The Annals of Applied Probability
1996, Vol. 6, No. 2, 370–398

HEDGING OPTIONS FOR A LARGE INVESTOR AND FORWARD–BACKWARD SDE'S

BY JAKŠA CVITANIĆ[1] AND JIN MA[2]

Columbia University and Purdue University

In the classical continuous-time financial market model, stock prices have been understood as solutions to linear stochastic differential equations, and an important problem to solve is the problem of hedging options (functions of the stock price values at the expiration date). In this paper we consider the hedging problem not only with a price model that is nonlinear, but also with coefficients of the price equations that can depend on the portfolio strategy and the wealth process of the hedger. In mathematical terminology, the problem translates to solving a forward–backward stochastic differential equation with the forward diffusion part being degenerate. We show that, under reasonable conditions, the four step scheme of Ma, Protter and Yong for solving forward–backward SDE's still works in this case, and we extend the classical results of hedging contingent claims to this new model. Included in the examples is the case of the stock volatility increase caused by overpricing the option, as well as the case of different interest rates for borrowing and lending.

1. Introduction and summary. In the usual continuous-time model of a stock market, going back to Merton [19], a stock price process P is modeled as a solution of a linear stochastic differential equation, with given drift and noise ("volatility") coefficients. An assumption that has long been viewed as standard is that the investor is "small" in the sense that his/her financial status and trading strategy should not affect the model of the market prices. Therefore, in the classical model the coefficients of the price equations are independent of the wealth and portfolio processes of the investor. In this paper we consider the case in which the influence of the investor's financial behavior is *not* a priori known to be irrelevant and the price model is *not* necessarily linear. In other words, we assume that the drift and the volatility terms can both be nonlinear in the price process and also depend on the wealth process X and the portfolio process π of the investor. Such a model is useful when the investor is "not-too-small"; we call him/her "large" in the sequel, since the "small investor" assumption is obviously removed. Natural

Received December 1994; revised January 1996.
[1]Research supported in part by NSF Grant DMS-95-03582 and Army Research Office Grant DAAH 04-95-1-0528.
[2]Research supported in part by NSF Grant DMS-93-01516.
AMS 1991 *subject classifications*. Primary 90A09, 60H30; secondary 90A12, 93A20.
Key words and phrases. Forward–backward stochastic differential equations, contingent claims, hedging strategy, large investor.

examples include a market with different interest rates for borrowing and lending (with a "small" investor) and models in which volatility of the prices can change due to the "strange" behavior of the large investor. We study a hedging problem for this investor on a finite time horizon $[0, T]$: given an initial stock price $p = P(0)$ and a desired terminal wealth value $g(P(T))$ (which we call an *option*, as a special case of contingent claims), the investor wants to find a portfolio process and an initial wealth $x = X(0)$, such that the corresponding wealth process satisfies $X(T) = g(P(T))$. Moreover, he/she wants to find the hedging portfolio process that goes with the smallest initial wealth possible. That smallest initial wealth x is then the upper bound for the price of the option $g(P(T))$; that is, no one should be willing to pay more than x at time $t = 0$ for the option worth $g(P(T))$ at time $t = T$. (For more on option pricing theory, refer to the famous Black and Scholes paper [2]; see also [11, 12, 13, 14] for the martingale theory.)

In mathematical terminology, the problem translates to finding a solution (P, X, π), with minimal $X(0)$, of a forward–backward stochastic differential equation (FBSDE), with P being the forward and X the backward component. We use the four step scheme of Ma, Protter and Yong [18] to solve the FBSDE (we refer the reader to that paper for more references on the relatively new notion and theory of FBSDE's). Historically, it was the special case of FBSDE's, called backward stochastic differential equations (BSDE's), that was first developed in a mathematical context by Pardoux and Peng [20], and independently by Duffie and Epstein [6] in finance; see also [21] and [22]. In our context, the term "pure backward case" will mean those cases in which the desired terminal value does not depend on the wealth or portfolio process, either because the price process does not, or because the value is equal to an a priori given $\mathscr{F}_T$-measurable random variable [which may be of a more general form than $g(P(T))$]. A problem similar to the one of this paper, but only in the pure backward case, with the volatility of stocks being independent of the investor's policy and using methods and assumptions completely different from ours, was studied by El Karoui, Peng and Quenez [8] and Cvitanić [3]. The forward–backward case as a model in finance is used by Duffie, Ma and Yong [7] for a different problem concerning the term structure of interest rates. For related work on the interactions between hedging strategies and market prices, in the (equilibrium) context of several agents, see the very interesting papers by Platen and Schweizer [23] and Grossman [10] and references therein.

We should note here that in the present model, the price equation is no longer linear. So, unlike the classical case, it is already questionable how to keep the price processes $P_i(\cdot)$, $i = 1, \dots, d$, from becoming negative (in such a case the market will be destroyed, by common sense). One seemingly simple way to treat this, say with $d = 1$, is to add a "local-time" term to the price equation, so that the process $P(\cdot)$ becomes diffusion reflected at zero. The disadvantage of doing this is that in such a model there would be an opportunity for "arbitrage," that is, a chance to make a positive amount of money out of zero initial wealth, with a positive probability. This has been commonly viewed as an undesirable property of a financial market. For more on this matter, refer to [5, 1] and the example in [15]. Another method, which is the one we shall employ in this paper, is to pose some conditions on the coefficients of the price equation so that the solution $P(\cdot)$, whenever it exists, will always stay inside the region $\mathbb{R}^d_+ \triangleq \{(x_1, \dots, x_d) \in \mathbb{R}^d | x_i > 0,\ i = 1, \dots, d\}$. In other words, the boundary $\partial \mathbb{R}^d_+$ is "natural." It turns out that one of the conditions is that the forward equation (price equation) will have to be degenerate; that is, the volatility function σ will vanish on the boundary $\partial \mathbb{R}^d_+$. On the other hand, however, this degeneracy will cause technical difficulties for us to obtain the unique adapted solution for the FBSDE, because the result in [18] cannot be applied directly. We show that, under certain conditions, such a conflict can be resolved in a satisfactory way, by extending the results of [18]. Some other possible methods that might lead to positive but weaker results, such as replacing the "price–wealth" pair by the "log-price–wealth" pair or using a suitable change of probability measure, will also be discussed.

In this paper we will only consider the problem of hedging an option. Namely, the terminal value of the wealth is specified as $g(P(T))$, where $P(T)$ is the value of stocks at the expiration date and g is some smooth function of linear growth. The complete resolution to the hedging problem for general contingent claims (i.e., arbitrary terminal conditions for the wealth) will require further development in the theory of FBSDE's and will be studied separately.

The paper is organized as follows. We describe the model and give definitions and some preliminary results in Section 2. In Section 3 we study the corresponding FBSDE's using the four step scheme and prove the admissibility of the solution as a hedging strategy. Section 4 is devoted to two comparison theorems. The first leads to the uniqueness of the FBSDE and implies that the solution to the FBSDE will indeed produce the optimal hedging strategy; that is, the corresponding wealth process has the smallest initial endowment $x = X(0)$ among those that do the hedging. The second theorem shows that the optimal strategy is monotone with respect to the terminal value; the higher the option value at the expiration date, the higher the premium. In Section 5 we give examples that motivate our model, including the case in which there is an increase in the stock's volatility if the option is overpriced and an example of a market with different interest rates for borrowing and lending. The former leads to a phenomenon unknown in the classical case: the hedging is guaranteed if one sells the option at the fair price (e.g., using the Black–Scholes formula); however, *if one sells the option for more than that, then one may not be able to do the hedging* because of the corresponding change of the volatility in the market. Finally, we give an example (in the Appendix) showing that the classical comparison theorem for a backward SDE need not hold in the present forward–backward case, supporting our argument in Section 4.

2. Problem formulation. Let us consider a market $\mathscr{M}$ in which $d+1$ assets are traded continuously. One is called a *bank account* which is "riskless" and the others are called *stocks*, which are assumed to be "risky." We consider an investor in this market and, contrary to the usual "small investor" hypothesis, we assume that both this investor's wealth and strategy, once exposed, might influence the prices of the financial instruments. More precisely, we assume that the price of the bank account evolves according to the differential equation

$$\begin{aligned} dP_0(t) &= P_0(t) r(t, X(t), \pi(t))\, dt, \qquad 0 \le t \le T, \\ P_0(0) &= 1, \end{aligned} \tag{2.1}$$

and that the price of the stocks evolves according to the stochastic differential equation, for $0 \le t \le T$,

$$\begin{aligned} dP_i(t) &= b_i(t, P(t), X(t), \pi(t))\, dt \\ &\qquad + \sum_{j=1}^{d} \sigma_{ij}(t, P(t), X(t), \pi(t))\, dW_j(t), \\ P_i(0) &= p_i > 0, \qquad i = 1, \dots, d, \end{aligned} \tag{2.2}$$

where $T > 0$ is the *maturity date* or *duration* and X is the *wealth process*, while $\pi = (\pi_1, \dots, \pi_d)$ is the *portfolio process* of the investor and $W = (W_1, \dots, W_d)$ is a d-dimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, \mathbf{P})$, with the filtration $\{\mathscr{F}_t\}_{t \ge 0}$, which is the $\mathbf{P}$-augmentation of the natural filtration $\mathscr{F}_t^W \triangleq \sigma\{W(s): 0 \le s \le t\}$ generated by W. We require now (and specify later) that the functions b and σ are such that the solution P_i's are positive processes.

Now let us suppose that the investor will start with an initial endowment $x \ge 0$ and try to allocate his wealth into the bank and stocks according to a certain strategy at each time $t \in [0, T]$. We define each process π_i to be the amount of money that the investor puts into the ith stock; thus the amount invested in the bank will be $X(t) - \sum_{i=1}^{d} \pi_i(t)$. Furthermore, if we allow the investor to consume a certain amount of money at each time t and denote the cumulative consumption up to time t by $C(t)$, then $C(\cdot)$ is a nondecreasing, $\{\mathscr{F}_t\}$-adapted process, $C(0) = 0$. It is intuitive that the change of the wealth in a small time increment $[t, t+h)$ can be described approximately by

$$\begin{aligned} &X(t+h) - X(t) \\ &\quad = \sum_{i=1}^{d} \frac{\pi_i(t)}{P_i(t)} (P_i(t+h) - P_i(t)) + \frac{X(t) - \sum_{i=1}^{d} \pi_i(t)}{P_0(t)} \\ &\qquad \times (P_0(t+h) - P_0(t)) - (C(t+h) - C(t)). \end{aligned} \tag{2.3}$$

This amounts to saying that the wealth process satisfies the stochastic differential equation

$$
\begin{aligned}
dX(t) &= \sum_{i=1}^{d} \frac{\pi_i(t)}{P_i(t)}\, dP_i(t) + \frac{\left(X(t) - \sum_{i=1}^{d}\pi_i(t)\right)}{P_0(t)}\, dP_0(t) - dC(t) \\
&= \sum_{i=1}^{d} \frac{\pi_i(t)}{P_i(t)} \Bigg\{ b_i(t, P(t), X(t), \pi(t))\, dt \\
&\qquad\qquad + \sum_{j=1}^{d} \sigma_{ij}(t, P(t), X(t), \pi(t))\, dW_j(t) \Bigg\} \\
&\quad + \frac{X(t) - \sum_{i=1}^{d}\pi_i(t)}{P_0(t)} P_0(t) r(t, X(t), P(t)\pi(t))\, dt - dC(t) \\
&= \hat{b}(t, P(t), X(t), \pi(t))\, dt \\
&\quad + \hat{\sigma}(t, P(t), X(t), \pi(t))\, dW(t) - dC(t), \\
X(0) &= x > 0,
\end{aligned}
\tag{2.4}
$$

where

$$
\begin{aligned}
\hat{b}(t, p, x, \pi) &= \left(x - \sum_{i=1}^{d} \pi_i\right) r(t, x, \pi) + \sum_{i=1}^{d} \frac{\pi_i}{p_i} b_i(t, p, x, \pi), \\
\hat{\sigma}_j(t, p, x, \pi) &= \sum_{i=1}^{d} \frac{\pi_i}{p_i} \sigma_{ij}(t, p, x, \pi), \qquad j = 1, \dots, d,
\end{aligned}
\tag{2.5}
$$

for $(t, p, x, \pi) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$. In this paper we shall use the following notation throughout: we denote the positive orthent by $\mathbb{R}_+^d = \{(x_1, \dots, x_d) \in \mathbb{R}^d | x_i > 0,\ i = 1, \dots, d\}$; the inner product in $\mathbb{R}^d$ by $\langle \cdot , \cdot \rangle$; the norm in $\mathbb{R}^d$ by $|\cdot|$ and that of $\mathbb{R}^{d\times d}$, the space of all $d \times d$ matrices, by $\|\cdot\|$ and the transpose of a matrix $A \in \mathbb{R}^{d\times d}$ (resp. a vector $x \in \mathbb{R}^d$) by A^{T} (resp. x^{T}). We also denote $\underline{\mathbf{1}}$ to be the vector $\underline{\mathbf{1}} \triangleq (1, \dots, 1) \in \mathbb{R}^d$ and define a (diagonal) matrix-valued function $\Lambda\colon \mathbb{R}^d \mapsto \mathbb{R}^{d\times d}$ by

$$
\Lambda(x) \triangleq \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_d \end{bmatrix}, \qquad x = (x_1, \dots, x_d) \in \mathbb{R}^d.
\tag{2.6}
$$

It is obvious that $\|\Lambda(x)\| = |x|$ for any $x \in \mathbb{R}^d$, and whenever $x \notin \partial\mathbb{R}_+^d$, $\Lambda(x)$ is invertible and $[\Lambda(x)]^{-1}$ is of the same form as $\Lambda(x)$ with $x_1, \dots, x_d$ being replaced by $x_1^{-1}, \dots, x_d^{-1}$. We can then rewrite the functions $\hat{b}$ and $\hat{\sigma}$ in (2.5) as

$$
\begin{aligned}
\hat{b}(t, p, x, \pi) &= x r(t, x, \pi) + \left\langle \pi, b^1(t, p, x, \pi) - r(t, x, \pi)\underline{\mathbf{1}} \right\rangle, \\
\hat{\sigma}(t, p, x, \pi) &= \left\langle \pi, \sigma^1(t, p, x, \pi) \right\rangle,
\end{aligned}
\tag{2.7}
$$

where

$$
(2.8)\quad \begin{aligned}
b^1(t,p,x,\pi) &\triangleq [\Lambda(p)]^{-1} b(t,p,x,\pi) = \left(\frac{b_1}{p_1},\dots,\frac{b_d}{p_d}\right)(t,p,x,\pi),\\
\sigma^1(t,p,x,\pi) &\triangleq [\Lambda(p)]^{-1} \sigma(t,p,x,\pi) = \left\{\frac{\sigma_{ij}}{p_i}\right\}_{i,j=1}^{d}(t,p,x,\pi).
\end{aligned}
$$

To be consistent with the classical model, we henceforth call b^1 the *appreciation rate* and σ^1 the *volatility matrix* of the stock market. We now give more precise definitions of the quantities appearing in (2.4).

DEFINITION 2.1. (i) A *portfolio process* $\pi = \{\pi(t);\ 0 \le t \le T\}$ is a real-valued, progressively $\{\mathscr{F}_t\}$-measurable process, such that $E\int_0^T |\pi(t)|^2\,dt < \infty$.

(ii) A *consumption process* $C = \{C(t);\ 0 \le t \le T\}$ is a real-valued, $\{\mathscr{F}_t\}$-adapted process, with nondecreasing and RCLL (right-continuous with left limits) paths, such that $C(0) = 0$ and $C(T) < \infty$, a.s. P.

(iii) For a given portfolio–consumption pair (π, C), the *price process* with the initial value $p > 0$ and the *wealth process* with initial capital $x \ge 0$ are the solutions to the SDE's (2.2) and (2.4), respectively, which will often be denoted by $P = P^{p,x,\pi,C}$ and $X = X^{p,x,\pi,C}$, whenever the dependence of the solution on p, x, π, C needs to be specified.

We will make use of the following *standing assumptions*:

(A1) The function $b, \sigma\colon [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ and $g\colon \mathbb{R}^d \mapsto \mathbb{R}$ are twice continuously differentiable. The functions b^1 and σ^1, together with their first order partial derivatives in p, x and π are bounded, uniformly in (t,p,x,π). Further, we assume that partial derivatives of b^1 and σ^1 in p satisfy

$$
(2.9)\qquad \sup_{(t,p,x,\pi)} \left\{ \left| p_k \frac{\partial b^1}{\partial p_k} \right|, \left| p_k \frac{\partial \sigma^1_{ij}}{\partial p_k} \right| \right\} < \infty, \qquad i,j,k = 1,\dots,d.
$$

(A2) The function σ satisfies $\sigma\sigma^{\mathsf{T}}(t,p,x,\pi) > 0$ for all (t,p,x,π) with $p \notin \partial\mathbb{R}^d_+$, and there exists a positive constant $\mu > 0$, such that

$$
(2.10)\qquad a^1(t,p,x,\pi) \ge \mu I \quad \text{for all } (t,p,x,\pi),
$$

where $a^1 = \sigma^1(\sigma^1)^{\mathsf{T}}$.

(A3) The function r is twice continuously differentiable and such that the following conditions are satisfied:

(a) For $(t,x,\pi) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d$, $0 < r(t,x,\pi) \le K$, for some constant $K > 0$.

(b) The partial derivatives of r in x and π, denoted by a generic function

ψ, satisfy

$$\limsup_{|x|,|\pi|\to\infty,} (|x|+|\pi|)^2 |\psi(t,x,\pi)| < \infty. \tag{2.11}$$

REMARK 2.2. (1) We note that Assumptions (A1) and (A2) obviously contain those cases in which $b(t,p,x,\pi) = \Lambda(p)b_1(t,x,\pi)$ and $\sigma(t,p,x,\pi) = \Lambda(p)\sigma_1(t,x,\pi)$, where b_1 and σ_1 are bounded, continuously differentiable functions with bounded first order partial derivatives, and $\sigma_1\sigma_1^{\mathsf{T}}$ is positive definite and bounded away from zero, as we often see in the classical model. In particular, our setting will contain the Black–Scholes model as a special case. Condition (A3) is somewhat restrictive, which is largely due to the generality of our setting. It also contains the classical case when $r(t) \equiv r$. Moreover, as we shall see in examples in Section 5, the method described below sometimes works even if the assumptions are far from being satisfied.

(2) The boundedness of the functions b^1 and σ^1 imply that b and σ will vanish on the set $\partial\mathbb{R}_+^d$, the boundary of $\mathbb{R}_+^d$. In other words, σ has to be *degenerate* on $\partial\mathbb{R}_+^d$. This requirement is to guarantee that the stock prices stay positive all the time so that the market is not destroyed, as we shall prove in the following lemma.

LEMMA 2.3. *Suppose that* (A1) *and* (A2) *hold. Then for any portfolio–consumption pair* (π, C) *and initial wealth* x, *the price process* P *satisfies* $P_i(t) > 0$, $i = 1,\dots,d$ *for all* $t \in [0,T]$, *almost surely*, *provided the initial prices* $p_1,\dots,p_d$ *are positive.*

PROOF. Let (π, C) and the initial values (p,x) be given. Let (P,X) denote the solutions to the (forward) SDE's (2.2) and (2.4). By definition of b^1 and σ^1, we can rewrite (2.2) in the form

$$\begin{aligned} P_i(t) = p_i + \int_0^t P_i(s)\big[b_i^1(s,P(s),X(s),\pi(s))\,ds \\ + \sigma_i^1(s,P(s),X(s),\pi(s))\,dW(s)\big], \end{aligned} \tag{2.12}$$

where σ_i^1 is the ith row vector of the matrix σ^1. Denoting $b_i^1(t) = b_i^1(t,P(t),X(t),\pi(t))$ and $\sigma^1(t) = \sigma^1(t,P(t),X(t),\pi(t))$, and recalling their boundedness, we see that the processes P_i can be written as stochastic exponentials (see, e.g., [16], [24] or [25]),

$$P_i(t) = p_i \exp\left\{ \int_0^t \Big[b_i^1(s) - \tfrac{1}{2}\|\sigma_i^1(s)\|^2 \Big]\,ds + \sigma_i^1(s)\,dW(s) \right\}, \qquad i = 1,\dots,d;$$

hence the conclusion follows. □

Lemma 2.3 now enables us to give the following definition of the "admissible portfolio–consumption" strategy.

DEFINITION 2.4. For a given $x \geq 0$, a pair of portfolio–consumption processes (π, C) is called admissible (with respect to x) if, for any $p > 0$, the corresponding price process $P(\cdot)$ and wealth process $X(\cdot)$ satisfy $P_i(t) > 0$, $i = 1, \dots, d$ [i.e., $P(t) \in \mathbb{R}_+^d$] and $X(t) \geq 0$, $\forall\ t \in [0, T]$, a.s. $\mathbf{P}$.

For each x, we denote the set of all admissible portfolio–consumption pairs by $\mathscr{A}(x)$. We claim that $\mathscr{A}(x) \neq \varnothing$ for all x. To see this, we first note that for any $x > 0$, $p \in \mathbb{R}_+^d$ and (π, C), we can solve a pair of (forward) SDE's for P and X. By Lemma 2.1, we know that it is always true that $P(\cdot) \in \mathbb{R}_+^d$, a.s. $\mathbf{P}$. Therefore, we need only show that for each $x > 0$ there exists a pair (π, C), such that for all $p \in \mathbb{R}_+^d$, $X^{p,x,\pi,C}(\cdot) \geq 0$, a.s. $\mathbf{P}$. This can be done by choosing $\pi \equiv 0$ and $C \equiv 0$. By the definition of $\hat{b}$ and $\hat{\sigma}$ (2.5), we see that the wealth process will satisfy

$$X(t) = x + \int_0^t X(s) r(s, X(s), 0)\, ds,$$

whence $X(t) = x \exp\{\int_0^t r(s, X(s), 0)\, ds\} > 0$ for all t. In other words, the trivial pair $(0, 0) \in \mathscr{A}(x)$.

To conclude this section, we give the following definition.

DEFINITION 2.5. An *option* is an $\mathscr{F}_T$-measurable random variable $B = g(P(T))$, where g is a real function. The *hedging price* of the option is defined by

$$h(B) \triangleq \inf H(B), \tag{2.13}$$

where

$$H(B) \triangleq \{x \in \mathbb{R}: \exists\, (\pi, C) \in \mathscr{A}(x), \text{ s.t. } X^{x,\pi,C}(T) \geq B \text{ a.s.}\}. \tag{2.14}$$

3. Forward–backward SDE's. In this section we study the FBSDE's that will play an important role in our future discussions. Consider the FBSDE given by

$$\begin{aligned} P(t) &= p + \int_0^t b(s, P(s), X(s), \pi(s))\, ds \\ &\qquad + \int_0^t \sigma(s, P(s), X(s), \pi(s))\, dW(s), \\ X(t) &= g(P(T)) - \int_t^T \hat{b}(s, P(s), X(t), \pi(s))\, ds \\ &\qquad - \int_t^T \hat{\sigma}(s, P(s), X(s), \pi(s))\, dW(s) + C(T) - C(t), \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} \hat{b}(t, p, x, \pi) &= x r(t, x, \pi) + \langle \pi, b_1(t, p, x, \pi) - r(t, x, \pi)\underline{\mathbf{1}} \rangle, \\ \hat{\sigma}(t, p, x, \pi) &= \langle \pi, \sigma_1(t, p, x, \pi) \rangle. \end{aligned} \tag{3.2}$$

We first give the definition of an *adapted solution* to the FBSDE (3.1).

DEFINITION 3.1. A quadruple (P, X, π, C) is called an adapted solution to the FBSDE (3.1), if the following hold:

(i) P, X and π are $\{\mathscr{F}_t\}$-adapted, square integrable processes.
(ii) C is an $\{\mathscr{F}_t\}$-adapted, RCLL, nondecreasing process, such that $C(0) = 0$ and $C(T) < \infty$.

In what follows we shall only consider the FBSDE (3.1) with $C \equiv 0$, namely, the FBSDE

$$
\begin{aligned}
P(t) &= p + \int_0^t b(s, P(s), X(s), \pi(s))\, ds \\
&\quad + \int_0^t \sigma(s, P(s), X(s), \pi(s))\, dW(s), \\
X(t) &= g(P(T)) - \int_t^T \hat{b}(s, P(s), X(s), \pi(s))\, ds \\
&\quad - \int_t^T \hat{\sigma}(s, P(s), X(s), \pi(s))\, dW(s).
\end{aligned} \tag{3.3}
$$

The existence of an adapted solution to such an FBSDE will lead to the nonemptyness of the set $H(g(P(T)))$ of (2.14), and in fact, to $h(g(P(T))) \le X(0)$, where X is the backward component of the solution to FBSDE (3.3). We shall also assume that the function g satisfies either one of the following two conditions:

(A4) (a) The function g is bounded, C^2 and nonnegative. Its partial derivatives up to second order are all bounded.
(b) The function g is nonnegative and $\lim_{|p| \to \infty} g(p) = \infty$. Moreover, g has bounded, continuous partial derivatives up to third order and there exist constants $K, M > 0$ such that

$$
|\Lambda(p) g_p(p)| \le K(1 + g(p)), \qquad \sup_{p \in \mathbb{R}_+^d} \left\| \Lambda^2(p) g_{pp} \right\| = M < \infty.
$$

Further, we assume that the partial derivatives of σ^1 in x and π satisfy

$$
\sup_{(t, p, x, \pi)} \left\{ \left| x \frac{\partial \sigma_{ij}^1}{\partial x} \right| + \left| x \frac{\partial \sigma_{ij}^1}{\partial \pi_k} \right| \right\} < \infty, \qquad i, j, k = 1, \dots, d. \tag{3.4}
$$

REMARK 3.2. Clearly, the first inequality in (A4)(b) holds for any g that behaves like a polynomial for $|p|$ large, but the second condition restricts it to one that has at most quadratic growth. The condition (3.4) can be called a "compatibility condition" to compensate for the unboundedness of g. We note that it also contains the classical models as special cases. An example of a function σ satisfying (A1), (A2) and (A4)(b) could be $\sigma(t, p, x, \pi) = p(\sigma(t) + \arctan(x^2 + |\pi|^2))$ with $\sigma(\cdot)$ satisfying (A2).

In order to prove the existence and uniqueness of the adapted solution to (3.3), we follow the four step scheme designed in [18]. The main idea of the four step scheme is based on the well-known Feynman–Kac formula, combined with a "decoupling procedure" for the FBSDE. We should note that in the present case, the function σ is degenerate on the boundary of $\mathbb{R}_+^d$, and the function g is allowed to be unbounded [condition (A4)(b)]. Thus the result in [18] does not apply directly. However, by using the special structure of the functions $\hat{b}$ and $\hat{\sigma}$ and some transformations, we show that the four step scheme will remain valid in the present case. For convenience of presentation, we shall discuss the existence part first in this section and defer the proof of uniqueness to the next section, as a corollary of our comparison theorem. Let us first review the four step scheme [18].

Four step scheme. Let us first keep in mind that the solution to (3.3), whenever it exists, will satisfy $P(t) \in \mathbb{R}_+^d$, for all t, a.s. $\mathbf{P}$, thanks to Lemma 2.3. So we might as well restrict ourselves to the region $(t, p, x, \pi) \in [0, T] \times \mathbb{R}_+^d \times \mathbb{R} \times \mathbb{R}^d \triangleq E$ without further specification and we proceed as follows.

STEP 1. In order to match diffusion terms, find a "smooth" mapping $z: [0, T] \times \mathbb{R}_+^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ so that

$$q^{\mathsf{T}}\sigma(t, p, x, z(t, p, x, q)) - \hat{\sigma}(t, p, x, z(t, p, x, q)) = 0 \qquad \forall\, (t, p, x, q) \in E. \tag{3.5}$$

In our case, (3.5) becomes

$$\begin{aligned} &q^{\mathsf{T}}\sigma(t, p, x, z(t, p, x, q)) \\ &\qquad - z^{\mathsf{T}}(t, p, x, q)[\Lambda(p)]^{-1}\sigma(t, p, x, z(t, p, x, q)) = 0, \end{aligned} \tag{3.6}$$

hence $z(t, p, x, q) = \Lambda(p)q$ since $\sigma\sigma^{\mathsf{T}} > 0$ by (A2) and $\Lambda(\cdot)$ is a diagonal, nonsingular matrix. One should note that in the present case we solve for the function $z(\cdot, \cdot, \cdot, \cdot)$ directly, which makes our solution more explicit than that in [18].

STEP 2. With the intention of setting $\pi = z(t, p, \theta, \theta_p) = \Lambda(p)\theta_p$, solve the quasilinear parabolic equation for $\theta(t, p)$:

$$\begin{aligned} 0 &= \theta_t + \tfrac{1}{2}\mathrm{tr}\big\{\sigma\sigma^{\mathsf{T}}\big(t, p, \theta, \Lambda(p)\theta_p\big)\theta_{pp}\big\} \\ &\qquad + \big\langle b\big(t, p, \theta, \Lambda(p)\theta_p\big), \theta_p\big\rangle - \hat{b}\big(t, p, \theta, \Lambda(p)\theta_p\big), \\ \theta(T, p) &= g(p), \qquad p \in \mathbb{R}_+^d. \end{aligned} \tag{3.7}$$

In our case, by an easy computation using (3.2), we have that

$$\begin{aligned} \hat{b}\big(t, p, \theta, \Lambda(p)\theta_p\big) &= r\big(t, \theta, \Lambda(p)\theta_p\big)\theta \\ &\qquad + \big\langle b\big(t, p, \theta, \Lambda(p)\theta_p\big) - r\big(t, \theta, \Lambda(p)\theta_p\big)p, \theta_p\big\rangle. \end{aligned}$$

Thus (3.7) becomes

$$
\begin{aligned}
0 &= \theta_t + \tfrac{1}{2}\mathrm{tr}\big\{\sigma\sigma^{\mathsf T}\big(t,p,\theta,\Lambda(p)\theta_p\big)\theta_{pp}\big\} \\
&\qquad + \big(\langle p,\theta_p\rangle - \theta\big) r\big(t,\theta,\Lambda(p)\theta_p\big), \\
\theta(T,p) &= g(p), \qquad p \in \mathbb{R}^d_+ .
\end{aligned}
\tag{3.8}
$$

STEP 3. Setting

$$
\begin{aligned}
\tilde b(t,p) &= b\big(t,p,\theta(t,p),\Lambda(p)\theta_p(t,p)\big), \\
\tilde\sigma(t,p) &= \sigma\big(t,p,\theta(t,p),\Lambda(p)\theta_p(t,p)\big),
\end{aligned}
\tag{3.9}
$$

solve the forward SDE

$$
P(t) = p + \int_0^t \tilde b(s,P(s))\,ds + \int_0^t \tilde\sigma(s,P(s))\,dW(s). \tag{3.10}
$$

STEP 4. Setting

$$
X(t) = \theta(t,P(t)), \qquad \pi(t) = \Lambda(P(t))\theta_p(t,P(t)), \tag{3.11}
$$

show that (P, X, π) is the *unique* adapted solution to (3.1).

Before we proceed any further, let us give a lemma which shows that if (P, X, π) is a solution to the FBSDE (3.3), then the pair $(\pi, 0) \in \mathscr{A}(X(0))$.

LEMMA 3.3. *Suppose that* (A1)–(A3) *hold and let* (P, X, π) *be an adapted solution to* (3.3), *with* $g(p) \geq 0$ *for all* $p > 0$. *Then the pair* $(\pi, 0)$ *is an admissible hedging strategy with respect to* $X(0)$, *in the sense of Definition* 2.4.

PROOF. Let (P, X, π) be an adapted solution to (3.3). The process $P(\cdot)$ will stay inside $\mathbb{R}^d_+$ for all $t \in [0,T]$, a.s. $\mathbf{P}$, thanks to Lemma 2.3. Since the $dt \times d\mathbf{P}$-square integrability of π is already contained in the definition of the adapted solution to FBSDE, it remains to show that the wealth process $X(t) \geq 0$ for all $t > 0$. To this end, let us define, for the given processes (P, X, π), a (random) function

$$
\begin{aligned}
f(t,x,z) &= r(t,X(t),\pi(t))\,x \\
&\quad + \Big\langle z, \big\{\sigma^1(t,P(t),X(t),\pi(t))\big\}^{-1} \\
&\qquad \times \big[b^1(t,P(t),X(t),\pi(t)) - r(t,X(t),\pi(t))\underline{\mathbf{1}}\big]\Big\rangle,
\end{aligned}
\tag{3.12}
$$

and consider the (linear) backward SDE

$$
x(t) = g(P(T)) + \int_t^T f(s,x(s),z(s))\,ds + \int_t^T \langle z(s), dW(s)\rangle. \tag{3.13}
$$

Comparing with (3.2) and (3.3), we see that the pair (x, z) defined by

$$x(t) = X(t), \quad z(t) = (\sigma^1)^{\mathsf{T}}(t, P(t), X(t), \pi(t))\pi(t), \qquad t \in [0, T], \tag{3.14}$$

is an adapted solution to (3.13). On the other hand, note that the function f is linear in x and z with bounded derivatives by assumptions (A1) and (A2), and that $f(t, 0, 0) \equiv 0$. We see that $-f$ is a *standard generator* for the BSDE (3.13), in the terminology of [8]. Therefore, a comparison theorem for the classical (linear) BSDE's (see [8]) leads to $X(t) = x(t) \geq 0\ \forall\ t$, a.s. $\mathbf{P}$, whenever $g(P(T)) \geq 0$, a.s. $\mathbf{P}$, proving the lemma. □

Our main result in this section is the following theorem.

THEOREM 3.4. *Suppose that the standing assumptions* (A1)–(A3) *and* (A4)(b) *hold. Then for any given* $p \in \mathbb{R}_+^d$, *the FBSDE* (3.3) *admits a unique adapted solution* (P, X, π), *given by* (3.11) *with* θ *being the solution of* (3.8).

PROOF. We follow the four step scheme mentioned above. Step 1 is obvious. For Step 2, we claim the following assertion which might be of interest in its own right:

There exists a unique classical solution $\theta(\cdot, \cdot)$ *to the PDE* (3.8), *defined on* $(t, p) \in [0, T] \times \mathbb{R}_+^d$, *which enjoys the following properties*:

(i) $\theta - g$ *is uniformly bounded for* $(t, p) \in [0, T] \times \mathbb{R}_+^d$.

(ii) *The partial derivatives of* θ *satisfy*, *for some constant* $K > 0$,

$$\begin{aligned} \left|\Lambda(p)\theta_p(t, p)\right| &\leq K(1 + |p|), \\ \sup_{(t,p) \in [0,T] \times \mathbb{R}_+^d} \left\|\Lambda(p)^2 \theta_{pp}(t, p)\right\| &= K < \infty. \end{aligned} \tag{3.15}$$

To prove the assertion, let us first consider the function $\hat{\theta} \triangleq \theta - g$. It is obvious that $\hat{\theta}_t = \theta_t$, $\hat{\theta}_p = \theta_p - g_p$ and $\hat{\theta}_{pp} = \theta_{pp} - g_{pp}$, and $\hat{\theta}$ satisfies the PDE

$$\begin{aligned} 0 = {} & \hat{\theta}_t + \tfrac{1}{2}\operatorname{tr}\Big\{\sigma\sigma^{\mathsf{T}}\Big(t, p, \hat{\theta} + g(p), \Lambda(p)\big(\hat{\theta}_p + g_p(p)\big)\Big)\big(\hat{\theta}_{pp} + g_{pp}\big)\Big\} \\ & + r\Big(t, \hat{\theta} + g(p), \Lambda(p)\big(\hat{\theta}_p + g_p(p)\big)\Big)\big[\langle p, \hat{\theta}_p + g_p\rangle - \big(\hat{\theta} + g\big)\big], \\ & \hat{\theta}(T, p) = 0, \qquad p \in \mathbb{R}_+^d. \end{aligned} \tag{3.16}$$

To simplify notation, let us set

$$\begin{aligned} \bar{\sigma}(t, p, x, \pi) &= \sigma\big(t, p, x + g(p), \pi + \Lambda(p) g_p(p)\big), \\ \bar{r}(t, p, x, \pi) &= r\big(t, x + g(p), \pi + \Lambda(p) g_p(p)\big). \end{aligned} \tag{3.17}$$

Thus, noting that for $p, q \in \mathbb{R}^d$, we have $r(t, p, x, \pi)\langle p, q\rangle = \langle r(t, p, x, \pi)\underline{\mathbf{1}}, \Lambda(p)q\rangle$, (3.16) becomes

$$
\begin{aligned}
0 = \hat{\theta}_t &+ \tfrac{1}{2}\mathrm{tr}\Big\{\overline{\sigma\sigma}^{\mathsf{T}}\big(t, p, \hat{\theta}, \Lambda(p)\hat{\theta}_p\big)\hat{\theta}_{pp}\Big\} \\
&+ \Big\langle \bar{r}\big(t, p, \hat{\theta}, \Lambda(p)\hat{\theta}_p\big)\underline{\mathbf{1}}, \Lambda(p)\theta_p\Big\rangle + \hat{r}\big(t, p, \hat{\theta}, \Lambda(p)\hat{\theta}_p\big), \\
&\hat{\theta}(T, p) = 0, \qquad p \in \mathbb{R}^d_+,
\end{aligned}
\tag{3.18}
$$

where

$$
\begin{aligned}
\hat{r}(t, p, x, \pi) = \tfrac{1}{2}\mathrm{tr}\big\{\overline{\sigma\sigma}^{\mathsf{T}}(t, p, x, \pi)g_{pp}(p)\big\} \\
+\big\langle \bar{r}(t, p, x, \pi)\underline{\mathbf{1}}, \Lambda(p)g_p(p)\big\rangle \\
- \bar{r}(t, p, x, \pi)(x + g(p)).
\end{aligned}
\tag{3.19}
$$

Next, we define a change of variable by setting $\xi = L^{-1}p$ and $\tilde{\theta}(t, \xi) \triangleq \hat{\theta}(t, L\xi)$, where L and L^{-1} are defined by $L(\xi_1, \dots, \xi_d) = (e^{\xi_1}, \dots, e^{\xi_d})$, $\xi \in \mathbb{R}^d$, $L^{-1}(p_1, \dots, p_d) = (\log p_1, \dots, \log p_d)$, $p \in \mathbb{R}^d_+$. Then it is easily checked that

$$
\begin{aligned}
\tilde{\theta}_t(t, \xi) &= \hat{\theta}_t(t, e^{\xi}), \\
\tilde{\theta}_{\xi}(t, \xi) &= \Lambda(L\xi)\hat{\theta}_p(t, L\xi), \\
\tilde{\theta}_{\xi\xi}(t, \xi) &= \Lambda(L\xi)^2\hat{\theta}_{pp}(t, L\xi) + \Lambda(L\xi)\Lambda\big(\hat{\theta}_p(t, L\xi)\big).
\end{aligned}
\tag{3.20}
$$

Plugging (3.20) into (3.18) and doing some computation, we obtain a quasilinear parabolic PDE for $\tilde{\theta}$:

$$
\begin{aligned}
0 = \tilde{\theta}_t &+ \tfrac{1}{2}\mathrm{tr}\Big\{\overline{\sigma\sigma}^{\mathsf{T}}\big(t, L\xi, \tilde{\theta}, \tilde{\theta}_{\xi}\big)\big[(\Lambda(L\xi))^{-1}\big]^2\big(\tilde{\theta}_{\xi\xi} - \Lambda(L\xi)\Lambda\big(\tilde{\theta}_{\xi}\big)\big)\Big\} \\
&+ \Big\langle \bar{r}\big(t, L\xi, \tilde{\theta}, \tilde{\theta}_{\xi}\big)\underline{\mathbf{1}}, \tilde{\theta}_{\xi}\Big\rangle + \hat{r}\big(t, L\xi, \tilde{\theta}, \tilde{\theta}_{\xi}\big), \\
= \tilde{\theta}_t &+ \tfrac{1}{2}\mathrm{tr}\Big\{\bar{\sigma}_0\bar{\sigma}_0^{\mathsf{T}}\big(t, \xi, \tilde{\theta}, \tilde{\theta}_{\xi}\big)\tilde{\theta}_{\xi\xi}\Big\} \\
&+ \Big\langle b_0\big(t, \xi, \tilde{\theta}, \tilde{\theta}_{\xi}\big), \tilde{\theta}_{\xi}\Big\rangle + \hat{b}_0\big(t, \xi, \tilde{\theta}, \tilde{\theta}_{\xi}\big), \\
&\tilde{\theta}(T, \xi) = g(e^{\xi}), \qquad \xi \in \mathbb{R}^d,
\end{aligned}
\tag{3.21}
$$

where

$$
\begin{aligned}
\bar{\sigma}_0(t, \xi, x, \pi) &= [\Lambda(L\xi)]^{-1}\bar{\sigma}(t, L\xi, x, \pi), \\
b_0(t, \xi, x, \pi) &= \bar{r}(t, L\xi, x, \pi)\underline{\mathbf{1}} - \tfrac{1}{2}\mathrm{diag}\big[\bar{\sigma}_0\bar{\sigma}_0^{\mathsf{T}}(t, \xi, x, \pi)\big], \\
\hat{b}_0(t, \xi, x, \pi) &= \hat{r}(t, L\xi, x, \pi).
\end{aligned}
\tag{3.22}
$$

Here $\hat{r}$ is defined by (3.19), and for a matrix A we denote by diag$[A]$ the vector composed of the diagonal elements of A.

Now by using condition (A2) [note (2.10)], one sees that there exists a constant $\mu > 0$ such that $\bar{a}_0 \triangleq \bar{\sigma}_0 \bar{\sigma}_0^{\mathsf{T}}(t, \xi, x, \pi) \geq \mu I > 0$ for all $(t, \xi, x, \pi) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$. Also, by definition (3.17), we see that for all $i, j, k = 1, \dots, d$, and (t, ξ, x, π), it holds (suppressing the variables) that

$$
\begin{aligned}
\frac{\partial (\bar{a}_0)_{ij}}{\partial \xi_k} &= \frac{\partial a^1_{ij}}{\partial p_k} e^{\xi_k} + \frac{\partial a^1_{ij}}{\partial x} \frac{\partial g}{\partial p_k} e^{\xi_k} \\
&\quad + \sum_{l=1}^{d} \frac{\partial a^1_{ij}}{\partial \pi_l} \left[\frac{\partial^2 g}{\partial p_k \, \partial p_l} e^{\xi_l} e^{\xi_k} + \delta_{kl} e^{\xi_l} \frac{\partial g}{\partial p_l} \right].
\end{aligned} \tag{3.23}
$$

Note that by (A4)(b) we have that for each k, $|(\partial q / \partial p_k) e^{\xi_k}| \leq |\Lambda(L\xi) g(L\xi)| \leq K(1 + g(L\xi))$ and $0 \leq g(L\xi) \leq x + g(L\xi)$. Thus conditions (A1), (2.9) and (A4)(b) imply that

$$
\sup_{(t, \xi, x, \pi)} \left| \frac{\partial a^1_{ij}}{\partial p_k} \big(t, L\xi, x + g(L\xi), \pi + \Lambda(L\xi) g_p(L\xi) \big) e^{\xi_k} \right| < \infty,
$$

$$
\begin{aligned}
&\sup_{(t, \xi, x, \pi)} \left| \frac{\partial a^1_{ij}}{\partial x} \big(t, L\xi, x + g(L\xi), \pi + \Lambda(L\xi) g_p(L\xi) \big) \frac{\partial g}{\partial p_k} e^{\xi_k} \right| \\
&\quad \leq K \sup_{(t, \xi, x, \pi)} \left\{ \left| \frac{\partial a^1_{ij}}{\partial x} \big(t, L\xi, x + g(L\xi), \pi + \Lambda(L\xi) g_p(L\xi) \big) \right| \right. \\
&\qquad\qquad \left. \times [1 + (x + g(L\xi))] \right\} < \infty,
\end{aligned}
$$

$$
\begin{aligned}
&\sup_{(t, \xi, x, \pi)} \left| \frac{\partial a^1_{ij}}{\partial \pi_k} \big(t, L\xi, x + g(L\xi), \pi + \Lambda(L\xi) g_p(L\xi) \big) \frac{\partial g}{\partial p_k} e^{\xi_k} \right| \\
&\quad \leq K \sup_{(t, \xi, x, \pi)} \left\{ \left| \frac{\partial a^1_{ij}}{\partial \pi_k} \big(t, L\xi, x + g(L\xi), \pi + \Lambda(L\xi) g_p(L\xi) \big) \right| \right. \\
&\qquad\qquad \left. \times [1 + (x + g(L\xi))] \right\} < \infty.
\end{aligned}
$$

That is, the function $\bar{a}_0$ has bounded first order partial derivatives in ξ, uniformly in (t, ξ, x, π). Note that the first order partial derivatives of $\bar{a}_0$ in x and π are the same as those of a^1 (with corresponding change of variables), and we conclude that $\bar{a}_0$ is uniformly Lipschitz in ξ, x and π.

To do a similar analysis for b_0 and $\hat{b}_0$, we note that for any $k = 1, \dots, d$

and (t, ξ, x, π), it holds by definition of $\bar{r}$ that

$$\frac{\partial \bar{r}}{\partial \xi_k} = \frac{\partial r}{\partial x}\frac{\partial g}{\partial p_k}e^{\xi_k} + \sum_{l=1}^{d}\frac{\partial r}{\partial \pi_l}\left[\frac{\partial^2 g}{\partial p_k\,\partial p_l}e^{\xi_l}e^{\xi_k} + \delta_{kl}e^{\xi_l}\frac{\partial g}{\partial p_l}\right],$$

$$\begin{aligned}\frac{\partial}{\partial \xi_k}&\langle \bar{r}(t, L\xi, x, \pi)\underline{\mathbf{1}}, \Lambda(L\xi)g_p(L\xi)\rangle\\ &= \left[\frac{\partial}{\partial \xi_k}\bar{r}(t, L\xi, x, \pi)\right]\sum_{l=1}^{d}e^{\xi_l}\frac{\partial g}{\partial p_l}(L\xi)\\ &\qquad + \bar{r}(t, L\xi, x, \pi)\sum_{l=1}^{d}\left[e^{\xi_l}e^{\xi_k}\frac{\partial^2 g}{\partial p_l\,\partial p_k} + \delta_{kl}e^{\xi_l}\frac{\partial g}{\partial p_l}\right],\end{aligned}$$

$$\begin{aligned}&\frac{\partial}{\partial \xi_k}[\bar{r}(t, L\xi, x, \pi)(x + g(L\xi))]\\ &\quad = \left[\frac{\partial}{\partial \xi_k}\bar{r}(t, L\xi, x, \pi)\right](x + g(L\xi)) + \bar{r}(t, L\xi, x, \pi)e^{\xi_k}\frac{\partial g}{\partial p_k}\end{aligned}$$

and that the function

$$\tfrac{1}{2}\mathrm{tr}\left\{\overline{\sigma\sigma}^{\mathsf{T}}(t, L\xi, x, \pi)g_{pp}(L\xi)\right\} = \tfrac{1}{2}\mathrm{tr}\left\{\bar{a}_0(t, \xi, x, \pi)[\Lambda(L\xi)]^2 g_{pp}(L\xi)\right\}$$

is uniformly bounded and Lipschitz in ξ, x and π by condition (A4)(b). Thus we obtain by a similar argument as before the uniform boundedness and Lipschitz property of the functions b_0 and $\hat{b}_0$ as well. Therefore, by Theorem 4.5 in [18] or by applying the results in [17] directly, we conclude that the PDE (3.18) has a unique classical solution $\tilde{\theta}$ in $C^{1+\alpha/2, 2+\alpha}$ [for any $\alpha \in (0, 1)$]. Furthermore, $\tilde{\theta}$, together with its first and second partial derivatives in ξ, is uniformly bounded throughout $[0, T] \times \mathbb{R}^d$. If we go back to the original variable and note the relations in (3.20), we obtain that the function variables $\hat{\theta}$ is uniformly bounded and its partial derivatives satisfy

$$\sup_{(t,p)}\left|\Lambda(p)\hat{\theta}_p(t, p)\right| < \infty, \qquad \sup_{(t,p)}\left\|\Lambda(p)\hat{\theta}_{pp}(t, p)\right\| < \infty.$$

This, together with the definition of $\hat{\theta}$ and condition (A4)(b), leads to the estimates (3.15); the assertion, and hence Step 2, is proved.

As for Step 3, we note that θ_p and θ_{pp} themselves are unbounded and will blow up when $p \downarrow 0$. Consequently, $\tilde{b}$ and $\tilde{\sigma}$ are only locally Lipschitz and not even defined at $p = 0$. Therefore a little bit more careful consideration is needed here, and we proceed as follows. First we observe that the local Lipschitz property of $\tilde{b}$ and $\tilde{\sigma}$ is enough for us to show the existence and uniqueness of the "local solution" of (3.10) inside $\mathbb{R}^d_+$ (with the possibility of *explosion*). However, by a similar argument as that in Lemma 2.3, one can "linearize" (3.10) and use Assumption (A1) to show that whenever the solution exists, it will neither leave $\partial\mathbb{R}^d_+$ nor explode before T. (In fact the solution is a square-integrable process.) Hence Step 3 is complete. Finally, Step 4 is trivial, and noting that the square integrability is the direct consequence of definition (3.11), property (3.15) and the square integrability of the process $P(\cdot)$, the existence is proved.

The uniqueness of the adapted solution will be proved in Corollary 4.2. The proof of the theorem is therefore complete. $\square$

We note from the proof of Theorem 3.4 that Condition (A4)(b) is only used to guarantee the boundedness of the partial derivatives of $\bar{a}_0$, which is unnecessary if g is bounded, because in such a case the intermediate solution $\hat{\theta}$ is not needed. For the same reason, condition (A3) can be relaxed to the following.

(A3′) The function r satisfies all the conditions of (A3) except that (2.11) is replaced by

$$\limsup_{|x|, |\pi| \to \infty} (|x| + |\pi|) |\psi(t, x, \pi)| < \infty. \tag{3.24}$$

Furthermore, the estimates (3.15) of the solution to the PDE (3.8) can be improved to

$$\sup_{(t,p)} |\Lambda(p)\theta(t,p)| < \infty; \qquad \sup_{(t,p)} \left\| \Lambda(p)^2 \theta_{pp}(t,p) \right\| < \infty. \tag{3.25}$$

In other words, we have the following corollary.

COROLLARY 3.5. *Under Assumptions* (A1), (A2), (A3′) *and* (A4)(a), *Theorem* 3.4 *remains valid*. *Furthermore*, *the classical solution* θ *of* (3.8) *satisfies the estimate* (3.25).

Discussion. A seemingly simpler way of proving the existence and uniqueness of FBSDE (3.3) can be carried out in the following way if some even stronger conditions are satisfied by the coefficients b and σ. We sketch the idea here, because it might be useful for some other applications.

Let us suppose that (A1)–(A3) hold and suppose that $d = 1$ and the function g is bounded and C^2 for simplicity. Let us also assume that the functions b and σ and their partial derivatives in x and π, denoted by a generic φ, satisfy the conditions

$$\limsup_{|\pi| \to \infty} \left| \frac{\pi}{p} \varphi(t, p, x, \pi) \right| < \infty \quad \text{uniformly in } (t, p, x), \tag{A5}$$

and the partial derivatives of b and σ in p, denoted by a generic ψ, satisfy

$$\limsup_{|\pi| \to \infty} |\pi \psi(t, p, x, \pi)| < \infty. \tag{A6}$$

(These conditions are obviously satisfied in the classical wealth–policy independent models.) By Lemma 2.3, we know that any solution to (3.3) must satisfy $P(t) > 0$ for all $t \in [0, T]$, a.s. **P**. Therefore, we can define a process $\xi(t) = \log P(t)$ for $t \in [0, T]$. A simple computation using Itô's formula will show that (ξ, X, π) satisfies the FBSDE

$$
\begin{aligned}
\xi(t) &= \xi(0) + \int_0^t b_1(s, \xi(s), X(s), \pi(s))\, ds \\
&\qquad + \int_0^t \sigma_1(s, \xi(s), X(s), \pi(s))\, dW(s), \\
X(t) &= g(\xi(T)) + \int_t^T \hat{b}_1(s, \xi(s), X(t), \pi(s))\, ds \\
&\qquad + \int_t^T \hat{\sigma}_1(s, \xi(s), X(s), \pi(s))\, dW(s),
\end{aligned}
\tag{3.26}
$$

where

$$
\begin{aligned}
b_1(t, \xi, x, \pi) &= \frac{b(t, e^{\xi}, x, \pi)}{e^{\xi}} - \frac{1}{2}\frac{\sigma^2(t, e^{\xi}, x, \pi)}{e^{2\xi}}, \\
\sigma_1(t, \xi, x, \pi) &= \frac{\sigma(t, e^{\xi}, x, \pi)}{e^{\xi}}, \\
\hat{b}(t, \xi, x, \pi) &= \hat{b}(t, e^{\xi}, x, \pi), \\
\hat{\sigma}_1(t, \xi, x, \pi) &= \hat{\sigma}(t, e^{\xi}, x, \pi).
\end{aligned}
$$

Obviously, the existence and uniqueness of the adapted solution to the FBSDE (3.3) is equivalent to that of (3.26). Since $\xi(\cdot) = \log P(\cdot)$, we will call (3.3) the "price–wealth equation" and (3.26) the "log-price–wealth equation." One can therefore study either one of them, whichever is easier.

It is fairly easy to check that under conditions (A1)–(A3), together with (A5) and (A6), the functions b_1 and σ_1 are bounded with bounded first order partial derivatives in p, x and π, while the functions $\hat{b}_1$ and $\hat{\sigma}_1$ are of linear growth in x and π, with bounded first order partial derivatives. Therefore by Theorem 4.5 in [17], the FBSDE (3.26) has a unique adapted solution; therefore so does (3.3) (with $d = 1$).

Comparing the above result to Theorem 3.4 or Corollary 3.5, we see that applying the result of [18] directly, say, to the log-price–wealth equation is not always the easiest way, because one would have to use conditions (A5) and (A6), which is obviously unnecessary, as we see in the theorem and its corollary. One of the main reasons for this to happen is that in our setting, the coefficients $\hat{b}$ and $\hat{\sigma}$ are explicitly related to b, σ and r, and the corresponding quasilinear parabolic PDE is simplified drastically so that the existence and uniqueness of classical solutions can be proved with fewer restrictions on the coefficients. Such a phenomenon can also be partially explained from a finance point of view. In fact, by the standard financial mathematics tool of the change of probability measure, we can in a sense replace the appreciation rate b_1 by the interest rate r, as in the classical case, which may facilitate some analysis. Such an equivalent probability measure corresponds to the "classical" *equivalent martingale measure*. However, we should point out here that this equivalent measure is now both wealth and policy dependent; therefore one cannot use it to simplify the FBSDE a priori to derive the adapted solution if the Brownian motion is not allowed to change.

4. Comparison theorems and main results. In this section we shall prove two comparison theorems. As corollaries of the first comparison theorem, we prove the uniqueness part of Theorem 3.4 and that the unique adapted solution to the FBSDE (3.3) is an optimal strategy among all admissible ones. The second comparison theorem shows that the optimal strategy is monotone in the option value at the expiration date; namely, the higher the value of the option, the higher the premium the buyer has to pay. We should note that due to the special feature of the FBSDE, these comparison theorems are much weaker than those of the pure backward case. Therefore, some new phenomena that are different from those in classical theory might be worth studying.

Let us first note that, in the forward–backward case, one *cannot* easily jump to a conclusion like $X_1(T) = g_1(P_1(T)) \geq g_2(P_2(T)) = X_2(T)$ from the assumption that $g_1(p) \geq g_2(p)$, $\forall$ p, because in the present situation, the price process is policy–wealth dependent and $P_1(T)$ and $P_2(T)$ are different if g_1 and g_2 are so. Thus, unlike the classical (pure backward) case, *no simple comparison can be made* for the processes X_1 and X_2, except for $t = 0$ (see Theorem 4.4 and the counterexample in the Appendix). We nonetheless have the following results, which will be enough for our purpose. Note that in what follows when we say "condition (A4) holds," we mean either (A4)(a) or (A4)(b) holds.

THEOREM 4.1 (Comparison theorem). *Suppose that* (A1)–(A4)(a) *or* (b) *hold. Let initial prices* $p \in \mathbb{R}_+^d$ *be given and let* (π, C) *be any admissible pair such that the corresponding price–wealth process* (P, Y) *satisfies* $Y(T) \geq g(P(T))$, *a.s. Then* $Y(\cdot) \geq \theta(\cdot, P(\cdot))$, *where* θ *is the solution to* (3.8). *In particular*, $Y(0) \geq \theta(0, p) = X(0)$, *where* X *is the solution to the FBSDE* (3.3) *starting from* $p \in \mathbb{R}_+^d$ *constructed by the four step scheme.*

PROOF. We only consider the case when condition (A4)(b) holds, since the other case, when (A4)(a) holds and g is bounded, is much easier and can be proved in a similar way. The method we use is similar in spirit to the method of linearizing the backward equation used in [8], except we have to be more careful in order to find "generators" which will be Lipschitz. Let (P, Y, π, C) be given such that $(\pi, C) \in \mathscr{A}(Y(0))$ and $Y(T) \geq g(P(T))$, a.s. We first define a change of probability measure as follows: let

$$\begin{aligned}
&\theta_0(t) = \big(\sigma^1(t,P(t),Y(t),\pi(t))\big)^{-1} \\
&\qquad\qquad \times\big[b^1(t,P(t),Y(t),\pi(t)) - r(t,X(t),\pi(t))\underline{\mathbf{1}}\big],
\end{aligned}$$

$$Z_0(t) = \exp\left\{-\int_0^t \theta_0(s)\,dW(s) - \frac{1}{2}\int_0^t |\theta_0(s)|^2\,ds\right\}, \tag{4.1}$$

$$\frac{d\mathbf{P}_0}{d\mathbf{P}} = Z_0(T),$$

so that the process $W_0(t) \triangleq W(t) + \int_0^t \theta_0(s)\,ds$ is a Brownian motion on the new probability space $(\Omega,\mathscr{F},\hat{\mathbf{P}}_0)$. Furthermore, the price–wealth FBSDE becomes [note the definitions of $\hat{b}$, $\hat{\sigma}$ and σ_1 in (2.7) and (2.8)]

$$\begin{aligned}
P(t) &= p + \int_0^t P(s) r(s,Y(s),\pi(s))\,ds \\
&\qquad + \int_0^t \sigma(s,P(s),Y(s),\pi(s))\,dW_0(s),
\end{aligned}$$

$$\begin{aligned}
Y(t) &= g(P(T)) - \int_t^T r(s,Y(t),\pi(s))Y(s)\,ds \\
&\qquad - \int_t^T \langle \pi(s), \sigma_1(s,P(s),Y(s),\pi(s))\,dW_0(s)\rangle \\
&\qquad + C(T) - C(t).
\end{aligned} \tag{4.2}$$

Note that in the present case the PDE (3.8) is degenerate, and the function g is not bounded, so the solution θ to (3.8) and its partial derivatives will blow up as p approaches $\partial\mathbb{R}_+^d$ and infinity. Therefore the usual estimates such as those in [18] will not work, and some more careful consideration will be needed. To overcome this difficulty, we proceed as follows. First we apply Itô's formula to the process $g(P(\cdot))$ from t to T to get

$$\begin{aligned}
g(P(t)) &= g(P(T)) - \int_t^T \Big\{\big\langle g_p(P(s)), r(s,Y(s),\pi(s))P(s)\big\rangle \\
&\qquad\qquad - \tfrac{1}{2}\mathrm{tr}\big\{\sigma_1\sigma_1^{\mathsf{T}}(s,P(s),Y(s),\pi(s))g_{pp}(P(s))\big\}\Big\}\,ds \\
&\qquad - \int_t^T \langle g_p(P(s)), \sigma(s,P(s),Y(s),\pi(s))\,dW_0(s)\rangle.
\end{aligned} \tag{4.3}$$

Then we define a process $\hat{Y}(t) = Y(t) - g(P(t))$, which obviously satisfies the (backward) SDE

$$\begin{aligned}
\hat{Y}(t) &= \hat{Y}(t) - \int_t^T \Big\{r(s,Y(s),\pi(s))\big[Y(s) - \langle g_p(P(s)), P(s)\rangle\big] \\
&\qquad\qquad - \tfrac{1}{2}\mathrm{tr}\big\{\sigma\sigma^{\mathsf{T}}(s,P(s),Y(s),\pi(s))g_{pp}(P(s))\big\}\Big\}\,ds \\
&\quad - \int_t^T \langle \pi(s) - \Lambda(P(s))g_p(P(s)), \sigma_1(s,P(s),Y(s),\pi(s))\,dW_0(s)\rangle \\
&\quad + C(T) - C(t).
\end{aligned} \tag{4.4}$$

We now use the notation $\hat{\theta} = \theta - g$ as that in the proof of Theorem 3.4. Then it suffices to show that $\hat{Y}(t) \geq \hat{\theta}(t, P(t))$ for all $t \in [0, T]$, a.s. $\mathbf{P}_0$. To this end, let us denote $\tilde{Y}(t) = \hat{\theta}(t, P(t))$, $\tilde{\pi}(t) = \Lambda(P(t))[\hat{\theta}_p(t, P(t)) + g_p(P(t))]$ and $\Delta_Y(t) = \hat{Y}(t) - \tilde{Y}(t)$, $\Delta_\pi(t) = \pi(t) - \tilde{\pi}(t)$. Applying Itô's formula to the process $\Delta_Y(t)$, we obtain

$$
\begin{aligned}
\Delta_Y(t) = \hat{Y}(T) - \int_t^T \Big\{ & r(s, Y(s), \pi(s)) \\
& \times \Big[Y(s) - \Big\langle g_p(P(s)) + \hat{\theta}_p(s, P(s)), P(s) \Big\rangle \Big] \\
& - \hat{\theta}_s(s, P(s)) - \tfrac{1}{2} \operatorname{tr}\Big\{ \sigma\sigma^\top(s, P(s), Y(s), \pi(s)) \\
& \qquad \times \Big[\hat{\theta}_{pp}(s, P(s)) + g_{pp}(P(s)) \Big] \Big\} \Big\} \, ds \\
- \int_t^T \Big\langle & \pi(s) - \Lambda(P(s)) \big[g_p(P(s)) + \hat{\theta}_p(s, P(s)) \big], \\
& \sigma_1(s, P(s), Y(s), \pi(s)) \, dW_0(s) \Big\rangle \\
= \hat{Y}(T) - \int_t^T & A(s) \, ds - \int_t^T \langle \Delta_\pi(s), \sigma_1(s, P(s), Y(s), \pi(s)) \, dW_0(s) \rangle \\
+ C(T) - C(t)&,
\end{aligned} \tag{4.5}
$$

where the process $A(\cdot)$ in the last term above is defined in the obvious way. We now recall that the function $\hat{\theta}$ satisfies PDE (3.16), recall the definition of $\tilde{\pi}$ and also note that $\langle p, q \rangle = \langle \Lambda(p)q, \underline{\mathbf{1}} \rangle$ and that $\hat{\theta}(t, P(t)) + g(P(t)) = Y(t) - \Delta_Y(t)$. We can easily rewrite $A(\cdot)$ as

$$
\begin{aligned}
A(s) &= r(s, Y(s), \pi(s))Y(s) - r(s, Y(s) - \Delta_Y(s), \tilde{\pi}(s))[Y(s) - \Delta_Y(s)] \\
&\quad - \Big\langle r(s, Y(s), \pi(s))\pi(s) - r\big(s, \hat{\theta}(s, P(s)), \tilde{\pi}(s)\big)\tilde{\pi}(s), \underline{\mathbf{1}} \Big\rangle \\
&\quad + \tfrac{1}{2} \operatorname{tr}\Big\{ \Delta\sigma_1\big(s, P(s), Y(s), \pi(s), \tilde{\pi}(s), \hat{\theta}(s, P(s))\big) \Theta(s, P(s)) \Big\} \\
&= I_1(s) + I_2(s) + I_3(s),
\end{aligned} \tag{4.6}
$$

where

$$
\begin{aligned}
\Theta(t, p) &\triangleq [\Lambda(p)]^2 \big(\hat{\theta}_{pp}(t, p) + g_{pp}(p) \big), \\
\Delta\sigma_1(t, p, x, \pi, \tilde{\pi}, q) &\triangleq \sigma_1\sigma_1^\top(t, p, q + g(p), \tilde{\pi}) - \sigma_1\sigma_1^\top(t, p, x, \pi),
\end{aligned}
$$

and I_i's are defined in the obvious way. Now noticing that

$$
\begin{aligned}
I_1(s) = {} & \big[r(s, Y(s), \pi(s))Y(s) - r(s, Y(s) - \Delta_Y(s), \pi(s))(Y(s) - \Delta_Y(s)) \big] \\
& + \big[r(s, Y(s) - \Delta_Y(s), \pi(s)) - r(s, Y(s) - \Delta_Y(s), \tilde{\pi}(s)) \big]
\end{aligned}
$$

$$
\begin{aligned}
&\times[Y(s)-\Delta_Y(s)]\\
&=\left\{\int_0^1 \frac{\partial}{\partial x}\{r(s,x,\pi(s))x\}\Big|_{x=(Y(s)-\lambda\Delta_Y(s))}\,d\lambda\right\}\Delta_Y(s)\\
&\quad+\left\langle\int_0^1 \frac{\partial r}{\partial \pi}(s,Y(s)-\Delta_Y(s),\pi(s)+\lambda\Delta_\pi(s))\right.\\
&\qquad\qquad\left.\times[Y(s)-\Delta_Y(s)]\,d\lambda,\Delta_\pi(s)\right\rangle\\
&=\alpha_1(s)\Delta_Y(s)+\langle\beta_1(s),\Delta_\pi(s)\rangle,
\end{aligned}
$$

we have from condition (A3) that both α_1 and β_1 are adapted processes and are uniformly bounded in (t, ω). Similarly, by conditions (A1), (A3) and (A4)(b), we see that the process $\Theta(\cdot, P(\cdot))$ is uniformly bounded and that there exist uniformly bounded, adapted processes α_2, α_3 and β_2, β_3 such that

$$
\begin{aligned}
I_2(s)&=\left\langle r(s,Y(s),\pi(s))\pi(s)-r\big(s,\hat\theta(s,P(s)),\pi(s)\big)\pi(s),\underline{\mathbf{1}}\right\rangle\\
&\quad+\left\langle r\big(s,\hat\theta(s,P(s)),\pi(s)\big)\pi(s)-r\big(s,\hat\theta(s,P(s)),\tilde\pi(s)\big)\tilde\pi(s),\underline{\mathbf{1}}\right\rangle\\
&=\alpha_2(s)\Delta_Y(s)+\langle\beta_2(s),\Delta_\pi(s)\rangle,\\
I_3(s)&=\alpha_3(s)\Delta_Y(s)+\langle\beta_3(s),\Delta_\pi(s)\rangle.
\end{aligned}
$$

Therefore, letting $\alpha=\sum_{i=1}^3\alpha_i$, $\beta=\sum_{i=1}^3\beta_i$, we obtain that

$$A(t)=\alpha(t)\Delta_Y(t)+\langle\beta(t),\Delta_\pi(t)\rangle,$$

where α and β are both adapted, uniformly bounded processes. In other words, we have

$$
\begin{aligned}
\Delta_Y(t)&=\hat Y(T)-\int_t^T\{\alpha(s)\Delta_Y(s)+\langle\beta(s),\Delta_\pi(s)\rangle\}\,ds\\
&\qquad-\int_t^T\langle\Delta_\pi(s),\sigma_1(s,P(s),Y(s),\pi(s))\,dW_0(s)\rangle\\
&\qquad+C(T)-C(t).
\end{aligned}
\tag{4.7}
$$

The remainder of the proof is similar to that in the comparison theorem for backward SDE's (see [8]). We include it for the sake of completeness. Define another change of probability measure similar to (4.1) by setting

$$
\begin{aligned}
\theta_1(t)&=\sigma_1^{-1}(t,P(t),Y(t),\pi(t))\beta(t),\\
Z_1(t)&=\exp\left\{-\int_0^t\theta_1(s)\,dW_0(s)-\tfrac12\int_0^t|\theta_1(s)|^2\,ds\right\},\\
\frac{d\mathbf{P}_1}{d\mathbf{P}_0}&=Z_1(T).
\end{aligned}
$$

Then it is clear that $Z_1(\cdot)$ is a $\mathbf{P}_0$-martingale by virtue of the boundedness of β. Hence $W_1(t) \triangleq W_0(t) + \int_0^t \theta_1(s)\,ds$ is a $\mathbf{P}_1$-Brownian motion. Moreover, we have

$$\begin{aligned} &\exp\left(-\int_0^T \alpha(s)\,ds\right)\Delta_Y(T) - \exp\left(-\int_0^t \alpha(s)\,ds\right)\Delta_Y(t) \\ (4.8)\quad &= -\int_t^T \exp\left(-\int_0^s \alpha(u)\,du\right)\left\langle \Delta_\pi(s), \sigma_1(s, P(s), Y(s), \pi(s))\,dW_1(s)\right\rangle \\ &\quad -\int_t^T \exp\left(-\int_0^s \alpha(u)\,du\right) dC(s). \end{aligned}$$

By definition of $\tilde{\pi}$ and the estimate (3.15), we see that $\tilde{\pi}$ is obviously square integrable (note also that π is square integrable by its admissibility). Recalling that α is bounded, we see that the first term in the right-hand side of (4.8) is a martingale. Taking conditional expectations on both sides above, we obtain that

$$\begin{aligned} \exp\left(-\int_0^t \alpha(s)\,ds\right)\Delta(t) &= E\left\{\exp\left(-\int_0^T \alpha(s)\,ds\right)\Delta(T)\right. \\ (4.9)\quad &\qquad \left. + \int_t^T \exp\left(-\int_0^s \alpha(u)\,du\right) dC(s) \,\middle|\, \mathscr{F}_t\right\} \end{aligned}$$

for all $t \in [0, T]$, a.s. $\mathbf{P}_1$. Thus $\Delta_Y(T) = Y(T) - g(P(T)) \geq 0$ implies that $\Delta_Y(t) \geq 0$, $\forall\ t \in [0, T]$, a.s. $\mathbf{P}_1$, whence a.s. $\mathbf{P}_0$ and finally a.s. $\mathbf{P}$. The proof is complete. $\square$

The main results of this paper are then contained in the following corollaries.

COROLLARY 4.2 (Uniqueness of FBSDE). *Suppose that conditions* (A1)–(A4)(a) *or* (b) *hold. Let* (P, Y, π) *be an adapted solution to the FBSDE* (3.3). *Then it must be the same as the one constructed from the four step scheme. In other words, the FBSDE* (3.3) *has a unique adapted solution and it can be constructed via* (3.10) *and* (3.11).

PROOF. Note that in this case $C \equiv 0$ and $\Delta(T) = Y(T) - g(P(T)) = 0$. The assertion follows from (4.9). $\square$

COROLLARY 4.3 (The optimal strategy). *Under Assumptions* (A1)–(A4)(a) *or* (b), *we have* $h(g(P(T))) = X(0)$, *where* P, X *are the first two components of the adapted solution to the FBSDE* (3.3). *Furthermore, the optimal hedging strategy is given by* $(\pi, 0)$, *where* π *is the third component of the adapted solution of FBSDE* (3.3).

PROOF. This is a direct consequence of Theorem 4.1 and Corollary 4.2. $\square$

We note that the uniqueness does not hold if we allow nonzero consumption in FBSDE (3.3). An intuitively obvious example is that one can first find a portfolio strategy to hedge $g(P(T))$ by solving (3.3) and then find another portfolio strategy to hedge $g(P(T)) + k$ [by solving (3.3) with $g(P(T))$ being replaced by $g(P(T)) + k$, $k > 0$] and consume k dollars at time $t = T$.

As mentioned before, we cannot always use the methods for comparison of backward SDE's in our case. In particular, the following comparison theorem is proved using comparison results for deterministic PDE's.

THEOREM 4.4 (Monotonicity in terminal condition). *Suppose that* (A1)–(A4)(a) *or* (b) *hold, and let* (P^i, X^i, π^i), $i = 1, 2$, *be the unique adapted solutions to* (3.3), *with the same initial prices* $p \in \mathbb{R}^d_+$ *but different terminal conditions* $X^i(T) = g^i(P^i(T))$, $i = 1, 2$, *respectively. If* g^1, g^2 *all satisfy the condition* (A4) *and* $g^1(p) \geq g^2(p)$ *for all* $p \in \mathbb{R}^d_+$, *then it holds that* $X^1(0) \geq X^2(0)$.

PROOF. By the existence and uniqueness of the adapted solution to the FBSDE (3.3), we know that X^1 and X^2 must have the form

$$X^1(t) = \theta^1(t, P^1(t)), \qquad X^2(t) = \theta^2(t, P^2(t)),$$

where θ^1 and θ^2 are the classical solutions to the PDE (3.8) with terminal conditions g^1 and g^2, respectively. We claim that the inequality $\theta^1(t, p) \geq \theta^2(t, p)$ must hold for all $(t, p) \in [0, T] \times \mathbb{R}^d_+$.

To see this, let us make the change of variable $p = L\xi$ again, where the mapping L is defined in the proof of Theorem 3.4, and define $u^i(t, \xi) = \theta^i(T - t, L\xi)$. It follows from the proof of Theorem 3.4 that u^1 and u^2 satisfy the PDE

$$\begin{aligned} 0 &= u_t - \tfrac{1}{2}\mathrm{tr}\left\{\bar{\sigma}_1\bar{\sigma}_1^{\mathsf{T}}(t, \xi, u, u_\xi)u_{\xi\xi}\right\} \\ &\quad - \left\langle b_0(t, \xi, u, u_\xi), u_\xi\right\rangle + u\bar{r}(t, u, u_\xi), \\ u(0, \xi) &= g^i(e^\xi), \qquad \xi \in \mathbb{R}^d, \end{aligned} \tag{4.10}$$

respectively, where

$$\begin{aligned} \bar{\sigma}_1(t, \xi, x, \pi) &= [\Lambda(L\xi)]^{-1}\sigma(T - t, L\xi, x, \pi), \\ b_0(t, \xi, x, \pi) &= r(T - t, x, \pi)\underline{\mathbf{1}} - \tfrac{1}{2}\mathrm{diag}\left[\bar{\sigma}_1\bar{\sigma}_1^{\mathsf{T}}(T - t, \xi, x, \pi)\right], \\ \bar{r}(t, x, \pi) &= r(T - t, x, \pi). \end{aligned}$$

Note that by a standard argument (cf. [17]) one can show that if we denote $u^i_R(\cdot, \cdot)$ to be the solution to the initial boundary value problem

$$\begin{aligned} 0 &= u_t - \tfrac{1}{2}\mathrm{tr}\left\{\bar{\sigma}_1\bar{\sigma}_1^{\mathsf{T}}(t, \xi, u, u_\xi)u_{\xi\xi}\right\} \\ &\quad - \left\langle b_0(t, \xi, u, u_\xi), u_\xi\right\rangle + ur(t, u, u_\xi), \\ u|_{\partial B_R}(t, \xi) &= g^i(e^\xi), \qquad |\xi| = R, \\ u(0, \xi) &= g^i(e^\xi), \qquad \xi \in B_R, \end{aligned} \tag{4.11}$$

$i = 1, 2$, respectively, where $B_R \triangleq \{\xi;\ |\xi| \le R\}$, then $u_R^i \to u^i$ uniformly on compacts, $i = 1, 2$, as $R \to \infty$. Therefore, we need only show that $u_R^1(t, \xi) \ge u_R^2(t, \xi)$ for all $(t, \xi) \in [0, T] \times B_R$ and $R > 0$.

For any $\varepsilon > 0$, consider the PDE,

$$\begin{aligned} u_t &= \tfrac{1}{2}\operatorname{tr}\left\{\bar{\sigma}_1\bar{\sigma}_1^{\mathsf{T}}(t, \xi, u, u_\xi)u_{\xi\xi}\right\} \\ &\qquad + \left\langle b_0(t, \xi, u, u_\xi), u_\xi \right\rangle - ur(t, u, u_\xi) + \varepsilon, \\ u|_{\partial B_R}(t, \xi) &= g^1(e^\xi) + \varepsilon, \qquad |\xi| = R, \\ u(0, \xi) &= g^1(e^\xi) + \varepsilon, \qquad \xi \in B_R, \end{aligned} \tag{4.11$'$}$$

and denote its solution by $u_{R,\varepsilon}^1$. It is not hard to check, using a standard technique (cf., e.g., [9]), that $u_{R,\varepsilon}^1$ converges to u_R^1, uniformly in $[0, T] \times \mathbb{R}^d$. Next, we define a function

$$F(t, \xi, x, q, \hat{q}) = \tfrac{1}{2}\operatorname{tr}\bar{\sigma}_1\bar{\sigma}_1^{\mathsf{T}}(t, \xi, x, q)\hat{q} + \left\langle b_0(t, \xi, q, \hat{q}), \hat{q}\right\rangle - x\bar{r}(t, x, q).$$

Clearly F is continuously differentiable in all variables, and $u_{R,\varepsilon}^1$ and u_R^2 satisfy

$$\begin{aligned} \frac{\partial u_{R,\varepsilon}^1}{\partial t} &> F\left(t, \xi, u_{R,\varepsilon}^1, (u_{R,\varepsilon}^1)_\xi, (u_{R,\varepsilon}^1)_{\xi\xi}\right), \\ \frac{\partial u_R^2}{\partial t} &= R\left(t, \xi, u_R^2, (u_R^2)_\xi, (u_R^2)_{\xi\xi}\right), \\ u_{R,\varepsilon}^1(t, \xi) &> u_R^2(t, \xi), \qquad (t, \xi) \in [0, T] \times B_T \cup \{0\} \times \partial B_R. \end{aligned}$$

Therefore by Theorem II.16 of [9], we have $u_{R,\varepsilon}^1 > u_R^2$ in B_R. By letting $\varepsilon \to 0$ and then $R \to \infty$, we obtain that $u^1(t, \xi) \ge u^2(t, \xi)$ for all $(t, \xi) \in [0, T] \times \mathbb{R}^d$, whence $\theta^1(\cdot, \cdot) \ge \theta^2(\cdot, \cdot)$. In particular, we have $X^1(0) = \theta^1(0, p) \ge \theta^2(0, p) = X^2(0)$, proving the theorem. $\square$

REMARK 4.5. From $\theta^1(t, p) \ge \theta^2(t, p)$ we cannot conclude that $X^1(t) \ge X^2(t)$ for all t, since in general there is no comparison between $\theta^1(t, P^1(t))$ and $\theta^2(t, P^2(t))$ (see the example in the Appendix).

5. Examples. In this section we provide some examples which motivate our model. We note that in some of the examples the standing assumptions of this paper are not actually satisfied, but because of the special features of these models, we can verify, by using some existing results, that our method will also derive the right answer to these problems. Therefore it would be interesting to develop our methodology further, to a wider class of problems with more general coefficients, although, technically, it will be much more challenging.

EXAMPLE 5.1 (Different interest rates for borrowing and lending). Suppose we want to model a market in which there are different interest rates for borrowing R and lending r, $0 < r < R$. In (2.2), we let

$$b(t, P(t), X(t), \pi(t)) = bP(t), \qquad \sigma(t, P(t), X(t)) = \sigma P(t)$$

for some real b and $\sigma > 0$. Recall that π is the amount that the investor puts in stock. In (2.1), let

$$r(t, X(t), \pi(t)) = r\mathbf{1}_{\{\pi(t) \leq X(t)\}} + R\mathbf{1}_{\{\pi(t) > X(t)\}}. \tag{5.1}$$

Also let $g(p) = (p - q)^+$; that is, we want to hedge a European call option with exercise price $q > 0$. We guess now (and verify later) that the hedging portfolio will always borrow (as is well known for a European call in the standard model) and, therefore, we shall work with $r(t, X(t), \pi(t)) = R$. Then (3.8) becomes

$$0 = \theta_t + \tfrac{1}{2}\theta_{pp}\sigma^2 p^2 + R(p\theta_p - \theta),$$

with $\theta(T, p) = (p - q)^+$. This is nothing else but a Black–Scholes PDE for the price of a European option on a stock with volatility σ and with the market's riskless interest rate equal to R. It has an explicit solution (see, e.g., [4]), which is easily seen to satisfy $p\theta_p > \theta$ and, therefore, θ also satisfies (3.8) with the function r given by (5.1). Now, by (3.11), we have $X(t) = \theta(t, P(t))$ and $\pi(t) = P(t)\theta_p(t, P(t))$, $C(t) \equiv 0$. We recover therefore the results of Example 9.5 of [4].

REMARK 5.2. In the previous example the smoothness assumptions in (A3) and (A4) on r and g are, in fact, not satisfied. Nevertheless, the results hold [again, see [4] for rigorous proofs which do not require assumptions like (A3), (A4), but require some concavity assumptions which are satisfied in the example]. Of course, regarding assumption (A3), we were "lucky," since the discontinuity in the function r disappears for the solution θ of (3.8). This also makes the PDE linear and allows weakening of the assumptions. The message is that one can try to construct a hedging strategy using the procedure of this paper, even if not all of the assumptions are satisfied, and see what happens.

REMARK 5.3 (Independence of drift). The PDE (3.8) does not depend on the drift function b, and neither does the price $X(0)$. This is a familiar fact from the standard Black–Scholes world, valid even in this general model, where the drift can be influenced by the portfolio strategy. Therefore, it is of no interest to look at examples in which b takes different forms.

EXAMPLE 5.4 (Large investor). We indicate here one of the possible models in our framework, not included in the standard theory. Suppose that our investor is really an important one, so that, if he/she invests too much in bonds, the government (or the market) decides to decrease the bond interest rate. For example, we can assume that $r(t, x, \pi)$ is a decreasing function of $x - \pi$, for $x - \pi$ large.

EXAMPLE 5.5 (Several agents—equilibrium model). In Platen and Schweizer [23] an SDE for the stock price is obtained from equilibrium considerations; both its drift and volatility coefficients depend on the hedging strategy of the agents in the market in a rather complex fashion. As the authors mention, "It is not clear at all how one should compute option prices in an economy where agents' strategies affect the underlying stock price process." Our results provide the price that would enable the seller to hedge against all the risk, that is, the upper bound for the price.

EXAMPLE 5.6 (Cheating does not always pay for the large investor). Suppose that, up to time $t = 0$, we have the standard Black–Scholes model; that is, the interest rate r is constant and the volatility function is given by $\sigma(t, p, x) = \sigma p$, $\sigma > 0$ (the drift function does not matter, as noted in Remark 5.3). Then, at time $t = 0$, the large investor sells the option worth $g(P(T))$ at time $t = T$, for the price of $\gamma(0, P(0))$, where γ is a pricing function. Typically, $\gamma(\cdot, P(\cdot)) \geq \rho(\cdot, P(\cdot))$, where ρ is the Black–Scholes pricing function, given by the solution to (3.8), with r and σ as above. Having a strict inequality means that the large investor is trying to "cheat," that is, to sell the option for more than its worth, the Black–Scholes price. Suppose that the investor, let us call him/her the seller, finds buyers for the option at this price. This would create instabilities in the market and arbitrage opportunities if the volatility of the stock were to remain the same. Let us assume that the effect is felt as a change in the volatility, so that the function $\sigma(t, p, x)$ is no longer equal to σp. A natural example would be $\sigma(t, p, x) = p[\sigma + f(\gamma(t, p) - \rho(t, p))]$, with f increasing. Also assume that $\sigma(t, p, x)$ remains equal to σp if $\gamma(\cdot, \cdot) = \rho(\cdot, \cdot)$; that is, if there is no cheating attempted at any time t, we remain in the Black–Scholes world and therefore hedging is possible if one starts with initial capital equal to the Black–Scholes price. If, on the contrary, the option sells for more, then the volatility increases and the minimal hedging price changes. For example, for a European call option the initial value of the solution to the PDE (3.8) increases with σ (at least for constant σ). Therefore, the "cheating" price (i.e., some price greater than the Black–Scholes price) might be smaller than the minimal hedging price in the market with the increased stock volatility, and hedging might not be possible, in which case cheating does not pay. On the other hand, there are cases for which hedging is possible and cheating would pay. From the point of view of the market, this model can indicate how the volatility has to change if the option is overpriced, in order for the seller not to make an arbitrage profit. For example, in a simple model in which $\gamma(\cdot, \cdot) = \rho(\cdot, \cdot) + \varepsilon$ and $\sigma(t, p, x) = p(\sigma + \delta_\varepsilon)$, it is easy to calculate (for standard European call options) what δ_ε would have to be in order that there be no arbitrage profit or, equivalently, in order to have $\gamma(0, P(0))$ equal to the Black–Scholes price of a stock with volatility $\sigma + \delta_\varepsilon$. If δ_ε is less than the critical value, cheating pays, and if it is larger, cheating does not pay. Here we have a phenomenon unknown in the classical models: *selling the option for its fair, Black–Scholes price ensures that hedging is possible, but selling for more than that price does not guarantee hedging.*

In general, if the seller of the option has an idea of how much cheating will affect the volatility of the stock then he/she should also have an idea of how much he/she can safely cheat, by solving the PDE (3.8) for different σ's (assuming, of course, that there are always buyers willing to buy the option).

Another example would be the case with $\sigma(t, p, x) = p[\sigma + f(x - \rho(t, p))]$, where, again, f is increasing and $f(z) = \sigma p$ for $z \le 0$. Assuming that the seller will always reinvest the profits rather than consume, $X(t) - \rho(t, P(t))$ could be thought of as a measure of arbitrage profit at time t. It is clear that the Black–Scholes pricing function ρ is a solution to (3.8), even with this modified volatility function $\sigma(t, p, x)$. Therefore, if $\sigma(t, p, x)$ is such that assumptions (A1)–(A4) are satisfied, then the Black–Scholes price $\rho(0)$ is the smallest price that still guarantees successful hedging (by Corollary 4.3). However, it is not clear that hedging is guaranteed if the investor sells the option for more than the minimal hedging price $\rho(0)$; since there is no consumption, we will have $X(t) > \rho(t, P(t))$ for small t, so that the volatility possibly increases. If the increase is significant compared to the price of the option, there might be no hedging portfolio. For example, one could have $\sigma(t, p, x) = p[\sigma + \arctan\{(x - \rho(t, p))^2\}]$. With $g(p) = p$ we have $\rho(t, p) = p$. Suppose, for example, that the seller sells for the price of $p + 1$ and invests this amount in the market, buying at least one whole share of the stock. If he/she does not consume, then the volatility will always be greater than σ.

APPENDIX

We now give an example showing that in the forward–backward case $g_1 \ge g_2$ does not necessarily imply that $X_1(\cdot) \ge X_2(\cdot)$, where X_1 and X_2 are the corresponding (backward components) of the FBSDE (3.3) with terminal conditions g_1 and g_2, respectively. Indeed, the reverse can happen with positive probability.

Let us assume that $d = 1$ and $r \equiv 0$. Let the price equation be

$$\text{(A.1)}\quad dP(t) = \frac{P(t)}{(\pi(t) - X(t))^2 + 1}\,dt + P(t)\,dW(t), \qquad P(0) = p;$$

hence the wealth equation is [see (2.4)]

$$dX(t) = \frac{\pi(t)}{(\pi(t) - X(t))^2 + 1}\,dt + \pi(t)\,dW(t), \qquad X(T) = g(P(T)).$$

We first assume that $g(p) = g_1(p) = p$ (i.e., one wants to hedge the price itself). The corresponding PDE (3.8) is now of the form

$$\text{(A.2)}\qquad 0 = \theta_t + \tfrac{1}{2}p^2\theta_{pp}, \qquad \theta(T, p) = g_1(p) = p.$$

Let us denote the solution of (A.2) by θ_1. Since $g_1(p) = p$ satisfies (A4), we easily see that $\theta_1(t, p) \equiv p$ is the unique classical solution to (A.2). Thus, by (3.11), $X_1(\cdot) \equiv \pi_1(\cdot) \equiv P_1(\cdot)$. Moreover, from (A.1), $X_1(t) = p\exp\{W(t) + t/2\}$.

We now let $g_2(p) = p + 1$ and denote by θ_2 the classical solution to (A.2), with the terminal condition being replaced by $g_2(p)$. Clearly, $\theta_2(t, p) = p + 1$. Therefore, $X_2(\cdot) \equiv P_2(\cdot) + 1 \equiv \pi_2(\cdot) + 1$. Moreover, from (A.1), $X_2(t) = 1 + p\exp\{W(t)\}$. Therefore, $X_1(t) - X_2(t) = pe^{W(t)}[e^{t/2} - 1] - 1$, which can be both positive or negative with positive probability for any $t > 0$.

Acknowledgments. The authors would like to thank Professors Jiongmin Yong of Fudan University, China, and Martin Schweizer of Technische Universität Berlin, Germany, for helpful discussions. In particular, Martin Schweizer and the two anonymous referees pointed out that our original, longer proof of Lemma 2.3 is not needed. Thanks are also due one of the referees for many useful suggestions, in particular, for simplifying some of the proofs by using results from [8].

REFERENCES

[1] ANSEL, J. P. and STRICKER, C. (1993). Lois de martingale, densités et décomposition de Föllmer-Schweizer. *Ann. Instit. H. Poincaré* **28** 375–392.

[2] BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *J. Polit. Economy* **81** 637–659.

[3] CVITANIĆ, J. (1996). Nonlinear financial markets: hedging and portfolio optimization. *Proceedings of the Isaac Newton Institute for Mathematical Sciences*. To appear.

[4] CVITANIĆ, J. and KARATZAS, I. (1993). Hedging contingent claims with constrained portfolios. *Ann. Appl. Probab.* **3** 652–681.

[5] DELBAEN, F. and SCHACHERMAYER, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300** 463–520.

[6] DUFFIE, D. and EPSTEIN, L. G. (1992). Stochastic differential utility. *Econometrica* **60** 353–394. (Appendix with Skiadas, C.)

[7] DUFFIE, D., MA, J. and YONG, J. (1995). Black's console rate conjecture. *Ann. Appl. Probab.* **5** 356–382.

[8] EL KAROUI, N., PENG, S. and QUENEZ, M. C. (1993). Backward stochastic differential equations in finance and optimization. Preprint.

[9] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ.

[10] GROSSMAN, S. J. (1988). An analysis of the implications for stock and futures price volatility of program trading and dynamic hedging strategies. *Journal of Business* **61** 275–289.

[11] HARRISON, J. M. and KREPS, D. M. (1979). Martingales and arbitrage in multiperiod security markets. *J. Econom. Theory* **20** 381–408.

[12] HARRISON, J. M. and PLISKA, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.* **11** 215–260.

[13] HARRISON, J. M. and PLISKA, S. R. (1983). A stochastic calculus model of continuous time trading: complete markets. *Stochastic Process. Appl.* **15** 313–316.

[14] KARATZAS, I. (1989). Optimization problems in the theory of continuous trading. *SIAM J. Control Optim.* **27** 1221–1259.

[15] KARATZAS, I., LEHOCZKY, J. P. and SHREVE, S. E. (1991). Equilibrium models with singular asset prices. *Math. Finance* **1** 11–29.

[16] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, New York.

[17] LADYZENSKAJA, O. A., SOLONNIKOV, V. A. and URAL'CEVA, N. N. (1968). *Linear and Quasilinear Equations of Parabolic Type*. Amer. Math. Soc., Providence, RI.

[18] MA, J., PROTTER, P. and YONG, J. (1994). Solving forward–backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields* **98** 339–359.

[19] MERTON, R. C. (1969). Lifetime portfolio selection under uncertainty: the continuous time case. *Rev. Econom. Statist.* **51** 247–257.

[20] PARDOUX, E. and PENG, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** 55–61.

[21] PARDOUX, E. and PENG, S. (1992). Backward stochastic differential equations and quasi-linear parabolic partial differential equations. *Stochastic Partial Differential Equations and Their Applications. Lecture Notes in Control and Inform. Sci.* **176** 200–217. Springer, Berlin.

[22] PENG, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep.* **37** 61–74.

[23] PLATEN, E. and SCHWEIZER, M. (1994). On smile and skewness. Preprint.

[24] PROTTER, P. (1990). *Stochastic Integration and Differential Equations. A New Approach.* Springer, Berlin.

[25] ROGERS, L. C. G. and WILLIAMS, D. (1987). *Diffusion, Markov Processes, and Martingales* **2**. *Itô Calculus*. Wiley, New York.

DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027
E-MAIL: cj@stat.columbia.edu

DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907-1395
E-MAIL: majin@math.purdue.edu