is defined in (6). Integrating between t_0 and t we obtain

$$\begin{aligned} y(t) = y(t_0) &+ \int_{t_0}^t (\mu_y - \sigma_y \cdot \kappa)y(s)\mathrm{d}s + \int_{t_0}^t \int_{-d}^0 y(s+\tau)\Phi(\mathrm{d}\tau)\mathrm{d}s \\ &+ \int_{t_0}^t \left[y(s)\sigma_y + \begin{pmatrix} \int_{-d}^0 y(s+\tau)\varphi_1(\mathrm{d}\tau) \\ \vdots \\ \int_{-d}^0 y(s+\tau)\varphi_n(\mathrm{d}\tau) \end{pmatrix}\right] \cdot \mathrm{d}\tilde{Z}(s)\,, \end{aligned} \tag{20}$$

and therefore, by taking the conditional expected value on both sides, we get

$$\begin{aligned} \tilde{\mathbb{E}}\left[y(t) \mid \mathcal{F}_{t_0}\right] =& y(t_0) + (\mu_y - \sigma_y \cdot \kappa)\tilde{\mathbb{E}}\left[\int_{t_0}^t y(s)\mathrm{d}s \mid \mathcal{F}_{t_0}\right] \\ &+ \tilde{\mathbb{E}}\left[\int_{t_0}^t \int_{-d}^0 y(s+\tau)\Phi(\mathrm{d}\tau)\mathrm{d}s \mid \mathcal{F}_{t_0}\right] \\ &+ \tilde{\mathbb{E}}\left[\int_{t_0}^t \left[y(s)\sigma_y + \begin{pmatrix} \int_{-d}^0 y(s+\tau)\varphi_1(\mathrm{d}\tau) \\ \vdots \\ \int_{-d}^0 y(s+\tau)\varphi_n(\mathrm{d}\tau) \end{pmatrix}\right] \cdot \mathrm{d}\tilde{Z}(s) \mid \mathcal{F}_{t_0}\right]. \end{aligned} \tag{21}$$

The following Lemma guarantees that the stochastic integral with respect to $\tilde{Z}$ is a martingale, and has zero mean. The proof is provided in Appendix A.2

Lemma 4.1. *It holds that*

$$\tilde{\mathbb{E}}\left[\int_{t_0}^t \left\| y(s)\sigma_y + \begin{pmatrix} \int_{-d}^0 y(s+\tau)\varphi_1(\mathrm{d}\tau) \\ \vdots \\ \int_{-d}^0 y(s+\tau)\varphi_n(\mathrm{d}\tau) \end{pmatrix}\right\|^2 \mathrm{d}s\right] < +\infty\,.$$

[3] Recall (see (3)) that our aim is to evaluate the expectation $\mathbb{E}\left[\int_{t_0}^{+\infty} \xi(s)y(s)\mathrm{d}s \mid \mathcal{F}_{t_0}\right]$. We will prove that $\int_{t_0}^{+\infty} \mathbb{E}\left[\xi(s)y(s) \mid \mathcal{F}_{t_0}\right] \mathrm{d}s$ is equal to the r.h.s. of (8) and then justify the equality $\mathbb{E}\left[\int_{t_0}^{+\infty} \xi(s)y(s)\mathrm{d}s \mid \mathcal{F}_{t_0}\right] = \int_{t_0}^{+\infty} \mathbb{E}\left[\xi(s)y(s) \mid \mathcal{F}_{t_0}\right] \mathrm{d}s$.

We thus obtain that

$$\tilde{\mathbb{E}}\left[\int_{t_0}^{t}\left[y(s)\sigma_y+\begin{pmatrix}\int_{-d}^{0} y(s+\tau)\varphi_1(\mathrm{d}\tau)\\ \vdots\\ \int_{-d}^{0} y(s+\tau)\varphi_n(\mathrm{d}\tau)\end{pmatrix}\right]\cdot \mathrm{d}\tilde{Z}(s)\mid \mathcal{F}_{t_0}\right]=0,$$

and, by definition of conditional mean and by Fubini's Theorem, the expression in (21) reduces to

$$\begin{aligned}\tilde{\mathbb{E}}\left[y(t)\mid \mathcal{F}_{t_0}\right]=y(t_0)+(\mu_y-\sigma_y\cdot\kappa)\int_{t_0}^{t}\tilde{\mathbb{E}}\left[y(s)\mid \mathcal{F}_{t_0}\right]\mathrm{d}s\\ +\int_{t_0}^{t}\int_{-d}^{0}\tilde{\mathbb{E}}\left[y(s+\tau)\mid \mathcal{F}_{t_0}\right]\Phi(\mathrm{d}\tau)\mathrm{d}s.\end{aligned}\tag{22}$$

Therefore, defining

$$M_{t_0}(t):=\tilde{\mathbb{E}}\left[y(t)|\mathcal{F}_{t_0}\right],\tag{23}$$

we have that M_{t_0} satisfies for $t\geq t_0$ the equation (with random initial conditions)

$$\begin{cases}\mathrm{d}M_{t_0}=\left[(\mu_y-\sigma_y\cdot\kappa)M_{t_0}(t)+\int_{-d}^{0}M_{t_0}(t+s)\,\Phi(\mathrm{d}s)\right]\mathrm{d}t,\\ M_{t_0}(t_0)=y(t_0),\\ M_{t_0}(t_0+s)=y(t_0+s), & s\in[-d,0).\end{cases}\tag{24}$$

Existence of a unique solution of the above system is guaranteed by the following generalization of [4, Part II, Chapter 4, Theorem 3.2] to random initial conditions.

Lemma 4.2. *Given any fixed $\mathcal{F}_{t_0}$-measurable $\mathbb{R}\times L^2([-d,0];\mathbb{R})$-valued random variable $m=(m_0,m_1)$, the Cauchy problem*

$$\begin{cases}\mathrm{d}m(t_0;t)=\left[(\mu_y-\sigma_y\cdot\kappa)m(t_0;t)+\int_{-d}^{0}m(t_0;t+s)\,\Phi(\mathrm{d}s)\right]\mathrm{d}t,\\ m(t_0;t_0)=m_0,\\ m(t_0;t_0+s)=m_1(s), & s\in[-d,0).\end{cases}\tag{25}$$

admits a unique absolutely continuous solution. Moreover, system (25) is equivalent to (24) when we choose $(m_0,m_1)=(y(t_0),y(t_0+\cdot))$.

4.2 Reformulation of the problem in an infinite-dimensional framework

We now reformulate the differential equation with delay (24) as an evolution equation with values in the so called Delfour-Mitter Hilbert space, defined as

$$\mathcal{H}:=\mathbb{R}\times L^2(-d,0;\mathbb{R}),$$

whose elements are denoted as $x=(x_0,x_1)$. $\mathcal{H}$ is a Hilbert space when endowed with the inner product $\langle(x_0,x_1),(y_0,y_1)\rangle_{\mathcal{H}}=x_0y_0+\langle x_1,y_1\rangle$, the latter being the usual inner product of $L^2(-d,0;\mathbb{R})$.

We define the operator $A:\mathcal{D}(A)\subset\mathcal{H}\to\mathcal{H}$ as

$$\mathcal{D}(A):=\left\{(x_0,x_1)\in\mathcal{H}:x_1(\cdot)\in W^{1,2}\big([-d,0];\mathbb{R}\big),\,x_0=x_1(0)\right\},$$

$$A(x_0,x_1):=\Big((\mu_0-\sigma_0\cdot\kappa)x_0+\int_{-d}^{0}x_1(s)\Phi(\mathrm{d}s),\frac{\mathrm{d}}{\mathrm{d}s}x_1\Big),\tag{26}$$

with Φ defined in (6). Here by $W^{1,2}\big([-d,0];\mathbb{R}\big)$ we denote the Sobolev space of weakly differentiable square-integrable functions.

We can then reformulate equation (24) as an evolution equation in $\mathcal{H}$.

Consider, again for any fixed $\mathcal{F}_{t_0}$-measurable $\mathcal{H}$-valued random variable $\mathbf{m} = (m_0, m_1)$, the $\mathcal{H}$-valued process $\mathbf{M}(t_0;\cdot)$ that is the solution on $[t_0, +\infty)$ of

$$\begin{cases} \mathrm{d}\mathbf{M}(t_0;t) = A\mathbf{M}(t_0;t)\mathrm{d}t, \\ \mathbf{M}(t_0;t_0) = \mathbf{m}. \end{cases} \tag{27}$$

We collect in the following Proposition some useful results about the above equation (for more details see e.g. [14, Appendix A]).

Proposition 4.3. *(i) The operator A generates a strongly continuous semigroup $\{S(t)\}_{t\geq 0}$ in $\mathcal{H}$.*

(ii) $S(t)$ is a compact operator for every $t \geq d$.

(iii) For every $\mathcal{F}_{t_0}$-measurable $\mathcal{H}$-valued random variable m the process

$$S(t-t_0)\mathbf{m}; \tag{28}$$

is the unique weak (in distributional sense) solution of (27); in particular

$$\mathbf{M}(t_0;t) = \mathbf{M}(0;t-t_0)\ . \tag{29}$$

(iv) The Cauchy problem (27) is equivalent to (25).

Proof. See Appendix A.3. □

As an immediate consequence of the above result we obtain the desired equivalence between equations (27) and (24).

Corollary 4.4. *Let y be a solution of (2) on $[0,t_0]$; when choosing $\mathbf{m}$ as $(m_0,m_1) = (y(t_0), y(t_0+\cdot))$, (27) is equivalent to (24) and in this case we have*

$$\mathbf{M}(t_0;t) = S(t-t_0)\mathbf{m} = (m(t_0;t), m(t_0;t+\cdot)) = \Big(M_{t_0}(t), \{M_{t_0}(t+s)\}_{s\in[-d,0]}\Big)\ .$$

From now on we thus will work with formulation (27). The spectral properties of the operator A, that appears in this infinite-dimensional formulation, will be crucial to prove our result. We devote the next Section to the analysis of these properties.

4.3 Spectral properties of A

In the present Section we collect some technical results concerning the spectral properties of the operator A. Proof of Theorem 2.3 is based on the Lemmas presented here. The technical proofs are postponed to the Appendix.

Lemma 4.5. *The spectrum of the operator A is given by*

$$\sigma(A) = \{\lambda \in \mathbb{C} : K(\lambda) = 0\},$$

where

$$K(\lambda) := \lambda - (\mu_y - \sigma_y \cdot \kappa) - \int_{-d}^{0} e^{\lambda\tau}\Phi(\mathrm{d}\tau)\,, \quad \lambda \in \mathbb{C}\,. \tag{30}$$

The spectrum $\sigma(A)$ is a countable set and every $\lambda \in \sigma(A)$ is an isolated eigenvalue of finite multiplicity.

The spectral bound of A is

$$\lambda_0 = \sup\{\mathrm{Re}\,\lambda :\ K(\lambda) = 0\}\,. \tag{31}$$

Proof. See [23, Chapter 7, Lemma 2.1 and Theorem 4.2] □

We ca explicitly compute the resolvent operator of A.

Lemma 4.6. *Let $\rho(A)$ denote the resolvent set of A and let $\lambda \in \mathbb{R} \cap \rho(A)$. The resolvent operator of A at λ, denoted by $R(\lambda, A)$ is given by*

$$R(\lambda, A)\,(m_0, m_1) = (u_0, u_1) \tag{32}$$

with

$$\begin{aligned} u_0 &= \frac{1}{K(\lambda)}\left[m_0 + \int_{-d}^{0}\int_{-d}^{s} e^{-\lambda(s-\tau)}\Phi(\mathrm{d}\tau)\, m_1(s)\mathrm{d}s\right], \\ u_1(s) &= u_0 e^{\lambda s} + \int_{s}^{0} e^{-\lambda(\tau-s)} m_1(\tau)\,\mathrm{d}\tau. \end{aligned} \tag{33}$$

Proof. See Appendix A.4. □

Lemma 4.7. *For every real λ such that $\lambda > \lambda_0$ and every $\mathbf{m} = (m_0, m_1) \in \mathcal{H}$ we have*

$$\int_{0}^{+\infty} e^{-\lambda t} S(t)\mathbf{m}\,\mathrm{d}t = R(\lambda, A)\mathbf{m}. \tag{34}$$

Proof. Identity (34) is well known to hold for all real λ larger than the type of $S(t)$. Since $S(t)$ is compact for every $t \geq d$, its type is actually equal to its spectral radius λ_0. For a reference see e.g. [4, Part II, Chapter 1, Corollary 2.5]. □

4.4 Deriving the explicit formula for H

In this Section we exploit the results derived in the above Sections to prove the following.

Proposition 4.8. *Assume $r > \lambda_0$, then for any $t_0 \geq 0$, the quantity $H(t_0)$ defined in (3) has the following explicit form*

$$H(t_0) = \frac{1}{K}\left(y(t_0) + \int_{-d}^{0} G(s) y(t_0 + s)\, ds\right), \quad \mathbb{P} - a.s.,$$

where $y(t_0)$ denotes the solution at time t_0 of equation (2),

$$K := r - \mu_0 + \sigma_y \cdot \kappa - \int_{-d}^{0} e^{r\tau}\Phi(\mathrm{d}\tau), \tag{35}$$

and G is given by

$$G(s) := \int_{-d}^{s} e^{-r(s-\tau)}\Phi(\mathrm{d}\tau). \tag{36}$$

Remark 4.9. *Notice that the statement of the above result is the same of Theorem 2.3, but the assumptions here are different: we assume $r > \lambda_0$ instead of Assumption 2.2. An explanation of why we do actually consider Assumption 2.2 in Theorem 2.3, will be provided in the next Section.*

Proof. Let $\mathbf{m} = (m_0, m_1) = (y(t_0), y(t_0 + \cdot))$. We denote here by Π the projection on the first (finite-dimensional) component of $\mathcal{H}$, i.e. $\Pi[\mathbf{m}] = \Pi[(m_0, m_1)] = m_0$.

Starting from (18), we have

$$\begin{aligned}
\frac{1}{\xi(t_0)}\int_{t_0}^{+\infty}\mathbb{E}\left[\xi(s)y(s)\mid\mathcal{F}_{t_0}\right]\mathrm{d}s &= e^{rt_0}\int_{t_0}^{\infty}e^{-rs}\tilde{\mathbb{E}}\left[y(s)|\mathcal{F}_{t_0}\right]\,\mathrm{d}s && \text{(by (18))}\\
&= e^{rt_0}\int_{t_0}^{\infty}e^{-rs}M_{t_0}(s)\,\mathrm{d}s && \text{(by (23)}\\
&= e^{rt_0}\int_{t_0}^{\infty}e^{-rs}\Pi\left[\mathbf{M}(t_0;s)\right]\,\mathrm{d}s && \text{(by Corollary 4.4)}\\
&= e^{rt_0}\int_{0}^{\infty}e^{-rt_0}e^{-rs}\Pi\left[\mathbf{M}(0;s)\right]\,\mathrm{d}s && \text{(by (29))}\\
&= \int_{0}^{\infty}e^{-rs}\Pi\left[S(s)\mathbf{m}\right]\,\mathrm{d}s && \text{(by (28))}\\
&= \Pi\left[R(r,A)\mathbf{m}\right] && \text{(by Lemma 4.7, since } r>\lambda_0)\\
&= \frac{1}{K(r)}\left[y(t_0)+\int_{-d}^{0}\int_{-d}^{s}e^{-r(s-\tau)}\Phi(\mathrm{d}\tau)\,y(t_0+s)\mathrm{d}s\right] && \text{(by Lemma 4.6).}
\end{aligned}\tag{37}$$

From the above equalities we infer, in particular, the $\mathbb{P}$-integrability of $\int_{t_0}^{+\infty}\mathbb{E}[\xi(s)y(s)\mid\mathcal{F}_{t_0}]\,\mathrm{d}s$, which justifies the equality

$$\mathbb{E}\left[\int_{t_0}^{+\infty}\xi(s)y(s)\mathrm{d}s\mid\mathcal{F}_{t_0}\right]=\int_{t_0}^{+\infty}\mathbb{E}\left[\xi(s)y(s)\mid\mathcal{F}_{t_0}\right]\mathrm{d}s. \tag{38}$$

In fact, by the characteristic property of the conditional mean, and Fubini's Theorem we have that, for any $F\in\mathcal{F}_{t_0}$

$$\begin{aligned}
\int_F\int_{t_0}^{+\infty}\mathbb{E}[\xi(s)y(s)\mid\mathcal{F}_{t_0}]\,\mathrm{d}s\,\mathrm{d}\mathbb{P} &= \int_{t_0}^{+\infty}\int_F\mathbb{E}[\xi(s)y(s)\mid\mathcal{F}_{t_0}]\mathrm{d}\mathbb{P}\,\mathrm{d}s\\
&= \int_{t_0}^{+\infty}\int_F\xi(s)y(s)\,\mathrm{d}\mathbb{P}\,\mathrm{d}s=\int_F\int_{t_0}^{+\infty}\xi(s)y(s)\,\mathrm{d}s\,\mathrm{d}\mathbb{P}\\
&= \int_F\mathbb{E}\left[\int_{t_0}^{+\infty}\xi(s)y(s)\,\mathrm{d}s\mid\mathcal{F}_{t_0}\right]\mathrm{d}\mathbb{P}.
\end{aligned}$$

Defining now $K:=K(r)$ and recalling (3), (35) and (36), by (37) and (38), the result immediately follows. □

4.5 Motivations for Assumption 2.2

In Proposition 4.8 we proved our main result under the Assumption $r>\lambda_0$. This requirement is difficult to check in practice, since it requires an explicit computation of the spectral bound λ_0. In the present Section we therefore look for some sufficient conditions easier to check.

Set for $\lambda\in\mathbb{C}$

$$\widetilde{K}(\lambda):=\lambda-(\mu_y-\sigma_y\cdot\kappa)-\int_{-d}^{0}e^{\lambda\tau}|\Phi|(\mathrm{d}\tau), \tag{39}$$

where by $|\Phi|$ we denote the total variation measure of Φ. Set

$$\widetilde{\lambda_0}=\sup\left\{\operatorname{Re}\lambda:\ \widetilde{K}(\lambda)=0\right\}. \tag{40}$$

We note that $\widetilde{\lambda_0}$ is the spectral radius of the operator $\widetilde{A}:\mathcal{D}(\widetilde{A})\subset\mathcal{H}\to\mathcal{H}$ defined as follows:

$$\begin{aligned}
\mathcal{D}(\widetilde{A}) &:= \left\{(x_0,x_1)\in\mathcal{H}:x_1\in W^{1,2}(-d,0;\mathbb{R}),\ x_1(0)=x_0\right\},\\
\widetilde{A}\,(x_0,x_1) &:= \left((\mu_y-\sigma_y\cdot\kappa)x_0+\int_{-d}^{0}x_1(s)|\Phi|(\mathrm{d}s),\frac{\mathrm{d}}{\mathrm{d}s}x_1\right).
\end{aligned}$$

Lemma 4.10. *The function $\widetilde{K}$, restricted to the real numbers, is strictly increasing and the spectral bound $\widetilde{\lambda_0}$ is the only real root of the equation $\widetilde{K}(\xi) = 0$. In particular,*

$$\widetilde{K}(r) > 0 \iff r > \widetilde{\lambda_0}. \tag{41}$$

Proof. See Appendix A.5. □

Recall the definition of K given in (30) and the definition of the spectral bound of A, λ_0, given (31).

Lemma 4.11. *It holds*

$$\widetilde{\lambda_0} \geq \lambda_0.$$

Proof. See Appendix A.6. □

Thanks to the above two Lemmas it becomes now clear why we work under Assumption 2.2 in Theorem 2.3. It provides a sufficient condition for the condition $r > \lambda_0$, imposed in Proposition 4.8, to hold. In fact, assume $\widetilde{K} > 0$ as in Assumption 2.2, then by Lemmas 4.10 and 4.11 we immediately get $r > \lambda_0$.

Remark 4.12. *Notice that, if Φ is a positive measure, then $\widetilde{K} \equiv K$, $\lambda_0 \equiv \widetilde{\lambda_0}$ and the condition $K > 0$ becomes also necessary, that is $K > 0 \iff r > \lambda_0$.*

5 Conclusion

In this paper we have considered a complete market model in the spirit of Samuelson's original contribution [36], in which security prices evolve as geometric Brownian motions. Despite working within a classical setup, we have obtained a novel explicit pricing formula for stream of payments with *delayed* dynamics, by using techniques from infinite-dimensional stochastic analysis. Our valuation formula results in an explicit expression demonstrating the importance of appreciating the past to quantify the current market value of the future. The approach followed in this paper highlights how tools from infinite-dimensional analysis can be successfully used to address valuation problems that are non-Markovian, and hence beyond the reach of conventional approaches. As highlighted in the applications discussed in Section 3, it is apparent how our results and the techniques developed here can be used successfully to explicitly solve interesting classes of infinite dimensional stochastic optimal control problems with nontrivial state constraints.

References

[1] Abowd, J. M., and D. Card (1989). On the covariance structure of earnings and hours changes, *Econometrica*, **57**(2), pp. 411-445.

[2] Arriojas, M. , Hu, Y. , Mohammed, S.A. and G. Pap (2007). A delayed Black and Scholes formula, *Stochastic Analysis and Applications*, **25**(2), pp. 471–492.

[3] Barattieri, A., Basu, S. and P. Gottschalk (2010). Some evidence on the importance of sticky wages, NBER Working Paper No. 16130.

[4] Bensoussan, A., Da Prato, G., Delfour, M.C. and S.K. Mitter, (2007). *Representation and control of infinite dimensional systems*, Second Edition, Birkhauser.

[5] Biagini, S., Biffis, E., Gozzi, F. and M. Zanella (2022). Wage rigidity and retirement in optimal portfolio choice, Working Paper: https://arxiv.org/abs/2101.09732.

[6] Biagini, S., Gozzi, F. and M. Zanella (2022). Robust portfolio choice with sticky wages, *SIAM Journal on Financial Mathematics*, **13**(3), pp. 1004–1039.

[7] Biffis, E., Gozzi, F. and C. Prosdocimi, (2020). Optimal portfolio choice with path dependent labor income: The infinite horizon case, *SIAM Journal on Control and Optimization*, **58**(4), pp. 1906-1938.

[8] Biffis, E., Blake, D., Pitotti, L. and A. Sun (2016). The cost of counterparty risk and collateralization in longevity swaps. *Journal of Risk and Insurance*, **83**(2), pp. 387-419.

[9] Bodie, Z. and Merton, R.C. and W.F. Samuelson (1992). Labor supply exibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control*, **16**(3), 427-449.

[10] Boucekkine, R., Licandro, O., Puch, L. and F. Del Rio (2005). Vintage capital and the dynamics of the AK model, *Journal of Economic Theory*, **120**(1), pp. 39–72.

[11] Brigo, D., Morini, M. and A. Pallavicini (2013). *Counterparty Credit Risk, Collateral and Funding: With Pricing Cases For All Asset Classes*, John Wiley & Sons.

[12] Brigo, D. and Pallavicini, A. (2014). Nonlinear consistent valuation of CCP cleared or CSA bilateral trades with initial margins under credit, funding and wrong-way risks, *Journal of Financial Engineering*, **1**(1), pp. 1–60.

[13] Campbell, J. Y., Viceira, L. M. (2001). *Strategic asset allocation: Portfolio choice for long-term investors*, Oxford University Press.

[14] Da Prato, G. and J. Zabczyk (1992). *Stochastic Equations in Infinite Dimensions*, Cambridge University Press.

[15] Dickens, W.T. , Goette, L., Groshen, E.L., Holden, S., Messina, J. , Schweitzer, M.E., Turunen, J. and M.E. Ward (2007). How wages change: Micro evidence from the International Wage Flexibility Project, *Journal of Economic Perspectives*, **21**(2), pp. 195–214.

[16] Djeiche, B., Gozzi, F., Zanco, G. and M. Zanella (2022). Optimal portfolio choice with path dependent benchmarked labor income: A mean field model, *Stochastic Processes and Their Applications*, **145**, pp.48-85.

[17] Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing, *Mathematische Annalen*, **300**(1), pp. 463–520.

[18] Duffie, D. (2001). *Dynamic asset pricing theory*, Princeton University Press.

[19] Dunsmuir, W., Goldys, B. and C.V. Tran (2015). Stochastic delay differential equations as weak limits of autoregressive moving average time series, Working Paper, University of New South Wales and University of Sydney.

[20] Dybvig, P. and H. Liu (2010). Lifetime consumption and investment: Retirement and constrained borrowing, *Journal of Economic Theory*, **145**(3), pp. 885–907.

[21] Fabbri, G. and F. Gozzi (2008). Solving optimal growth models with vintage capital: The dynamic programming approach, *Journal of Economic Theory*, **143**(1), pp. 331–373.

[22] Guvenen, F. (2007). Learning your earning: Are labor income shocks really very persistent?, *American Economic Review*, **97**(3), 687-712.

[23] Hale, J.K. and S.M. Verduyn Lunel (1993). *Introduction to functional differential equations*, Springer-Verlag.

[24] Harrison, J.M. and D.M. Kreps (1979). Martingales and arbitrage in multiperiod securities markets, *Journal of Economic theory*, **20**(3), pp. 381–408.

[25] Harrison, J.M. and S.R. Pliska, (1981). Martingales and stochastic integrals in the theory of continuous trading, *Stochastic Processes and Their Applications*, **11**(3), pp. 215–260.

[26] Hubbard, R.G., Skinner, J. and S.P. Zeldes (1995). Precautionary saving and social insurance, *Journal of Political Economy*, **103**(2), pp. 360-399.

[27] Karatzsas, I. and S.E. Shreve (1991). *Brownian motion and stochastic calculus*, Springer-Verlag.

[28] Khan, S. (1997). Evidence of nominal wage stickiness from microdata, *American Economic Review*, **87**(5), pp. 993–1008.

[29] Lorenz, R. (2005). *Weak approximation of stochastic delay differential equations with bounded memory by discrete time series*, PhD Thesis, Humboldt University, Berlin.

[30] Le Bihan, H., Montornès, J. and T.Heckel (2012). Sticky wages: Evidence from quarterly microeconomic data, *American Economic Journal: Macroeconomics*, **4**(3), pp. 1–32.

[31] MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis, *Journal of Econometrics*, **18**(1), 83-114.

[32] Mao, X. and S. Sabanis, (2013). Delay geometric Brownian motion in financial option valuation, *Stochastics An International Journal of Probability and Stochastic Processes*, **85**(2), pp. 295–320.

[33] Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity, *Econometrica*, **72**(1) pp. 1–32.

[34] Mohammed, S.A. (1998). *Stochastic Differential Systems with Memory:Theory, Examples and Applications*, Southern Illinois University Carbondale.

[35] Reiss, M., (2002). *Nonparametric estimation for stochastic delay differential equations*, PhD Thesis, Humboldt University, Berlin.

[36] Samuelson, P.A., (1965). Rational theory of warrant pricing, *Industrial Management Review*, 6, pp. 13-32.

[37] Samuelson, P.A., (1973). Mathematics of speculative price, *SIAM Review*, **15**(1), pp. 1-42.

A Further technical details

A.1 Proof of Proposition 2.1

The existence and uniqueness result for (2) is not covered by the extant literature. When the initial datum $x = (x_0, x_1)$, seen as a function on $[-d, 0]$, is continuous, existence and uniqueness of the strong solution to the SDDE for y is proved by [34, Theorem I.2]. When the initial datum $x \in \mathbb{R} \times L^2([-d, 0), \mathrm{d}t; \mathbb{R})$, with the additional requirement that ϕ and $(\varphi_1, ...\varphi_n)$ are absolutely continuous w.r.t the Lebesgue measure, that is $\mathrm{d}\phi = \varphi \mathrm{d}t$, $\mathrm{d}(\varphi_1, ...\varphi_n) = (\phi_1, ...\phi_n)\,\mathrm{d}t$ and the Radon-Nikodym densities $\varphi, \phi_1, ...\phi_n \in L^2([-d, 0), \mathrm{d}t)$, the existence and uniqueness result follows by [34, Remark I.3(iv)]. We need to extend this latter result to the case in which ϕ and $(\varphi_1, ...\varphi_n)$ are signed measures of bounded variation on $[-d, 0]$. We will prove the result by means of the same procedure employed for the proof of [6, Proposition B.2]. There the authors prove the existence and uniqueness of the solution for an equation similar to (2), under more general assumptions on the measure ϕ, but with no delay in the diffusion term.

Let us start by introducing the standard notation for the past path at t of a (deterministic) function $h : [-d, T] \to \mathbb{R}$, for $0 \le t \le T$, that is the function h_t

$$h_t(s) := h(t + s) \quad \text{for } -d \le s \le 0.$$

The past path of y at t for the realization ω is thus $y_t(s, \omega) := y(t + s, \omega)$ $s \in [-d, 0]$. As usual, we hide the dependence of the process on ω and write the delay terms in the drift and in the diffusion as follows:

$$\int_{-d}^{0} y(t+s)\phi(\mathrm{d}s) = \int_{-d}^{0} y_t(s)\phi(\mathrm{d}s) \tag{42}$$

and

$$\begin{pmatrix} \int_{-d}^{0} y(t+s)\varphi_1(\mathrm{d}s) \\ \vdots \\ \int_{-d}^{0} y(t+s)\varphi_n(\mathrm{d}s) \end{pmatrix} = \begin{pmatrix} \int_{-d}^{0} y_t(s)\varphi_1(\mathrm{d}s) \\ \vdots \\ \int_{-d}^{0} y_t(s)\varphi_n(\mathrm{d}s). \end{pmatrix} \tag{43}$$

The delay parts in (2), given by (42) and (43) can be then expressed in terms of (an extension of) the following linear operators ok kernel type:

$$L : C([-d, 0]; \mathbb{R}) \to \mathbb{R}, \qquad Lf := \int_{-d}^{0} f(s)\,\phi(\mathrm{d}s), \tag{44}$$

and

$$G : C([-d, 0]; \mathbb{R}) \to \mathbb{R}^n, \qquad Gf := \begin{pmatrix} \int_{-d}^{0} f(s)\varphi_1(\mathrm{d}s) \\ \vdots \\ \int_{-d}^{0} f(s)\varphi_n(\mathrm{d}s) \end{pmatrix}. \tag{45}$$

Since the operators L and G are well-defined only on the space of continuous functions $C([-d, 0]; \mathbb{R})$, when the initial datum does not belongs to $C([-d, 0]; \mathbb{R})$ but just to $L^2([-d, 0); \mathbb{R})$ problems may arise. In fact, consider the initial datum $(x_0, x_1) \in \mathbb{R} \times L^2(-d, 0; \mathbb{R})$ and proceed by assuming that the solution to (2) exists. Denote the past path on the window $[t - d, t]$ by $y_t : [-d, 0] \to \mathbb{R}$, $y_t(s) := y(t + s)$ a.e. $t \ge 0$, $s \in [-d, 0]$. Then, for $0 \le t < d$, the past path is

$$y_t(s) = \begin{cases} y(t+s) & \text{if } -t \le s < 0 \\ x_1(s) & \text{if} -d \le s < -t. \end{cases}$$

which, in general, is not a continuous function, but only square integrable. Therefore, the operators L and G introduced in (44) and (45) cannot be applied to y_t since the integrals in (42) and (43) may not be well defined.

The first issue is thus to show that L and G admit continuous extensions to the square integrable functions on $[-d, 0]$, as made precise in the following lemma.

Lemma A.1. *Let $L : C([-d,0];\mathbb{R}) \to \mathbb{R}$ and $G : C([-d,0];\mathbb{R}) \to \mathbb{R}^n$ be the continuous and linear maps given in (44) and (45) respectively. Fix $T > 0$ and define on $C([-d,T];\mathbb{R})$ the delay operators*

$$\mathcal{L}(y)(t) := Ly_t, \qquad 0 \le t \le T,$$

$$\mathcal{G}(y)(t) := Gy_t, \qquad 0 \le t \le T.$$

Then,

1. *the maps $\mathcal{L} : C([-d,T];\mathbb{R}) \to L^2([0,T];\mathbb{R})$ and $\mathcal{G} : C([-d,T];\mathbb{R}) \to L^2([0,T];\mathbb{R}^n)$ satisfy, respectively, the inequalities*

$$\|\mathcal{L}(y)\|_{L^2([0,T];\mathbb{R})} \le |\phi|([-d,0])\|y\|_{L^2([-d,T];\mathbb{R})}, \qquad \forall y \in C([-d,T];\mathbb{R}). \tag{46}$$

$$\|\mathcal{G}(y)\|_{L^2([0,T];\mathbb{R}^n)} \le \left(\sum_{i=1}^{n}[|\varphi_i|([-d,0])]^2\right)^{\frac{1}{2}} \|y\|_{L^2([-d,T];\mathbb{R})}, \qquad \forall y \in C([-d,T];\mathbb{R}). \tag{47}$$

2. *$\mathcal{L}$ and $\mathcal{G}$ have L^2-norm continuous, linear extensions (still denoted by $\mathcal{G}$ and $\mathcal{L}$, respectively) to $L^2([-d,T];\mathbb{R})$.*

Proof. The proof follows the lines of [6, Lemma B.1] (see also [4, Part II, Chapter 4, Theorem 3.3], but for the sake of completeness, we prove the result for the operator $\mathcal{G}$. For the operator $\mathcal{L}$ one follows the same reasoning.

1.

$$\begin{aligned}
\|\mathcal{G}(y)\|_{L^2([0,T];\mathbb{R}^n)} = \|Gy_\cdot\|_{L^2([0,T];\mathbb{R}^n)} &= \sup_{h \in L^2([0,T];\mathbb{R}^n), \|h\|_{L^2} \le 1} \int_0^T Gy_r \cdot h(r)\,\mathrm{d}r \\
&= \sup_{h \in L^2([0,T];\mathbb{R}^n), \|h\|_{L^2} \le 1} \int_0^T \begin{pmatrix} \int_{-d}^0 y_r(s)\varphi_1(\mathrm{d}s) \\ \vdots \\ \int_{-d}^0 y_r(s)\varphi_n(\mathrm{d}s) \end{pmatrix} \cdot h(r)\,\mathrm{d}r \\
&= \sup_{h \in L^2([0,T];\mathbb{R}^n), \|h\|_{L^2} \le 1} \sum_{i=1}^n \int_0^T h_i(r) \int_{-d}^0 y_r(s)\varphi_i(\mathrm{d}s)\,\mathrm{d}r.
\end{aligned}$$

We estimate the i-th component ($i = 1...n$) of the above expression exploiting the Fubini Theorem and the Hölder inequality.

$$\begin{aligned}
\int_0^T h_i(r) \int_{-d}^0 y_r(s)\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r &\le \int_0^T |h_i(r)| \int_{-d}^0 |y(r+s)|\,|\varphi_i|(\mathrm{d}s)\,\mathrm{d}r \\
&= \int_{-d}^0 \int_0^T |h_i(r)||y(r+s)|\,\mathrm{d}r\,|\varphi_i|(\mathrm{d}s) \\
&\le \int_{-d}^0 \|h_i\|_{L^2([0,T];\mathbb{R})}\|y\|_{L^2([s,s+T];\mathbb{R})}\,|\varphi_i|(\mathrm{d}s) \\
&\le |\varphi_i|([-d,0])\|h_i\|_{L^2([0,T];\mathbb{R})}\|y\|_{L^2([-d,T];\mathbb{R})},
\end{aligned}$$

where for the last inequality we exploit the inclusion $[s, s+T] \subseteq [-d,T]$. Therefore, by means of the Hölder inequality we obtain

$$
\begin{aligned}
\|\mathcal{G}(y)\|_{L^2([0,T];\mathbb{R}^n)} &\leq \sup_{h\in L^2([0,T];\mathbb{R}^n),\|h\|_{L^2}\leq 1} \sum_{i=1}^n |\varphi_i|([-d,0])\|h_i\|_{L^2([0,T];\mathbb{R})}\|y\|_{L^2([-d,T];\mathbb{R})}\\
&\leq \|y\|_{L^2([-d,T];\mathbb{R})} \sup_{h\in L^2([0,T];\mathbb{R}^n),\|h\|_{L^2}\leq 1} \left(\sum_{i=1}^n [|\varphi_i|([-d,0])]^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n \|h_i\|^2_{L^2([0,T];\mathbb{R})}\right)^{\frac{1}{2}}\\
&= \|y\|_{L^2([-d,T];\mathbb{R})} \sup_{h\in L^2([0,T];\mathbb{R}^n),\|h\|_{L^2}\leq 1} \left(\sum_{i=1}^n [|\varphi_i|([-d,0])]^2\right)^{\frac{1}{2}} \|h\|^2_{L^2([0,T];\mathbb{R}^n)}\\
&\leq \left(\sum_{i=1}^n [|\varphi_i|([-d,0])]^2\right)^{\frac{1}{2}} \|y\|_{L^2([-d,T];\mathbb{R})}.
\end{aligned}
$$

2. The existence of the bounded linear extension of $\mathcal{L}$ and $\mathcal{G}$ to $L^2([-d,t];\mathbb{R})$ is a consequence of inequalities (46) and (47) and the fact that $C([-d,T];\mathbb{R})$ is dense in $L^2([-d,T];\mathbb{R})$.

□

We are now ready to prove Proposition 2.1. We will use the following notation: if functions $a,b\geq 0$ satisfy the inequality $a\leq C(A)b$ with a constant $C(A)>0$ depending on the expression A, we write $a\lesssim_A b$.

Proof of Proposition 2.1. The proof of the result relies on a contraction type argument. The same argument has been used in the proof of [6, Proposition B.2]. There the authors consider a SDDE of type (2) with no delay in the diffusion term. On the other hand they work in a more general setting considering a measure valued process ϕ in the delay integral of the drift term.
We provide here a sketch of the proof referring to [6] for more details. We will give just the details of the estimates concerning the delay integral in the diffusion term that is missing in [6].

Let us fix the initial condition $(x_0,x_1)\in\mathbb{R}^2\times L^2([-d,0];\mathbb{R})$. Let $T>0$, we introduce the space

$$S_T := \{y\in C([0,T];\mathbb{R}) : y(0)=x_0\},$$

endowed with the sup norm

$$\|y\|_{S_T} = \sup_{t\in[0,T]} |y(t)|.$$

We consider the space $L^p(\Omega;S_T)$, $p\geq 2$, endowed with the norm

$$\|y\|_{L^p(\Omega;S_T)} = \left(\mathbb{E}\left[\|y\|^p_{S_T}\right]\right)^{\frac{1}{p}} = \left(\mathbb{E}\left[\sup_{t\in[0,T]}|y(t)|^p\right]\right)^{\frac{1}{p}}.$$

In the sequel we will denote by $p' := \frac{p}{p-1}$ the conjugate exponent to p and by $p^* := \frac{p}{p-2}$ the conjugate exponent to $\frac{p}{2}$.
Given $y\in L^p(\Omega;S_T)$, let

$$F(y)(t) := x_0 + \mu_y\int_0^t y(r)\,\mathrm{d}r + \int_0^t \mathcal{L}(\bar{y}^{x_1})\,\mathrm{d}r + \int_0^t y(r)\sigma_y\cdot\mathrm{d}Z(r) + \int_0^t \mathcal{G}(\bar{y}^{x_1})\cdot\mathrm{d}Z(r), \qquad 0\leq t\leq T. \tag{48}$$

Here $\mathcal{L}$ and $\mathcal{G}$ are the continuous linear operators introduced in Lemma A.1 and $\bar{y}^{x_1}\in L^p(\Omega;L^2([-d,T];\mathbb{R}))$ is defined as follows:

$$\bar{y}^{x_1}(t) = \begin{cases} x_1(t), & \text{if } -d\leq t<0;\\ y(t), & \text{if } \ 0\leq t\leq T.\end{cases} \tag{49}$$

We aim at proving that F maps $L^p(\Omega; S_T)$ into itself for any $p \geq 2$ and that it is a contraction on the same space when $p > 4$.

Let us start by proving that F maps $L^p(\Omega, S_T)$, $p \geq 2$, into itself. We write

$$\|F(y)\|_{L^p(\Omega;S_T)} \leq |x_0| + |\mu_y| \left\| \int_0^\cdot y(r)\,\mathrm{d}r \right\|_{L^p(\Omega;S_T)} + \left\| \int_0^\cdot \mathcal{L}(\bar{y}^{x_1})\,\mathrm{d}r \right\|_{L^p(\Omega;S_T)} + \left\| \int_0^\cdot y(r)\sigma_y \cdot \mathrm{d}Z(r) \right\|_{L^p(\Omega;S_T)} + \left\| \int_0^\cdot \mathcal{G}(\bar{y}^{x_1}) \cdot \mathrm{d}Z(r) \right\|_{L^p(\Omega;S_T)}. \tag{50}$$

The boundedness of the terms that appears in the r.h.s. of (50), except the last one, can be proved following the lines of [6, Proposition B.2]. We estimate the last term in the r.h.s. of (50) by means of the Burkholder-Davies-Gundy inequality

$$\begin{aligned}
\left\| \int_0^\cdot \mathcal{G}(\bar{y}^{x_1}) \cdot \mathrm{d}Z(r) \right\|^p_{L^p(\Omega;S_T)} &= \mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_0^t \mathcal{G}(\bar{y}^{x_1}) \cdot \mathrm{d}Z(r) \right|^p \right] \lesssim \mathbb{E}\left[\left| \int_0^T \|\mathcal{G}(\bar{y}^{x_1})\|^2\,\mathrm{d}r \right|^{\frac{p}{2}} \right] \\
&= \mathbb{E}\left[\|\mathcal{G}(\bar{y}^{x_1})\|^p_{L^2([0,T];\mathbb{R}^n)} \right] \lesssim \mathbb{E}\left[\|\bar{y}^{x_1}\|^p_{L^2([-d,T];\mathbb{R})} \right] \\
&= \|\bar{y}^{x_1}\|^p_{L^p(\Omega;L^2([-d,T];\mathbb{R}))} < \infty,
\end{aligned}$$

where in the last inequality we exploited (47) of Lemma A.1.

Let us now prove that, for $p > 4$, F defines a contraction in $L^p(\Omega, S_T)$. We endow this space by the equivalent norm

$$\|y\|_\alpha := \left(\mathbb{E}\left[\sup_{t\in[0,T]} \left(e^{-\alpha t}|y(t)| \right)^p \right] \right)^{\frac{1}{p}}, \tag{51}$$

where $\alpha > 0$ will be chosen later on. Once we proved that F defines a contraction, by the Banach fixed point Theorem, we can infer the existence of a unique $y \in L^p(\Omega; S_T)$ such that $y = F(y)$, i.e.

$$y(t) = x_0 + \mu_y \int_0^t y(r)\,\mathrm{d}r + \int_0^t \mathcal{L}(\bar{y}^{x_1})\,\mathrm{d}r + \int_0^t y(r)\sigma_y \cdot \mathrm{d}Z(r) + \int_0^t \mathcal{G}(\bar{y}^{x_1}) \cdot \mathrm{d}Z(r), \qquad 0 \leq t \leq T, \qquad \mathbb{P}-a.s.,$$

and this will conclude the proof.

Given $y, z \in L^p(\Omega; S_T)$, from (48) and (51), we have

$$\begin{aligned}
\|F(z) - F(y)\|^p_\alpha \lesssim_p &\mathbb{E}\left[\sup_{t\in[0,T]} e^{-p\alpha t} \left(|\mu_y| \left| \int_0^t (z(r) - y(r))\mathrm{d}r \right|^p + \left| \int_0^t \mathcal{L}(\bar{z}^{x_1} - \bar{y}^{x_1})\,\mathrm{d}r \right|^p \right) \right] \\
&+ \mathbb{E}\left[\sup_{t\in[0,T]} e^{-p\alpha t} \left| \int_0^t (z(r) - y(r))\sigma_y \cdot \mathrm{d}Z(r) \right|^p \right] \\
&+ \mathbb{E}\left[\sup_{t\in[0,T]} e^{-p\alpha t} \left| \int_0^t \mathcal{G}(\bar{z}^{x_1} - \bar{y}^{x_1}) \cdot \mathrm{d}Z(r) \right|^p \right].
\end{aligned} \tag{52}$$

We can estimate the first three terms in the r.h.s. of (52) proceeding as in [6, Proposition B.2] [4]. For the first term we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]} e^{-p\alpha t} |\mu_y| \left| \int_0^t (z(r) - y(r))\mathrm{d}r \right|^p \right] \leq |\mu_y| T \left(\frac{1}{\alpha p'} \right)^{\frac{p}{p'}} \|z - y\|^p_\alpha \lesssim_{\mu_y,T,p} C_1(\alpha)\|z - y\|^p_\alpha. \tag{53}$$

[4] For more details on the estimates, the interested reader can consult that paper.

For the second term we get

$$\mathbb{E}\left[\sup_{t\in[0,T]} e^{-p\alpha t}\left|\int_0^t \mathcal{L}(\bar{z}^{x_1}-\bar{y}^{x_1})\,\mathrm{d}r\right|^p\right] \le \left(\frac{|\phi|([-d,0])}{\alpha p'}\right)^{\frac{p}{p'}} T|\phi|([-d,0])\|z-y\|_\alpha^p \lesssim_{|\phi|,p,T} C_2(\alpha)\|z-y\|_\alpha^p. \tag{54}$$

For the third term, by means of the so called factorization method (see e.g. [?, Section 5.3]), for a given $\delta \in \left(\frac{1}{p},\frac{1}{2}\right)$[5], we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} e^{-p\alpha t}\left|\int_0^t (z(r)-y(r))\sigma_y\cdot \mathrm{d}Z(r)\right|^p\right]\\
&\lesssim_{p,\delta} \left(\int_0^T u^{p'(\delta-1)}e^{-p'\alpha u}\,\mathrm{d}u\right)^{\frac{p}{p'}} T\|\sigma_y\|^p \left(\sup_{u\in[0,T]}\int_0^u (u-r)^{-2\delta}e^{-2\alpha(u-r)}\,\mathrm{d}r\right)^{\frac{p}{2}}\|z-y\|_\alpha^p\\
&\lesssim_{p,\delta,T,\|\sigma_y\|} C_3(\alpha)\|z-y\|_\alpha^p.
\end{aligned} \tag{55}$$

Let us now come to the estimate of the fourth term in (52). Exploiting the factorization method, for $\eta\in\left(\frac{1}{p},\frac{p-2}{2p}\right)$ [6] we rewrite that stochastic integral as follows

$$\int_0^t \mathcal{G}(\bar{z}^{x_1}-\bar{y}^{x_1})\cdot\,\mathrm{d}Z(r) = c_\eta\int_0^t (t-u)^{\eta-1}Y(u)\,\mathrm{d}u,$$

with

$$\frac{1}{c_\eta} := \int_r^t (t-u)^{\eta-1}(u-r)^{-\eta}\,\mathrm{d}u = \frac{\pi}{\sin(\pi\eta)},$$

and

$$Y(u) = \int_0^u (u-r)^{-\eta}\mathcal{G}(\bar{z}^{x_1}-\bar{y}^{x_1})\cdot\mathrm{d}Z(r).$$

Thanks to the Hölder inequality we estimate

$$\begin{aligned}
e^{-\alpha t}\left|\int_0^t \mathcal{G}(\bar{z}^{x_1}-\bar{y}^{x_1})\cdot\mathrm{d}Z(r)\right| &= c_\eta e^{-\alpha t}\left|\int_0^t (t-u)^{\eta-1}Y(u)\,\mathrm{d}u\right|\\
&= c_\eta\left|\int_0^t e^{-\alpha(t-u)}(t-u)^{\eta-1}e^{-\alpha u}Y(u)\,\mathrm{d}u\right|\\
&\le c_\eta\left(\int_0^t e^{-\alpha p'(t-u)}(t-u)^{p'(\eta-1)}\,\mathrm{d}u\right)^{\frac{1}{p'}}\left(\int_0^t e^{-\alpha p u}|Y(u)|^p\,\mathrm{d}u\right)^{\frac{1}{p}}.
\end{aligned}$$

Therefore we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} e^{-\alpha p t}\left|\int_0^t \mathcal{G}(\bar{z}^{x_1}-\bar{y}^{x_1})\cdot\mathrm{d}Z(r)\right|^p\right]\\
&\le c_\eta^p\mathbb{E}\left[\sup_{t\in[0,T]}\left(\int_0^t e^{-\alpha p'(t-u)}(t-u)^{p'(\eta-1)}\,\mathrm{d}u\right)^{\frac{p}{p'}}\left(\int_0^t e^{-\alpha p u}|Y(u)|^p\,\mathrm{d}u\right)\right]\\
&\le c_\eta^p\left(\int_0^T e^{-\alpha p' u}u^{p'(\eta-1)}\,\mathrm{d}u\right)^{\frac{p}{p'}}\int_0^T e^{-\alpha p u}\mathbb{E}\left[|Y(u)|^p\right]\,\mathrm{d}u.
\end{aligned}$$

[5] Notice that this condition require to work with $p>2$.

[6] This condition is made in order to guarantee the convergence of the integrals that will appear in what follows. Notice that this condition require to work with $p>4$.

Now, recalling the definition of $\mathcal{G}$ and that, when $r < d$, $\bar{z}^{x_1}_r(s) - \bar{y}^{x_1}_r(s) = 0$ for $s \in [-d, -r)$ (see (49)), by means of the Burkholder-Davis-Gundy (BDG) and the Hölder (H) inequalities, we obtain for all $u \in [0, T]$,

$$
\begin{aligned}
e^{-\alpha pu}\mathbb{E}\left[|Y(u)|^p\right] &= e^{-\alpha pu}\mathbb{E}\left[\left|\int_0^u (u-r)^{-\eta}\mathcal{G}(\bar{z}^{x_1} - \bar{y}^{x_1}) \cdot \mathrm{d}Z(r)\right|^p\right] \\
&\overset{BDG}{\lesssim_p} e^{-\alpha up}\mathbb{E}\left[\left|\int_0^u (u-r)^{-2\eta}\|\mathcal{G}(\bar{z}^{x_1} - \bar{y}^{x_1})\|^2 \,\mathrm{d}r\right|^{\frac{p}{2}}\right] \\
&= e^{-\alpha up}\mathbb{E}\left[\left|\int_0^u (u-r)^{-2\eta}\sum_{i=1}^n \left|\int_{-d}^0 (\bar{z}^{x_1}_r - \bar{y}^{x_1}_r)(s)\,\varphi_i(\mathrm{d}s)\right|^2 \mathrm{d}r\right|^{\frac{p}{2}}\right] \\
&= e^{-\alpha up}\mathbb{E}\left[\left|\int_0^u (u-r)^{-2\eta}\sum_{i=1}^n \left|\int_{-d\vee -r}^0 ((z(r+s) - y(r+s))\ \varphi_i(\mathrm{d}s)\right|^2 \mathrm{d}r\right|^{\frac{p}{2}}\right] \\
&= \mathbb{E}\left[\left|\int_0^u (u-r)^{-2\eta}e^{-2\alpha(u-r-s)}e^{-2\alpha(r+s)}\sum_{i=1}^n \left|\int_{-d\vee -r}^0 (z(r+s) - y(r+s))\ \varphi_i(\mathrm{d}s)\right|^2 \mathrm{d}r\right|^{\frac{p}{2}}\right] \\
&\overset{H}{\leq} \mathbb{E}\left[\left|\int_0^u (u-r)^{-2\eta}e^{-2\alpha(u-r-s)}e^{-2\alpha(r+s)}\sum_{i=1}^n |\varphi_i|([-d,0])\int_{-d\vee -r}^0 |(z(r+s) - y(r+s)|^2\ \varphi_i(\mathrm{d}s)\,\mathrm{d}r\right|^{\frac{p}{2}}\right] \\
&\overset{H}{\leq} \left(\sum_{i=1}^n (|\varphi_i|([-d,0]))^{p^*}\int_0^u\int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}} \\
&\qquad \mathbb{E}\left[\sum_{i=1}^n\int_0^u\int_{-d\vee -r}^0 e^{-\alpha p(r+s)}|z(r+s) - y(r+s)|^p\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right] \\
&\leq \left(\sum_{i=1}^n (|\varphi_i|([-d,0]))^{p^*}\int_0^u\int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}} \\
&\qquad \sum_{i=1}^n\int_0^u\int_{-d\vee -r}^0 \mathbb{E}\left[\sup_{(r+s)\in[0,u]}\left(e^{-\alpha p(r+s)}|z(r+s) - y(r+s)|^p\right)\right]\varphi_i(\mathrm{d}s)\,\mathrm{d}r \\
&\leq \left(\sum_{i=1}^n (|\varphi_i|([-d,0]))^{p^*}\int_0^u\int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}} \\
&\qquad u\sum_{i=1}^n |\varphi_i|([-d,0])\mathbb{E}\left[\sup_{(r+s)\in[0,u]}\left(e^{-\alpha p(r+s)}|z(r+s) - y(r+s)|^p\right)\right].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\int_0^T e^{-\alpha p u}\mathbb{E}\left[|Y(u)|^p\right]\,\mathrm{d}u \\
&\quad\lesssim_{|\varphi_i|,p} \int_0^T u\left(\sum_{i=1}^n \int_0^u \int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}} \\
&\qquad\qquad \mathbb{E}\left[\sup_{(r+s)\in[0,u]}\left(e^{-\alpha p(r+s)}|z(r+s)-y(r+s)|^p\right)\right]\,\mathrm{d}u \\
&\quad\lesssim_{|\varphi_i|,p} \int_0^T u\left(\sum_{i=1}^n \int_0^u \int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}} \\
&\qquad\qquad \mathbb{E}\left[\sup_{(r+s)\in[0,T]}\left(e^{-\alpha p(r+s)}|z(r+s)-y(r+s)|^p\right)\right]\,\mathrm{d}u \\
&\quad=\left(\int_0^T u\left(\sum_{i=1}^n \int_0^u \int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}}\,\mathrm{d}u\right)\ \|z-y\|_\alpha^p \\
&\quad\lesssim_{|\varphi_i|,T,p}\left(\int_0^T r^{-2p^*\eta}e^{-2\alpha p^* r}\,\mathrm{d}r\right)^{\frac{p}{2p^*}}\ \|z-y\|_\alpha^p
\end{aligned}
$$

where the last inequality is obtained as follows:

$$
\begin{aligned}
&\int_0^T u\left(\sum_{i=1}^n \int_0^u \int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r-s)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}}\,\mathrm{d}u \\
&\quad\le T\int_0^T \left(\sum_{i=1}^n \int_0^u \int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r)}e^{2\alpha p^* s}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}}\,\mathrm{d}u \\
&\quad\le T\int_0^T \left(\sum_{i=1}^n \int_0^u \int_{-d\vee -r}^0 (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r)}\,\varphi_i(\mathrm{d}s)\,\mathrm{d}r\right)^{\frac{p}{2p^*}}\,\mathrm{d}u \\
&\quad\le T\sup_{u\in[0,T]}\left(\sum_{i=1}^n |\varphi_i|([-d,0])\int_0^u (u-r)^{-2p^*\eta}e^{-2\alpha p^*(u-r)}\,\mathrm{d}r\right)^{\frac{p}{2p^*}} \\
&\quad\lesssim_{|\varphi_i|,T}\left(\int_0^T r^{-2p^*\eta}e^{-2\alpha p^* r}\,\mathrm{d}r\right)^{\frac{p}{2p^*}}.
\end{aligned}
$$

Putting together the above estimates we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} e^{-\alpha p t}\left|\int_0^t \mathcal{G}(\bar{z}^{x_1}-\bar{y}^{x_1})\cdot\mathrm{d}Z(r)\right|^p\right] \\
&\quad\lesssim_{T,|\varphi_i|,\eta,p}\left(\int_0^T e^{-\alpha p' u}u^{p'(\eta-1)}\,\mathrm{d}u\right)^{\frac{p}{p'}}\left(\int_0^T r^{-2p^*\eta}e^{-2\alpha p^* r}\,\mathrm{d}r\right)^{\frac{p}{2p^*}}\ \|z-y\|_\alpha^p \\
&\quad\lesssim_{T,|\varphi_i|,\eta,p} C_4(\alpha)\|z-y\|_\alpha^p, \qquad (56)
\end{aligned}
$$

Finally, from (53), (54), (55) and (56) we infer

$$
\|F(z)-F(y)\|_\alpha^p \lesssim_{\mu_y,T,p,|\phi|,|\varphi_i|,\|\sigma_y\|,\delta,\eta}\sum_{i=1}^4 C_i(\alpha)\|z-y\|_\alpha^p,
$$

where $C_i(\alpha) \to 0$ as $\alpha \to \infty$, for $i = 1, ..., 4$. So, by taking $\alpha > 0$ sufficiently large, this proves that F is a contraction, thus there exists a unique fixed point of it. In this way we prove the existence and uniqueness of the solution in the space $L^p(\Omega, S_T)$ for $p > 4$. Since, for such p, $L^p(\Omega, S_T) \subset L^2(\Omega, S_T)$, such solution also belongs to $L^2(\Omega, S_T)$. To get uniqueness in the space $L^2(\Omega, C([0,T];\mathbb{R}))$ one can take two solutions y and $\tilde{y}$ in this space and take their difference. Using the fact that both are fixed points of F, by means of the Gronwall Lemma one gets $\sup_{t\in[0,T]} \mathbb{E}\left[|y(t) - \tilde{y}(t)|^2\right] = 0$ and this concludes the proof. □

A.2 Proof of Lemma 4.1

Proof. Let us denote with σ_y^i the i-th component of σ_y, and let us show that

$$\mathbb{E}\left[\int_{t_0}^t \left|y(s)\sigma_y^i + \int_{-d}^0 y(s+\tau)\varphi_i(\mathrm{d}\tau)\right|^2 \mathrm{d}s\right] < +\infty.$$

By the trivial inequality $(a+b)^2 \le 2(a^2+b^2)$, it is sufficient to show that

$$\mathbb{E}\left[\int_{t_0}^t |y(s)\sigma_y^i|^2 \mathrm{d}s\right] < +\infty, \tag{57}$$

and

$$\mathbb{E}\left[\int_{t_0}^t \left|\int_{-d}^0 y(s+\tau)\varphi_i(\mathrm{d}\tau)\right|^2 \mathrm{d}s\right] < +\infty. \tag{58}$$

We immediately see that (57) holds true thanks to Proposition 2.1.

To show (58), we use the Hölder inequality and the Fubini Theorem to estimate

$$\begin{aligned}
\mathbb{E}\left[\int_{t_0}^t \left|\int_{-d}^0 y(\tau+s)\,\varphi_i(\mathrm{d}\tau)\right|^2 \mathrm{d}s\right] &\le |\varphi_i|([-d,0])\mathbb{E}\left[\int_{t_0}^t \int_{-d}^0 |y(\tau+s)|^2\,|\varphi_i|(\mathrm{d}\tau)\,\mathrm{d}s\right] \\
&= |\varphi_i|([-d,0])\mathbb{E}\left[\int_{-d}^0 \int_{t_0}^t |y(\tau+s)|^2\,\mathrm{d}s\,|\varphi_i|(\mathrm{d}\tau)\right] \\
&= |\varphi_i|([-d,0])\mathbb{E}\left[\int_{-d}^0 \int_{t_0+\tau}^{t+\tau} |y(r)|^2\,\mathrm{d}r\,|\varphi_i|(\mathrm{d}\tau)\right] \\
&\le |\varphi_i|([-d,0])\mathbb{E}\left[\int_{-d}^0 \int_{t_0-d}^{t} |y(r)|^2\,\mathrm{d}r\,|\varphi_i|(\mathrm{d}\tau)\right] \\
&= \left(|\varphi_i|([-d,0])\right)^2 \mathbb{E}\left[\int_{t_0-d}^{t} |y(r)|^2\,\mathrm{d}r)\right],
\end{aligned}$$

which is finite, thanks to Proposition 2.1. □

A.3 Proof of Proposition 4.3

Proof. (*i*) The operator A can be written in the form

$$A\left(x_0, x_1\right) = \left(\int_{-d}^0 x_1(\theta)a(\mathrm{d}\theta), \frac{\mathrm{d}}{\mathrm{d}s}x_1\right), \tag{59}$$

where

$$a(\mathrm{d}\theta) = \mu_y\delta_0(\mathrm{d}\theta) + \Phi(\mathrm{d}\theta),$$

and δ_0 is the delta-Dirac measure at zero. The measure a defines a finite measure on $[-d,0]$. The result is thus an immediate consequence of [14, Proposition A.27].

(*ii*) See e.g. [23, Chapter 7, Lemma 1.2].

(*iii*) Existence and uniqueness of a weak solution given by (28) for deterministic $\mathbf{m}$ is a classical result (see [14, Proposition A.5]). One can then easily generalize the result to random $\mathbf{m}$. Property (29) follows from uniqueness of the solution.

(*iv*) If $m(t_0;\cdot)$ is the unique solution to (25) then the $\mathcal{H}$-valued process $(m(t_0;t), m(t_0;t+\cdot))_{t\geq t_0}$ solves (27) by [4, Part II, Chapter 4, Theorem 4.3]. Since also the latter has a unique solution, its first component must be the solution to (27).

□

A.4 Proof of Lemma 4.6

Proof. If $\lambda \in \mathbb{R} \cap \rho(A)$ then $K(\lambda) \neq 0$ by Lemma 4.5. To compute $R(\lambda, A)$, we will consider for a fixed $\mathbf{m} = (m_0, m_1) \in \mathcal{H}$ the equation

$$(\lambda - A)(u_0, u_1) = (m_0, m_1)\,, \tag{60}$$

in the unknown $(u_0, u_1) \in \mathcal{D}(A)$, that by definition of A is equivalent to

$$\begin{cases} (\lambda - (\mu_y - \sigma_y \cdot \kappa))u_0 - \displaystyle\int_{-d}^{0} u_1(\tau)\Phi(\mathrm{d}\tau) = m_0 \\ \lambda u_1 - \dfrac{\mathrm{d}u_1}{\mathrm{d}s} = m_1. \end{cases}$$

Then

$$u_1(s) = e^{\lambda s}u_0 + \int_s^0 e^{-\lambda(\tau - s)} m_1(\tau)\,\mathrm{d}\tau, \quad s \in [-d, 0],$$

and u_0 is determined by the equation

$$\big(\lambda - (\mu_y - \sigma_y \cdot \kappa)\big)u_0 = \left[m_0 + \int_{-d}^{0}\left(e^{\lambda\tau}u_0 + \int_\tau^0 e^{-\lambda(s-\tau)} m_1(s)\,\mathrm{d}s\right)\Phi(\mathrm{d}\tau)\right],$$

yielding

$$K(\lambda)u_0 = m_0 + \int_{-d}^{0}\int_{-d}^{s} e^{-\lambda(s-\tau)}\Phi(\mathrm{d}\tau)\, m_1(s)\mathrm{d}s.$$

Then the result immediately follows. □

A.5 Proof of Lemma 4.10

Proof. It is immediate to check that the function $\widetilde{K} : \mathbb{R} \to \mathbb{R}$ is continuous and differentiable with

$$\widetilde{K}'(\xi) = 1 + \int_{-d}^{0} e^{\xi\tau}|\tau|\,|\Phi|(\mathrm{d}\tau) > 0,$$

and that

$$\lim_{\xi \to \pm\infty} \widetilde{K}(\xi) = \pm\infty\,.$$

Equation $\widetilde{K}(\xi) = 0$ has thus exactly one real solution $\bar{\xi}$. Let us now show that $\bar{\xi} = \widetilde{\lambda_0}$. By the definition of $\widetilde{\lambda_0}$, clearly we have that $\bar{\xi} \leq \widetilde{\lambda_0}$. To show the opposite inequality, $\bar{\xi} \geq \widetilde{\lambda_0}$, we consider an arbitrary $\lambda = a + ib \in \mathbb{C}$ such that $\widetilde{K}(\lambda) = 0$. Then

$$\begin{aligned} 0 = \mathrm{Re}(\widetilde{K}(\lambda)) &= a - \mu_y + \sigma_y \cdot \kappa - \int_{-d}^{0} e^{a\tau}\cos(b\tau)\,|\Phi|(\mathrm{d}\tau) \\ &\geq a - \mu_y + \sigma_y \cdot \kappa - \int_{-d}^{0} e^{a\tau}\,|\Phi|(\mathrm{d}\tau) =: \widetilde{K}(a)\,. \end{aligned}$$

Since $\widetilde{K}$ is an increasing function, we can infer $\mathrm{Re}\lambda \leq \bar{\xi}$ and taking the supremum in the definition of $\widetilde{\lambda_0}$ we obtain $\widetilde{\lambda_0} \leq \bar{\xi}$. By the same argument, the relation $\widetilde{K}(r) > 0 \iff r > \widetilde{\lambda_0}$ immediately follows. $\square$

A.6 Proof of Lemma 4.11

Proof. Since by Lemma 4.10 we know that $\widetilde{K}$ is an increasing function and $\widetilde{K}(\widetilde{\lambda_0}) = 0$, we just need to prove that $\widetilde{K}(\lambda_0) \leq 0$. For every $\lambda = a + ib \in \mathbb{C}$ we have

$$\mathrm{Re}(K(\lambda)) = a - (\mu_y + \sigma_y \cdot \kappa) - \int_{-d}^{0} e^{a\tau} \cos(b\tau)\, \Phi(\mathrm{d}\tau).$$

Recalling the definition of λ_0 it is enough to show that, for every $\lambda = a+ib \in \mathbb{C}$ such that $K(\lambda) = 0$, it holds $\widetilde{K}(a) \leq 0$. We have that

$$\begin{aligned}
\tilde{K}(a) &= a - (\mu_y + \sigma_y \cdot \kappa) - \int_{-d}^{0} e^{a\tau} |\Phi|(\mathrm{d}\tau) \\
&= a - (\mu_y + \sigma_y \cdot \kappa) - \int_{-d}^{0} e^{a\tau} \cos(b\tau)\, \Phi(\mathrm{d}\tau) - \int_{-d}^{0} e^{a\tau} |\Phi|(\mathrm{d}\tau) + \int_{-d}^{0} e^{a\tau} \cos(b\tau)\, \Phi(\mathrm{d}\tau) \\
&\leq \mathrm{Re}(K(\lambda)) + \int_{-d}^{0} e^{a\tau}\, [\Phi - |\Phi|](\mathrm{d}\tau) \leq 0.
\end{aligned}$$

This concludes the proof. $\square$