Job Market Paper

CONTINUOUS-TIME CONTRACTING WITH AMBIGUOUS PERCEPTIONS

MARTIN DUMAV
EUROPEAN UNIVERSITY INSTITUTE

Abstract. This paper examines a principal-agent problem in continuous time with ambiguous information. A problem of this nature arises in an employment relationship where there is limited knowledge, or ambiguity, about the technology that governs performance. To address this problem, this contribution connects models of the contracting problems in continuous time with models of decision making under ambiguity in continuous time. The connection preserves their tractability in analysis. By means of computed examples, I show that taking ambiguity into consideration results in compensation schemes that feature robustness and durability. Intuitively, when expectations about future outcomes are highly pessimistic, as in the worst-case scenario, the certainty of immediate payments becomes relatively more attractive. The unique optimal way for this preference for certainty to be mutually beneficial to parties with conflicting interests is to bind them together in a durable contract with little sensitivity to outcomes. By systematically incorporating ambiguity, the formulation proposed provides a possible rationale for simpler contracts, thereby responding to criticisms levelled at existing theories of contracts that predict compensation schemes that are unrealistically sensitive to performance.

1. Introduction

In contractual relationships, parties often do not have precise knowledge about the economic environment. Understanding of the role of contracts in such realistic environments is a central theme in economics. How do firms structure the payments and payments to workers when there is uncertainty about how their effort relates to performance? How do entrepreneurs find finance for their unique ideas? How do banks design venture capital contracts to motivate scientists towards a promising yet untested technology? Why do high-tech firms show high tolerance for failure? Why are real world contracts so simple? This paper develops methods and results that shed light on such problems.

Date: November 10, 2014
JEL Classification: D81, D86, M52.
Key words and phrases. Ambiguity, Contracts, Tolerance, Multiple Priors, Continuous-time stochastic methods.
Many thanks to Eduardo Faingold, Takashi Hayashi, Urmee Khan, David Levine, Ramon Marimon, Andrea Mattozzi, Marcin Peski, Frank Riedel, Mathias Staudlig and Gordan Zitkovic for help with this paper. I also thank Yuliy Sannikov and Borys Grochulski for sharing their computer codes. This research developed during my stay at the European University Institute as a Max Weber postdoctoral fellow: the hospitality of this institution is gratefully acknowledged. All remaining errors are mine.
Department of Economics, European University Institute, Villa San Paolo, Via della Piazzuola 43, 50133 Firenze, Email: martin.dumav at eui.eu.

In this paper, I formulate a continuous-time principal-agent problem that incorporates a rich structure for uncertainty and examines the dynamics of incentives. I identify the ways uncertainty influences incentives to the agent over time. I analyze the relative strength of immediate payments and delayed rewards in motivating agents. I also investigate the sensitivity of payments and rewards to the experience under an optimal contract. Furthermore, I analyze how the costs of creating incentives and the dynamic properties of the optimal contract depend on the nature of uncertainty.

I consider a dynamic contracting model where a principal and an agent engage in a contractual relationship with unobservable effort. The agent likes consumption and dislikes effort and the principal enjoys profits from output. The agent's unobservable effort input influences output with uncertainty. For a given effort, output evolves according to a Brownian motion with a drift term that belongs to a *set*. This is the main departure from most of the principal-agent literature. Rather than assuming that the map from effort to the evolution of output is known, as is common, I assume that this map is known only imprecisely. The parties know instead that the effort maps onto a set of probability distributions over output. I assume that each party does not have a prior belief about which outcomes are more likely, and I call this uncertainty 'ambiguity' (the use of terminology in relation to the decision theory literature is clarified below). Furthermore, the principal and the agent evaluate the technology according to their own worst case. One may view my model of ambiguity as a model of robustness in the presence of model uncertainty. My motivation for this choice is related to the model of robustness consideration developed by Hansen and Sargent [33, 34].

The principal and the agent share the same understanding of this ambiguous technology and each evaluates the technology according to his/her own worst-case scenario, which reflects robustness in their choice criteria. The special case in which the set has a single element reduces to the standard agency problem. At the beginning of the contracting relationship, the principal commits to a history-dependent contract that specifies wages to the agent at every moment in time contingent on the entire past output realizations. The principal only offers a contract that ex ante generates him positive profits under his worst case. The agent requires at least his outside value to participate in the contract at every moment and after every past history. I show that, under broad conditions, the unique optimal contract features simplicity: the agent's wage does not vary too sensitively with output. Moreover, the optimal contract features tolerance on the part of the principal: the contracting relationship has durability and does not end after early failures.

The main intuition for the new results, in broad terms, is that the sensitivity of delayed payments to output realization loses its incentive effects when parties in a contracting relationship are faced with ambiguity over technology. Pessimism over the expectation of uncertain future outcomes according to the worst case makes a certainty of immediate values in the relationship relatively more attractive. The unique optimal way for this preference for certainty to be beneficial to parties with conflicting interests is to bind them together into a durable contract with little sensitivity to outcomes.

The importance of the finding can be viewed in several different ways. First, the formulation proposed addresses a long-standing problem in contract theory, i.e., why, in the real world are simple contracts so common? Commonly studied agency models predict that performance-sensitive and contingent compensation should be common. However, contrary to these predictions real world contracts are simple.[1] Moreover, when tasks are complex and difficult to execute and describe, standard mechanisms are ineffective.[2] The pay-for-performance nature of standard agency contracts does not reflect the durability and tolerance of early failures that are commonly observed, especially in innovative firms with uncertain technologies.[3] The model presented here is rich enough to account for the tolerance for early failures and the rewards for long-term success that are commonly observed in firms operating in non-quantifiable uncertainty. Therefore, the modeling approach with ambiguity presented here thus offers a realistic description of these important problems and provides new explanations to help resolve them.

Secondly, the paper introduces a simple mathematical result that provides a methodological tool to address more complex dynamic incentive problems with tractability. I generalize the standard moral hazard problem into a technology with ambiguity by introducing a single simple additive term. This simplicity is due to the analytical advantage of the connection established here in a tractable way between two models in continuous time: the moral hazard problem with risk and decision theory under ambiguity. This connection uses the intuitive distinction made by Knight [45] between different forms of uncertainty: risk and ambiguity. Risk, as in the classical case, refers to the situation where there is a known probability distribution associated with each action. Ambiguity, on the other hand, as in our case, refers to a situation where the information is too imprecise to summarize likelihoods into a single probability distribution, and instead there is a *set* of probability distributions associated with each action. The classic work of Knight [45] and Keynes [43] emphasized the role of ambiguity in the real world decision making. In my model, I assume ambiguity about technology and consider behavioral response to it. Thirdly, the assumptions in the model that allow broader uncertainty are quite mild. In this way, the framework is flexible enough to be applied to the practical design of incentive contract where non-quantifiable uncertainty defines an important aspect of reality.

In the present contribution, I develop the framework in continuous time because the methods using it enables me to formulate the agency problem in a convenient

[1] See, for example, Bhattacharyya and Lafontaine [6], Chu and Sappington[12] and the references therein.

[2] See, for example, in the context of venture capital financing, Kaplan and Strömberg [42].

[3] The effective role of this dynamic property and the ineffectiveness of performance-sensitive incentive schemes in encouraging innovations was first raised by Holmstrom [39]. See Manso [47] for a recent treatment, and Tian and Wang [61] and the references therein for recent evidence on venture capital.

recursive manner and I can implement simple computational procedures to analyze the properties of the optimal contract relative to the different sources of uncertainty. My formulation leads to a simple ordinary differential equation that incorporates ambiguity as one additive term and discounts values for the worst case. This equation incorporates the key elements that determine optimal consumption, effort, and the duration of the relationship relative to the strength of ambiguity. I then extend the familiar methods of dynamic programming and show that the evolution of the optimal contract is completely characterized by using the drift and volatility of the agent's value in the relationship. As in familiar results in the literature on dynamic contracts, in my analysis I can completely characterize the optimal contract using the agent's continuation value as a single state variable. The agent's continuation value in a contract is the sum of future expected utilities. Unlike in the previous literature, ambiguity adds the consideration that the principal and the agent can a priori disagree on the worst case. In my analysis I show that wages are increasing in output and therefore both parties agree on the same worst-case. Finally, the solution offers insights into how the nature of the optimal contract varies according to the nature of the uncertainty.

The present study contributes to the literature on moral hazard problems. It is most closely related to Sannikov [54]'s model of moral hazard in continuous time and into his framework introduces ambiguity about technology. Using the formulation of ambiguous information and worst-case objectives from Chen and Epstein [11] the present model offers a tractable formulation of the agency problem. Similar as in the classical moral hazard problem the dynamics of contractual relationship are characterized by two terms: the drift and volatility of the agent's value. The former is related to the allocation of payments over time. The first novel finding is that the presence of ambiguity tempers the incentive effects of back-loaded payments. This is because ambiguity aversion has a direct effect, introducing one added term, which reduces the drift according to the worst case objective. The intuition for this result is that the agent prefers the certainty of wages received today over the expectations of uncertain future payments since the expectations are valued according to the worst case. This is a dynamic version of the intuition in static model of Gilboa and Schmeidler [28] that formalizes as a worst-case objective people's tendency to prefer known outcomes over ambiguous ones, which is illustrated in Ellsberg [25]'s classic experiment.

Ambiguity also operates on incentives through its effect on the volatility of the value. As in the classic case, a strong volatility of continuation value with effort incentivizes the agent to work hard. The second novel finding in the present work is that the presence of ambiguity and the agent's aversion to it, however, reduces incentive effect of this volatility by making smaller the marginal expected benefits of effort since the expectation is evaluated according to worst case. Therefore, ambiguity aversion acts like a effort cost. However, unlike the latter, the effect of ambiguity aversion is forward looking as it is jointly determined in the optimal contract by the worst-case scenario *and* the continuation value of the contract.

In the optimal contract, ambiguity averse principal trades-off volatility of future payments against certainty of immediate wages. In addition to the classic dynamic moral hazard problem concern for ambiguity introduces two terms in to the principals optimization problem in resolving this trade-off. The first is the penalty in output due to the worst case and the second is the cost of providing incentives to the agent through volatility in the continuation utility. The former has a direct effect on the profit and reduces the drift of profits according to the worst case due to the principal's ambiguity aversion. The latter has an indirect effect through the volatility in payments to the agent.

The trade-off in general depends on the nature of ambiguity, in particular whether ambiguity increases or decreases with effort. It turns out that in either case the direct effect dominates the indirect and principal's profits are lower under ambiguity relative to the case without ambiguity. Moreover in the optimal contract the compensation scheme becomes flatter and volatility smaller relative to the classical case. Intuitively, each party prefers the certainty of immediate payments and steady flow of values. Standard intuition in contract theory suggests that this form of scheme does not have strong incentive properties. Durability in a dynamic contract here provides incentives: a long relationship with stable payments to the agent for his steady effort that generates steady profits to the ambiguity averse principal, which he prefers. This is the main new explanation the current work offers.

The present paper is also related to a growing literature on dynamic contracting problems in continuous time,[4] and to the microeconomic literature that examines contracting problems and mechanism design in static settings (see Bergemann and Schlag [5], Bodoh-Creed [9], Gottardi *et al.* [30] Bergemann and Morris [4] and the references therein.) The latter work typically uses static models with adverse selection rather than moral hazard. Miao and Rivera [48] introduces robustness considerations into a dynamic contracting problem in continuous time. They focus on the principal's concern for robustness and their modeling of ambiguity builds on a model of multiplier preferences proposed by Anderson et al. [2] and Hansen et al. [35], while it differs from the model in Chen and Epstein [11] adopted here. Szydlowski [59] examines a dynamic contracting problem in continuous time with ambiguity. His model assumes ambiguity regarding the agent's effort cost and his preference representation differs from the model here. Discussion of the closer relationship to the literature on the foundations of contracts is delayed until after the full model and analysis of it is presented.

[4]Holmstrom and Milgrom [40], Schaettler and Sung [55], Ou-Yang [49], DeMarzo and Sannikov [19], Biais et al. [7], Biais et al. [8], Sannikov [54], Grochulski and Yuzhe [31, 32] He [37], He [38], Williams [62], Zhang [63], Piskorski and Tchistyi [51], Prat and Jovanovic [52], De Marzo et al. [18], Cvitanic and Zhang [15], Zhu [64], and Szydlowski [59]. This literature complements and extends the vast literature on dynamic contracts in discrete time including Spear and Srivastava [57], Thomas and Worrall [60], Atkeson and Lucas [3], Albuquerque and Hopenhayn [1], Clementi and Hopenhayn [13], Quadrini [53], DeMarzo and Fishman [17], and DeMarzo and Fishman [16].

The remainder of the paper proceeds as follows. Section 2 specifies the contracting problem in continuous time with ambiguous information. Section 3 formalizes the ambiguous information. Sections 4 specifies the utility values associated with a contract. Section 5 presents a derivation of a tractable incentive compatibility constraint under ambiguity. Section 6 derives the optimal contract and characterizes its properties through parametric examples. Section 7 discusses the close relationship of the new findings here to the existing literature. Section 8 concludes. Technical details are relegated to appendices in Section 9.

2. The contracting problem with ambiguous information

I present a model of a continuous-time principal/agent problem with ambiguous information. The agent chooses effort at each instant of time a_t from a compact set A_t. Following Sannikov [54], the choice of action process (a_t) determines the realization of output $\{X_t\}$ over time in a stochastic manner. Formally, I assume that the total output X_t produced up to time t evolves according to a diffusion process

$$dX_t = \theta_t^{a_t} dt + \sigma dB_t, \tag{1}$$

where $B = \{B_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is a standard Brownian motion under a reference measure P; as in standard moral hazard problems, the agent's choice of effort level a_t is privately observed; and the drift term $\theta_t^{a_t}$ belongs to a set $\Theta_t^{a_t}$. The latter is the main departure from the literature that studies dynamic contracting problems and I refer to it as ambiguous information. This generalizes the classical case to allow for imprecise information regarding technology: the productivity of actions is not perfectly known; rather, each party only knows that it lies in a set Θ_t^a.

To illustrate, consider a special case of stationary and independent set of drift terms so that $\theta^a \in \mu(a) + \kappa(a)U$, where μ maps effort into a set of drift terms as in the classic case, U is the unit interval $[-1, 1]$, and $\kappa(a)$ is the size of the set of drifts. From this special case, one can see that there is ambiguity about which value in U will be realized after each time t and history of output realization. In particular, any distribution over the interval U is possible and $\kappa(a)$ determines the strength of ambiguity. One interpretation of this specification is that the parties to the contract are aware of the possibility that they have erroneous probabilistic beliefs about the true technology $\mu(a)$ and seek robustness. The dependency of $\kappa(a)$ on the effort choice allows for general ways through which effort impacts the strength of ambiguity. Taking $\kappa(a) = 0$ for all a reduces to a singleton drift term and specializes to the classical contracting problem examined by Sannikov [54]. More generally, $\kappa(\cdot)$ can depend on time and history in arbitrary ways subject to mild technical restrictions due to measurability requirements. In this way, one can model large set of technological possibilities, including, for example, technologies in which ambiguity about actions repeatedly selected over time decreases as a result of experience and hence reflecting learning.

The problem I address here is the design of a contract by the principal when there is ambiguous information, in the sense as defined here regarding the contracting environment. Moreover, I assume that the agent and the principal have the same knowledge of the technology.[5] Our aim is to solve and characterize the optimal contract problem in this environment.

In the rest of this section I formulate the optimal choice of the contract by the principle as an optimization problem. First I represent the ambiguous information as a set of probabilities. Then using this set I specify an criterion for the evaluation of the contract by the parties.

Following Epstein and Chen [11] ambiguous information is equivalently formulated as a set of priors. The key observation is that a drift process (θ_t) is a density generator: it induces a probability measure Q^θ under which $\theta_t dt + \sigma dZ_t$ is a Brownian motion. This probability measure Q^θ is determined by its density with respect to the reference measure P using Girsanov exponential as follows

$$Q^\theta(\omega) = \exp\left\{\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t |\theta_s|^2 ds\right\} \cdot P(\omega). \tag{2}$$

Each effort process (a_t) induces a set of drift terms $\Theta^a = \{(\theta_t) : \theta_t \in \Theta^{a_t}\}$ and the corresponding set of priors is

$$\mathcal{P}^a = \{Q^\theta : \theta \in \Theta^a \text{ and } Q^\theta \text{ is defined by (2)}\}. \tag{3}$$

In other words, ambiguity concerns the drift of the diffusion process for output and parties to the contract do not know precisely which drift term governs the output today. From Girsanov transformation one can see that ambiguity about drift can equivalently be seen as specifying a set of probabilities for output realizations, as in the classic theme of ambiguity in atemporal models.

The principal offers a contract to the agent, which specifies a stream of payments for consumption (C_t) and an incentive-compatible advice of effort (a_t), both contingent only on the entire history of publicly observed output realization. Effort input generates output with ambiguity so that the set of priors $\mathcal{P}^a$ determines output realizations in a stochastic manner. I assume that each party to the contract evaluate the contract using the worst-case criterion, on which a priori parties do not have to agree. Accordingly, the principal contract offer maximizes his expected profit under his worst-case criterion

$$F = \min_{Q\in\mathcal{P}^a} E^Q\left[r\int_0^\infty e^{-rt}dX_t - r\int_0^\infty e^{-rt}C_t dt\right] \geq 0 \tag{4}$$

subject to delivering the agent a required outside value of at least $\hat{W}$ for all t

$$V_t(C,a) = \min_{Q\in\mathcal{P}^a} E^Q\left[r\int_t^\infty e^{-rt}\left(u(C_t) - h(a_t)\right)dt|\mathcal{F}_t\right] \geq \hat{W}$$

and subject to incentive-compatibility

$$V_t(C,a) \geq V_t(C,\widetilde{a})$$

for all t and for all alternative actions $\widetilde{a} \neq a$.

The interest is in contracts that generate non-negative expected profits for the principal. To gain tractability by exploiting the power of continuous-time formulation the next section sets up keys features of the set of multiple priors.

[5]We discuss briefly the role of this modeling choice in the conclusion.

3. The Set of Priors

The contracting problem as posed in a general form in (4) in Section 2 is difficult to solve. It allows for arbitrary dependence of uncertainty on the history of actions and outcomes. However, little is known about solution methods applicable at such generality. The tractability of the analysis of relies on representing the values in contracting problem in a recursive manner. To do so, the formulation in the present paper uses the decision theoretic model of choice under ambiguity in continuous time presented in Chen and Epstein [11]. Their recursive multiple-prior preference formulation in continuous time is a (non-axiomatic) generalization of that by Epstein and Schneider [26] in discrete-time, which generalizes atemporal model of maxmin expected utility representation of Gilboa and Schmeidler [28]. In this approach, utility functions are modeled as recursive ambiguity averse preferences, and, under mild assumptions on the set of priors, is represented recursively using continuation value as a state variable. Additionally, the concern for ambiguity is reflected in one additional term added to the martingale representation in the classical case and that term is determined by the worst-case. In this section I will introduce dependence of the set of multiple priors $\mathcal{P}_a$ on any effort process a and gradually establish that this can be done in a tractable manner.

following Chen and Epstein, I model the set of one-step-ahead densities for any effort process (a_t) via a process (Θ_t^a) of correspondences from Ω into its range $R^A \subset R$, that is, for each t

$$\Theta_t^A : \Omega \rightsquigarrow R^A.$$

The set of all measures that can be constructed by some selection from these sets of one-step-ahead densities is defined using the following set of density generators:

$$\Theta^a = \{(\theta_t) : \theta_t \in \Theta_t^a(\omega)\ dt \otimes dP_a\ a.e.\}. \tag{5}$$

Fixing an effort process to a constant, say zero, for each time t and ω specializes to the formulation in Chen and Epstein. This property on the multiple priors $\mathcal{P}$ in continuous time is a generalization of rectangularity introduced in Epstein and Schneider [26] in discrete time The interpretation of rectangularity is that at each instant of time, the increments in the diffusion process θ_t are independently drawn from a family Θ_t. As shown by Epstein and Schneider [26] in discrete time rectangularity of the multiple priors is consistent with the suitable axiomatization of preferences that satisfy Gilboa-Schmeidler [28] axioms conditional on each history realizations and the conditional preferences satisfy dynamic consistency. In their framework and in its generalization to continuous time by Chen and Epstein [11] rectangularity enables recursive relation in utilities by ensuring that repeated local minimizations over the set of one-step-ahead conditional measures replaces global minimization over $\mathcal{P}$.

To present a recursive formulation, the framework for the contracting relationship in this paper adopts the modeling approach of specifying ambiguous information as a rectangular set of multiple priors and generalize it to cases in which ambiguity varies with action choices. The latter is important for incentive compatibility as, in general, different efforts induce different multiple priors and the contracting parties might seek robustness to erroneous probabilities. In the formulation here, one-step-ahead densities for any action process are taken as the primitives for the analysis and can be specified arbitrarily. For any action process (a_t), as shown in Chen and Epstein, any rectangular set of priors $\mathcal{P}^a$ is uniquely generated by the set of densities of the form given in (5).

For the contracting problem at hand, unobservability of effort introduces a consideration for incentive compatibility which affects the structure of the multiple-priors that is not in general accommodated in Chen and Epstein's framework. In particular, the set of multiple priors changes if the agent deviates from the effort process that the principal desires to implement with the contract. This section therefore constructs a suitable generalization of the rectangularity of the set of multiple priors to an arbitrary effort process. For the tractability, this construction is done in three stages, reflecting the various ways in which effort effects the set of priors. These are variation in the base-line measure P_a, the variable interval size Θ_t^a, and both, as (a_t) changes.

The main results of this section show, using mainly Girsanov's theorem for changes of measures, these generalizations are possible while preserving the key properties of the set of priors, namely regularity that guarantees that the contracting problem is well-defined, and "dynamic consistency" that enables recursive representation. For reasons of simplicity we first consider fixing base-line measure P and changing the interval size. For the sake of easing the illustration I start with a particular case in which the set of drift terms are time and state invariant, is centered around zero, and depends on the effort: $\Theta_t^a(\omega) = [-\kappa(a), \kappa(a)]$ for each t and ω. Following the terminology introduced by Chen and Epstein we denote this case as κ ignorance with variable interval.

3.1. **kappa-ignorance with variable interval.** On the standard Wiener space $(\Omega, \mathcal{F}, P)$ the process (X_t) governing the agent's output is a Brownian motion. The agent's technology is described by the set of drifts induced by his choices. In a particular case examined by Chen and Epstein [11], the technology is characterized by $\kappa-$ignorance. In particular, the base-line measure is augmented by a family of measures using a process $\theta = (\theta_t)$ that determines the size of the interval which in the present sense captures (interpreted as) ambiguity associated with each choice of action. This is the simplest case that generalizes Chen and Epstein [11] while connecting with Sannikov [54]'s contracting problem. Consider first a base-line ambiguity with base-line action of no effort. That is, under P the process $dX_t = \sigma dW_t$ is a Brownian motion. Uncertainty is modeled as a family of Brownian motions following analogous ideas in Chen and Epstein [11]. In particular, drift terms belong in a time-invariant set and are represented by a process $\theta = (\theta_t)$ with $\theta_t \in [-\kappa, \kappa] = \mu(0) + [-\kappa, +\kappa]$. Incentive-compatibility consideration of a contract requires a comparison in a one-stage deviation sense and accordingly I consider a more specialized set for the drift terms that: $\Theta_0 := (\Theta_t)$ with $\Theta_\tau = [-\kappa, +\kappa]$ for $\tau \leq t$ and $\Theta_\tau = [-\widetilde{\kappa}, +\widetilde{\kappa}]$ for $\tau < t \leq T$. That is, up to a fixed time τ drift term is in set $[-\kappa, +\kappa]$ and after then in $[-\widetilde{\kappa}, \widetilde{\kappa}]$. This interpretation is based on the following characterization. Taking the supermartingale $Z_t = \exp\left(-\int_0^t \theta_s dW_s\right)$ and noticing that $\int_0^T ||\theta_t||^2 dt < \infty$ and that

$E[Z_T] = 1 < \infty$ so that the supermartingale is actually a martingale (under the measure P), by Girsanov's the change of variable [29] give that

$$\widetilde{W}_t = W_t - \int_0^t \theta_s ds, \quad \mathcal{F}_t, \quad 0 \le t \le T$$

is a Brownian motion that possibly has different drift after τ.

The set of probability measures $\mathcal{P}^{\Theta_0} := \{\widetilde{P}^\theta : \theta \in \Theta_0\}$ is equivalent to the base-line measure P, that is absolutely continuous with respect to it. Conversely, adapting arguments in Duffie (1996, pg. 289) any set of equivalent probabilities can be constructed in this fashion. The interest is in showing that the set of probability measures $\mathcal{P}^{\Theta_0}$ satisfies rectangularity or "time-consistency." More formally, let $\Theta_t : \Omega \rightsquigarrow \mathcal{R}_{t,A,\widetilde{A}}$ be the progressively measurable correspondence that maps paths to the drift terms where $\mathcal{R}_{t,A,\widetilde{A}} = \Theta_\tau(\Omega) = [-\kappa, +\kappa]$ for $\tau < t$ and $\mathcal{R}_{t,A,\widetilde{A}} = \Theta(\Omega) = [-\widetilde{\kappa}, +\widetilde{\kappa}]$ and take Θ to be the collection of all progressively measurable selections from Θ_t $0 \le t \le T$, that is, $\Theta = \{(\theta_t) : \theta_t(\omega) \in \mathcal{R}_{t,A,\ \widetilde{A}}\ dt \otimes dP_A \text{ w.p.1}\}$. The set $\mathcal{P}^\Theta$ therefore contains all the probability measures equivalent to P constructed using (2) for all $\theta \in \Theta_0$.

When the base-line measure is fixed, Girsanov's theorem therefore shows that the change of actions transforms the set probability measures equivalent (in the sense of absolute continuity) to a base-line measure into another set of measures equivalent to the same base-line measure. Since the set of all measures are constructed by some selection from the set Θ_0 of one-step-ahead densities, adopting the terminology from Chen and Epstein this establishes the rectangularity of multiple priors

Lemma 1. The set $\mathcal{P}^{\Theta_0}$ of probability measures, under which the progressive measurable processes $X_t = \theta dt + \sigma dB_t$ for $\theta \in [-\kappa, +\kappa]$ are Brownian Motions with drift θ, is rectangular.

This property ensures that the probability measures on the sample paths induced by any change in action process at a later date is recognized from the current period's perspective. Its significance lies in representation of utility functions recursively and thereby making the analysis tractable without having to impose any restrictions as to which probabilistic models are more relevant for the decision maker. In the latter sense, it corresponds to a situation where the agents have learned "everything" relevant for the contractual relationship.

Next, I address a normalization assumed by Chen and Epstein [11], which in our formalization of multiple priors corresponds to a particular choice of effort process $a_t = 0$ for all t and ω. We show that the normalization can be replaced by an arbitrary alternative probability measure while preserving time consistency of the multiple priors.

3.2. **kappa-ignorance with variable base measure.** An important aspect of the contracting problem is that different choices of effort by the agent not only affect the set of drifts through the change in the size of interval but also through a change in the base measure. The main result of this section in Proposition 1 shows that the normalization used by Chen and Epstein to a base measure P

with respect to which the output process is a Brownian motion without a drift and can be made to an arbitrary base measure $\widetilde{P}$.

To fully formalize the structure of the set of multiple priors as the base measure varies, first consider the case that different actions induce different base-line measures but they have the same interval. Formally, let us assume that changing the effort process from (A_t) to $(\widetilde{A}_t)$ changes the family of probability measures that gives Brownian motions $(A_t + \theta_t)dt + \sigma dB_t$ with drift in $\theta_t + [-\kappa, +\kappa]$ under each of P^θ_A measures corresponding to Z^θ_A to the family that gives Brownian Motions $(\widetilde{A}_t + \theta_t)dt + \sigma dB_t$ with drift in $\theta_t + [-\kappa, +\kappa]$ under each of $P^\theta_{\widetilde{A}}$ measures corresponding to $Z^\theta_{\widetilde{A}}$ by transforming A to $\widetilde{A}$ through Girsanov theorem. Thus the family obtained has measures each equivalent to P_A. The family hereby is rectangular set of probability measures. The rest of the section fills in the formal details.

Take Girsanov exponential

$$Z^\theta_t = \exp\left(-\int_0^t \theta_s dB^A_s - \frac{1}{2}\int_0^t ||\theta_s||^2 ds\right) \tag{6}$$

and denote Θ_A as the set of Girsanov exponentials Z^θ_t associated with the processes (θ_t): $\theta_\tau \in [-\kappa, +\kappa]$ for $0 \leq \tau \leq t$ and $\theta_\tau \in (\widetilde{A}_t - A_t) + [-\kappa, +\kappa]$. The corresponding set of probability measures $\mathcal{P}^{\Theta_A}$ contains all possible measures that can be generated using the probability densities. As in the case with variable interval in kappa-ignorance Sect. 3.1, since the set of all measures are constructed by some selection from the set Θ_0 of one-step-ahead densities, I establish:

Lemma 2. The set of probability measures $\mathcal{P}^{\Theta_A}$ is a rectangular set of probability distributions with a base-line probability measure P_A.

Nothing in the previous line of reasoning depends on the processes (θ_t) except that it satisfies weak regularity conditions in applying Girsanov and that it has a bounded second moment, namely that $E_P(\int_0^T ||\theta_t||^2 dt) < \infty \ P - a.e.$ By construction, these conditions hold in the form of the ambiguity the contracting problem deals with. In particular, taking a function for the drift term μ, which is measurable and has a bounded second moment, in the role of θ the previous analysis goes through and therefore:

Proposition 1. The set of probability measures $\mathcal{P}^{\Theta_A, \widetilde{A}, \kappa}$ is a rectangular set of probability distributions with a base-line probability measure P_A.

Notice that in this case the base-line measure P_A is the one that makes the process $dX_t = \mu(A_t)dt + \sigma dB_t$ Brownian motion. Note also that the base-line probability measure P_A is the center of the ambiguous set of probability distributions associated with the action A. It captures the notion of ambiguity formalized by Dumav and Stinchcombe [24] and Siniscalchi [56] in which a set of multiple priors is represented as the sum of its center and a set centered at zero.

In summary, the analysis in this section verifies that various sets of probability distributions $\mathcal{P}^a$ that correspond to the set of drift terms Θ^a 5 and hence arise in formulating the contracting problems in continuous time satisfy time consistency. This property will play an important role in the recursive representation of the contracting problem below. Before this, I turn to examine the regularity properties of the multiple priors.

3.3. **Regularity Properties of the Set of Priors.** The sets of priors that arise in the contracting problem are ones that vary in base-line measures and in the interval around the base-line by different choices of effort process. I examine in this section whether these extensions preserve regularity properties (defined below) so that the contracting problem is well-posed and admits a solution. I show that the contracting problems with ambiguity satisfy required regularity properties.

Chen and Epstein's [11] formulation for decision problems uses $0 \in \Theta_t(\omega)\ dt \otimes dP\ a.e.$ In our case, this corresponds to taking $\kappa(A_t) = 0$ and setting base-line measure to P: for each $t \in (0, T]$ so that $\mu(A_t) \in \Theta_t(\omega)\ dt \otimes dP_A\ a.e.$ Intuitively, the agents consider the base-line measure to be the one that corresponds to the center of the interval for the values of drift. Our main departure is to allow dependence of the drift on the effort process. By Girsanov's theorem [29] for changes of measures, our base-line measure is equivalent to the base-line. Therefore, this difference is but in looking at the processes equivalently. Second, the measurability follows from from the fact that the correspondence $(t, \omega) \mapsto \Theta_t(\omega)$ which defines the set of priors through (3) when restricted to $[0, s] \times \Omega$ is $\mathcal{B}([0, s]) \otimes \mathcal{F}_s-$measurable for any $0 < s \leq T$. The remaining regularity properties of the set are its compactness and convexity, and follow from standard arguments. I collect these in the following result and its proof in the appendix fills in the remaining details.

Proposition 2. The set of priors $\mathcal{P}^{\widetilde{\Theta}}$ satisfies:

(a) $P_A \in \mathcal{P}^{\widetilde{\Theta}}$.
(b) $\mathcal{P}^{\widetilde{\Theta}}$ is absolutely continuous with respect to P and each measure in $\mathcal{P}^{\Theta}$ is equivalent to P.
(c) $\mathcal{P}^{\widetilde{\Theta}}$ is convex.
(d) $\mathcal{P}^{\widetilde{\Theta}} \subset ca^1_+(\Omega, \mathcal{F}_T)$ is compact in the weak topology.
(e) For every $\xi \in L^2(\Omega, \mathcal{F}_T, P)$, there exists $Q^* \in \mathcal{P}^{\widetilde{\Theta}}$ such that

$$E_{Q^*}[\xi|\mathcal{F}_t] = \min_{Q \in \mathcal{P}^{\widetilde{\Theta}}} E_Q[\xi|\mathcal{F}_t], \quad 0 \leq t \leq T$$

Parts (a)-(c) are self-explanatory. By (d), min exists for any $\xi \in L^1(\Omega, \mathcal{F}_T, P)$, a fortiori in $L^2(\Omega, \mathcal{F}_T, P)$. Part (e) extends the existence of a minimum to the process of conditional expectations.

Having established rectangularity and regularity of the set of multiple priors that arise we next move to give a recursive representation of the contracting problem.

4. Recursive utility in the contracting problem

The tractability of analysis in contracting problems in continuous time relies on recursive representation of values to the contracting parties. This section shows that the recursive utility formulation of Chen and Epstein [11] for decision problem under ambiguity generalizes to the contracting problem. The key observation that allows for the generalization is that the contract variables consumption and effort processes takes an analogous role of consumption processes in the analysis of Chen and Epstein. The main difference is that the effort choice made in the current period is imperfectly observed. The latter concern is not present in Chen and Epstein. Insights from Sannikov's [54] formulation relevant for the contracting problem further our analysis in this case and I represent the utility from each consumption and effort process specified in a recursive manner.

The main elements in our analysis builds on Chen and Epstein's recursive formulation which in turn uses recursive utility formulation in Duffie and Epstein [22, 23]. Duffie and Epstein showed that, under suitable Lipschitz conditions on contemporaneous utility function f, the recursive utility solves a Backward Stochastic Differential Equation (BSDE) and satisfies the usual properties of standard utilities (e.g., concavity with respect to consumption if the BSDE is concave). Their analysis makes powerful use of the Martingale Representation theorem and Girsanov's theorem for change of measures. Our construction of recursive formulation rests on these ideas.

The main result of this section shows that the value processes in the contracting problem (4) under the minmax criterion has an equivalent recursive representation. As a preliminary step that specializes to Duffie and Epstein [23]'s formulation, fix a contract (c_t), take $a_t = 0$ and assume no ambiguity. In this case the consumption process is measurable only with respect to the standard Brownian motion under the reference measure P. Following Duffie and Epstein [23] the expected utility process of any given consumption process (c_t) is then defined by

$$V_t^P = E_P\left[\int_t^T f(c_s, V_s^P)ds|\mathcal{F}_t\right]. \tag{7}$$

where f is an aggregator function that in general allows for non-separability over temporal composition of utility flow. In the special case of our main interest we assume the standard expected discounted utility $f(c, a, v) = u(c) - h(a) - \beta v$. In this case, the value process is given by

$$V_t^P = E_P\left[\int_t^T e^{-\beta(s-t)}\left(u(c_s) - h(a_s)\right)ds|\mathcal{F}_t\right].$$

Under ambiguity given any action process $a = (a_t)$ there is a set of priors $\mathcal{P}^{\Theta^a}$ associated with set of Girsanov exponentials Θ^a induced by the action process and the minmax criterion implies the following value to the agent

$$V_t^a = \min_{\theta\in\Theta^a} E_{Q^\theta}\left[\int_t^T e^{-\beta(s-t)}u(c_s) - h(a_s)ds|\mathcal{F}_t\right]. \tag{8}$$

Our goal is to represent this value process recursively in a tractable manner. To develop the analysis consider first with Duffie and Epstein [23] that the process in the standard expected utility specification is rewritten in a simpler recursive form:

$$V_t^P + \int_0^t f(c_s, V_s^P) ds = E_P \left[\int_0^T f(c_s, V_s^P) ds | \mathcal{F}_t \right]$$

which is a martingale under P. The recursive formulation of value in this case follows from the martingale representation theorem:

$$dV_t^P = -f(c_t, a_t, V_t^P) dt + \sigma_t^P \cdot dB_t, \quad V_T^P = 0 \tag{9}$$

with the unique solution for the value process $\left(V_t^P\right)$ and the volatility process $\left(\sigma_t^P\right)$, where the dependence on the reference measure is noted by superscript P. Using the fact that $\int_0^t \sigma_s^P dW_s$ is a martingale and reversing the arguments establish that the solution to the BSDE (9) for $\left(V_t^P\right)$ is the expected utility process for (c_t) in (7).

Using Girsanov Theorem one can change the measure from P to Q^θ for each θ and hence obtains the analogous representation of the utility process. By the representation in Chen and Epstein [11] the value process solves the following BSDE

$$dV_t^\theta = \left[-f(c_t, V_t^\theta) + \theta_t \cdot \sigma_t^\theta\right] dt + \sigma_t^\theta \cdot dB_t, \quad V_T^\theta = 0 \tag{10}$$

The additional additive term in drift relative to that in (9) accounts for the change in measure.[6] In the contracting problem by Sannikov [54], there is no ambiguity and a contract induces a costly effort process (a_t) that generates an outcome process with a drift $\mu(a_t) = a_t$. Taking the latter in the role of (θ_t) in (10) and using the standard aggregator, $f(c, a, V) = u(c) - h(a) + \beta V$, gives the recursive representation for the agent's utility process in Sannikov [54] as a (weakly) unique solution to the following BSDE

$$dV_t^a = \left[-f(c_t, a_t, V_t^a) + \mu(a_t) \cdot \sigma_t^a\right] dt + \sigma_t^a \cdot dB_t, \quad V_T^a = 0 \tag{11}$$

Building on this representation I introduce the notion of ambiguity (IID and symmetric) for any effort process (a_t) which gives rise to a set of drift terms $\mu(a_t) + \Theta^a$. Our main representation result is that under the minmax criterion the expected utility process can similarly be represented as a diffusion process by generalizing the representation of recursive utility in Chen and Epstein [11] to allow for a family of drift terms that depends on the action process:

Proposition 3. Fix a contract $(c_t, a_t) \in D$ and let Θ_a be the corresponding set of measures. Then:

(a) There exists unique processes (V_t^a) and (σ_t^a) solving the BSDE

$$dV_t^a = \left[-f(c_t, a_t, V_t^a) + \mu(a_t) \cdot \sigma_t^a - \min_{\theta \in \Theta^a} \theta_t \cdot \sigma_t^a\right] dt + \sigma_t^a \cdot dB_t, \quad V_T^a = 0. \tag{12}$$

[6] For details on the use of the Martingale representation theorem and Girsanov thereom, see the manuscript Duffie [21].

(b) For each $Q_a \in \mathcal{P}^{\Theta^a}$, let $(V_t^{Q_a})$ be the unique solution to (11). Then V_t^a defined in (a) is the unique solution to (8) and there exists $Q^* \in \mathcal{P}^{\Theta^a}$ such that

$$V_t^a = V_t^{Q^*}, \quad 0 \le t \le T. \tag{13}$$

(c) The process (V_t^a) is the unique solution to $V_T^a = 0$ and

$$V_t^a = \min_{Q \in \mathcal{P}^{\Theta^a}} E_Q \left[\int_t^\tau f(c_s, a_s, V_s) ds + V_\tau | \mathcal{F}_t \right], \quad 0 \le t < \tau \le T. \tag{14}$$

The formulation of the recursive utility is related to Chen and Epstein [11] due to ambiguity and to Sannikov [54] due to the contracting problem. The main difference from the former is in generalizing the recursive utility formulation given by Chen and Epstein to a family of drift terms that varies with the effort process chosen by the agent motivated by Sannikov. The generalization follows from using Girsanov's Theorem that allows the changes of measures. and specializes to the case examined by Sannikov by choosing $\Theta_t^a(\omega) = \mu(a_t)$ for each t and ω. We have shown earlier in Proposition 2 that the notion of ambiguity that is modeled as IID and symmetric between the principal and agent gives rise to the set of priors which satisfies time-consistency and regularity conditions as defined in Chen and Epstein. Accordingly, the regularity ensures that various value processes that arise in contracting problem are well-defined and the rectangularity of the set of priors allows us to replace the agent's optimization under the entire contract with a sequence of temporal optimization problems. This, together with two powerful results from stochastic analysis, namely the martingale representation theorem and Girsanov's theorem for changes of measures, yield a recursive formulation for the agent's expected utility in a similar manner as in Sannikov. The interpretation with maxmin criterion is that the agent evaluates a given contract under the worst-case scenario which corresponds to the lowest drift induced by her effort choice.

The last piece among the analytical results represents the primitive set (Θ_t) in an equivalent functional form, using its support functions and this form is more convenient in the theoretical development. Because each correspondence Θ_t is convex-valued, its structure, by Hanh-Banach theorem in its supporting functions form, can be represented by its support functions defined by

$$e_t(x)(\omega) = \max_{y \in \Theta_t^a(\omega)} y \cdot x, \quad x \in R^d. \tag{15}$$

The difference from Chen and Epstein [11] is the renormalization to the baseline drift to $\mu(a_t)$ under the action process (a_t). Under this renormalization, the support function is still Lipschitz continuous, convex and linear; and the joint measurability holds: the map $(t, \omega) \to e_t(x)(\omega)$ is $\mathcal{B}([0,s]) \times \mathcal{F}_s-$measurable on $[0, s] \times \Omega$ on $(0, T] \times R^d$. However, unlike in Chen and Epstein it need not be non-negative as the normalization is not the origin for each effort choice but a principal does not implement such an effort as an outside option that yields non-negative value is always feasible. With these elements in place the proposition follows from the following observations.

Proof. (of Proposition 3)

(a) Since the support function e and the utility function are Lipschitz continuous and satisfy progressive measurability, the unique existence of the solution to (11) and (12) follows from Pardaoux and Peng [50, Theorem 4.1.]

(b) With the renormalization, the set of density generators Θ_a to the base-line measure $\mu(a_t)$ under the action process (a_t), as I show in Propositions 1 and 2, satisfy dynamic consistency and regularity defined in Chen and Epstein. Furthermore, the Comparison Theorem applies the same way as it does not depend on the structure of the set of density generators. It therefore follows from analogous arguments as in Chen and Epstein [11, Theorem 2.2 (b)].

(c) The analogous arguments from Chen and Epstein go through as they do not depend on the particular choice of the normalization used as we established in Propositions 1 and 2.

□

The analysis thus far has used a fixed terminal time T. This is mainly done to bring forth the key elements in the analysis in a simple way. The contracting problem, however, does not necessarily have a relationship for a predetermined period of time. In particular, the continuation of a contract depends on the performance within the relationship and there can be termination following sufficiently many observations of poor performance or retirement when the continuation of a contract becomes costly after good performances. The extension of the results to allowing a stopping-time instead of a deterministic time horizon follows from virtually the same way as it is done in Duffie and Epstein [22] for stochastic differential utility.

Having established the recursive representation of value induced by any effort process (a_t) I move to derive a tractable incentive-compatibility condition and using it characterize the optimal contract. In the next section I formulate a "one-shot deviation" principle from discrete-time dynamic games to verify incentive compatibility of effort process given a contract.

5. Incentive compatibility under ambiguity

An effort process (a_t) is implementable if there is a contract that specifies transfers (c_t) to the agent given observable output realizations and that (a_t) is compatible with the agent's incentives, that is he chooses effort (a_t). We use this standard definition for implementability of effort to determine the feasible set of implementable contracts for the principal. We specialize the implementability to the ambiguity regarding the drift term. Assume that for each effort process (A_t) the associated multiple set of priors $\mathcal{P}^A$ is equivalently characterized by the set of drift terms Θ^A using the formulation in (2) and (3). A useful characterization for implementability follows below from representing agent's value from a contract as a diffusion process.

Proposition 4. (Representation of the agent's value as a diffusion process) For any contract (C_t) and any effort process (A_t) with its associated set of drift terms Θ^A there exists a progressively measurable process (Z_t) such that

$$W_t = W_0 + \int_0^t r\left(W_s - u(C_s) + h(A_s) + \min_{\theta_t \in \Theta_t^A} \theta_s |Z_s|\right) ds + \int_0^t r Z_t (dX_s - \mu(A_s) ds) \tag{16}$$

for every $t \in [0, \infty)$.

Proof. For a given pair of processes (C_t) and (A_t) for transfers to the agent and effort, respectively, define the valuation process V by

$$V_t = r \int_0^t e^{-rs}(u(c_s) - h(A_s)) ds + e^{-rt} W_t(C, A) \tag{17}$$

where $W_t(C, A)$ is the continuation value defined by

$$W_t = \min_{Q \in \mathcal{P}^{\Theta^A}} E^Q \left[\int_t^\infty e^{-rs} \left(u(C_s) - h(A_s)\right) ds \mid \mathcal{F}_t \right]$$

By rectangularity of the multiple-priors, the valuation process (V_t) is a $g-$martingale. Using the $g-$martingale representation theorem in Chen and Epstein [11], there exits a measurable process Z_t such that

$$-dV_t = -\kappa_t^* r e^{-rt} |Z_t| dt - \sigma r e^{-rt} Z_t dB_t^A \tag{18}$$

where B_t is a Brownian motion under the reference measure P; $\kappa_t^* \min_{\theta_t \in \Theta_t^A} \theta_t |Z_t|$ is the worst-case drift; and the factor $re^{-rt}\sigma$ is a convenient rescaling. On the other hand, differentiating (17) with respect to t one finds that

$$dV_t = re^{-rt}(u(C_t) - h(A_t))dt - re^{-rt} W_t dt + e^{-rt} dW_t \tag{19}$$

Together (18) and (19) imply that

$$re^{-rt}(u(C_t) - h(A_t))dt - re^{-rt} W_t dt + e^{-rt} dW_t = \kappa_t^* r e^{-rt} \kappa |Z_t| dt + \sigma r e^{-rt} Z_t dB_t^A$$

$$\Longrightarrow \; dW_t = r\kappa_t^* |Z_t| dt + \sigma r Z_t dB_t^A - r(u(C_t) - h(A_t))dt + rW_t dt$$

This further implies

$$W_t = W_0 + \int_0^t r\left(W_s - u(C_s) + h(A_s) + \min_{\theta_s \in \Theta_s^A} \theta_s |Z_s|\right) ds + \int_0^t r Z_t dB_s^A$$

□

The analysis here is closely related to Sannikov's representation. Compared to the formulation in Sannikov [54] the analysis with ambiguous information introduces a term $\kappa^*(A_s)|Z_s|$ which is interpreted as capturing the effect introduced by ambiguity. The agent uses the worst case to evaluate a contract. Here the worst case corresponds to the drift terms that yield the minimum value to the agent. Using this observation I next present a tractable incentive compatibility condition that characterizes the agent's effort choice for a given contract in an environment with ambiguity.

Proposition 5. (The Agent's incentives) For a given strategy $A = (A_t)$, let (Z_t) be the volatility process from Proposition 4. Then A is optimal if and only if

$$\forall a \in \mathcal{A} \quad Z_t\mu(A_t) - h(A_t) + \min_{\theta_t \in \Theta^A} \theta_t |Z_t| \geq Z_t\mu(a_t) - h(a_t) + \min_{\theta_t \in \Theta^a} \theta_t |Z_t| \quad dt \otimes dP \ a.e. \tag{20}$$

Remark: Since the Q^A is equivalent to P by Girsanov's Theorem, and hence has the same zero-sets, without any loss in generality $dt \otimes dP$ $a.e.$ replaces $dt \otimes dQ^A$ $a.e.$ in Sannikov [54].

The characterization uses analogous ideas from Sannikov, generalizes to ambiguous information using the representation I develop earlier and finally applies a version of the Comparison Theorem that has been helpful in establishing principle of optimality in stochastic analysis. The following fills in the details.

Proof. Consider an arbitrary alternative strategy A' that follows possibly different actions A'_τ up to t and afterwards continues with A_t. The effort process A' induces a set of densities Θ^A satisfying the regularity conditions as specified earlier. The corresponding set of multiple priors $\mathcal{P}^{\mathcal{Q}'_A}$ is rectangular. The agent's expected payoff from this action process is well defined by

$$V'_t = \min_{Q \in \mathcal{P}^{\Theta_{A'}}} V_t^Q,$$

where V_t^Q is unique solution (ensured by Duffie and Epstein [23]) to BSDE

$$V_t^Q = E^Q \left[\int_t^\infty f(C_s, A'_s, V_s^Q) \right],$$

where in our formulation I use the standard aggregator, i.e., $f = u(C) - h(A) - \beta V$. By [11, Theorem 2.2], V' is equivalently uniquely characterized as follows:

$$dV'_t = \left[-f(C_t, A'_t, V'_t) + \max_{\theta \in \Theta^{A'}} \theta_t Z'_t \right] dt + Z'_t dB'_t$$

for a unique volatility process Z'.

More generally, V' and Z' uniquely solves a BSDE of the following form

$$dV_t = g'(V_t, Z_t, \omega, t)dt + Z_t dB'_t, \tag{21}$$

with terminal condition ξ. In the special case relevant for our analysis, I have

$$g'(V, Z, \omega, t) = -f(C_t(\omega), A'_t(\omega), V) + \max_{\theta \in \Theta^{A'}} \theta(\omega) Z \tag{22}$$

Under the action process A_t the value process V_t and volatility Z_t solve (21) for A and $g(\cdot)$.

Suppose that the condition (20) holds. Since the terminal conditions are the same under A and A', by the Comparison Theorem [25, Theorem 2.2]

$$g(V, Z, \omega, t) \leq g'(V, Z, \omega, t) \quad dt \otimes dP \ a.e. \tag{23}$$

or equivalently the condition (20) in our model

$$\mu(A_t) Z_t - h(A_t) - \max_{\theta \in \Theta^A} \theta(\omega) Z_t \geq \mu(A'_t) Z_t - h(A'_t) - \max_{\theta \in \Theta^{A'}} \theta(\omega) Z_t \quad dt \otimes dP \ a.e.$$

implies that $V \geq V'$ for almost every t.

Suppose now that the condition (20) fails on a set of positive measures, choose A' that maximizes $\mu(A')Z_t - h(A') - \max_{\theta \in \Theta^{A'}} \theta(\omega)Z$ for all $t \geq 0$. Then $g(V, Z, \omega, t) \leq g'(V, Z, \omega, t) \quad dt \otimes dP$ *a.e.*. Since A' specifies the same action as A after t, by the Comparison theorem $V' > V$. Therefore, A is suboptimal. $\square$

If the volatility process is written as $-Z_t$ the minimum replaces the maximum in (23). Notice that the Proposition 5 is formulated for any generating process Θ^A. To illustrate the intuition that the presence of ambiguity introduces into contract design I specialize the formulation to a simpler case. Taking $\Theta^A := \{(\theta)_t : \mu(A_t) + |\theta_t| \leq \kappa(A_t)\}$ in the set up of Proposition 4, Proposition 5 specializes the result to κ-ignorance model (that features symmetry around the base-line drift) and the corresponding necessary and sufficient incentive compatibility condition is given by

Lemma 3. For a given strategy A, let (Z_t) be the volatility process from Proposition 4 for $\kappa(A)$-ignorance. Then A is incentive compatible if and only if

$$\forall a \in \mathcal{A} \quad Z_t\mu(A_t) - \kappa(A_t)|Z_t| - h(A_t) \geq Z_t\mu(a_t) - \kappa(a_t)|Z_t| - h(a_t) \quad dt \otimes dP \ a.e. \tag{24}$$

Notice that setting $\kappa \equiv 0$ removes ambiguity and specializes the condition to the incentive compatibility condition in the classical case formulated in Sannikov [54] without ambiguity. Compared to this case the presence of ambiguity introduces added additive terms in the middle of both sides of the incentive compatibility comparison. These additional terms have negative signs and discounts for the worst case using the minimum drift relative to the continuation value. Note also that for effort levels $a, \widetilde{a}$ with $\kappa(a) = \kappa(\widetilde{a})$ the incentive compatibility condition does not depend on ambiguity, and the analysis of the agent's incentives reduces to that in the classical case.

The simplicity of the incentive compatibility condition further allows one to examine the effects of sensitivity of payments with output depending on how effort relates to ambiguity. The extent of this effect is directly reflected by the strength of the one additional term in the incentive-compatibility condition. One natural case is to consider a technology in which higher effort leads to higher drift terms and higher imprecision, in other words, $\kappa(a)$ increases in a. Since the process (Z_t) reflects from (16) how the agent values the variation in the continuation value, I see that higher ambiguity, reduces the value of process more drastically. In the standard contracting problem, higher effort is incentivized through larger variation in the continuation value that is sensitive to output realizations. This effect is still present in the first terms as (Z_t) measures the utility consequence to the agent of this variation. However, the incentive effect of variation in the continuation value is now tempered by the presence of ambiguity, which acts as a cost and penalizes high variations in the continuation value. Therefore, everything else being equal, the presence of ambiguity limits the incentive effects of variation in the continuation value through output realizations.

To illustrate further, consider the opposite case for the uncertainty on technology so that working harder leads to higher drift term and also reduces the imprecision so that for $a' > a$ $\mu(a') > \mu(a)$ and $\kappa(a') < \kappa(a)$. In this case, one can see from the incentive compatibility condition in (24) that everything else being equal higher effort levels become easier to implement with variation in the promised value and hence enhances the contracting possibilities for the principal.

Note also that, as in the classic contracting problem, the variation (Z_t) in continuation value is an endogenous object and depends on the contract offered. Therefore, in the contract design the principal optimally resolves the trade-off between high effort and high variation. Using the tractable incentive compatibility condition presented in this section, the next section formulates the optimal contract under ambiguity and characterizes its properties.

6. The Optimal Contract

Trading off the benefit of higher effort against its effort cost and ambiguity aversion the principal designs the optimal contract. The principal offers a contract to the agent that specifies a stream of consumption (C_t) contingent on the realized output and an incentive-compatible advice of effort (A_t) that maximizes the principal's expected profit under minmax criterion

$$F = \min_{Q \in \mathcal{P}^{\Theta^A}} E^Q \left[r \int_0^\infty e^{-rt} dX_t - r \int_0^\infty e^{-rt} C_t dt \right]$$

subject to delivering the agent a required initial value of at least $\hat{W}$

$$V_t(C, A) = \min_{Q \in \mathcal{P}^{\Theta^A}} E^Q \left[r \int_0^\infty e^{-rt} \left(u(C_t) - h(A_t) \right) dt | \mathcal{F}_t \right] \geq \hat{W}$$

Implicit in this formulation are the termination and retirement clauses of a contract. These events are explicitly characterized below within the set of consumption streams. To illustrate briefly, termination is captured as follows. After a sufficiently long period of low enough output realizations consumption a contract is terminated and the consumption stream is set to a low level. The interest is in contracts that generate non-negative expected profits for the principal. Derivation of the optimal contract uses the techniques of Sannikov [54] in a continuous-time moral hazard problem while introducing ambiguity similar to Chen and Epstein [11].

One possible option for the principal is to retire the agent with any value $W \in [0, u(\infty)$, where $u(\infty) = \lim_{c \to \infty} u(c)$. To retire the agent with value $u(c)$, the principal offers him constant consumption c and allows him to choose zero effort. Denote the principal's profit from retiring the agent by

$$F_0(u(c)) = -c.$$

Since the agent can always guarantee himself non-negative utility by taking effort 0, the principal cannot deliver any value less than 0. The only way to deliver value 0 is through retirement. To see this, notice that the future payments to the agent are not always 0, the agent can guarantee himself a strictly positive value by putting effort 0. I call F_0 the principal's *retirement profit*.

Given the agent's consumption $c(W)$ and recommended effort $a(W)$, the evolution of the agent's continuation value W_t can be written as

$$dW_t = r\left(W_t - u(c(W_t)) + h(a(W_t)) + \kappa(a(W_t))|Z(W_t)|\right)dt + rZ(W_t)\sigma dY_t \tag{25}$$

where $\sigma dY_t := (dX_t - \mu(a(W_t))dt)$ and $rZ(W)$ is the sensitivity of the agent's continuation value to output and follows from the representation given in the previous section. When the agent takes the recommended effort, the second term $dX_t - \mu(a(W_t))dt$ has mean 0, and so drift of the agent's expected continuation value is given by the first term $r\left(W_t - u(c(W_t)) + h(a(W_t)) + \kappa(a(W_t))|Z(W_t)|\right)$. To account for the value that the principal owes to the agent, W_t grows at the interest rate r and falls due to the flow of repayments $r\left(u(c(W_t)) - h(a(W_t))\right)$ and *additionally* due to aversion to ambiguity this fall is reduced by $\kappa(a(W_t))|Z(W_t)|$ to account for the worst case. The latter is the main direct effect that ambiguity aversion introduces to the design of dynamic contracts.

In a dynamic contract the allocation of payments over time determines the drift of the agent's expected value. The reduction in the drift due to the presence of ambiguity makes front-loaded wages appealing because the agent prefers the certainty of current wages over the expectation of delayed payments that are evaluated according to the worst-case. To see this more clearly, notice that when the agent's current wage is small relative to the expectation of his future earnings, that is when his wages are back-loaded, the agent's value from the contract has an upward-drift. Analogously, front-loaded wages move the drift of the agent's value downwards. Since the agent evaluates his future expectations according to his worst-case scenario, he prefers the certainty of current wages over the uncertainty of future payments. Therefore, ambiguity aversion induces a preference in favor of front-loaded wages. From incentive-compatibility condition one can see that, as usual in agency problems, to induce effort the agent's wages over time must be made sensitive to output realizations. However, with a preference for the certainty of current wages, the agent's ambiguity aversion limits the incentive effect of intertemporal distortions in wages over time. In particular, this result does not depend on the form of ambiguity in technology and applies generally.

The other important component of the agent's value in a contract is its volatility as described in the second term in (25). The contract determines the sensitivity $Z(W_t)$ of the agent's value to output, which in turn affects the agent's incentives. If the agent deviates to a different effort level, his actual effort affects only the drift of X_t and his incentive compatible choice is characterized by (24). Notice from the agent's value that the agent's effort choice maximizes

$$Z(W_t)\mu(a) - h(a) - |Z(W_t)|\kappa(a).$$

As in the analysis of the standard moral hazard problem, the agent's effort depends strongly on output realizations. In the equation This is seen in the first two terms that describe the expected change tomorrow in the continuation value in response to output net of the effort cost incurred now. Ambiguity aversion, however, reduces the incentive effects of such back-loaded payments by discounting the value according to the worst case, depending on the technology as seen in the last term. In particular, for technologies with κ increasing in effort, see, for example Figure 1, the incentive effects of back-loaded payment that vary with output realization are reduced. This is because of the fact that with such ambiguity in technology the expected benefits becomes less sensitive to effort and reduces the incentive effect of the variation in the continuation payments. More formally, the marginal expected benefits of effort $(\mu'(a)-\kappa'(a))|Z(W_t)|$, assuming differentiable functions to illustrate the workings of the model, decreases with increases in effort due to the agents' aversion that reduces increments in drifts $\mu'(a)$ by $\kappa'(a)$ according to the expectations evaluated under the worst case. On the other hand, for technologies in which uncertainty becomes smaller with increases in effort, i.e. those with $\kappa'(a) < 0$ as in Figure 3, the expected benefits becomes more sensitive to effort and this helps with incentives.

The optimal mix of short-run and long-run payments depends on the nature of uncertainty and are determined by the principal. The optimal contract offered by the ambiguity averse principal describes the choice of payments $c(W)$ and effort recommendations $a(W)$. Let $F(W)$ be the highest profit that the principal can obtain when he delivers the agent value W. Function $F(W)$ together with the optimal choices of $a(W)$ and $c(W)$ satisfy the Hamiltonian-Jacobi-Bellman (HJB) equation

$$\begin{aligned} rF(W) = \max_{(a>0,c>0)} r[\mu(a)-\kappa(a)-c] &+ F'(W)r[W-u(c)+h(a)+\kappa(a)|Z(a)|] \\ &+ \frac{F''(W)}{2}r^2\sigma^2 Z(a)^2 \end{aligned} \tag{26}$$

In this formulation, the principal is maximizing the expected current flow of profit $r\left(\mu(a)-\kappa(a)-c\right)$ discounted according to the worst-case drift plus the expected change of future profit due to the drift and volatility of the agent's continuation value that reflects the agent's ambiguity aversion.

The equation (26) is rewritten in the following form suitable for computation

$$F''(W) = \min_{(a>0,c>0)} \frac{F(W)-a+c+\kappa(a)-F'(W)\left(W-u(c)+h(a)+\kappa(a)|Z(a)|\right)}{r\sigma^2 Z^2(a)/2} \tag{27}$$

The optimal contract is characterized as a solution to this differential equation by setting

$$F(0) = 0 \tag{28}$$

and choosing the largest slope $F'(0) \geq F_0'(0)$ such that the solution F satisfies

$$F(W_{gp}) = F_0(W_{gp}) \quad \text{and} \quad F'(W_{gp}) = F_0'(W_{gp}) \tag{29}$$

at some point $W_{gp} \geq 0$, where $F'(W_{gp}) = F_0'(W_{gp})$ is called the *smooth-pasting condition*. Let functions $c : (0, W_{gp}) \to [0,\infty)$ and $a : (0, W_{gp}) \to \mathcal{A}$ be the minimizers in equation (27). A typical form of the value function F together with $a(W)$, $c(W)$ and the drift of the agent's continuation value is shown in Figure 2. Theorem 1, which is proved formally in the Appendix, characterizes the optimal contracts.

Theorem 1. *The unique concave function $F \geq F_0$ that satisfies* (27)*,* (28)*, and* (29) *characterizes any optimal contract with positive profit to the principal. For the agent's starting value of $W_0 > W_{gp}$, $F(W_0) < 0$ is an upper bound on the principal's profit. If $W_0 \in [0, W_{gp}]$, then the optimal contract attains profit $F(W_0)$. Such a contract is based on the agent's continuation value as a state variable, which starts at W_0 and evolves according to*

$$dW_t = r\left(W_t - u(C_t) + h(A_t) + \kappa(A_t)|Z(A_t)|\right)dt + rZ(W_t)\sigma dY_t \tag{30}$$

where $\sigma dY_t := dX_t - (\mu(A_t) - \kappa(A_t))dt$ under payments $C_t = c(W_t)$ and effort $A_t = a(W_t)$, until the retirement time τ. Retirement occurs when W_τ hits 0 *or W_{gp} for the first time. After retirement the agent gets constant consumption of $-F_0(W_\tau)$ and puts effort* 0*.*

As in discrete time continuation-value W_t summarizes the past history in the optimal contract. Replacing the continuation contract, while leaving the continuation value the same, does not affect the incentives governing the choice of effort in the current period. Therefore, to maximize the principal's profit after any history, the continuation contract must be optimal given W_t. It follows that the agent's continuation value W_t completely determines the continuation contract. This logic does not necessarily follow when there are additional state variables, for example, when hidden savings by the agent are allowed. I abstract from the latter to focus on the implication of ambiguous information.

Turning to the discussion of optimal effort and consumption using (26) notice that the optimal effort maximizes

$$r(\mu(a) - \kappa(a)) + r\left(h(a) + \kappa(a)|Z(a)|\right)F'^2(W)\sigma^2 Z(a)^2 \frac{F''(W)}{2} \tag{31}$$

where $r(\mu(a) - \kappa(a))$ is the expected flow of output according to the worst-case scenario, $rF'(W)(h(a) + \kappa(a)|Z(a)|)$ is the cost of compensating the agent for his effort, and $r^2\sigma^2\frac{\gamma(a)^2}{2}F''(W)$ is the cost of exposing the agent to income uncertainty to provide incentives. The presence of ambiguity introduces a worst-case scenario $\kappa(a)$ and changes the sensitivity of the continuation value $Z(a)$ relative to the case without ambiguity. These two costs typically work in opposite directions, creating a complex effort profile (see Figure 2). While $F'(W)$ decreases in W because F is concave, $F''(W)$ increase over some ranges of W. It turns out that in the optimal contract the introduction of ambiguity in the contracting problem reduces the sensitivity of the optimal incentive scheme to output realizations.

The optimal choice of consumption maximizes

$$-c - F'(W)u(c)$$

Thus the agent's consumption is 0 when $F'(W) \geq -1/u'(0)$ in the probationary interval $[0, W^{**}]$, and it is increasing in W according to $F'(W) = -1/u'(c)$ above W^{**}. Intuitively, $1/u'(c)$ and $-F'(W)$ are the marginal costs of giving the agent value through current consumption and through his continuation payoff, respectively. Those marginal costs must be equal under the optimal contract, except in the probationary interval. There, consumption zero is optimal because it maximizes the drift of W_t away from the inefficient low retirement point.

The HJB formulation describes the solution to the optimal contract in a compact form. The analytical solution to the equation is not in general available. We therefore numerically solve it in various parametric examples to illustrate the effects of ambiguity on the optimal contract. To provide a clear comparison, I take the classical example and add the set of drift terms. As in the classical case, in our examples I set $u(c) = \sqrt{c}$, $h(a) = \frac{1}{2}a^2 + 0.4a$, $r = 0.1$, and $\sigma = 1$. Our analytical framework does not restrict how effort influences the set of drift terms. In our example, I focus on the sets of drift terms of the form $[a - \kappa(a), a + \kappa(a)]$. The set is symmetric, centered at $\mu(a) = a$ corresponding to the drift term in the classical case without ambiguity, and has a width $\kappa(a)$.

To ease the illustration, I consider two natural cases for uncertainty about the technology. In the first case, higher effort increases the strength of ambiguity. In particular, to ease the illustration, in the numerical example I set $\kappa(a) = \frac{1}{4}a$, while the qualitative features of the optimal contract do not sensitively depend on this particular choice of the function. Figure 1 graphically illustrates the nature of this experiment.

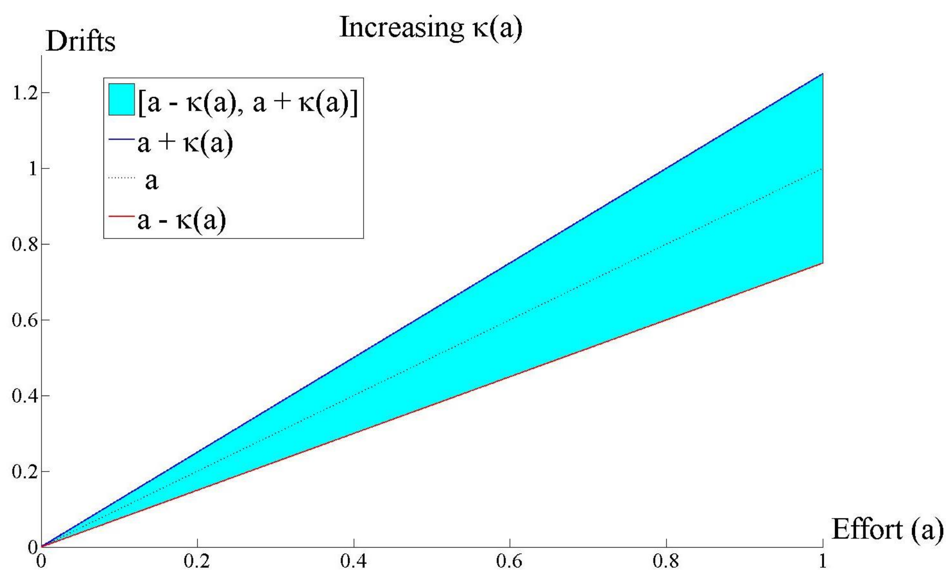

FIGURE 1. Technology with ambiguity - case of increasing κ

In this experiment, increasing effort increases the set of drift terms but ambiguity, as measured by the size of the intervals, also increases. The lower bound of the set corresponds the lowest drift, which, as I have shown, represents the worst-case scenario. Given this structure for the uncertainty governing technology, I continue analyzing the properties of the optimal contract. The main characteristics of the contract regarding profits, efforts and consumption schedule as a function of the continuation value is illustrated below in Figure 2.

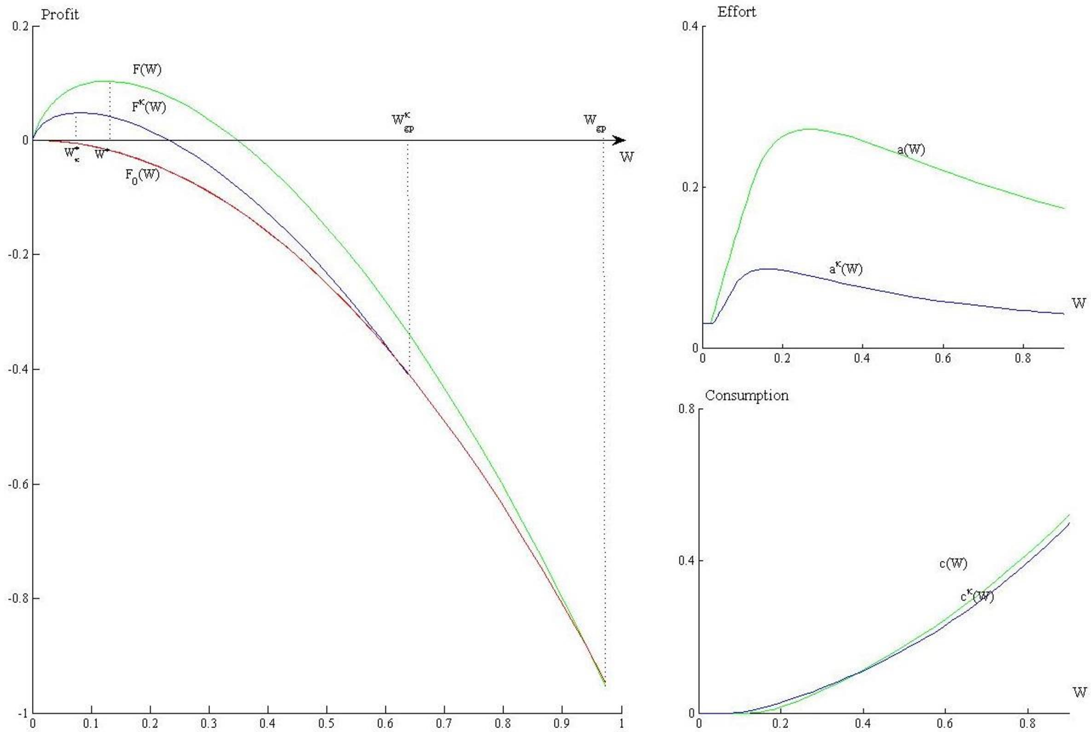

FIGURE 2. The Optimal Contract with increasing κ

Notice that in the optimal contract under ambiguity the principal makes lower expected profits, and offers a flatter incentive scheme and less back-loaded compensations for the agent as ambiguity increases with effort, i.e, $\kappa(a)$ increases in effort. To examine closely the rationale for these results that the current model offers I turn to analyze the implication of the optimal contract for the compensation scheme over the career path of the agent.

Lemma 4. *The consumption stream to the agent under ambiguity* c^κ *is less back-loaded relative to the optimal consumption stream* c *without ambiguity. Moreover, the ratios of volatilities of the agent's consumption and continuation value is greater in the optimal contract under ambiguity. Therefore, the contract associated with ambiguity relies less on short-term incentives.*

Formally, this conclusion mainly follows from noticing that the introduction of ambiguity modifies HJB in a tractable way and applying in my framework a general result from Sannikov that systematically compares the optimal contract under various HJBs. The following fills in the required details.

Proof. Notice that with ambiguity in the optimal contract the profit function to the principal F^κ is lower than that without ambiguity F. From Theorem 4(a) in Sannikov [54] that holds in general HJB, the contract associated with the ambiguous environment involves less back-loaded payments $c^\kappa(w) \geq c(w)$. Analogous arguments from Sannikov's Theorem 4(b) establishes the stated relationship on the volatilities of consumption and the continuation value. $\square$

Turning to the interpretation of the result in Lemma 4 notice that similar as in the classical case, from the incentive compatibility condition (24) we see that the variation in the continuation value with output realizations is still present. The usual intuition is still applicable here that to solve agency conflict the principal's design of contract aligns interests by letting agent's compensation to positively vary with his profits. One can also see from the Lemma that the presence of ambiguity makes back-loaded incentives less effective relative to the classical case. Intuitively, the ambiguity averse agent prefers certainty of payments today over uncertainty of future transfers. Moreover, the sensitivity in the continuation value is not too effective to incentivize the agent due to his pessimism captured by the worst-case criterion. In the optimal contract, the principal offers a contract that has certainty and steadyness to induce efforts. This is not immediate to interpret in a standard agency problem as it would have adverse incentive effects. In the current model, the principal has an additional tools that decide termination of the contract depending on the agent's continuation value which reflect the history of his performance.

In the optimal contract, faced with unknown technology and unobservable effort the principal chooses a higher continuation value for the termination and lower value for the retirement relative to one without ambiguity, as I see in Figure 2. The weaker guarantee in the contractual relationship in this form helps to incentivize the agent. To interpret this result start noticing that in the optimal contract low continuation values lead to termination of the contract, and higher values result in the retirement of the agent. Intuition is similar as in the classic case that termination incentivizes agent to work hard while high continuation value makes it costly to the principal to induce high effort from the agent due to the income effect. To the literature on the agency, the new explanation this work offers is that the durability of the contracting relation can be seen as a part of incentive scheme. One consequence of lower volatility of the continuation value in Lemma 4 is that the duration of the relationship is longer under ambiguity. The agent receives his preferred compensation scheme that features certainty and steadyness in his wage stream in return for his continued steady effort. Since such a flat scheme leads to steady profits, ambiguity averse principal finds it optimal to implement. The lowers profits results from the discount in the principal's expectations that his pessimism according to the worst case of technology. A flatter compensation scheme means that the wage profile of the worker does not change sensitively with the output realizations and therefore my model with higher ambiguity predicts simpler contract.

One possible behavioral interpretation of the result is that an ambiguity averse principal is tolerant. Despite possible low output realizations he continues with the contractual relationship attributing them to the bad luck due to the worst-case of uncertain technology rather than to the lack of due diligence by the worker. The optimality of such delay also means longer contractual relations for the agent to receive steady flow of payments. This helps with incentives to provide higher effort levels when the continuation value is low as shown in Figure 2. On the other hand, for higher values of continuation value, since the back-loaded payments as I have observed in the analysis of the incentive compatibility condition under ambiguity losses its incentive effects due to the familiar income effect, the effort levels are lower relative to the schedule without ambiguity. The combination of two properties leads to wage structure that has a narrow range and, therefore, generates a flat compensation scheme. The result on tolerance has an empirical corroboration. It is consistent with the empirical evidence from the venture capital contracts that find more successful innovative firms have higher tolerance for early failures as documented in Tian and Wang [61].

Next I turn to characterize optimal contract under technology in which ambiguity decreases with effort. Figure 3 below graphically illustrates the nature of this experiment. The parameter values except that on ambiguity is fixed as above. In the second experiment, I consider a technology in which higher effort increases the strength of ambiguity. In particular, to ease the illustration, in the numerical example I set $\kappa(a) = -\frac{1}{4}a$, while the qualitative features of the optimal contract do not sensitively depend on this particular choice of the function. Figure 3 graphically illustrates the nature of this experiment.

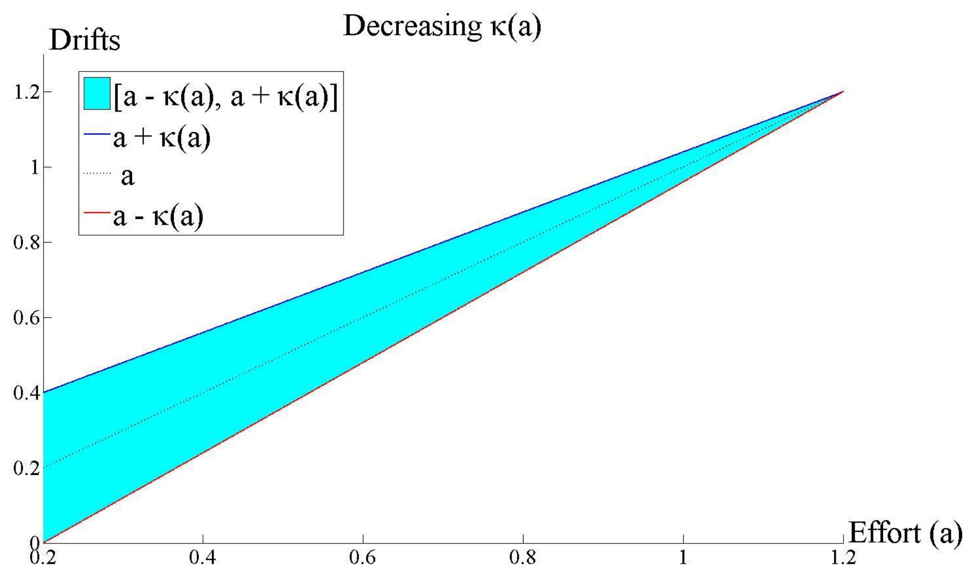

FIGURE 3. Technology with ambiguity - case of decreasing κ

In this experiment, increasing effort increases, similar as before, the set of drift terms but ambiguity, as measured by the size of the intervals, also decreases. Given this structure for the uncertainty governing technology, I continue analyzing the properties of the optimal contract. The main characteristics of the contract regarding profits, efforts and consumption schedule as a function of the continuation value is illustrated below in Figure 4.

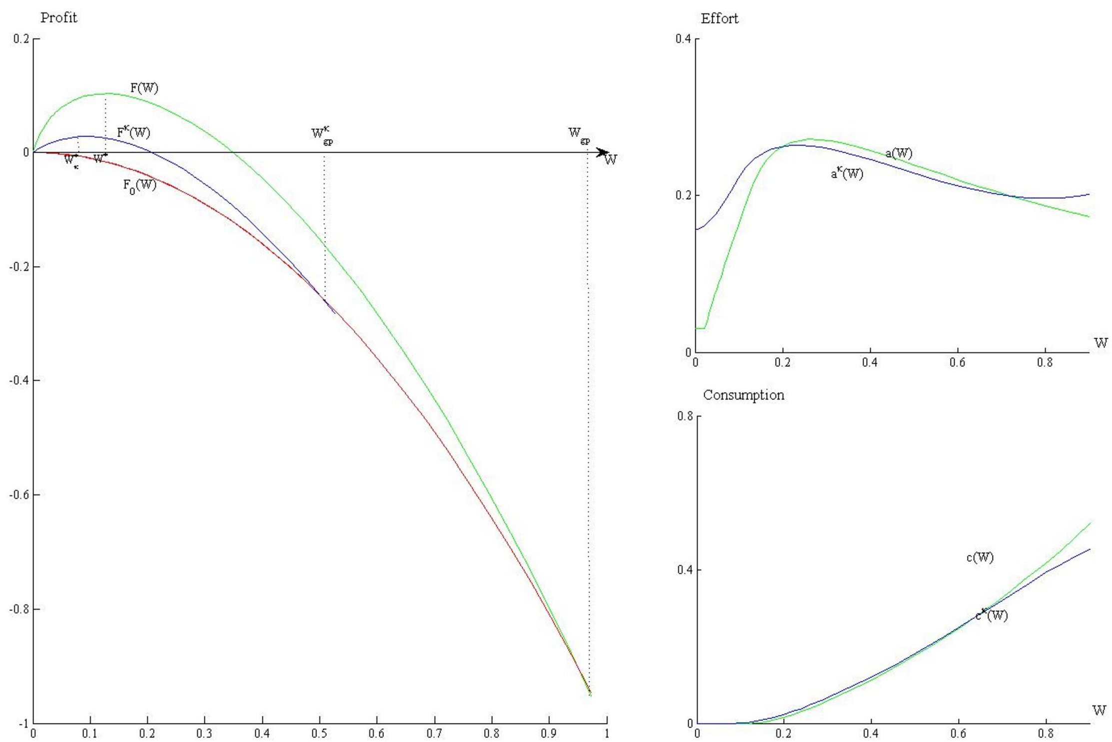

FIGURE 4. The Optimal Contract with decreasing κ

The qualitative features of the optimal contract in this case is very similar to that with increasing ambiguity. In particular, in the optimal contract under ambiguity the principal earns lower expected profits, and offers a flatter incentive scheme and less back-loaded compensations for the agent as ambiguity decreases with effort, i.e, $\kappa(a)$ decreases in effort. The main difference in this case is that, increasing effort reduces uncertainty and this helps with incentives as can analytically be seen from the incentive compatibility constraint (24) and from Figure 4.

7. DISCUSSION

This work has presented a dynamic principal-agent model with ambiguity about the technology. The model has depicted the simplicity and durability of the contracts in that environment. When parties have common ambiguous understanding of the technology, pessimism over the expectation of uncertain future outcomes according to the worst case makes the certainty of immediate values in the relationship relatively more attractive. The unique optimal way for this preference for certainty to be beneficial to the parties with conflicting interests is to tie them in a durable contract with little sensitivity to outcomes.

As noticed in the introduction, many scholars have observed that relative to the simplicity of real world contracts, the theoretical predictions of the agency models are contracts that feature sensitivity to the details of the environment and contingencies in the contractual relationship. Holmstrom and Milgrom [40] offered an early model to illustrate the simplicity of linear contracts. Their model is in continuous time, similar to the one here. Unlike the model here, the agent controls the drift of a diffusion process in a precisely known way, and both the principal and the agent has CARA preferences. The key benefit of these assumptions is that the optimal incentives is independent of the history in the relationship and this leads to simple linear contracts. These assumptions are probably strong in seeking robustness to the details of the economics environment. The framework here allows for general preferences and general structure for uncertainty about the technology. Despite there is in principle dependence of the optimal incentives looking forward on the history of realizations, the optimal contract features little sensitivity to the history.

The part of the intuition in this paper is particularly close to the argument made by Carroll [10] that generalizes Diamond [20]'s intuition for the unique optimality of linear contracts. He considers a static model in which the agent is free to choose an action from his feasible set and the principal has partial information about this set. The agent knows his feasible set of actions, while the principal has full ambiguity about this set and knows only the worst-case associated with each of his contract offering. In this environment Carroll shows that linearity is the unique way for the principal to turn this information into a guarantee for himself. In my model, although both the principal and the agent has symmetric ambiguous information about technology, the knowledge of worst-case still ties the principal's expected profits to the agent's expected compensation and provides him a guarantee. In the dynamic contracting problem here, dependence of compensations on the output is in general non-linear reflecting the history dependence in the contracting relationship. When the dynamic contract is viewed as a sequence of static contracts tied over time with the continuation values, one can see that at each instant the contract is linear in utilities. This mainly follows from an analogous argument made by Carroll and the fact, as I have shown, that the parties agree on the worst case. Moreover, the dynamic model features simplicity as a path property. Although the optimal contract is history dependent and can vary sensitively with past outputs, ambiguity about technology and the parties' aversion to it as captured in the worst-case criterion leads to incentive schemes that have little variation over time.

This paper is also related to the literature on the design of robust contracts with moral hazard problem in continuous time. Miao and Rivera [48] consider a dynamic agency problem in which the principal has ambiguity about the technology, while the agent has precise knowledge of it; and they also consider an alternative specification in which parties have common precise knowledge of technology but the principal has ambiguity about the agent's beliefs. Their focus is different and on the analysis of capital structure a firm and asset pricing. In a similar model Szydlowski [59] introduces the principal's ambiguity about the agent's effort cost rather than about technology and shows that his model implies highly sensitive compensation scheme as a path property due to the worst case regarding effort costs. Both of these papers have strong assumptions regarding ambiguity averse preferences and that the agent has a binary choice for an action. The present work considers a general structure for ambiguous information and allows for general ambiguity averse preferences.

For reasons of simplicity, in modeling aversion to ambiguity I adopt the maximin criterion in the recursive formulation. In particular, in my model of preference representation I assumed a standard aggregator for the function f that represent how in total current and future utilities are valued. For the main application considered here with IID ambiguity, my modeling choice is motivated by the results on the general nature of preferences, as shown by Strzalecki [58], that the standard aggregator together with maxmin expectation operator is the only form among a large family of dynamic ambiguity preferences that has the property of the indifference to the timing of the resolution of uncertainty. The model here is flexible enough to incorporate, for example, a class of recursive smooth ambiguity preferences recently developed in Hayashi and Miao [36], and Klibanoff, Marinacci and Mukerji [44]. Similarly, the model is general enough to incorporate general specifications for ambiguous information, including those that feature various aspects of learning. To illustrate applications of the framework, in the present work I use here specification of $\kappa-$ model of ambiguity with its dependency on effort, which allows for both the constraint on the set of priors and the incentive-compatibility constraint to be analyzed separately. A working in progress attempt to analyze the presence of differential information about the technology, building on the decision theoretic work on the updating of ambiguous information in discrete time by Epstein and Schneider [27] and its generalization in Ju and Miao [41], and in continuous-time Leippold, Trojani, and Vanini [46].

8. Concluding remarks

Contracting parties interact with imprecise information about the environment they interact in. In this paper, I focused one a particular form of information imprecision: both contracting parties have common ambiguity about the productive technology. This assumption has allowed us to: (1) apply and extend the decision-theoretic model of Chen and Epstein [11] to continuous-time setting relevant for the dynamic contracting problem; and (2) tractably generalize the principal-agent problem proposed by Sannikov [54] to incorporate richer uncertainty. Pursuing the latter we have found that our model of ambiguity illustrates a new trade-off between effort and variation of compensation and that the optimal resolution of the trade-off favors simple contract structures.

For tractability I have abstracted our analysis from differential information between the contracting parties on the technology. It is left to future research to extend our model to incorporate the richer nature of ambiguous information that, for instance, could allow for learning and experimenting in the design of contracts.

9. Appendix for the proofs

To establish the unique optimal contract I follow the analogous steps as in Sannikov [54]. Using the HJB I formulate a conjecture for an optimal contract. We show that the HJB satisfies appropriate regularity properties and that it has a unique solution. From that solution I form a conjecture for an optimal contract and then verify its optimality.

For regularity I consider a version of HJB

$$F''(W) = \min_{(a,Z)\in\Gamma, c\in[0,\overline{C}])} \frac{F(W) - a + c + \kappa(a) - F'(W)(W - u(c) + h(a) + \kappa(a)|Z(a)|)}{r\sigma^2 Z^2(a)/2} \tag{32}$$

where the sensitivity parameter Z is bounded from below by γ_0, and consumption is bounded from above by the level $\overline{C}$ such that $u'(\overline{C}) = \gamma_0$. The existence and uniqueness of solutions to the HJB equation (27) satisfying the boundary conditions (29) follows from analogous arguments made by Sannikov [54] since the right-hand side of (27) is Lipschitz continuous in all of its arguments.

9.1. **Conjecture of a contract.** We conjecture an optimal contract from the solution of equation HJB just constructed.

Proposition 6. Consider the unique solution $F(W) \geq F_0(W)$ that satisfies boundary conditions (29) for some $W_{gp} \in [0, W^*_{gp}]$. Let $a : [0, W^*_{gp}] \to \mathcal{A}$, $Y : [0, W^*_{gp}] \to [\gamma_0, \gamma_1]$ and $c : [0, W^*_{gp}] \to [0, \overline{C}]$ be the minimizers in (27). For any starting condition $W_0 \in [0, W_{gp}]$ there is a unique solution, in the sense of weak probability law, to the following equation

$$\begin{aligned} dW_t &= r(W_t - u(c(W_t)) + h(a(W_t)) + \kappa(a(W_t))|Z(W_t)|)dt \\ &+ rZ(W_t)(dX_t - [a(W_t) - \kappa(a(W_t))]\, dt) \end{aligned}$$

where the last term is a Brownian Motion: $\sigma dB_t^{a(W)} = dX_t - a(W_t)Z(W_t)dt$ until the time τ. The contract (C, A) defined by

$$\begin{aligned} C_t &= c(W_t), \text{ and } A_t = a(W_t), \text{ for } t \in [0, \tau]) \\ C_t &= -F_0(W_\tau), \text{ and } A_t = 0, \text{ for } t \geq \tau \end{aligned}$$

is incentive-compatible, and it has a value W_0 to the agent and profit $F(W_0)$ to the principal.

Proof. From the representation of $W_t(C, A)$ in Proposition 4, I have

$$\begin{aligned} d(W_t(C, A) - W_t) &= r(W_t(C, A) - W_t)dt + r(Y_t - Y(W_t))\sigma dB_t^A \\ &+ r\kappa(A)(|Z_t| - |Z(W_t)|)dt \end{aligned}$$

where the changes of measures are conducted under the worst-case measures. This implies that

$$E_t[W_{t+s}(C, A) - W_{t+s}] = e^{rs}(W_t(C, A) - W_t) + e^{rs}E_t\kappa(A)(|Z_t| - |Z(W_t)|)$$

Notice that the left hand side must remain bounded, because both W and $W(A, C)$ (since C_t is bounded) are bounded, and the processes Z_t and $Z(W_t)$ are bounded by the representation theorem. It follows that $W_t = W_t(C, A)$ for all $t \geq 0$, and in particular, the agent gets value $W_0 = W_0(C, A)$ from the entire contract. Also, the contract (C, A) is incentive compatible, since $(A_t, Z_t) \in \Gamma$ for all t.

To see that the principal gets profit $F(W_0)$, consider

$$G_t = r\int_0^t e^{-rs}(A_s - \kappa(A_s) - C_s)ds + e^{-rt}F(W_t).$$

By Ito's lemma, the drift of G_t is

$$\begin{aligned} re^{-rt}\left((A_t - \kappa(A_t) - C_t - F(W_t))\right. &+ F'(W_t)(W_t - u(C_t) + h(A_t) + \kappa(A_t)|Z_t|) \\ &+ \left. r\sigma^2 Z_t^2 F''(W_t)/2\right). \end{aligned}$$

The value of this expression is 0 before time τ by the HJB equation. Therefore, G_t is a bounded martingale until τ and the principal's profit from the entire contract is

$$\min_{Q^A\in\mathcal{P}^A} E^{Q^A}\left[r\int_0^\tau e^{-rs}(A_s - C_s)ds + e^{-r\tau}F_0(W_\tau)\right]$$

$$= E\left[e\int_0^\tau e^{-rs}(A_s - \kappa(A_s) - C_s)ds + e^{-r\tau}F_0(W_\tau)\right] = E[G_\tau] = G_0 = F(W_0),$$

since $F(W_\tau) = F_0(W_\tau)$. □

9.2. **Verification.** Our last step is to verify that the contract presented in Proposition 6 is optimal. We start with a lemma that bounds from above the principal's profit from contracts that give the agent a value higher than W_{gp}^*.

Lemma 5. *The profit from any contract (C, A) with the agent's value $W_0 \geq W_{gp}^*$ is at most $F_0(W_0)$*

Proof. Define c by $u(c) = W_0$. Then $W_0 \geq W_{gp}^*$ implies that $u'(c) \leq \gamma_0$. I have For any Q I have

$$\begin{aligned} E^Q\left[r\int_0^\infty e^{-rt}(u(C_t) - h(A_t))\right] &\leq E^Q\left[r\int_0^\infty e^{-rt}(u(c) + (C_t - c)u'(c) - \gamma_0 A_t)dt\right] \\ &\leq u(c) - u'(c)\left(E^Q\left[r\int_0^\infty e^{-rt}(A_t - C_t)dt\right] + c\right), \end{aligned}$$

In particular,

$$\begin{aligned} W_0 &= \min_Q E^Q\left[r\int_0^\infty e^{-rt}(u(C_t) - h(A_t))\right] \\ &\leq u(c) - u'(c)\left(E\left[r\int_0^\infty e^{-rt}(A_t - \kappa(A_t) - C_t)dt\right] + c\right) \end{aligned}$$

where $u(c) = W_0$ and $c = -F_0(W)$. It follows that the profit from this contract is at most $F_0(W)$. □

Next, note that function F from which the contract is constructed satisfies

$$\begin{aligned} &\min_{W'\in[0,\infty)} F(W) - F_0(W') - F'(W)(W - W') \\ &= \min_{c\in[0,\infty)} F(W) + c - F'(W)(W + u(c)) \qquad \geq 0 \end{aligned} \tag{33}$$

for all $W \geq 0$. For any such solution, the optimizers in the HJB equation satisfy $a(W) > 0$ and $c(W) < \overline{C}$. If either of these conditions failed, (33) would imply that $F''(W) \geq 0$. Also I have that $Z(W) = \gamma(a(W))$.

Proposition 7. Consider a concave solution F to the HJB equation that satisfies (33). Any incentive-compatible contract (C, A) achieves profit of at most $F(W_0(C, A))$.

Proof. Denote the agent's continuation value by $W_t = W_t(C, A)$, which is represented by (16) using the process Z_t. By the Lemma, the profit is at most $F_0(W_0) \leq F(W_0)$ if $W_0 \geq W^*_{gp}$. If $W_0 \in [0, W^*_{gp}]$, define

$$G_t = r \int_0^t re^{-rs}(A_s - \kappa(A_t) - C_s)ds + e^{-rt}F(W_t)$$

as in Proposition 3. By Ito's lemma, the drift of G_t is

$$\begin{aligned} re^{-rt}\big((A_t - \kappa(A_t) - C_t - F(W_t)) \quad &+ \quad F'(W_t)(W_t - u(C_t) + h(A_t) + \kappa(A_t)|Z_t|) \\ &+ \quad r\sigma^2 Z_t^2 F''(W_t)/2\big) \end{aligned}$$

which is computed under the worst-case scenario.

Let us show that the drift of G_t is always non-positive. If $A_t > 0$ then Proposition 2 and the definition of γ imply that $Y_t \geq \gamma(A_t)$. Then equation HJB and together with $F''(W_t) \leq 0$ imply that the drift if G is non-positive. If $A_t = 0$, then $F''(W) < 0$ and (33) imply that the drift of G_t is non-positive.

It follows that G_t is a bounded supermartingale until the stopping time τ' (possibly ∞) when W_t reaches W^*_{gp}. At time τ' the principal's future profit is less than or equal to $F_0(W^*_{gp}) \leq F(W^*_{gp})$ by Lemma 4. Therefore, the principal's expected profit at time 0 is less than or equal to

$$E^A\left[\int_0^{\tau'} e^{-rt}(dX_t - C_t dt) + e^{-r\tau'}F(W_{\tau'})\right] = E^A[G_{\tau'}] \leq G_0 = F(W_0).$$

□

9.3. **On the set of multiple priors.**

Proof. of Proposition 2

(b) The process $dW^A = \mu(A_t)dt + \sigma dB_t$ is a Brownian motion under the baseline measure $Q^A = Z^A P$.

Fix $B \in \mathcal{F}_t$ and $Q^\theta_A \in \mathcal{P}^\Theta_A$. By Girsanov's Theorem, $Q^\theta_A(B|\mathcal{F}_t) = y_t$, where (y_t, σ_t) is the unique solution to

$$dy_t = \sigma_t\left(\mu(A_t)dB_t\right), \qquad y_T = 1_B$$

By the bounding inequality in El Karoui, Peng, and Quenez [25] and Uniform Boundedness, there exits $k > 0$ such that

$$(Q^\theta_A(B))^2 \leq kE_{Q_A}(1_B) = kQ_A(B),$$

where k is independent of θ. This delivers uniform absolute continuity. Equivalence obtains because $Z^\theta_T > 0$ for each θ.

(c) Follows from replacing P with $Q_A^{\theta=0}$ in the proof by CE. For $i = 1, 2$, let Q^i be the measure corresponding to $\theta^i \in \Theta_A$ and the martingale $Z_A^{\theta_i}$ as in (6). Define $\theta = (\theta_t)$ by

$$\theta_t = \frac{(\theta_t^1 + \theta_t^2)}{z_t^1 + z_t^2}$$

It thus follows that $\theta \in \Theta_A$ and $d\left(z_t^1 + z_t^2\right) = -\left(z_t^1 + z_t^2\right)\theta_t \cdot dW_t^A$, which implies that $\left(z_T^1 + z_T^2\right)/2$ is the density for $\left(Q^1 + Q^2\right)/2$. This shows that the latter measure lies in $\mathcal{P}_A^\Theta$.

(d) By the analogous arguments in Cuoco and Cvitanic [14], using the weak compactness of Θ_A by Lemma 3, $Z^\Theta = \left\{z_T^\theta : \theta \in \Theta_A\right\}$ is norm closed in $L^1(\Omega, \mathcal{F}_T, P)$. Moreover, because Z is convex, it is also weakly closed. Since $E_A\left(|z_T^\theta|\right) = 1$ for all θ, Z is norm-bounded. Therefore, by the Alaoglu Theorem, Z is weakly compact. Finally, Z^Θ is homeomorphic to $\mathcal{P}^\Theta$ when weak topologies are used in both cases.

(e) follows from the properties of Θ established in Lemma 4. □

References

[1] Rui Albuquerque and Hugo A Hopenhayn. Optimal lending contracts and firm dynamics. *The Review of Economic Studies*, 71(2):285–315, 2004.

[2] Evan W Anderson, Lars Peter Hansen, and Thomas J Sargent. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association*, 1(1):68–123, 2003.

[3] Andrew Atkeson and Robert E Lucas. On efficient distribution with private information. *The Review of Economic Studies*, 59(3):427–453, 1992.

[4] Dirk Bergemann and Stephen Morris. Robust implementation in general mechanisms. *Games and Economic Behavior*, 71(2):261–281, 2011.

[5] Dirk Bergemann and Karl Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146(6):2527–2543, 2011.

[6] Sugato Bhattacharyya and Francine Lafontaine. Double-sided moral hazard and the nature of share contracts. *The RAND Journal of Economics*, pages 761–781, 1995.

[7] Bruno Biais, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet. Dynamic security design: Convergence to continuous time and asset pricing implications. *The Review of Economic Studies*, 74(2):345–390, 2007.

[8] Bruno Biais, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78(1):73–118, 2010.

[9] Aaron L Bodoh-Creed. Ambiguous beliefs and mechanism design. *Games and Economic Behavior*, 75(2):518–537, 2012.

[10] Gabriel Carroll. Robustness and linear contracts. *Working Paper*, 2013.

[11] Zengjing Chen and Larry Epstein. Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 70(4):1403–1443, 2002.

[12] Leon Yang Chu and David EM Sappington. Simple cost-sharing contracts. *The American Economic Review*, 97(1):419–428, 2007.

[13] Gian Luca Clementi and Hugo A Hopenhayn. A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics*, 121(1):229–265, 2006.

[14] Domenico Cuoco and Jakša Cvitanić. Optimal consumption choices for a largeinvestor. *Journal of Economic Dynamics and Control*, 22(3):401–436, 1998.

[15] Jaksa Cvitanic and Jianfeng Zhang. *Contract Theory in Continuous-Time Models*. Springer, 2012.

[16] Peter M DeMarzo and Michael J Fishman. Agency and optimal investment dynamics. *Review of Financial Studies*, 20(1):151–188, 2007.
[17] Peter M DeMarzo and Michael J Fishman. Optimal long-term financial contracting. *Review of Financial Studies*, 20(6):2079–2128, 2007.
[18] Peter M DeMarzo, Michael J Fishman, Zhiguo He, and Neng Wang. Dynamic agency and the q theory of investment. *The Journal of Finance*, 67(6):2295–2340, 2012.
[19] Peter M DeMarzo and Yuliy Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *The Journal of Finance*, 61(6):2681–2724, 2006.
[20] Peter Diamond. Managerial incentives: on the near linearity of optimal compensation. *Journal of Political Economy*, 106(5):931–957, 1998.
[21] Darrell Duffie. *Dynamic asset pricing theory*. Princeton University Press, 2010.
[22] Darrell Duffie and Larry G Epstein. Asset pricing with stochastic differential utility. *Review of Financial Studies*, 5(3):411–436, 1992.
[23] Darrell Duffie and Larry G Epstein. Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, pages 353–394, 1992.
[24] Martin Dumav and Maxwell B. Stinchcombe. The von neumann/morgenstern approach to ambiguity. *Institute of Mathematical Economics Working Paper No. 480.*
[25] Nicole El Karoui, Shige Peng, and Marie Claire Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.
[26] Larry G Epstein and Martin Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31, 2003.
[27] Larry G Epstein and Martin Schneider. Learning under ambiguity. *The Review of Economic Studies*, 74(4):1275–1303, 2007.
[28] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, 18(2):141–153, 1989.
[29] Igor Vladimirovich Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3):285–301, 1960.
[30] Piero Gottardi, Jean-Marc Tallon, and Paolo Ghirardato. Flexible contracts. *European University Institute Working Paper*, 2014.
[31] Borys Grochulski and Yuzhe Zhang. Optimal risk sharing and borrowing constraints in a continuous-time model with limited commitment. *Journal of Economic Theory*, 146(6):2356–2388, 2011.
[32] Borys Grochulski and Yuzhe Zhang. Market-based incentives. 2013.
[33] Lars Peter Hansen and Thomas J Sargent. Robust control and model uncertainty. *American Economic Review*, pages 60–66, 2001.
[34] Lars Peter Hansen and Thomas J Sargent. *Robustness*. Princeton university press, 2008.
[35] Lars Peter Hansen, Thomas J Sargent, Gauhar Turmuhambetova, and Noah Williams. Robust control and model misspecification. *Journal of Economic Theory*, 128(1):45–90, 2006.
[36] Takashi Hayashi and Jianjun Miao. Intertemporal substitution and recursive smooth ambiguity preferences. *Theoretical Economics*, 6(3):423–472, 2011.
[37] Zhiguo He. Optimal executive compensation when firm size follows geometric brownian motion. *Review of Financial Studies*, 22(2):859–892, 2009.
[38] Zhiguo He. A model of dynamic compensation and capital structure. *Journal of Financial Economics*, 100(2):351–366, 2011.
[39] Bengt Holmstrom. Agency costs and innovation. *Journal of Economic Behavior & Organization*, 12(3):305–327, 1989.
[40] Bengt Holmstrom and Paul Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328, 1987.
[41] Nengjiu Ju and Jianjun Miao. Ambiguity, learning, and asset returns. *Econometrica*, 80(2):559–591, 2012.

[42] Steven N Kaplan and Per Strömberg. Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *The Review of Economic Studies*, 70(2):281–315, 2003.
[43] John Maynard Keynes. The general theory of interest, employment and money, 1936.
[44] Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. Recursive smooth ambiguity preferences. *Journal of Economic Theory*, 144(3):930–976, 2009.
[45] Frank Knight. Risk, uncertainty and profit, 1921.
[46] Markus Leippold, Fabio Trojani, and Paolo Vanini. Learning and asset prices under ambiguous information. *Review of Financial Studies*, 21(6):2565–2597, 2008.
[47] Gustavo Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011.
[48] Jianjun Miao and Alejandro Rivera. Robust contracts in continuous time. *Boston University*, 2013.
[49] Hui Ou-Yang. Optimal contracts in a continuous-time delegated portfolio management problem. *Review of Financial Studies*, 16(1):173–208, 2003.
[50] Etienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.
[51] Tomasz Piskorski and Alexei Tchistyi. Optimal mortgage design. *Review of Financial Studies*, 23(8):3098–3140, 2010.
[52] Julien Prat and Boyan Jovanovic. Dynamic incentive contracts under parameter uncertainty. Technical report, National Bureau of Economic Research, 2010.
[53] Vincenzo Quadrini. Investment and liquidation in renegotiation-proof contracts with moral hazard. *Journal of Monetary Economics*, 51(4):713–751, 2004.
[54] Yuliy Sannikov. A continuous-time version of the principal-agent problem. *The Review of Economic Studies*, 75(3):957–984, 2008.
[55] Heinz Schaettler and Jaeyoung Sung. The first-order approach to the continuous-time principal–agent problem with exponential utility. *Journal of Economic Theory*, 61(2):331–371, 1993.
[56] Marciano Siniscalchi. Vector expected utility and attitudes toward variation. *Econometrica*, 77(3):801–855, 2009.
[57] Stephen E Spear and Sanjay Srivastava. On repeated moral hazard with discounting. *The Review of Economic Studies*, 54(4):599–617, 1987.
[58] Tomasz Strzalecki. Temporal resolution of uncertainty and recursive models of ambiguity aversion. *Econometrica*, 81(3):1039–1074, 2013.
[59] Martin Szydlowski. Ambiguity in dynamic contracts. Technical report, Discussion Paper, Center for Mathematical Studies in Economics and Management Science, 2012.
[60] Jonathan Thomas and Tim Worrall. Income fluctuation and asymmetric information: An example of a repeated principal-agent problem. *Journal of Economic Theory*, 51(2):367–390, 1990.
[61] Xuan Tian and Tracy Yue Wang. Tolerance for failure and corporate innovation. *Review of Financial Studies*, 27(1):211–255, 2014.
[62] Noah Williams. Persistent private information. *Econometrica*, 79(4):1233–1275, 2011.
[63] Yuzhe Zhang. Dynamic contracting with persistent shocks. *Journal of Economic Theory*, 144(2):635–675, 2009.
[64] John Y Zhu. Optimal contracts with shirking. *The Review of Economic Studies*, 80(2):812–839, 2013.

European University Institute