Robust Contracts in Continuous Time*

Jianjun Miao† Alejandro Rivera‡

March 5, 2016

Abstract

We study a continuous-time contracting problem under hidden action, where the principal has ambiguous beliefs about the project cash flows. The principal designs a robust contract that maximizes his utility under the worst-case scenario subject to the agent's incentive and participation constraints. Robustness generates endogenous belief heterogeneity and induces a tradeoff between incentives and ambiguity sharing so that the incentive constraint does not always bind. We implement the optimal contract by cash reserves, debt, and equity. In addition to receiving ordinary dividends when cash reserves reach a threshold, outside equity holders also receive special dividends or inject cash in the cash reserves to hedge against model uncertainty and smooth dividends. The equity premium and the credit yield spread generated by ambiguity aversion are state dependent and high for distressed firms with low cash reserves.

JEL Classification: D86, G12, G32, J33

Keywords: robustness, ambiguity, moral hazard, principal-agent problem, capital structure, equity premium, asset pricing

*We thank Patrick Bolton, Zhiguo He, Li Li, Juan Ortner, Yuliy Sannikov, Tom Sargent, Neng Wang, Noah Williams, the coeditor, and the referees for helpful comments. We have also benefited from comments by the seminar participants at Boston University, University of Wisconsin at Madison, CICF, NYU Conference on Ambiguity and Robustness in Macroeconomics and Finance, Conference on Ambiguity and Robustness in Economics at LAEF of University of California at Santa Barbara, NSF Decentralization Conference at Northwestern University, and AFR Summer Institute of Economics and Finance at Zhejiang University. First version: December 2012.

†Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA; CEMA, Central University of Finance and Economics, and AFR, Zhejiang University, China. Tel.: 617-353-6675. Email: miaoj@bu.edu. Homepage: http://people.bu.edu/miaoj

‡Naveen Jindal School of Management, University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080, Tel.: 857-998-2329, Email: axr150331@utdallas.edu

1 Introduction

Uncertainty and information play an important role in principal-agent problems. Consistent with the rational expectations hypothesis, the traditional approach to these problems typically assumes that both the principal and the agent share the same belief about the uncertainty underlying an outcome, say output. The agent can take unobservable actions to influence the output distribution. The distribution given the recommended action in a contract is common and known to both the principal and the agent. This approach has generated important economic implications and found increasingly widespread applications in practice, e.g., managerial compensation, insurance contracts, and lending contracts, etc.

However, there are several good reasons for us to think about departures from the traditional approach. First, the Ellsberg (1961) paradox and related experimental evidence demonstrate that there is a distinction between risk and uncertainty (or ambiguity). Risk refers to the situation where there is a known probability distribution over the state of the world, while ambiguity refers to the situation where the information is too vague to be adequately summarized by a single probability distribution (Gilboa an Schmeidler (1989)). Second, as Anderson, Hansen and Sargent (2003) and Hansen and Sargent (2001, 2008) point out, economic agents view economic models as an approximation to the true model. They believe that economic data come from an unknown member of a set of unspecified models near the approximating model. Concern about model misspecification induces a decision maker to want robust decision rules that work over that set of nearby models.[1]

The goal of this paper is to study how to design robust contracts with hidden action in a dynamic environment. We adopt a continuous-time framework to address this question. More specifically, our model is based on DeMarzo and Sannikov (2006) and Biais et al (2007). The continuous-time framework is analytically convenient for several reasons. First, it allows us to represent belief distortions by perturbations of the drift of the Brownian motion using the powerful Girsanov Theorem (see Karatzas and Shreve (1991)). Second, it allows us to adapt and extend the martingale approach to the dynamic contracting problems recently developed by DeMarzo and Sannikov (2006), Sannikov (2008), and Williams (2009, 2011). Third, it allows us to express solutions in terms of ordinary differential equations (ODEs) which can be numerically solved tractably. Finally, it allows us to conduct capital structure implementation so that we can analyze the impact of robustness on asset pricing transparently.

When formulating robust contracting problems, we face two important issues. The first issue is that we have to consider who faces model ambiguity in our two-party contracting

[1]There is a growing literature on the applications of robustness and ambiguity to finance and macroeconomics, e.g., Epstein and Wang (1994), Epstein and Miao (2003), Maenhout (2004), Hansen (2007), Hansen and Sargent (2010), and Ju and Miao (2012), among others.

problems, unlike in the representative agent models. As a starting point, it is natural to assume that the agent knows the output distribution chosen by himself. Due to the lack of information, the principal faces model uncertainty in the sense that he believes that there may be multiple distributions surrounding the output distribution chosen by the agent.

The second issue is how to model decision making under ambiguity. There are several approaches in decision theory. A popular approach is to adopt the maxmin expected utility model of Gilboa and Schmeidler (1989). Chen and Epstein (2002) formulate this approach in a continuous-time framework. We adopt the approach proposed by Hansen and Sargent (2001), Anderson, Hansen, and Sargent (2003), and Hansen et al (2006).[2] This approach is especially useful for our analysis since model discrepancies are measured by entropy, which is widely used in statistics and econometrics for model detection. Moreover, it is analytically tractable.

We assume that the principal copes with model uncertainty by designing a robust contract that maximizes his utility in the worst-case scenario subject to the agent's incentive and participation constraints. The principal's utility is modeled as the multiplier preferences proposed by Hansen and Sargent (2001) and Anderson, Hansen and Sargent (2003) and axiomatized by Maccheroni, Marinacci and Rustichini (2006a,b) and Strzalecki (2011). The principal solves a maxmin problem, which is related to the zero-sum differential game literature (e.g., Fleming and Souganidis (1989)). Our key insight is that the principal's aversion to model uncertainty generates an endogenous belief distortion in that he pessimistically puts more weight on worse outcomes. Since the agent is assumed not to face ambiguity, there is endogenous belief heterogeneity. This belief heterogeneity generates a tradeoff between incentives and ambiguity sharing and has important implications for contract dynamics and asset pricing.

We find the following main novel results. First, unlike DeMarzo and Sannikov (2006) and Biais et al (2007), our robust contract implies that the optimal sensitivity of the agent's continuation value to the cash flow uncertainty is not always at the lower bound to ensure incentive compatibility. The intuition is the following. The principal is ambiguous about the probability distribution of the project cash flows. He wants to remove this ambiguity and transfer uncertainty to the agent. But he does not want the agent to bear too much uncertainty since this may generate excessive volatility and a high chance of liquidation. When the agent's continuation value is low, the principal is more concerned about liquidation and hence the optimal sensitivity is at the lower bound so that the incentive constraint just binds. But when the agent's continuation value is high, the principal is more concerned about model uncertainty and hence the optimal contract allows the agent to bear more uncertainty.

[2]See Hansen and Sargent (2008) for a textbook treatment of this approach in discrete time.

In this case the optimal sensitivity of the agent's continuation value to the cash flow is state dependent and exceeds its lower bound.

Second, we show that the robust contract can be implemented by cash reserves, debt, and equity as in Biais et al (2007).[3] Unlike their implementation, the equity payoffs consist of regular dividends (paid only when the cash reserves reach a threshold level) and special dividends (or cash injections if negative). The special dividends or cash injections are used as a hedge against model uncertainty. They ensure that cash reserves track the agent's continuation value so that the payout time and the liquidation time coincide with those in the robust contract. Special dividends or cash injections occur only when the firm builds up sufficiently high cash reserves. In this case, when the project performs well, outside equity holders inject cash to raise cash reserves through new equity issues. But when the project performs bad, outside equity holders receive special dividends so that total dividends are smoothed. This result provides an explanation of dividend smoothing widely documented in the corporate finance literature dated back to Lintner (1956). Our model prediction is consistent with the empirical evidence documented by Leary and Michaely (2011) that dividend smoothing is most common among firms that are cash cows.

Third, incorporating model uncertainty has important asset pricing implications. The shadow prices of model uncertainty for the principal are revealed by the endogenously determined worst-case belief and contribute to the uncertainty premium and hence the equity premium. The uncertainty premium lowers the stock price and debt value and hence makes some profitable projects unfunded. It also raises the credit yield spread. Importantly, the equity premium and the credit yield spread increase with the degree of ambiguity aversion. They are state dependent and high for distressed firms with low cash reserves. This also implies that the equity premium and the credit yield spread are high in recessions since cash reserves are low in bad times.[4] By contrast, there is no equity premium in DeMarzo and Sannikov (2006) and Biais et al (2007).

Fourth, unlike DeMarzo and Sannikov (2006) and Biais et al (2007), we show that the stock price in the robust contracting problem is convex for low levels of cash reserves and concave for high levels of cash reserves when investors (the principal) are sufficiently ambiguity averse. Thus shareholders will benefit from a gamble involving the gain or loss from an initial position of low cash reserves. This result implies that the asset substitution or risk-shifting problem is more likely to occur for financially distressed firms or newly established firms with low cash reserves (Jensen and Meckling (1976)). This prediction is consistent with

[3]See DeMarzo et al. (2012) for a related implementation. We can also implement the robust contracts by credit lines, debt and equity as in DeMarzo and Sannikov (2006). We have not pursued this route in this paper.

[4]Biais et al (2007) also show that the credit yield spread is high for distressed firms with low cash reserves.

the empirical evidence reported by Eisdorfer (2008) that, unlike healthy firms, financially distressed firms speed up investments as uncertainty increases to transfer wealth from creditors to shareholders.

Our paper is related to a fast growing literature on dynamic contracting problems in continuous time.[5] Our paper is most closely related to the seminal contributions by DeMarzo and Sannikov (2006) and Biais et al (2007). Our main contribution is to introduce robustness into their models and study capital structure implementation and asset pricing implications. Szydlowski (2012) introduces ambiguity into a dynamic contracting problem in continuous time. He assumes that the principal is ambiguous about the agent's effort cost. His modeling of ambiguity is quite different from ours and can be best understood as a behavioral approach. His utility model cannot be subsumed under the decision-theoretic setting of Gilboa and Schmeidler (1989) and its continuous time version by Chen and Epstein (2002). Carroll (2015) studies a static contracting problem in which the principal has ambiguous beliefs about the agent's possible actions. He shows that the optimal contract is linear, but he does not study the questions related to security design and asset pricing.

Our modeling of robust contracting problems is inspired by Hansen and Sargent (2012) who classify four types of ambiguity in robust monetary policy problems in which a Ramsey planner faces private agents. They argue that "a coherent multi-agent setting with ambiguity must impute possibly distinct sets of models to different agents, and also specify each agent's understanding of the sets of models of other agents." This point is particularly relevant for contracting problems because such problems must involve at least two parties. Ambiguity aversion generates endogenous belief heterogeneity and delivers interesting contract dynamics and asset pricing implications.

The remainder of the paper proceeds as follows. Section 2 lays out the model. Section 3 presents the solution to the robust contract. Sections 4 analyzes capital structure implementation and asset pricing implications. Section 5 concludes. Technical details are relegated to appendices.

2 The Model

We first introduce the model setup which is based on DeMarzo and Sannikov (2006) and Biais et al (2007). We then introduce belief distortions and incorporate a concern for model uncertainty.

[5] See Sannikov (2013) and Cvitanic and Zhang (2013) for surveys of this literature.

2.1 Setup

Consider a continuous-time infinite-horizon environment. An agent (or entrepreneur) owns a technology (or project) that can generate a cumulative cash-flow process represented by (Y_t). The project needs initial capital $K > 0$ to be started. The agent has no initial wealth and needs financing from outside investors (the principal). Once the project is started, the agent affects the technology performance by taking an action or effort level $a_t \in [0, 1]$, which changes the distribution of cash flows. The cash flows (Y_t) follow the process

$$dY_t = \mu a_t dt + \sigma dB_t^a, \tag{1}$$

where $\mu, \sigma > 0$ and B^a is a standard Brownian motion defined on the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P^a\right)$.[6] Note that the triple (Y, B^a, P^a) is a weak solution to the preceding stochastic differential equation.

The agent can derive private benefits $\lambda\mu(1 - a_t)\,dt$ from the action a_t, where $\lambda \in (0, 1)$. Due to linearity, this modeling is also equivalent to the binary effort setup where the agent can either shirk, $a_t = 0$, or work, $a_t = 1$. Hence, we adopt this simple assumption throughout the paper. Alternatively, we can interpret $1 - a_t$ as the fraction of cash flow that the agent diverts for his private benefit, with λ equal to the agent's net consumption per dollar diverted. In either case, λ represents the severity of the agency problem. The choice of the agent's action is unobservable to the principal, creating the moral hazard issue. The principal only observes past and current cash flows and his information set is represented by the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by (Y_t).

Both the principal and the agent are risk neutral and discount the future cash flows according to r and γ respectively. Assume that $r < \gamma$ so that the agent is more impatient than the principal. The technology can be liquidated. If it is liquidated, the principal obtains L and the agent gets outside value zero.

The principal offers to contribute capital K in exchange for a contract (C, τ, a) that specifies a termination (stopping) time τ, a cash compensation $C = \{C_t : 0 \leq t \leq \tau\}$ to the agent, and a suggested effort choice $a = \{a_t \in \{0, 1\} : 0 \leq t \leq \tau\}$. Assume that C and a are adapted to $(\mathcal{F}_t)$ and that C is a right continuous with left limits and increasing process satisfying[7]

$$E^{P^a}\left[\left(\int_0^t e^{-\gamma s} dC_s\right)^2\right] < \infty,\ t \geq 0,\ C_0 \geq 0.$$

[6] All processes in the paper are assumed to be progressively measurable with respect to $\{\mathcal{F}_t\}$. Inequalities in random variables or stochastic processes are understood to hold almost surely.

[7] The square integrability is imposed to ensure W_t defined in (2) has a martingale representation (see Cvitanic and Zhang (2013), Chapter 7).

The monotonicity requirement reflects the fact that the agent has limited liability.

Fix a contract (C, τ, a) and assume that the agent follows the recommended choice of effort. His continuation value W_t at date t is defined as

$$W_t = E_t^{P^a}\left[\int_t^\tau e^{-\gamma(s-t)}(dC_s + \lambda\mu(1-a_s)ds)\right], \tag{2}$$

where $E_t^{P^a}$ denotes the conditional expectation operator with respect to the measure P^a given the information set $\mathcal{F}_t$. His total expected utility at date 0 is equal to W_0.

DeMarzo and Sannikov (2006) and Biais et al (2007) study the following contracting problem in which the principal maximizes his expected utility value without a concern for model uncertainty.

Problem 2.1 *(benchmark model)*

$$\max_{(C,\tau,a)} \quad E^{P^a}\left[\int_0^\tau e^{-rs}(dY_s - dC_s) + e^{-r\tau}L\right], \tag{3}$$

subject to

$$E^{P^a}\left[\int_0^\tau e^{-\gamma s}(dC_s + \lambda\mu(1-a_s)ds)\right] \geq E^{P^{\hat{a}}}\left[\int_0^\tau e^{-\gamma s}(dC_s + \lambda\mu(1-\hat{a}_s)ds)\right], \tag{4}$$

$$E^{P^a}\left[\int_0^\tau e^{-\gamma s}(dC_s + \lambda\mu(1-a_s)ds)\right] = W_0, \tag{5}$$

where $\hat{a}_s \in \{0,1\}$ *and* $W_0 \geq 0$ *is given.*

In this problem, consistent with the rational expectations hypothesis, both the principal and the agent use the measure P^a to evaluate expected utility. Inequality (4) is the incentive constraint and equation (5) is the promising-keeping or participation constraint. Assume that the agent and the principal cannot save and both the principal and the agent have full commitment.

Let $F^b(W_0)$ denote the principal's value function for Problem 2.1. Then the project can be funded if and only if $\max_{w\geq 0} F^b(w) \geq K$. If the agent has all bargaining power due to competition of principals, he extracts the maximal W_0 such that $F^b(W_0) = K$. If the principal has all bargaining power due to competition of agents, he delivers the agent W^* such that W^* solves $\max_w F^b(w)$.

2.2 Robustness and Belief Distortions

We now consider the possibility of belief distortions due to concerns about model misspecifications or model ambiguity. Both the principal and the agent view the probability measure P^a as an approximating model. Suppose that the principal does not trust this model and considers alternative models as possible.

Suppose that all distorted beliefs are described by mutually absolutely continuous measures with respect to P^a over any finite time intervals. This assumption allows us to use the Girsanov Theorem for changing measures so that model ambiguity is only about the drift of the diffusion processes.[8] Define a *density generator* associated with an effort process a as a real-valued process (h_t) satisfying $\int_0^t h_s^2 ds < \infty$ for all $t > 0$ such that the process (z_t) defined by

$$z_t = \exp\left(\int_0^t h_s dB_s^a - \frac{1}{2}\int_0^t h_s^2 ds\right) \tag{6}$$

is a $(P^a, \mathcal{F}_t)$-martingale (see Hansen et al (2006) for the technical details). Denote the set of density generators by $\mathcal{H}^a$. By the Girsanov Theorem, there is a measure Q^h corresponding to h defined on $(\Omega, \mathcal{F})$ such that z_t is the Radon-Nikodym derivative of Q^h with respect to P^a when restricted to $\mathcal{F}_t$, $dQ^h/dP^a|_{\mathcal{F}_t} = z_t$, and the process $\left(B_t^h\right)$ defined by

$$B_t^h = B_t^a - \int_0^t h_s ds,$$

is a standard Brownian motion under the measure Q^h. Under measure Q^h, cash flows follow dynamics

$$dY_t = \mu a_t dt + \sigma\left(dB_t^h + h_t dt\right). \tag{7}$$

Following Anderson, Hansen and Sargent (2002), Hansen et al (2006), and Hansen and Sargent (2012), we use discounted relative entropy to measure the discrepancy between Q^h and P^a,

$$rE^{P^a}\left[\int_0^\infty e^{-rt} z_t \ln z_t dt\right] = \frac{1}{2}E^{P^a}\left[\int_0^\infty e^{-rt} z_t h_t^2 dt\right],$$

where the equality follows from (6) and integration by parts. To incorporate a concern for robustness of belief distortions, we represent the principal's preferences by multiplier utility

$$\inf_h \; E^{Q^h}\left[\int_0^\tau e^{-rt}(dY_t - dC_t) + e^{-r\tau}L\right] + \frac{\theta}{2}E^{P^a}\left[\int_0^\tau e^{-rt} z_t h_t^2 dt\right], \tag{8}$$

where the last term penalizes belief distortions. This utility model can be viewed as a

[8] See Epstein and Ji (2013) for a continuous-time model of ambiguity about both the drift and volatility without the mutually absolute continuity assumption.

continuous-time version of the dynamic variational utility model studied by Maccheroni, Marinacci, and Rustichini (2006b). The parameter $\theta > 0$ describes the degree of concern for robustness. We may interpret $1/\theta$ as an ambiguity aversion parameter. A small θ implies a large degree of ambiguity aversion or a large degree of concern for robustness. When θ approaches infinity, the preceding utility reduces to expected utility in (3).

3 Robust Contract

We formulate the robust contracting problem with agency as follows:

Problem 3.1 *(robust contract with agency)*

$$\sup_{(C,\tau,a)} \inf_{h} \quad E^{Q^h}\left[\int_0^\tau e^{-rt}(dY_t - dC_t) + e^{-r\tau}L\right] + \frac{\theta}{2}E^{P^a}\left[\int_0^\tau e^{-rt} z_t h_t^2 dt\right], \tag{9}$$

subject to (4), (5), and (6).

Mathematically, Problem 3.1 is a combined singular control and stopping problem (see Fleming and Soner (1993)). As Hansen et al (2006) point out, it is also related to the zero-sum stochastic differential game problem (e.g., Fleming and Souganidis (1989)). We shall proceed heuristically to derive the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation for optimality and then provide a formal verification theorem. We finally analyze several numerical examples to illustrate economic intuition. It is technically challenging and quite involved to provide a rigorous derivation of the HJBI equation. Such an analysis is beyond the scope of this paper.

3.1 First-Best Robust Contract

Before analyzing the robust contract with agency, we start with the first-best case in which the principal observes the agent's effort choice and hence the incentive constraint (4) in problem 3.1 is not valid. The derivations of the HJBI equation consist of several steps. First, we ignore the incentive constraint (4) and keep the participation constraint. Using Girsanov's Theorem and the Martingale Representation Theorem, (W_t) in (2) satisfies

$$dW_t = \gamma W_t dt - dC_t - \lambda\mu\left(1 - a_t\right)dt + h_t\phi_t dt + \phi_t dB_t^h, \tag{10}$$

where B_t^h is a standard Brownian motion under the measure Q^h.

Second, we write the principal's utility in (3) under the measure Q^h as

$$\begin{aligned} & E^{Q^h}\left[\int_0^\tau e^{-rt}(dY_t - dC_t) + e^{-r\tau}L\right] + \frac{\theta}{2}E^{P^a}\left[\int_0^\tau e^{-rt} z_t h_t^2 dt\right] \\ = \quad & E^{Q^h}\left[\int_0^\tau e^{-rt}(\mu a_t dt + \sigma h_t dt - dC_t) + e^{-r\tau}L\right] + \frac{\theta}{2}E^{Q^h}\left[\int_0^\tau e^{-rt} h_t^2 dt\right], \end{aligned}$$

where we have used the fact that $dY_t = (\mu a_t + \sigma h_t)\, dt + \sigma dB_t^h$.

Third, define $F(W_0)$ as the value function for Problem 3.1 without the incentive constraint (4) when we vary the promised value W_0 to the agent. We use the dynamic programming principle to write an approximate Bellman equation:

$$rF(W_t)\, dt = \sup_{dC_t, \phi_t, a_t \in \{0,1\}} \inf_{h_t} \ \mu a_t dt + \sigma h_t dt - dC_t + \frac{\theta}{2} h_t^2 dt + E_t^{Q^h}\left[dF(W_t)\right], \tag{11}$$

subject to (10). This equation has an intuitive economic interpretation. The left-hand side represents the mean return required by the principal. The right-hand side represents the total return expected by the principal. It consists of the cash flow plus the expected capital gain or loss $E_t^{Q^h}\left[dF(W_t)\right]$. Optimality requires the expected return equal the required mean return. Note that all expected values are computed using the measure Q^h.

Now we use Ito's Lemma and (10) to derive

$$E_t^{Q^h}\left[dF(W_t)\right] = F'(W_t)(\gamma W_t dt - dC_t - \lambda\mu(1 - a_t)\, dt + h_t\phi_t dt) + \frac{F''(W_t)}{2}\phi_t^2 dt.$$

Plugging this equation into (11) yields

$$\begin{aligned} rF(W_t)\, dt \quad = \quad & \sup_{a_t \in \{0,1\}, dC_t, \phi_t} \inf_{h_t} \ \mu a_t dt + \sigma h_t dt - dC_t + \frac{\theta}{2} h_t^2 dt \\ & + F'(W_t)(\gamma W_t dt - dC_t - \lambda\mu(1 - a_t)\, dt + h_t\phi_t dt) + \frac{F''(W_t)}{2}\phi_t^2 dt. \end{aligned}$$

Suppose that $dC_t = c_t dt$, where $c_t \geq 0$. Removing the time subscripts and cancelling dt, we obtain the HJBI equation

$$\begin{aligned} rF(W) \quad = \quad & \sup_{a \in \{0,1\}, c \geq 0, \phi} \inf_{h} \ \mu a + \sigma h - \left(1 + F'(W)\right) c \\ & + F'(W)(\gamma W + h\phi - \lambda\mu(1 - a)) + \frac{F''(W)}{2}\phi^2 + \frac{\theta h^2}{2}. \end{aligned} \tag{12}$$

Clearly, for this problem to have a finite solution, we must have $F'(W) \geq -1$. We then get $c > 0$ if and only if $F'(W) = -1$. Define $\bar{W}$ as the lowest level such that $F'(W) = -1$.

This illustrates the feature of the singular control problem: the principal makes payments to the agent if and only if W_t reaches the point $\bar{W}$. The payments make the process (W_t) reflects at this point.

The objective function in (12) is convex in h. Solving for the worst-case density generator yields

$$h = -\frac{\phi F'(W) + \sigma}{\theta}. \tag{13}$$

Substituting it back into (12) yields

$$rF(W) = \sup_{a\in\{0,1\},\phi} \mu a + F'(W)\left(\gamma W - \lambda\mu\left(1-a\right)\right) + \frac{\phi^2}{2}F''(W) - \frac{\left[\phi F'(W) + \sigma\right]^2}{2\theta}. \tag{14}$$

Assuming that

$$\theta F''(W) - F'(W)^2 < 0, \tag{15}$$

so that the expression on the right-hand side of equation (14) is concave in ϕ, we can derive the optimal sensitivity

$$\phi^*(W) = \frac{F'(W)\sigma}{\theta F''(W) - F'(W)^2}. \tag{16}$$

Note that the concavity of F is sufficient but not necessary for (15) to hold. Since $\lambda \in (0,1)$ and $F'(W) \geq -1$, it follows that $\lambda F'(W) + 1 \geq 0$ and hence implementing high effort $a_t = 1$ is optimal.

The following result characterizes the first-best robust contract.

Proposition 1 *Consider the first-best robust contracting problem. Suppose that*

$$L < \frac{\mu}{r} - \frac{\sigma^2}{2r\theta}, \tag{17}$$

and that there is a unique twice continuously differentiable solution F to the ODE on $[0,\bar{W}]$,

$$rF(W) = \mu + F'(W)\gamma W - \frac{\left[F'(W)\sigma\right]^2}{2\theta\left[\theta F''(W) - F'(W)^2\right]} - \frac{\sigma^2}{2\theta}, \tag{18}$$

with the boundary conditions,

$$\begin{aligned} F(0) &= \frac{\mu}{r} - \frac{\sigma^2}{2r\theta}, \\ F'\left(\bar{W}\right) &= -1,\ F''\left(\bar{W}\right) = 0, \end{aligned} \tag{19}$$

such that condition (15) holds and $F'(W) > -1$ on $[0,\bar{W})$. Then:

(i) When $W \in [0,\bar{W}]$, the principal's value function is given by $F(W)$, the first-best

sensitivity $\phi^*(W)$ *is given by (16), the worst-case density generator is given by*

$$h^*(W) = -\frac{\phi^*(W) F'(W) + \sigma}{\theta}, \tag{20}$$

and the agent always exerts high effort $a^*(W) = 1$. *The contract initially delivers* $W \in \left[0, \bar{W}\right] \geq 0$ *to the agent whose continuation value* (W_t) *follows the dynamics*

$$dW_t = \gamma W_t dt - dC_t^* + \phi^*(W_t) dB_t^1, \; W_0 = W, \tag{21}$$

for $t \geq 0$, *where the optimal payments are given by*

$$C_t^* = \int_0^t \mathbf{1}_{\left\{W_s = \bar{W}\right\}} dC_s^*, \tag{22}$$

and the project is never liquidated.

(ii) When $W > \bar{W}$, *the principal's value function is* $F(W) = F\left(\bar{W}\right) - \left(W - \bar{W}\right)$. *The principal pays* $W - \bar{W}$ *immediately to the agent and the contract continues with the agent's new initial value* $\bar{W}$.

The intuition behind this proposition is as follows. The principal is ambiguity averse and would like to transfer uncertainty to the agent when designing a contract. Ideally, the risk-neutral agent should insure the principal by making the principal's payoff flows constant. This means that the agent should absorb all risk from the project cash flows. However, this contract is not feasible due to limited liability. The project cash flows can be negative and the agent can incur losses. With limited liability, uncertainty sharing is limited. The net marginal cost to the principal from delivering an additional unit of value to the agent is $1 + F'(W) \geq 0$. The principal makes payments to the agent when and only when the net marginal cost is equal to zero at some point $\bar{W}$. The tradeoff is the following: On the one hand, the principal wants to make payments to the agent earlier because the agent is more impatient. On the other hand, the principal wants to delay payments, allowing the agent's continuation value W_t to get larger. This benefits the principal because if W_t is closer to zero, the principal has to bear more of the project cash flow uncertainty. In particular, when $W_t = 0$, the principal bears full uncertainty and his value is given by (19). The term $\sigma^2/(2r\theta)$ represents the discount due to model uncertainty. It increases with volatility σ and ambiguity aversion parameter $1/\theta$. Assumption (17) implies that liquidation is never optimal in the first-best robust contract.

Proposition 1 shows that the worst-case density generator and the sensitivity of the agent's continuation value to the cash flow are state dependent. The agent bears large cash flow uncertainty, but he does not absorb all uncertainty due to limited liability. Because the

principal also bears uncertainty, his value function F is nonlinear and the last two nonlinear terms in the ODE reflect the value discount due to model ambiguity.

We emphasize that in two-party contracting problems, model ambiguity generates endogenous belief heterogeneity. Specifically, the agent trusts the approximating model P^a and his continuation value (W_t) follows the dynamics (21) under P^a. However, the principal has doubt about the approximating model P^a and the agent's continuation value under the principal's worst-case model Q^{h^*} follows the dynamics

$$dW_t = \gamma W_t dt - dC_t^* + \phi^*(W_t) h^*(W_t) dt + \phi^*(W_t) dB_t^{h^*}. \tag{23}$$

This point has important pricing implications when we implement the robust contract with agency later. The endogenous belief heterogeneity is a useful interpretation and not intended as an ex ante statement of uncertainty.

3.2 Robust Contract with Agency

Turn to the case with moral hazard in which the principal does not observe the agent's effort choice and hence the incentive constraint (4) must be imposed in Problem 3.1. Without risk of confusion, we still use $F(W_0)$ to denote the value function for Problem 3.1 when we vary the promised value W_0 to the agent. Suppose that implementing high effort is optimal. Then DeMarzo and Sannikov (2006) show that the incentive constraint is equivalent to

$$\phi_t \geq \sigma\lambda. \tag{24}$$

Using a similar argument to that in the previous subsection, we can proceed heuristically to derive the HJBI equation for optimality. Imposing constraint (24) and setting $a_t = 1$ in the associated equations in the previous subsection, we can show that the HJBI equation is given by

$$\begin{aligned} rF(W) &= \sup_{c\geq 0, \phi \geq \sigma\lambda} \inf_h \mu + \sigma h - \left(1 + F'(W)\right) c \\ &\quad + F'(W)(\gamma W + h\phi) + \frac{F''(W)}{2}\phi^2 + \frac{\theta h^2}{2}. \end{aligned} \tag{25}$$

Thus the worst-case density generator is still given by (13) and there is a boundary point $\bar{W}$ such that $c > 0$ if $F'\left(\bar{W}\right) = -1$ and $c = 0$ if $F'(W) > -1$ for $W \in [0, \bar{W})$. We can then

rewrite the HJBI equation as

$$rF(W) = \sup_{\phi \geq \sigma\lambda} \mu + F'(W)\gamma W + \frac{\phi^2}{2}F''(W) - \frac{\left[\phi F'(W) + \sigma\right]^2}{2\theta}. \tag{26}$$

Under condition (15), the optimal sensitivity is given by

$$\phi^*(W) = \max\left\{\frac{F'(W)\,\sigma}{\theta F''(W) - F'(W)^2}, \sigma\lambda\right\}. \tag{27}$$

The last term in (26) reflects the cost of model uncertainty. When $\theta \to \infty$, that term disappears and $\phi^*(W) \to \sigma\lambda$. We then obtain the solution as in DeMarzo and Sannikov (2006).

The following result characterizes the robust contract with agency.

Proposition 2 *Consider the robust contracting problem with agency. Suppose that implementing high effort is optimal and that condition (17) holds. Suppose that there exists a unique twice continuously differentiable solution F to the ODE (26) on $\left[0, \bar{W}\right]$ with boundary conditions*

$$F'\left(\bar{W}\right) = -1,\ F''\left(\bar{W}\right) = 0,\ F(0) = L,$$

such that condition (15) holds and $F'(W) > -1$ on $[0, \bar{W})$. Then:

(i) When $W \in [0, \bar{W}]$, $F(W)$ is the value function for Problem 3.1, the optimal sensitivity $\phi^*(W)$ is given by (27), and the worst-case density generator is given by (20). The contract delivers the value $W \in \left[0, \bar{W}\right]$ to the agent whose continuation value (W_t) follows the dynamics (21) for $t \in [0, \tau]$, where the optimal payments are given by (22). The contract terminates at time $\tau = \inf\{t \geq 0 : W_t = 0\}$.

(ii) When $W > \bar{W}$, the principal's value function is $F(W) = F\left(\bar{W}\right) - \left(W - \bar{W}\right)$. The principal pays $W - \bar{W}$ immediately to the agent and the contract continues with the agent's new initial value $\bar{W}$.

Unlike in the first-best case, the incentive constraint requires that the sensitivity ϕ_t be at least as large as a lower bound $\sigma\lambda$ as in the DeMarzo-Sannikov model. In their model, the choice of ϕ_t reflects the following tradeoff: a large ϕ_t is needed to provide incentives to the agent. But a large ϕ_t also raises the volatility of the agent's continuation value and hence raises the chance of liquidation. The optimal sensitivity just achieves the lower bound $\sigma\lambda$ when the principal's value function is concave. However, this lower bound does not always bind in the presence of model ambiguity. The reason is that there is an uncertainty and incentive tradeoff. The robust contract should transfer uncertainty from the ambiguity averse

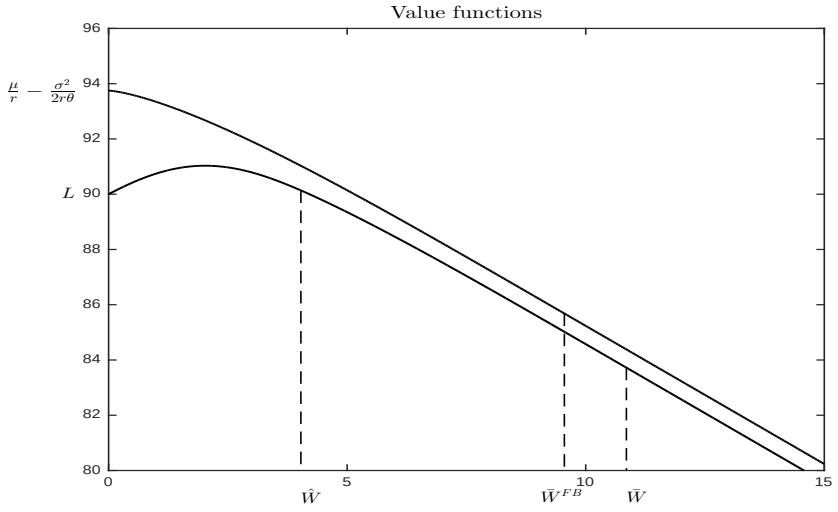

Figure 1: **Value functions for the robust contracting problem**. The upper curve is the first-best value function and the payout boundary is $\bar{W}^{FB}$. The lower curve is the value function with agency and the payout boundary is $\bar{W}$. The optimal sensitivity changes value at $\hat{W}$. Parameter values are $\mu = 10$, $r = 0.10$, $\gamma = 0.15$, $\lambda = 0.20$, $\sigma = 5$, $L = 90$, and $\theta = 20$.

principal to the risk neutral agent as much as possible. Thus the agent should be exposed more to the uncertainty so that the optimal sensitivity may exceed the lower bound.

Under what situation does this happen? For a low value of W, the principal is more concerned about inefficient liquidation. Thus the optimal contract will set ϕ_t at the lower bound. When W is large and close to the payout boundary $\bar{W}$, the principal is more concerned about model uncertainty and hence he would like the agent to be exposed more to the cash flow uncertainty by providing him more incentives so that

$$\phi^*(W) = \frac{F'(W)\sigma}{\theta F''(W) - F'(W)^2} > \sigma\lambda. \tag{28}$$

From the analysis above, the agent is more likely to be overincentivized when his continuation value is high.

Figure 1 plots the value functions for the robust contracting problem with and without agency. The payout boundary is given by $\bar{W}^{FB}$ for the first-best case. It is lower than that for the contract with agency, implying that moral hazard generates inefficient delay in payout. Both value functions are concave and become linear after the payout boundaries with a slope -1. Figure 2 plots the worst-case density generator h^* and the optimal sensitivity ϕ^* for the contract with agency. Consistent with the previous intuition, the figure shows that there is a cutoff value $\hat{W}$, such that the sensitivity $\phi^*(W)$ reaches the lower bound $\sigma\lambda$ for all $W \in [0, \hat{W}]$

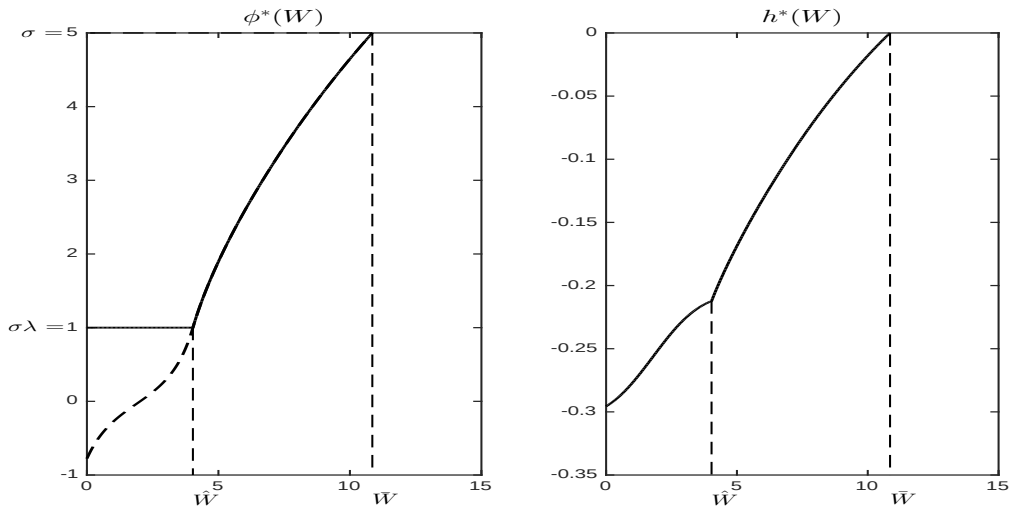

Figure 2: **Optimal sensitivity and the worst-case density generator for the robust contracting problem.** Parameter values are $\mu = 10$, $r = 0.10$, $\gamma = 0.15$, $\lambda = 0.20$, $\sigma = 5$, $L = 90$, and $\theta = 20$.

and it is given by (28) for all $W \in [\hat{W}, \bar{W}]$. We conjecture that this cutoff property holds more generally because it is supported by our numerical examples for a wide range of parameter values. Figure 2 also shows that $h^*(W)$ increases with W and $h^*(W) < 0$ for all $W \in [0, \bar{W})$. Intuitively, the principal's aversion to model uncertainty leads to his pessimistic behavior. The local mean of the Brownian motion is shifted downward under the principal's worst-case belief. At $W = \bar{W}$, the boundary conditions $F'\left(\bar{W}\right) = -1$ and $F''\left(\bar{W}\right) = 0$ imply that $h^*\left(\bar{W}\right) = 0$ and $\phi^*\left(\bar{W}\right) = \sigma$.

Figure 3 illustrates that the value function may not be globally concave. In particular, it is convex when the agent's continuation value is close to the liquidation boundary. To see why this can happen, we rewrite (26) as

$$\frac{\phi^*(W)^2}{2}F''(W) = \left[rF(W) - \mu - F'(W)\gamma W\right] + \frac{\left[\phi^*(W)F'(W) + \sigma\right]^2}{2\theta}.$$

When $\theta \to \infty$, the second expression on the right-hand side of the above equation vanishes and the model reduces to the DeMarzo-Sannikov model so that the first square bracket expression is negative and $F''(W) < 0$. However, when the principal is sufficiently ambiguity averse (i.e., $1/\theta$ is sufficiently large), the second expression may dominate so that $F''(W) > 0$. This case can happen when W is sufficiently small for the first square bracket expression to be small. In this case, public randomization in the sense of stochastic termination of the project is optimal

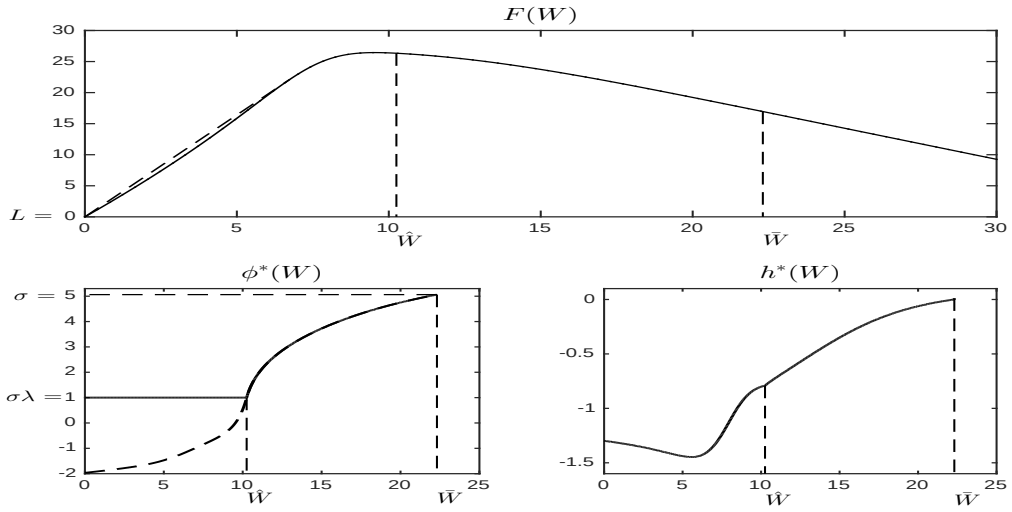

Figure 3: **Value function, optimal sensitivity, and the worst-case density generator for the robust contracting problem.** Parameter values are $\mu = 5$, $r = 0.10$, $\gamma = 0.15$, $\lambda = 0.20$, $\sigma = 5$, $L = 0$, and $\theta = 6$.

to the principal, as illustrated by the dashed line from the origin in Figure 3.[9] By contrast, when W is sufficiently large, $F'(W)$ is close to -1 and $\phi^*(W)$ is close to σ so that the second square bracket expression is close to zero. Thus $F''(W) < 0$ when W is sufficiently large.

Intuitively, the marginal value to the principal $F'(W)$ for a small W consists of three components. First, an increase in W pushes the agent's continuation value away from the liquidation boundary. This marginal benefit decreases with W. Second, an increase in W pushes the continuation value closer to the payout boundary. This marginal cost increases with W. Third, an increase in W pushes the continuation value closer to the cutoff $\hat{W}$, allowing the principal to have better ambiguity sharing. This marginal benefit increases with W when W is small. This component is unique in our robust contracting problem and may dominate the other two components when W is sufficiently small. In this case $F'(W)$ increases with W so that F is convex. But when W is sufficiently large, the first two components dominate so that F is concave.

The following proposition gives a necessary and sufficient condition for implementing high effort, which is satisfied in all our numerical examples.

Proposition 3 *Implementing high effort is optimal at all times for Problem 3.1 if and only*

[9]Stochastic liquidation is common in the discrete-time models, e.g., Clementi and Hopenhayn (2006), Biais et al. (2007), and DeMarzo and Fishman (2007a,b). Since such an analysis is standard, we omit it here.

if

$$rF(W) \geq \max_{\phi \leq \sigma\lambda} F'(W)\left(\gamma W - \lambda\mu\right) + \frac{\phi^2}{2}F''(W) - \frac{\left[\phi F'\left(W\right) + \sigma\right]^2}{2\theta}, \tag{29}$$

for $W \in \left[0, \bar{W}\right]$, *where* F *is given in Proposition 2 and satisfies condition (15).*

When $\theta \to \infty$, the condition in (29) reduces to that in Proposition 8 in DeMarzo and Sannikov (2006) for expected utility. If we adopt the recursive multiple-priors utility model of Chen and Epstein (2002) with the κ-ignorance specification of the set of priors, the HJBI equation is given by

$$rF(W) = \max_{dC \geq 0, \phi \geq \sigma\lambda} \min_{|h| \leq \kappa} \mu + \sigma h - (1 + F'\left(W\right))dC + F'(W)(\gamma W + h\phi) + \frac{F''(W)}{2}\phi^2.$$

In the online appendix we show that the incentive constraint binds when the agent's continuation value is sufficiently low. This result is similar to that in our model with multiplier preferences. The intuition follows from the same incentive and ambiguity sharing tradeoff. In contrast to our model, the value function for the Chen-Epstein model is globally concave and the corner solution $h^*\left(W\right) = -\kappa$ for all W achieves the minimum. It is much more difficult to analyze the HJBI equation for the Chen-Epstein model both analytically and numerically because there are two constrained optimization problems involved.

The following proposition shows that the value function F decreases if the degree of concern for robustness or the degree of ambiguity aversion increases, i.e., $1/\theta$ increases. The intuition is that model uncertainty is costly to the principal and hence reduces his value. The last term in (26) gives this cost, which is the local entropy $\theta h^*\left(W\right)^2/2$.

Proposition 4 *The value function* $F(W)$ *on* $\left[0, \bar{W}\right]$ *in Problem 3.1 increases with the parameter* θ.

3.3 Bellman-Isaacs Condition

As explained by Hansen et al (2006), the Bellman-Isaacs condition is important for various interpretations of robust control. This condition allows for the exchange of orders of extremization. Given this condition, the decision maker also optimizes with respect to a different probability specification of the original Brownian motion without period-by-period feedback onto an endogenous state variable. One can then offer an ex-post Bayesian interpretation of robust control. Hansen et al (2006) provide several examples (no binding inequality constraints, separability, and convexity) to verify this condition. Their examples do not fit in our model because the control problem in our model involves inequality constraints and control of the diffusion or volatility term (ϕ_t). For our robust contracting problem, one can check that if

F is concave, then the max and min operators in (12) or (25) can be exchanged (Fan (1953)) and hence the Bellman-Isaacs condition is satisfied. But condition (15) in our verification theorem (Proposition 2) does not imply global concavity of F and our numerical examples show that F may be convex in some region for the robust contracting problem. In this case, if one extremized ϕ first in (25), there would be no solution (ϕ is infinity). By contrast, we extremize h first and ϕ next, the optimal solution for ϕ in problem (26) is given by (27) as long as the second-order condition (15) holds. To conclude, whenever F is convex in some region, the Bellman-Issacs condition will not hold in our model.

Since the Bellman-Isaacs condition may not hold in our robust contracting problem, we cannot apply the ex-post Bayesian interpretation. Our interpretation of the worst-case probability measure as an endogenous belief is based on the decision theoretic axiomatization in Maccheroni, Marinacci, and Rustichini (2006a,b) and Strzalecki (2011). The multiplier utility given in (8) can be viewed as a continuous-time version of the discrete-time model in Maccheroni, Marinacci, and Rustichini (2006b). According to their axiomatization, the set of mutually absolutely continuous probability measures described in Section 2.2 represents the decision maker's ambiguous beliefs about all possible model specifications and the minimizing measure represents the worst-case belief. Since the order of extremization may not be exchanged in our model, we cannot drop the penalty term in (8) when performing maximization given the worst-case belief. Thus we cannot interpret the ambiguity averse principal in our model as an ex-post Bayesian decision maker with expected utility. But we can interpret the worst-case probability measure as his endogenous belief with non-expected utility.

4 Implementation and Asset Pricing

4.1 Capital Structure

We use cash reserves, debt, and equity to implement the optimal contract characterized in Proposition 2. We will show that ambiguity aversion generates some new insights into asset pricing. As in Biais et al (2007), the firm has a bank account that holds cash reserves M_t with interest rate r. The project payoffs are put in the account. Outside investors (principal) hold debt with coupon payment $[\mu - (\gamma - r) M_t] dt$ and a fraction $1-\lambda$ of equity. The entrepreneur (agent) holds a fraction λ of equity. Equity pays total regular dividends dC_t^*/λ. The cash reserves follow dynamics

$$dM_t = rM_t dt + dY_t - \underbrace{dC_t^*}_{\text{inside dividends}} - d\Psi_t, \; M_0 = W_0/\lambda, \tag{30}$$

for $M_t \in \left[0, \bar{W}/\lambda\right]$, where

$$d\Psi_t = \underbrace{\left[\mu - (\gamma - r)\, M_t\right] dt}_{\text{coupon}} + \underbrace{\frac{1-\lambda}{\lambda} dC_t^*}_{\text{outside dividends}} + \underbrace{\left[\sigma - \frac{\phi^*(\lambda M_t)}{\lambda}\right] dB_t^1}_{\text{special dividends}},$$

and $\bar{W}$, C^* and ϕ^* are given in Proposition 2.

Unlike the implementation in Biais et al (2007), there is a new term in the cash reserve dynamics (30)

$$\left[\sigma - \phi^*(\lambda M_t)/\lambda\right] dB_t^1 = \left[\sigma - \phi^*(W_t)/\lambda\right](dY_t - \mu dt)/\sigma.$$

The interpretation of the other terms are the same as in the implementation of Biais et al (2007). We interpret the new term as special dividends paid only to the outside equity holders. Note that this term can be negative and we interpret it as cash injection through equity issues as in Leland (1994) style models.[10] The expected value of special dividends is equal to zero under the agent's belief P^1.

We can rewrite the cash reserves dynamics as

$$dM_t = \gamma M_t dt + \frac{\phi^*(\lambda M_t)}{\lambda} dB_t^1 - \frac{1}{\lambda} dC_t^*,$$

and use (21) to show that $M_t = W_t/\lambda$. We can also check that $W_t = E_t^{P^1}\left[\int_t^\tau e^{-\gamma(s-t)} dC_s^*\right]$. Thus the above capital structure is incentive compatible and implements the robust contract.

By Proposition 2, when the agent's continuation value W_t is small, $\phi^*(W_t) = \sigma\lambda$. But when W_t is large, $\phi^*(W_t) > \sigma\lambda$. Thus special dividends occur only when cash reserves $M_t = W_t/\lambda$ are sufficiently large. In this case, when the project performs well (i.e., $dB_t^1 > 0$), outside equity holders inject cash in the firm in order to raise cash reserves.[11] But when the project performs bad (i.e., $dB_t^1 < 0$), outside equity holders receive positive special dividends. This payout policy is used to hedge against model uncertainty so that cash reserves track the agent's continuation value, i.e., $M_t = W_t/\lambda$. This ensures that the liquidation time and the payout time coincide with those in the robust contract.

Following Bolton, Chen, and Wang (2013), we can interpret positive special dividends as equity repurchases because our model does not have taxes or other features to distinguish between these two. Because the firm's market value is high (low) when $dB_t^1 > (<)\, 0$, our model predicts that the firm times the market by issuing equity when its market value is

[10] In the model of DeMarzo et al. (2012), dividends can be negative and they provide the same interpretation.

[11] In the Leland (1994) model, equity holders inject capital through new equity issues for the purpose of avoiding costly bankruptcy.

high and repurchasing equity when its market value is low. This result is consistent with the empirical evidence in Baker and Wurgler (2002) and references cited therein.

4.2 Asset Prices

To price securities, we consider a representative agent equilibrium setting. The principal represents a representative investor who trades equity and debt with the payoffs given in the previous subsection, while the agent represents insiders restricted from trading. Suppose that the risk-free rate is exogenously given at r. We price securities using the principal's pricing kernel which is based on his worst-case belief Q^{h^*}. Specifically, equity value per share is given by

$$S_t = E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} \frac{1}{\lambda} dC_s^* + \frac{1}{1-\lambda} \int_t^\tau e^{-r(s-t)} \left(\sigma - \frac{\phi^* (\lambda M_s)}{\lambda} \right) dB_s^1 \right],$$

where $\tau = \inf \{ t \geq 0 : M_t = 0 \}$ is the liquidation time. By a similar analysis in Anderson, Hansen and Sargent (2003), we can show that the principal's fear of model misspecification generates a market price of model uncertainty. This market price of model uncertainty is given by $-h^* (\lambda M_t)$, where $h^* (\lambda M_t)$ is the worst-case density generator in (20).

Proposition 5 *The local expected equity premium under the measure P^1 is given by*

$$\frac{1}{1-\lambda} \left[\sigma - \frac{\phi^* (\lambda M_t)}{\lambda} \right] \frac{-h^* (\lambda M_t)}{S(M_t)} + \frac{\phi^* (\lambda M_t)}{\lambda} \frac{S'(M_t)}{S(M_t)} \left[-h^* (\lambda M_t) \right], \tag{31}$$

for $M_t \in \left[0, \bar{W}/\lambda \right]$, where h^* and ϕ^* are given in Proposition 2 and $S_t = S(M_t)$ for a function S given in Appendix B1.

The equity premium contains two components. The first component is due to the exposure of special dividends to the Brownian motion uncertainty. This component is negative because the factor loading $\left[\sigma - \phi^* (\lambda M_t) / \lambda \right] / (1-\lambda) < 0$ and special dividends are intertemporal hedges. The second component is due to the exposure of the stock price to the Brownian motion uncertainty. This component is positive whenever $S'(M) > 0$. Since the first component is zero for sufficiently small values of M, the equity premium is positive for these values. In all our numerical examples below, we find that the hedge component is small so that the equity premium is also positive for high values of M and approaches zero when $M = \bar{W}/\lambda$ (since $h^* \left(\bar{W} \right) = 0$).

Debt value satisfies

$$D_t = E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} \left[\mu - (\gamma - r) M_s \right] ds + e^{-r(\tau - t)} L \right].$$

The credit yield spread Δ_t is defined as

$$\int_t^\infty e^{-(r+\Delta_t)(s-t)}ds = E_t^{Q^{h^*}}\left[\int_t^\tau e^{-r(s-t)}ds\right].$$

Solving yields $\Delta_t = rT_t/(1-T_t)$, where $T_t = E_t^{Q^{h^*}}\left[e^{-r(\tau-t)}\right]$ for all $t \in [0,\tau]$ represents the Arrow-Debreu price at time t of one unit claim paid at the time of default.

The following result is similar to Proposition 6 in Biais et al (2007).

Proposition 6 *At any time $t \geq 0$, the following holds:*

$$D_t + (1-\lambda)S_t = F(W_t) + M_t - \frac{\theta}{2}E_t^{P^1}\left[\int_t^\tau e^{-r(s-t)}z_s h^*(W_s)^2\,ds\right]. \tag{32}$$

The left-hand side of (32) is the market value of outside securities, i.e., the present value of the cash flows these securities will distribute. The right-hand side of (32) represents the assets generating these cash flows. The last term is the entropy cost, which is subtracted to obtain the operating cash flows allocated to the principal (outside investors),

$$E_t^{Q^{h^*}}\left[\int_t^\tau e^{-r(s-t)}(dY_s - dC_s^*) + e^{-r(\tau-t)}L\right].$$

4.3 Empirical Implications

In Appendix B we show that the stock price, equity premium, debt value, and credit yield spreads are functions of the state variable, the level of cash reserves M. Figure 4 plots these functions for three values of θ. The benchmark model corresponds to $\theta = \infty$. The figure shows that the stock price is an increasing function of M, while the equity premium and the credit yield spread are decreasing functions of M. Moreover, the payout boundary $\bar{M}$ (or $\bar{W}$) increases with ambiguity aversion, implying that a more ambiguity averse principal wants to delay dividend payout in order to reduce the likelihood of liquidation. The principal's aversion to model uncertainty generates a positive equity premium, which approaches infinity as M goes to zero and decreases to zero as M rises to the payout boundary. The equity premium is high for financially distressed or recently established firms with low cash reserves. This also implies that the equity premium is high in recessions since cash reserves are low in bad times. Intuitively, when M is low, the incentive constraint binds and the ambiguity averse principal bears more uncertainty and hence demanding a higher equity premium. But when M is large, the agent can share the principal's uncertainty since the optimal sensitivity ϕ_t^* is state dependent. This leads the principal to bear less uncertainty, thereby reducing the equity premium.

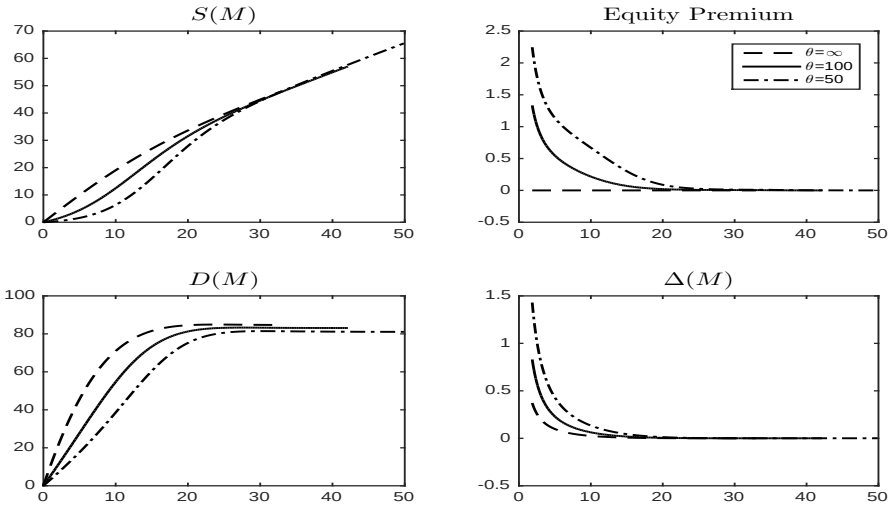

Figure 4: **Stock prices, equity premiums, debt values, and credit yield spreads for the robust contracting problem**. The parameter values are $\mu = 10$, $r = 0.10$, $\gamma = 0.15$, $\lambda = 0.20$, $\sigma = 5$, and $L = 0$.

Our model predicts that the equity premium and the credit yield spread are positively related: both are high for distressed firms with low cash reserves. This result is useful to test the relation between cross-sectional excess stock returns and distress risk (Fama and French (1993, 1995, 1996)). Our model prediction is consistent with the empirical evidence documented by Friewald, Wagner and Zechner (2014) that firms' equity premiums increase with credit yield spreads estimated from CDS spreads. It merits emphasis that one has to be cautious to interpret our model's cross-sectional asset pricing implications, since our implementation is based on a one-on-one contract. While our key insights may carry over to an equilibrium setting with many firms, a thorough analysis should be an interesting future research topic and is beyond the scope of this paper.

Figure 4 also shows that debt value decreases with the ambiguity aversion parameter $1/\theta$, while the equity premium and the credit yield spread increase with $1/\theta$. Thus our model can generate both a high equity premium and a high credit yield spread in a unified framework by raising the ambiguity aversion parameter $1/\theta$. Though our model is too stylized to be confronted with the data, it can help explain the credit spread puzzle and the equity premium puzzle. Chen, Collin-Dufresne, and Goldstein (2009) study these two puzzles by comparing the Epstein-Zin preferences and the Campbell-Cochrane habit formation preferences. Our paper contributes to this literature by introducing robustness and ambiguity aversion.

To generate time-varying equity premium or credit yield spread, the existing literature typically introduces one of the following assumptions: time-varying risk aversion as in the

habit formation model of Campbell and Cochrane (1999), time-varying economic uncertainty combined with Epstein-Zin preferences as in the long-run risk model of Bansal and Yaron (2004), or regime-switching consumption and learning under ambiguity as in Ju and Miao (2012). By contrast, in our contracting model, investors are risk neutral with distorted beliefs, dividends are endogenous, and the driving state process is identically and independently distributed.

Interestingly, unlike in the benchmark model, here the equity price may not be a concave function of the cash reserves. This happens when $1/\theta$ is sufficiently large. In this case the stock price is convex for low levels of cash reserves and concave for high levels of cash reserves. This result is related to the non-concavity of the principal's value function.

To gain intuition, we rewrite the equity valuation equation (B.1) in Appendix B as

$$\frac{\left[\phi^{*}\left(\lambda M\right)\right]^{2}}{2\lambda^{2}}\frac{S''\left(M\right)}{S\left(M\right)}=r-\gamma M\frac{S'\left(M\right)}{S\left(M\right)}$$
$$+\frac{-h^{*}\left(\lambda M\right)}{\left(1-\lambda\right)S\left(M\right)}\left[\sigma-\frac{\phi^{*}\left(\lambda M_{s}\right)}{\lambda}\right]+\frac{-h^{*}\left(\lambda M\right)\phi^{*}\left(\lambda M\right)}{\lambda}\frac{S'\left(M\right)}{S\left(M\right)}.$$

The expression on the second line is the expected equity premium. When $\theta\to\infty$, $\phi^{*}\left(\lambda M\right)\to\sigma\lambda$ and $h^{*}\left(\lambda M\right)\to 0$ so that the expected equity premium goes to zero. Then the first line above reduces to the ODE in Biais et al (2007). Proposition 7 of their paper shows that equity value $S\left(M\right)$ is concave. However, when investors are ambiguity averse with a sufficiently small $\theta>0$, the equity premium is positive and large for a small M. This positive value on the second line above can make $S''\left(M\right)>0$ for a small M.

Intuitively, the marginal value to equity $S'\left(M\right)$ consists of three components. First, an increase in M pushes cash reserves away from the liquidation boundary. The associated marginal benefit decreases with M. Second, an increase in M pushes cash reserves closer to the payout boundary. This raises the frequency of dividend payout, but also depletes cash reserves. The associated marginal benefit decreases with the level of cash reserves M. Third, an increase in M from a low level pushes cash reserves closer to $\hat{W}/\lambda$, allowing equity holders to have better ambiguity sharing. This marginal benefit increases with M when M is small. This component is unique in our robust contracting problem and may dominate the other two components when M is sufficiently small. In this case $S'\left(M\right)$ increases with M so that S is convex. In addition, the equity premium decreases with M because investors bear less ambiguity. But when M is sufficiently large, the first two components dominate so that S is concave.

Finally, we examine the impact of the parameter λ, which describes the severity of the agency problem. We focus on the implications for the stock market performance such as stock

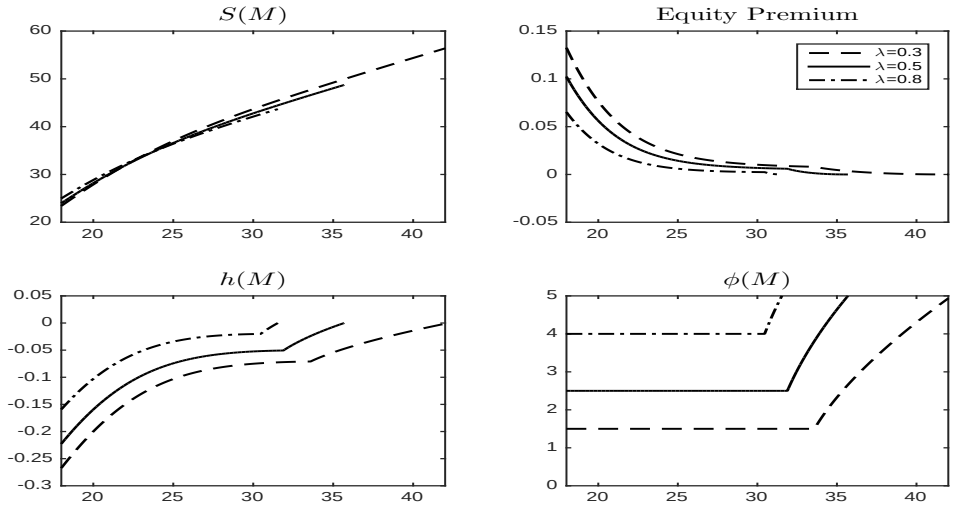

Figure 5: **Stock prices, equity premiums, belief distortions, and optimal sensitivities for the robust contracting problem**. The parameter values are $\theta = 50$, $\mu = 10$, $r = 0.10$, $\gamma = 0.15$, $\sigma = 5$, and $L = 0$. Low values of M are not displayed for a better view of graphs.

prices and returns. Figure 5 shows that there is no monotonic relation between stock prices and λ, but the expected equity premium decreases with λ. Intuitively, when λ is higher, the agency problem is more severe. To incentivize the agent, the principal should pay the agent more frequently, resulting in a smaller payout boundary. A more frequent dividend payout raises stock prices, but also depletes cash reserves and raises default risk, thereby reducing stock prices. The net effect is ambiguous. This result is different from that in Biais et al (2007) where stock prices decrease monotonically with λ. Figure 5 shows that their result holds in our model when cash reserves are sufficiently high. But when cash reserves are sufficiently low, the preceding positive effect dominates so that stock prices increase with λ. For empirical work, one has to be cautious to interpret the parameter λ since it describes both the exogenous severity of the agency problem and the endogenous insider ownership. In an empirical study, Demsetz and Villalonga (2001) show that there is no systematic relation between ownership structure and stock prices once ownership structure is treated as an endogenous variable.

Why does the expected equity premium decrease with λ? As Figure 5 shows, for a higher λ, the agent is exposed more to the project cash flow uncertainty (i.e., $\phi(M)$ is larger). Thus investors (the principal) are exposed less to uncertainty and hence the expected equity premium and the market price of uncertainty ($-h(M)$) are lower. We can interpret firms with a smaller λ as those with better corporate governance. Then our result is consistent with the evidence, documented by Gompers, Isshi and Metrick (2003), that firms with better

corporate governance earn higher returns.

5 Concluding Remarks

Contracting problems involve at least two parties. Introducing ambiguity or robustness into such problems must consider which party faces ambiguity and what it is ambiguous about. In this paper we have focused on the case where the principal does not trust the distribution of the project cash flow chosen by the agent. But the agent trusts it. The principal is averse to model ambiguity. This case is particularly interesting because it generates time-varying equity premium and has interesting asset pricing implications. In particular, the equity premium and the credit yield spread are high for distressed firms with low cash reserves.

One may argue that ambiguity aversion is similar to risk aversion. As in Hansen et al (2006), we establish a limited observational equivalence result in the online appendix, which states that our robust contracting problem delivers similar results to those in a contracting problem where the principal's preferences are represented by a special class of recursive utility (Epstein and Zin (1989) and Duffie and Epstein (1992)). Although many models of ambiguity in decision theory admit some form of observational equivalence to standard expected utility, we believe that the interpretation based on ambiguity aversion and robustness helps us understand many empirical puzzles in corporate finance and asset pricing. For example, from a quantitative point of view, an implausibly high risk aversion parameter is often needed to explain a high equity premium. But the high equity premium could be due to ambiguity aversion instead of high risk aversion (Chen and Epstein (2002), Hansen and Sargent (2010), and Ju and Miao (2012)). In our model, the principal is risk neutral, but his concerns about robustness generate an incentive for him to share model uncertainty with the agent. Unlike risk sharing induced by risk aversion, ambiguity sharing is caused essentially by the endogenous belief heterogeneity between the principal and the agent (also see Epstein and Miao (2003)).

In the online appendix we also study a contracting problem with risk aversion. We suppose that the principal has time-additive expected exponential utility and show that the solution to this contracting problem is similar to ours under some special assumptions. In future research it would be interesting to incorporate both risk aversion and ambiguity aversion and study how they interact with each other. Such an analysis would be mathematically complicated because the principal's wealth level will be an additional state variable in addition to the agent's continuation value. For analytical tractability, special assumptions are needed to remove one state variable (e.g., Maenhout (2004)).

Another direction for future research is to consider other types of ambiguity. For example,

the agent may face ambiguity about the project cash flows or both the principal and the agent may face ambiguity. Our paper focuses on contracting problems under moral hazard with binary actions. It would be interesting to generalize our analysis to a more general principal-agent problem such as that in Sannikov (2008). Finally, it would also be interesting to extend our approach to dynamic contracts with hidden information and study robust mechanism design problems in continuous time.

Appendices

A Proofs

Proof of Proposition 1: Define $\mathcal{H}^a$ as the set of density generators associated with an effort process a. Let $Q^h \in \mathcal{P}^a$ be the measure induced by $h \in \mathcal{H}^a$. Define $\Gamma(w)$ as the set of progressively measurable processes (ϕ, C, a) such that (i) ϕ satisfies

$$E^{Q^h}\left[\int_0^t \left(e^{-\gamma s}\phi_s\right)^2 ds\right] < \infty \text{ for all } t > 0,$$

(ii) C is increasing, right continuous with left limits and satisfies

$$E^{Q^h}\left[\left(\int_0^t e^{-rs} dC_s\right)^2\right] < \infty \text{ for all } t > 0,$$

(iii) $a_t \in \{0, 1\}$, and (iv) (W_t) satisfies (10), with boundary conditions $W_0 = w$ and $W_t = 0$ for $t \geq \inf\{t \geq 0 : W_t \leq 0\}$. For any $(\phi, C, a) \in \Gamma(w)$ and $h \in \mathcal{H}^a$, define the principal's objective function as

$$\begin{aligned} & J(\phi, C, a, h; w) \\ = \; & E^{Q^h}\left[\int_0^\tau e^{-rt}(dY_t - dC_t) + e^{-r\tau}L\right] + \frac{\theta}{2}E^{P^a}\left[\int_0^\tau e^{-rt} z_t h_t^2 dt\right] \\ = \; & E^{Q^h}\left[\int_0^\tau e^{-rt}(\mu a_t dt + h_t\sigma dt - dC_t) + e^{-r\tau}L\right] + \frac{\theta}{2}E^{Q^h}\left[\int_0^\tau e^{-rt} h_t^2 dt\right], \end{aligned} \quad (A.1)$$

where we have used the fact that $dY_t = \mu a_t dt + h_t\sigma dt + \sigma dB_t^h$ and B_t^h is a standard Brownian motion under the measure Q^h. We can then write the first-best robust contracting problem as

$$F(w) = \sup_{(\phi,C,a)\in\Gamma(w)} \inf_{h\in\mathcal{H}^a} J(\phi, C, a, h; w), \; w \geq 0. \quad (A.2)$$

Define an operator as

$$\mathcal{D}^{(\phi,a,h)}F(W) = \mu a + \sigma h + F'(W)(\gamma W + h\phi - \lambda\mu(1-a)) + \frac{F''(W)}{2}\phi^2 + \frac{\theta h^2}{2}. \quad (A.3)$$

We can describe the optimality conditions stated in the proposition as variational inequalities:

$$0 = \min\left\{rF(W) - \sup_{a\in\{0,1\},\phi\in\mathbb{R}} \inf_{h\in\mathbb{R}} \mathcal{D}^{(\phi,a,h)}F(W), \; F'(W) + 1\right\}, \quad (A.4)$$

for all $W \geq 0$ and the boundary conditions are given in the proposition. In particular, we can check that under condition (15), the policies (ϕ^*, a^*, h^*) stated in the proposition satisfy

$$rF(W) = \sup_{a\in\{0,1\},\phi\in\mathbb{R}} \inf_{h\in\mathbb{R}} \mathcal{D}^{(\phi,a,h)} F(W) = \mathcal{D}^{(\phi^*,a^*,h^*)} F(W),$$

for $W \in \left[0, \bar{W}\right]$ and $F'(W) = -1$ for $W \geq \bar{W}$. By (18) and the boundary conditions, we can show that $rF\left(\bar{W}\right) = \mu - \gamma\bar{W}$. Thus, for $W \geq \bar{W}$,

$$\begin{aligned} rF(W) - \mathcal{D}^{(\phi^*,a^*,h^*)} F(W) &= rF\left(\bar{W}\right) - r\left(W - \bar{W}\right) - (\mu - \gamma W) \\ &= \mu - \gamma\bar{W} - r\left(W - \bar{W}\right) - (\mu - \gamma W) \\ &= (\gamma - r)\left(W - \bar{W}\right) \geq 0. \end{aligned}$$

We now show that F is the value function in five steps. **Step 1.** Define the following process:

$$G_t^{(\phi,C,a,h)} = \int_0^t e^{-rs}(dY_s - dC_s) + \theta\int_0^t e^{-rs}\frac{h_s^2}{2}ds + e^{-rt}F(W_t), \tag{A.5}$$

where (W_t) satisfies (10).

Step 2. Fix a process $h^\phi = (h_t^\phi)$ defined as

$$h_t^\phi \equiv -\frac{\phi_t F'(W_t) + \sigma}{\theta}.$$

By (20), $h_t^* \equiv h^*(W_t) = h_t^{\phi^*}$, where $\phi_t^* \equiv \phi^*(W_t)$. Consider any candidate choice $(\phi, C, a) \in \Gamma(w)$. By Ito's Lemma under Q^{h^*},

$$\begin{aligned} e^{rt}dG_t^{\left(\phi,C,a,h^\phi\right)} &= \mu a_t dt + \sigma h_t^\phi dt + \sigma dB_t^{h^\phi} - dC_t^c + \frac{\theta(h_t^\phi)^2}{2}dt \\ &\quad + F'(W_t)\left[\gamma W_t dt - dC_t^c - \lambda\mu(1-a_t)dt + h_t^\phi\phi_t dt + \phi_t dB_t^{h^\phi}\right] \\ &\quad + \frac{1}{2}F''(W_t)\phi_t^2 dt - rF(W_t)dt + \Delta F(W_t) - \Delta C_t \\ &= \left[\mathcal{D}^{\left(\phi_t,a_t,h_t^\phi\right)}F(W_t) - rF(W_t)\right]dt - \left(1 + F'(W_t)\right)dC_t^c \\ &\quad + \left(\sigma + \phi_t F'(W_t)\right)dB_t^{h^\phi} + \Delta F(W_t) - \Delta C_t, \end{aligned}$$

where C^c is the continuous part of C, $\Delta C_t = C_t - C_{t-}$ is the jump, and $\Delta F(W_t) = F(W_t) -$

$F(W_{t-})$. By the variational inequalities (A.4) and $dC_t^c \geq 0$,

$$\begin{aligned} \left(1 + F'(W_t)\right) dC_t^c &\geq 0, \\ \mathcal{D}^{\left(\phi_t, a_t, h_t^\phi\right)} F(W_t) - rF(W_t) &\leq \mathcal{D}^{\left(\phi_t^*, a_t^*, h_t^{\phi^*}\right)} F(W_t) - rF(W_t) \leq 0, \end{aligned}$$

where $a_t^* \equiv a^*(W_t) = 1$ and the first inequality on the second line follows from condition (15). Since $F'(W_s)$ is bounded on $\left[0, \bar{W}\right]$ and $F'(W_s) = -1$ on $[\bar{W}, \infty)$,

$$E^{Q^{h^\phi}}\left[\int_0^t e^{-rs}\left(\sigma + \phi_s F'(W_s)\right) dB_s^{h^\phi}\right] = 0.$$

Since $F'(W) \geq -1$ for $W \geq 0$,

$$\Delta F(W_t) - \Delta C_t = F(W_t - \Delta C_t) - F(W_t) - \Delta C_t = -\int_{W_t - \Delta C_t}^{W_t} \left[F'(c) + 1\right] dc \leq 0.$$

It follows that $G_t^{\left(\phi, C, a, h^\phi\right)}$ is a $\left(Q^{h^\phi}, \mathcal{F}_t\right)$-supermartingale. This implies that $G_0^{\left(\phi, C, a, h^\phi\right)} \geq E^{Q^{h^\phi}}\left[G_{t\wedge\tau}^{\left(\phi, C, a, h^\phi\right)}\right]$ for any finite time $t \geq 0$. Taking limit as $t \to \infty$, we have

$$G_0^{\left(\phi, C, a, h^\phi\right)} \geq E^{Q^{h^\phi}}\left[G_\tau^{\left(\phi, C, a, h^\phi\right)}\right] \geq \inf_{h \in \mathcal{H}^a} E^{Q^h}\left[G_\tau^{(\phi, C, a, h)}\right].$$

Taking supremum for $(\phi, C, a) \in \Gamma(w)$ and using (A.5), we obtain

$$\begin{aligned} F(w) &= F(W_0) = G_0^{\left(\phi, C, a, h^\phi\right)} \geq \sup_{(\phi, C, a) \in \Gamma(w)} E^{Q^{h^\phi}}\left[G_\tau^{\left(\phi, C, a, h^\phi\right)}\right] \\ &\geq \sup_{(\phi, C, a) \in \Gamma(w)} \inf_{h \in \mathcal{H}^a} E^{Q^h}\left[G_\tau^{(\phi, C, a, h)}\right]. \end{aligned}$$

Step 3. Fix (ϕ^*, C^*, a^*) and consider any process $(h_t) \in \mathcal{H}^{a^*}$. Use Ito's Lemma to derive

$$\begin{aligned} e^{rt} dG_t^{(\phi^*,C^*,a^*,h)} &= \mu a_t^* dt + \sigma h_t dt + \sigma dB_t^h - dC_t^{*c} + \frac{\theta h_t^2}{2} dt \\ &\quad + F'(W_t)\left[\gamma W_t dt - dC_t^{*c} - \lambda\mu(1-a_t^*)dt + h_t\phi_t^* dt + \phi_t^* dB_t^h\right] \\ &\quad + \frac{1}{2}F''(W_t)\phi_t^{*2} dt - rF(W_t)dt + \Delta F(W_t) - \Delta C_t^* \\ &= \left[\mathcal{D}^{(\phi_t^*,a_t^*,h_t)}F(W_t) - rF(W_t)\right]dt - \left(1+F'(W_t)\right)dC_t^{*c} \\ &\quad + \left(\sigma + \phi_t^* F'(W_t)\right)dB_t^h + \Delta F(W_t) - \Delta C_t^* \\ &\geq \left[\mathcal{D}^{(\phi_t^*,a_t^*,h_t^*)}F(W_t) - rF(W_t)\right]dt - \left(1+F'(W_t)\right)dC_t^{*c} \\ &\quad + \left(\sigma + \phi_t^* F'(W_t)\right)dB_t^h + \Delta F(W_t) - \Delta C_t^*. \end{aligned}$$

Note that $\mathcal{D}^{(\phi_t^*,a_t^*,h_t^*)}F(W_t) - rF(W_t) = 0$. In addition, by (22),

$$\int_0^t e^{-rs}\left(1+F'(W_s)\right)dC_s^{c*} = \int_0^t e^{-rs}\left(1+F'(W_s)\right)\mathbf{1}_{\{W_s=\bar{W}\}}dC_s^{c*} = 0.$$

We also know that C^* jumps only at $t=0$ when $w > \bar{W}$ so that

$$\Delta F(W_0) - \Delta C_0^* = F\left(\bar{W}\right) - F(w) - \left(w - \bar{W}\right) = 0.$$

Thus $G_t^{(\phi^*,C^*,a^*,h)}$ is a $\left(Q^h, \mathcal{F}_t\right)$-submartingale. This implies that $G_0^{(\phi^*,C^*,a^*,h)} \leq E^{Q^h}\left[G_{t\wedge\tau}^{(\phi^*,C^*,a^*,h)}\right]$ for any finite time t. Taking limit as $t \to \infty$ yields

$$F(w) = G_0^{(\phi^*,C^*,a^*,h)} \leq E^{Q^h}\left[G_\tau^{(\phi^*,C^*,a^*,h)}\right].$$

Taking infimum for $h \in \mathcal{H}^{a^*}$ yields

$$F(w) \leq \inf_{h\in\mathcal{H}^{a^*}} E^{Q^h}\left[G_\tau^{(\phi^*,C^*,a^*,h)}\right] \leq \sup_{(\phi,C,a)\in\Gamma(w)} \inf_{h\in\mathcal{H}^a} E^{Q^h}\left[G_\tau^{(\phi,C,a,h)}\right].$$

Step 4. By Steps 2 and 3, we deduce that

$$\begin{aligned} F(w) &= \sup_{(\phi,C,a)\in\Gamma(w)} E^{Q^{h^\phi}}\left[G_\tau^{\left(\phi,C,a,h^\phi\right)}\right] = \inf_{h\in\mathcal{H}^{a^*}} E^{Q^h}\left[G_\tau^{(\phi^*,C^*,a^*,h)}\right] \\ &= \sup_{(\phi,C,a)\in\Gamma(w)} \inf_{h\in\mathcal{H}^a} E^{Q^h}\left[G_\tau^{(\phi,C,a,h)}\right]. \end{aligned}$$

Since $F(W_\tau) = L$, it follows from (A.5) that

$$\begin{aligned} F(w) &= \sup_{(\phi,C,a)\in\Gamma(w)} J\left(\phi, C, a, h^\phi; w\right) = \inf_{h\in\mathcal{H}^{a^*}} J(\phi^*, C^*, a^*, h; w) \\ &= \sup_{(\phi,C,a)\in\Gamma(w)} \inf_{h\in\mathcal{H}^a} J(\phi, C, a, h; w). \end{aligned}$$

Step 5. Evaluating at the processes (ϕ^*, C^*, a^*, h^*) induced by the policies described in Proposition 1, we show that

$$F(w) = J(\phi^*, C^*, a^*, h^*; w).$$

Consider first $w \in \left[0, \bar{W}\right]$. In this case C^* has no jump. As in Step 2, we can easily check that $G_t^{(\phi^*,C^*,a^*,h^*)}$ is a $\left(Q^{h^*}, \mathcal{F}_t\right)$-martingale. Thus

$$F(w) = E^{Q^h}\left[G_\tau^{(\phi^*,C^*,a^*,h^*)}\right] = J(\phi^*, C^*, a^*, h^*; w).$$

Consider next $w > \bar{W}$. In this case $C_0^* = w - \bar{W}$ and the agent's initial continuation value jumps to $\bar{W}$. By definition of J and the previous case, we can derive

$$J(\phi^*, C^*, a^*, h^*; w) = -\left(w - \bar{W}\right) + J\left(\phi^*, C^*, a^*, h^*; \bar{W}\right) = -\left(w - \bar{W}\right) + F\left(\bar{W}\right) = F(w).$$

We conclude that (ϕ^*, C^*, a^*, h^*) is optimal for the first-best robust contracting problem. The boundary condition (19) follows from the fact that the principal can deliver $W = 0$ to the agent who always exerts high effort and never gets paid. In this case $W = 0$ is an absorbing state and the principal obtains the value $\mu/r - \sigma^2/(2r\theta)$. If condition (17) holds, the project is never liquidated, i.e., $\tau = \infty$. Q.E.D.

Proof of Propositions 2 and 3: Define J as in (A.1). We modify condition (iii) in the definition of the feasible set $\Gamma(w)$ to incorporate the incentive constraint as follows: if $a_t = 0$, then $\phi_t \leq \sigma\lambda$ and if $a_t = 1$, then $\phi_t \geq \sigma\lambda$. The optimality condition described in Propositions 2 and 3 can be summarized by the following variational inequalities:

$$0 = \min\left\{rF(W) - \sup_{(a,\phi)\in\Lambda} \inf_{h\in\mathbb{R}} \mathcal{D}^{(\phi,a,h)} F(W),\ F'(W) + 1\right\}, \tag{A.6}$$

for all $W \geq 0$, where

$$\Lambda = \{(0,\varphi) : \varphi \leq \sigma\lambda\} \cup \{(1,\varphi) : \varphi \geq \sigma\lambda\}.$$

The boundary conditions are given in Proposition 2. It is easy to verify that under conditions (15) and (29), $a^*(W) = 1$, $\phi^*(W)$, and $h^*(W)$ described in Proposition 2 achieves the preceding maxmin. Unlike in the first-best contract, delivering the agent $W = 0$ who always exerts high effort is not incentive compatible. Thus the principal will liquidate the project at $W = 0$ so that $F(0) = L$. The rest of the proof is the same as that for Proposition 1. Q.E.D.

Proof of Proposition 4: We adapt Lemma 6 in DeMarzo and Sannikov (2006). We use the Envelope Theorem to differentiate ODE (26) with respect to θ to obtain

$$\begin{aligned} r\frac{\partial F(W)}{\partial \theta} &= \frac{\partial F'(W)}{\partial \theta}\gamma W + \frac{\phi^*(W)^2}{2}\frac{\partial F''(W)}{\partial \theta} + \frac{\left[\phi^*(W)F'(W)+\sigma\right]^2}{2\theta^2} \\ &\quad -\frac{\left[\phi^*(W)F'(W)+\sigma\right]\phi^*(W)}{\theta}\frac{\partial F'(W)}{\partial \theta}. \end{aligned}$$

Under measure Q^{h^*}, it follows from (20) and (21) that (W_t) satisfies

$$dW_t = \gamma W_t dt - dC_t^* - \frac{\left[\phi^*(W_t)F'(W_t)+\sigma\right]\phi^*(W_t)}{\theta}dt + \phi^*(W_t)\,dB_t^{h^*},$$

where $\left(B_t^{h^*}\right)$ is a standard Brownian motion under Q^{h^*}. Using the Feynman-Kac formula, we obtain that the solution to the above ODE for $\partial F(W)/\partial\theta$ is

$$\frac{\partial F(W)}{\partial \theta} = E^{Q^{h^*}}\left[\int_t^\tau e^{-r(s-t)}\frac{\left[\phi^*(W_s)F'(W_s)+\sigma\right]^2}{2\theta^2}ds|W_t = W\right] \geq 0,$$

as desired. Q.E.D.

Proof of Proposition 5: The equity premium is defined as

$$\frac{1}{S_t}\left(\underbrace{\frac{dC_t^*}{\lambda} + \frac{1}{1-\lambda}\left[\sigma - \frac{\phi^*(W_t)}{\lambda}\right]dB_t^1}_{\text{dividends}} + \underbrace{dS_t}_{\text{capital gains}} - rS_t dt\right). \tag{A.7}$$

By Ito's Lemma,

$$\begin{aligned} dS_t &= dS(M_t) = S'(M_t)\gamma M_t dt + S'(M_t)\frac{\phi^*(\lambda M_t)}{\lambda}dB_t^1 \\ &\quad -\frac{S'(M_t)}{\lambda}dC_t^* + \frac{\left[\phi^*(\lambda M_t)\right]^2}{2\lambda^2}S''(M_t)\,dt. \end{aligned} \tag{A.8}$$

Plugging (A.8) and (B.1) into (A.7) and noting the fact that C_t increases only when $S'(M_t) = 1$, we can compute the local expected equity premium under measure P^1 given in the propo-

sition. Q.E.D.

Proof of Proposition 6: It follows from (21) and Girsanov's Theorem that

$$dW_t = \gamma W_t dt - dC_t^* + \phi^*(W_t) h^*(W_t) dt + \phi^*(W_t) dB_t^{h^*}.$$

By Ito's Lemma,

$$\begin{aligned} e^{-rT\wedge\tau} W_{T\wedge\tau} &= e^{-rt} W_t + \int_t^{T\wedge\tau} e^{-rs} (\gamma - r) W_s ds + \int_t^{T\wedge\tau} e^{-rs} \phi^*(W_s) dB_s^{h^*} \\ &\quad - \int_t^{T\wedge\tau} e^{-rs} dC_s^* + \int_t^{T\wedge\tau} e^{-rs} \phi^*(W_s) h^*(W_s) ds, \end{aligned}$$

for any $T > t$, where $\tau = \inf\{t \geq 0 : W_t = 0\}$. Taking expectations with respect to Q^{h^*} and letting $T \to \infty$, we use $M_t = W_t/\lambda$ and $W_\tau = 0$ to derive

$$M_t = E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} \left(\frac{1}{\lambda} dC_s^* - (\gamma - r) M_s ds - \frac{\phi^*(W_s) h^*(W_s)}{\lambda} ds \right) \right].$$

It follows that

$$\begin{aligned} & D_t + (1-\lambda) S_t \\ = \; & E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} (\mu - (\gamma - r) M_s) ds + e^{-r(\tau-t)} L \right] \\ & + E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} \frac{1-\lambda}{\lambda} dC_s^* \right] + E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} \left(\sigma - \frac{\phi^*(W_s)}{\lambda} \right) dB_s^1 \right] \\ = \; & E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} (dY_s - dC_s^*) + e^{-r(\tau-t)} L \right] \\ & + E_t^{Q^{h^*}} \left[\int_t^\tau e^{-r(s-t)} \left(\frac{1}{\lambda} dC_s^* - (\gamma - r) M_s ds - \frac{\phi^*(W_s) h^*(W_s)}{\lambda} ds \right) \right], \end{aligned}$$

where we have used the fact that $dY_t = \mu dt + \sigma h^*(W_t) + \sigma dB_t^{h^*}$. The expression on the last last line is M_t. We then obtain the desired result. Q.E.D.

B Asset Pricing Formulas

In this appendix we follow DeMarzo and Sannikov (2006) and Biais et al (2007) to represent asset prices as ODEs. We use the cash reserves M as a state variable and write debt value, equity price and credit yield spreads as functions of M. Under the worst-case belief Q^{h^*}, we

use Girsanov's Theorem to write the cash reserve dynamics as

$$dM_t = \gamma M_t dt + \frac{\phi^*(\lambda M_t)\, h^*(\lambda M_t)}{\lambda} dt + \frac{\phi^*(\lambda M_t)}{\lambda} dB_t^{h^*} - \frac{1}{\lambda} dC_t^*,$$

where ϕ^*, h^*, and C^* are given by (27), (20), and (22), respectively. Thus the equity price $S_t = S(M_t)$ satisfies the ODE

$$\begin{aligned} rS(M) &= \frac{1}{1-\lambda}\left[\sigma - \frac{\phi^*(\lambda M_s)}{\lambda}\right] h^*(\lambda M) \\ &\quad + \left(\gamma M + \frac{\phi^*(\lambda M)\, h^*(\lambda M)}{\lambda}\right) S'(M) + \frac{[\phi^*(\lambda M)]^2}{2\lambda^2} S''(M), \end{aligned} \tag{B.1}$$

with the boundary conditions $S(0) = 0$ and $S'\left(\bar{W}/\lambda\right) = 1$.

The bond price $D_t = D(M_t)$ satisfies the ODE

$$rD(M) = \mu - (\gamma - r) M + \left(\gamma M + \frac{\phi^*(\lambda M)\, h^*(\lambda M)}{\lambda}\right) D'(M) + \frac{[\phi^*(\lambda M)]^2}{2\lambda^2} D''(M),$$

with boundary conditions $D(0) = L$ and $D'\left(\bar{W}/\lambda\right) = 0$. The Arrow-Debreu price of one unit claim paid at the time of default, $T_t = T(M_t)$, satisfies the ODE

$$rT(M) = \left(\gamma M + \frac{\phi^*(\lambda M)\, h^*(\lambda M)}{\lambda}\right) T'(M) + \frac{[\phi^*(\lambda M)]^2}{2\lambda^2} T''(M),$$

subject to the boundary conditions $T(0) = 1$ and $T'\left(\bar{W}/\lambda\right) = 0$.

References

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent, 2003, A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection, Journal of the European Economic Association, 1, 68-123.

Baker, M., and J. Wurgler, 2002, Market Timing and Capital Structure, Journal of Finance 57, 1-32.

Bansal, Ravi and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, Journal of Finance 59, 1481-1509.

Biais, Bruno, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet, 2007, Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications, Review of Economic Studies 74, 345-390.

Bolton, Patrick, Hui Chen and Neng Wang, 2013, Market Timing, Investment, and Risk Management, Journal of Financial Economics 109, 40-62.

Campbell, John Y., and John H. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, Journal of Political Economy 107, 205-251.

Carroll, Gabriel, 2015, Robustness and Linear Contracts, American Economic Review 105, 536-563.

Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein, 2009, On the Relation between Credit Spread Puzzles and the Equity Premium Puzzle, Review of Financial Studies 22, 3367-3409.

Chen, Zengjing and Larry G. Epstein, 2002, Ambiguity, Risk, and Asset Returns in Continuous Time, Econometrica 70, 1403–1443.

Clementi, Gian Luca, and Hugo A. Hopenhayn, 2006, A Theory of Financing Constraints and Firm Dynamics, Quarterly Journal of Economics 121, 229-265.

Cvitanic, Jaksa and Jianfeng Zhang, 2013, Contract Theory in Continuous-Time Models, Springer-Verlag, Berlin.

DeMarzo, Peter and Michael Fishman, 2007a, Agency and Optimal Investment Dynamics, Review of Financial Studies 20, 151-188.

DeMarzo, Peter and Michael Fishman, 2007b, Optimal Long-Term Financial Contracting, Review of Financial Studies 20, 2079-2128.

DeMarzo, Peter M., Michael Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic Agency and the q Theory of Investment, Journal of Finance 67, 2295-2340.

DeMarzo, Peter M. and Yuliy Sannikov, 2006, Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model, Journal of Finance 61, 2681–2724.

Demsetz, Harold, and Belen Villalonga, 2001, Ownership Structure and Corporate Performance, Journal of Corporate Finance 7, 209-233.

Duffie, Darrell and Larry G. Epstein, 1992, Stochastic Differential Utility, Econometrica 60, 353-394.

Eisdorfer, Assaf, 2008, Empirical Evidence of Risk Shifting in Financially Distressed Firms, Journal of Finance 63, 609-637.

Ellsberg, Daniel, 1961, Risk, Ambiguity and the Savage Axiom, Quarterly Journal of Economics 75, 643-669.

Epstein, Larry G. and Shaolin Ji, 2013, Ambiguous Volatility and Asset Pricing in Continuous Time, Review of Financial Studies 26, 1740-1786.

Epstein, Larry G. and Jianjun Miao, 2003, A Two-Person Dynamic Equilibrium under Ambiguity, Journal of Economic Dynamics and Control 27, 1253-1288.

Epstein, Larry G. and Tan Wang, 1994, Intertemporal Asset Pricing under Knightian Uncertainty, Econometrica 62, 283-322.

Epstein, Larry G. and Stanley Zin, 1989, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica 57, 937-969.

Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, Journal of Finance 50, 131-155.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, Journal of Finance 51, 55–84.

Fan, Ky, 1953, Minimax Theorems, Proceedings of National Academy of Science 39, 42–47.

Fleming, Wendell H. and H. Mete Soner, 1993, Controlled Markov Processes and Viscosity Solutions, Springer-Verlag, New York.

Fleming, W. H. and Panagiotis E. Souganidis, 1989, On the Existence of Value Functions of Two-player, Zero-sum Stochastic Differential Games, Indiana University Mathematics Journal 38, 293–314.

Friewald, Nils, Christian Wagner, and Josef Zechner, 2014, The Cross-Section of Credit Risk Premia and Equity Returns, Journal of Finance 69, 2419-2469.

Gilboa, Itzhak and David Schmeidler, 1989, Maxmin Expected Utility with Non-unique Prior, Journal of Mathematical Economics 18, 141–153.

Gompers, Paul, Joy Ishii, and Andrew Metrick, 2003, Corporate Governance and Equity Prices, Quarterly Journal of Economics 118, 107-156.

Hansen, Lars Peter, 2007, Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk, American Economic Review 97, 1-30.

Hansen, Lars Peter and Thomas J. Sargent, 2001, Robust Control and Model Uncertainty, American Economic Review 91, 60-66.

Hansen, Lars Peter and Thomas J. Sargent, 2008, Robustness, Princeton University Press, Princeton and Oxford.

Hansen, Lars Peter and Thomas J. Sargent, 2010, Fragile Beliefs and the Price of Uncertainty, Quantitative Economics 1, 129–162.

Hansen, Lars Peter and Thomas J. Sargent, 2012, Three Types of Ambiguity, Journal of Monetary Economics 59, 422-445.

Hansen, Lars Peter, Thomas J. Sargent, Gauhar A. Turmuhambetova, and Noah Williams, 2006, Robust Control and Model Misspecification, Journal of Economic Theory 128, 45-90.

Jensen, Michael C., and William H. Meckling, 1976, Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, Journal of Financial Economics 3, 305–360.

Ju, Nengjiu and Jianjun Miao, 2012, Ambiguity, Learning, and Asset Returns, Econometrica 80, 559-591.

Karatzas, Ioannis and Steven E. Shreve, 1991, Brownian Motion and Stochastic Calculus, 2nd Edition, Springer-Verlag, New York.

Leary, Mark T. and Roni Michaely, 2011, Determinants of Dividend Smoothing: Empirical Evidence, Review of Financial Studies 24, 3197-3249.

Leland, Hayne, 1994, Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, Journal of Finance 49, 1213–1252.

Lintner, John, 1956, Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes, American Economic Review 46, 97–113.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006a, Ambiguity Aversion, Robustness and the Variational Representation of Preferences, Econometrica 74, 1447-1498.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006b, Dynamic Variational Preferences, Journal of Economic Theory 128, 4-44.

Maenhout, Pascal J., 2004, Robust Portfolio Rules and Asset Pricing, Review of Financial Studies 17, 951-983.

Sannikov, Yuliy, 2008, A Continuous-Time Version of the Principal-Agent Problem, Review of Economic Studies 75, 957–984.

Sannikov, Yuliy, 2013, Dynamic Security Design and Corporate Financing, in George M. Constantinides, Milton Harris, and Rene M. Stulz eds, Handbook of the Economics of Finance, Volume 2A, North Holland.

Strzalecki, Tomasz, 2011, Axiomatic Foundations of Multiplier Preferences, Econometrica 79, 47-73.

Szydlowski, Martin, 2012, Ambiguity in Dynamic Contracts, working paper, Northwestern University.

Williams, Noah, 2009, On Dynamic Principal-Agent Problems in Continuous Time, working paper, University of Wisconsin at Madison.

Williams, Noah, 2011, Persistent Private Information, Econometrica 79, 1233-1274.