Optimal and robust contracts for a risk-constrained principal

L. C. G. Rogers[*]
University of Cambridge
First version: April 2008

September 2, 2009

Abstract

The theory of risk measurement has been extensively developed over the past ten years or so, but there has been comparatively little effort devoted to using this theory to inform portfolio choice. One theme of this paper is to study how an investor in a conventional log-Brownian market would invest to optimize expected utility of terminal wealth, when subjected to a bound on his risk, as measured by a coherent law-invariant risk measure. Results of Kusuoka lead to remarkably complete expressions for the solution to this problem.

The second theme of the paper is to discuss how one would actually manage (not just measure) risk. We study a principal/agent problem, where the principal is required to satisfy some risk constraint. The principal proposes a compensation package to the agent, who then optimises selfishly ignoring the risk constraint. The principal can pick a compensation package that induces the agent to select the principal's optimal choice.

Keywords: Principal, agent, contract, risk measure, optimal investment

AMS Subject Classifications:

1   Introduction

The study of risk measurement in the mathematical finance literature can be said to date from the seminal paper [2], and has since grown into one of the biggest branches of the

subject. While a lot of effort has been expended on framing axioms for risk measures and their consequences, less attention has focused on how one might apply these notions to more practical matters. A particularly interesting question is how one might optimally invest if constrained by a bound on some risk measure. Early contributions to this literature were [10], [7], [3], which dealt with the problem of optimizing subject to a VaR constraint, among other questions. The use of VaR as a risk measure remains commonplace in the industry, though it has been comprehensively and justly discredited in the academic literature. Coherent or convex risk measures are generally preferred as ways of expressing the riskiness of a position, and we shall discuss only these. A further natural restriction to be placed on a risk measure is *law invariance*, which is to say that two contingent claims with the same law are considered equally risky[1]. Law-invariant risk measures have been characterised by [16] and [11]. In Section 2, we shall work with a standard complete log-Brownian market, and solve the problem of investing to optimize the expected utility of terminal wealth, subject to a *general* constraint on the law of the terminal wealth. The key observation is that the optimal terminal wealth must be a decreasing function of the state-price density[2]. We obtain the simple but appealing result (Proposition 1) that the solution to the constrained optimization problem is the same as the solution to an *unconstrained* optimization problem for a different utility. In Section 4 we study the situation where the constraint on the law comes from some law-invariant coherent risk measure. The form of the solution is more concrete, and numerical solution is possible in some cases.

But beyond the characterisation of the optimal policy, we are interested in the contracting problem which would arise between a trader and an investment bank which employs him. The investment bank will have its own utility, and will in addition be subject to regulatory constraints on risk. The question we next investigate is what contract the bank might offer the trader so that *the trader acting in his own self-interest implements the bank's preferred solution.* The theory of contracts is a difficult area of economic theory; see the excellent survey of Stole [14]. We shall obviate many of the difficulties, which can arise in a situation where the agent may misreport effort or outcomes, or where the level of effort of the agent is not verifiable by the principal; the paper of Palomino & Prat [12], and further papers cited there, provide examples of the kinds of problems that are tackled in this literature. In the context of a trader in an investment bank, his every action is recorded electronically and subject to daily inspection, so misreporting is virtually impossible. Likewise, the trader has no opportunity to slack; he sits in an open-plan workspace with his boss within a few feet most of the day. Moreover, competitive pressure from other traders means that it is reasonable to assume that the trader will give maximum effort regardless. And in any case, in the story we are telling, there is no scope for cleverness or innovation on the part of the trader; whatever he may believe about his god-like insights, he is working in a complete market, and the only thing

---

[1]While initially appealing, there are reasons why this may not be desirable; a trader would not necessarily be indifferent between two contingent claims with the same law, one of which was the market portfolio, the other being negatively correlated with the market portfolio. In a similar spirit, Cherny & Grigoriev [4] point out that law invariance would not be a natural assumption when an agent has an existing position to be offset against the proposed contingent claim.

[2]This is a result of Dybvig [5], [6].

he has to do is to implement the trades that replicate his desired terminal wealth. In such a situation, it is not clear why the principal would employ an agent at all, since the principal could theoretically simply place the desired trades himself. However, the management of the portfolio still requires attention, and therefore would place demands on the principal, and we propose that this is the reason he would be willing to pay someone to act for him.

We turn in Section 3 to the contract which the principal would offer the agent. We suppose that the principal has solved his optimization problem, starting from a given initial wealth, and respecting some risk-measure constraints, and then consider three different contracts which might be offered to the agent. The first is the *first-best contract*; this is the cheapest contract whose payoff would satisfy the agent's participation constraint, while delivering to the principal his desired terminal wealth, assuming that the principal can observe and contract upon the precise actions of the agent. We show that the agent's reward takes a familiar form, namely, the inverse marginal utility of some multiple of the state-price density. This is not particularly surprising; nevertheless, this allows us to express the contract, the function of terminal wealth which is paid as a wage to the agent, in a reasonably explicit way. Notice that we think that the principal commits a fixed initial wealth to the creation of his desired portfolio, and then pays whatever extra cost is required to get the agent to act for him. This is technically easier than (but conceptually equivalent to) taking an initial wealth which gets split into two parts, one of which ultimately provides the agent's wage, the other of which ultimately becomes the principal's payoff. The second form of contract considered is the *second best* contract, where the agent is given the contract and then acts without supervision to generate the wealth from trading; this is a more realistic form of contract than the first best. The final contract considered is what we refer to as a *robust* contract. The idea here is that the contract between principal and agent is chosen in such a way that *their objectives are perfectly aligned*. This appears to be a notion outside the usual realm of contracting theory, but it seems to have interest and merit. There are many different ways to align the objectives of principal and agent, corresponding to affine transformations of the agent's utility, so what we propose is that the principal chooses the cheapest one which guarantees the agent his participation utility. Of course, this contract will be more expensive than the second best contract, but we are able to compare the three contracts, both in terms of their cost, and also in terms of the wage function which results. One reason that robust contracts might in practice be preferred is that these contracts survive arbitrary distortions of the distributional assumptions[3]; whatever the distribution of the underlying random processes, the interests of principal and agent are the *same*. This embodies a feature common in solutions to optimal contracting problems, where the optimal contract prevents the agent from cheating by ensuring that if he cheats he only takes from himself.

---

[3]Nevertheless, the principal's optimal contract was derived under precise distributional assumptions, so it is not the case that the robust contract we derive will deliver the principal his optimal solution if the distributional assumptions are varied.

2 Optimal investment under a constraint on the terminal law.

We shall suppose that there is a complete market with a single risky asset[4], where the dynamics of the wealth process w take the familiar form

$$dw_t = rw_t dt + \theta_t(\sigma dW_t + (\mu - r)dt), \tag{2.1}$$

for some given initial value w_0 of wealth, where W is a standard Brownian motion. The previsible portfolio process θ should be chosen in such a way that the terminal wealth w_T generated at the fixed terminal time T should attain

$$\sup_{\theta \in \mathcal{A}} EU(w_T), \tag{2.2}$$

where $\mathcal{A}$ denotes the collection of all admissible portfolio processes θ, which have the property that the wealth process w generated remains non-negative for all time; see [9] for a full discussion of the mathematical framework and the notions used here. The utility function $U : [0, \infty) \to [-\infty, \infty)$ is supposed to be concave, and strictly increasing where finite, and to satisfy the condition[5]

$$\lim_{x \to \infty} U'(x) = 0. \tag{2.3}$$

We write I for the inverse marginal utility,

$$I(y) = \inf\{x : U'(x) < y\}. \tag{2.4}$$

The market price of risk

$$\kappa \equiv \sigma^{-1}(\mu - r) \tag{2.5}$$

and the state-price density process ζ defined by

$$d\zeta_t = \zeta_t(-rdt - \kappa dW_t), \quad \zeta_0 = 1, \tag{2.6}$$

play a key rôle in the solution of optimal investment problems in this context; it is well known (see [9] again) that the optimal terminal wealth w_T^* satisfies

$$w_T^* = I(\lambda \zeta_T) \tag{2.7}$$

for some $\lambda > 0$ chosen to match the budget constraint

$$w_0 = E[\zeta_T w_T]. \tag{2.8}$$

We now modify the problem; the objective remains to maximise $EU(w_T)$ over admissible portfolio processes θ, but now subject to further constraints *which may be expressed solely in*

---

[4]There would be no difficulty in extending the formulation to multiple risky assets, with suitably bounded processes for volatility and growth rate; the key assumption is that the market is complete. We take the unvariate constant coefficient case only to allow us to do explicit calculations for a number of examples.

[5]Notice that we do not assume the Inada conditions, for reasons that will be explained later.

terms of the law of the terminal wealth[6]. Such further constraints may often be interpreted as some form of risk-management constraint. The optimization problem is therefore to select a *distribution* for the terminal wealth, subject to the budget constraint, so as to maximize the expected utility of terminal wealth. As Dybvig [5] proves[7], if the law of the terminal wealth is specified, then the cheapest way to achieve a terminal wealth with that law is to take $w_T = \psi(\zeta_T)$ for some decreasing function ψ defined by the property that w_T has the desired distribution. Thus the problem is to select a decreasing function ψ to maximise the objective $EU(\psi(\zeta_T))$ subject to the budget constraint $E\zeta_T\psi(\zeta_T) = w_0$, and whatever further constraints on the law of w_T have been imposed. We shall make the following assumption:

ASSUMPTION A: $\quad \lim_{z\to\infty} \psi(z) = 0.$

The point of this assumption is that if it fails, then the optimal terminal wealth $w_T^* = \psi(\zeta_T)$ is bounded below by some positive constant. We may interpret this as a subsistence level of wealth which must be ensured, by investment in the riskless bank account. Once the required wealth to guarantee subsistence is set aside, the problem of investing the remainder under the corresponding constraint on the law of the terminal wealth leads to optimal terminal wealth equal to $\psi(\zeta_T) - \psi(\infty)$, so Assumption A does not make a substantive change to the optimal policy.

The conclusion of this argument is summarized as the following proposition, which is easy to state and understand, and provides a useful characterization of the optimal solution; see Takahashi & Yamamoto [15] for a similar result.

Proposition 1 *Assume Assumption A. The terminal wealth* $w_T^* = \psi(\zeta_T)$ *which optimizes* (2.2) *subject to constraints on the law of* w_T *has the same form as the* unconstrained *solution to the problem*

$$\max_{w_T} Eu(w_T) \quad \text{subject to} \quad w_0 = E[\zeta_T w_T], \tag{2.9}$$

where the utility function u satisfies

$$\psi = (u')^{-1}. \tag{2.10}$$

PROOF. Because $\lim_{z\to\infty} \psi(z) = 0$, the function ψ is onto $(0, a)$ for some $a \leq \infty$, and its inverse function is well defined and decreasing from $(0, \infty)$ to $[0, \infty)$. Clearly the optimal solution $w_T^* = \psi(\zeta_T)$ to the original (constrained) optimization problem is budget-feasible. The optimal w_T^{**} for the problem (2.9) must satisfy budget feasibility along with the optimality condition that $u'(w_T^{**})$ is a multiple of the state-price density ζ_T. However, from the way that u was defined, $u'(w_T^*) = \zeta_T$, so $w_T^{**} = w_T^*$. □

REMARKS. The point of this result will be amplified in Section 4, when we study the special case where the constraint on the law is given by a fairly general law-invariant coherent risk

---

[6]This is quite restrictive, in that it rules out all path-dependent constraints, such as drawdown constraints, or constraints that the portfolio lie always in some convex set. Nevertheless, for some important interesting situations, this assumption applies.

[7]Dybvig works under restrictive assumptions, but his result is true in general. In effect, he rediscovers an old result of Hoeffding which characterizes the extremal values of correlation for a joint distribution with given square-integrable marginals. For a contemporary treatment, see for example [13], pp 107–109, or Theorem 2.76 in [8].

measure. We show that the numerical optimization problem can be solved quite effectively. However, the contracting story to be discussed in Section 3 does not require that we assume this form for the constraint on the law of w_T.

3 Contracting.

We now take the constrained optimization problem solved in Section 2, and treat this as the problem of a *principal*. We shall write U_P for the utility of the principal, and u_P for the *derived* utility of the principal, related to the constrained optimum w_T^* as given by Proposition 1. The agent has utility U_A. We suppose that U_P and U_A are strictly increasing and concave. The principal has initial wealth w_0 to invest so as to maximize

$$EU_P(w_T)$$

over admissible portfolios θ, while satisfying the given constraint on the law of w_T. As we argued in Section 2, the optimal wealth w_T^* has the form $w_T^* = \psi(\zeta_T)$ for some decreasing function ψ. For reasons that will become apparent, we shall assume that

$$U_P(0) = 0, \quad U_P'(0) < \infty, \tag{3.1}$$

and we shall make the additional assumptions that

$$u_P(0) = 0, \quad u_P'(0) < \infty. \tag{3.2}$$

This already implies that the optimal terminal wealth is zero with positive probability, and that for some positive z^* the events $\{w_T^* = 0\}$ and $\{\zeta_T > z^*\}$ are equal[8].

We now suppose that the principal is going to employ an agent to do the trading for him. This agent has utility[9] U_A, and will be paid Y at time T. He will only agree to work if

$$EU_A(Y) \geq \underline{u}, \tag{3.3}$$

where $\underline{u}$ is his reservation utility level. The principal's aim is that he should receive terminal wealth $w_T^* = \psi(\zeta_T)$, which has time-0 value w_0; in order for the wealth generated by trading to provide for this and for the payment Y to the agent, an initial wealth $w_0 + c$ must be made available[10], and when constructing contracts, the principal will be attempting to minimise c subject to the participation constraint (3.3).

We shall assume also that
$$U_A(0) = 0. \tag{3.4}$$

This mild condition is imposed to simplify some of the following arguments.

---

[8]We shall justify these assumptions in the context of law-invariant risk-measure constraints in Section 4, Proposition 3.

[9]U_A is concave and strictly increasing.

[10]Of course, $c = E\zeta_T Y$.

3.1 The first best contract.

The cheapest reward Y which can be offered which would induce the agent to work will solve

$$\inf_Y E[\zeta_T Y] \quad \text{subject to (3.3).}$$

Taking the Lagrangian form of the problem, it is clear that the optimal Y must satisfy the equation

$$U_A'(Y) = \lambda \zeta_T \tag{3.5}$$

for some $\lambda > 0$ chosen to ensure the participation constraint (3.3) holds with equality. Finding this value $\lambda = \lambda^*$ tells us two things. Firstly, it tells us the cost of appointing the agent to act for the principal; this is

$$c_1 \equiv E[\zeta_T I_A(\lambda^* \zeta_T)]. \tag{3.6}$$

Secondly, it tells us the wage schedule φ_1, and this is defined via the equation

$$I_A(\lambda^* z) = \varphi_1(\psi(z) + I_A(\lambda^* z)) \quad \forall z > 0. \tag{3.7}$$

This is because the total amount of wealth generated at time T is $\psi(\zeta_T) + I_A(\lambda^* \zeta_T)$, of which $I_A(\lambda^* \zeta_T)$ is paid to the agent.

What we have just presented is the so-called *first best* contract, where the principal is able to observe the agent's information and actions, and punish any departure from the actions which the principal has specified, thereby enforcing his desired outcome. It is worth noting that in this case we do *not* obtain the Borch rule, where the ratio of marginal utilities of principal and agent is constant; the point of course is that the principal is not seeking an unconstrained maximum. However, if we replace the principal's marginal utility $U_P(w_T^*)$ with the marginal utility $u_P'(w_T^*)$ defined at (2.10), then we do indeed get the Borch rule, as one would expect.

3.2 The second best contract.

The first-best solution is somewhat unrealistic; it is not likely in practice that the principal would wish to scrutinize the actions of the agent, and a so-called *second best* contract would be offered instead. This would simply specify the payment $Y = \varphi_2(X)$ to be made to agent in the event that he produced an outcome X. The agent receives the contract, and is then left to get on with investing. In this case, the agent's optimization problem becomes

$$\sup E[U_A(\varphi_2(X))] \quad \text{subject to } E[\zeta_T X] = w_0. \tag{3.8}$$

Writing $v \equiv U_A \circ \varphi_2$, the first-order conditions for this problem are simply

$$v'(X) \equiv U_A(\varphi_2(X))\varphi_2'(X) \equiv U_A(Y)\varphi_2'(X) = \gamma \zeta_T \tag{3.9}$$

for some $\gamma > 0$, contrasting with (3.5). The principal wants a contract which will leave him with terminal wealth $X - \varphi_2(X)$ equal to $\psi(\zeta)$, and will be least costly subject to the

participation constraint of the agent. The criterion that the principal's terminal wealth should equal $\psi(\zeta)$ is

$$X - \varphi_2(X) \equiv X - U_A^{-1}(v(X)) = \psi(\zeta_T) = \psi\left(\frac{v'(X)}{\gamma}\right) \tag{3.10}$$

which yields the first-order differential equation for v:

$$x - U_A^{-1}(v(x)) = \psi\left(\frac{v'(x)}{\gamma}\right). \tag{3.11}$$

Numerical solution of this requires an initial condition, which is not easy to specify in general; this was the reason for making the assumptions (3.1), (3.4). We know that under the assumption (3.1) there is a critical value z^* such that if $\zeta \geq z^*$ then $w_T^* = \psi(\zeta_T) = 0$; accordingly, we solve the ODE (3.11) with the initial condition

$$v(0) = 0, \quad v'(0) = \gamma z^*. \tag{3.12}$$

This gives a solution v depending on the parameter γ, and we adjust the parameter γ to minimize the cost subject to the agent's participation constraint.

3.3 The robust contract.

We present here a notion different from (but related to) the second best contract. The idea is that the agent receives the contract from the principal, and then optimizes from his own point of view (just as in the second best contract), but the contract is designed so that *the objectives of the principal and agent coincide.* The way that this is achieved is extremely simple; for some $\gamma > 0$, we propose that

$$\frac{U_A(\varphi(x))}{\gamma} = u_P(x - \varphi(x)), \tag{3.13}$$

where we suppose[11] that $\varphi(0) = 0$. The agent will generate total wealth X by time T, of which he gets $Y = \varphi(X)$, and the remaining $Z = X - \varphi(X)$ goes to the principal. We want that the principal's share Z should be the optimal solution w_T^* to his law-constrained optimization problem, which might have unbounded utility $u_P(Z)$; in order to avoid problems with the definition (3.13) therefore, we make the assumption:

ASSUMPTION B: U_A is onto $[0, \infty)$

How should the constant γ in (3.13) be chosen? This is actually very easy. The participation constraint (3.3) gives us

$$EU_A(Y) = \underline{u} = \gamma E u_P(Z) \tag{3.14}$$

and since the optimal terminal wealth Z and the derived utility u_P are already known from the solution to the principal's law-constrained problem, the constant γ is determined.

---

[11]This normalization is natural under our standing assumption that $U_P(0) = U_A(0) = 0$; more general situations can be handled similarly, but we prefer here to keep the exposition simple.

The relationship (3.13) defines the wage schedule φ, and the resulting contract has some good properties, as the following result shows.

Proposition 2 *Assuming that U_A is strictly increasing, the function $\varphi : [\underline{x}, \infty) \to \mathbb{R}^+$ is well defined by (3.13) and $\varphi(0) = 0$. It is increasing, and $U_A \circ \varphi$ is concave.*

PROOF. For any $x > 0$, the function

$$y \mapsto \gamma u_P(x - y) - U_A(y)$$

is continuous and strictly decreasing on $[0, x]$, from a positive value at $y = 0$ to a negative value at $y = x$. Thus there is a unique $y = \varphi(x) \in (0, x)$ at which the function is zero. The monotonicity of φ is obvious.

Turning to the concavity of $f(x) \equiv U_A(\varphi(x))$, suppose that concavity *fails*. Thus there exist $x_1, x_2 \geq \underline{x}$, and $p \equiv 1 - q \in (0, 1)$ such that (with $x = px_1 + qx_2$)

$$
\begin{aligned}
f(x) &= U_A(\varphi(x)) \\
&< pf(x_1) + qf(x_2) \\
&= pU_A(\varphi(x_1)) + qU_A(\varphi(x_2)) \\
&\leq U_A(p\varphi(x_1) + q\varphi(x_2)),
\end{aligned}
\tag{3.15}
$$

and so $\varphi(x) < p\varphi(x_1) + q\varphi(x_2)$, using the fact that U_A is strictly monotone. Hence

$$
\begin{aligned}
f(x) = U_A(\varphi(x)) &= u_P(x - \varphi(x)) \\
&\geq u_P(x - p\varphi(x_1) - q\varphi(x_2)) \\
&\geq pu_P(x_1 - \varphi(x_1)) + qu(x_2 - \varphi(x_2)) \\
&= pU_A(\varphi(x_1)) + qU_A(\varphi(x_2)) \\
&= pf(x_1) + qf(x_2),
\end{aligned}
$$

contradicting (3.15).

$\square$

REMARKS. The relationship defining φ is (3.13), equivalently,

$$x - \varphi(x) = u_P^{-1}\left(\frac{(U_A \circ \varphi)(x)}{\gamma}\right).$$

Compare this with the relationship (3.11) defining φ_2:

$$x - \varphi_2(x) = (u_P')^{-1}\left(\frac{(U_A \circ \varphi_2)'(x)}{\gamma}\right)$$

with a possibly different γ.

9

4   Law-invariant coherent risk-measure constraints.

We now make the general discussion of Sections 2 and 3 more explicit, by supposing that the constraints to be imposed on the law of w_T are that *a given coherent law-invariant risk measure does not exceed some required threshold.* Similar problems with other constraints of a risk-management flavour have been solved by [10], [7], [3]; what we do here tackles a general *class* of problems (which does not, however, include those cited which deal with VaR constraints). Law-invariant coherent risk measures have been characterised by [11], [1], [16], who show that any coherent risk measure takes the form

$$\rho(X) = \sup\{\rho^\mu(X) : \mu \in \mathcal{M}\} \tag{4.1}$$

where $\mathcal{M}$ is a collection of probability measures on $[0, 1]$, and

$$\rho^\mu(X) \equiv \int \rho_a(X) \, \mu(da), \tag{4.2}$$

where

$$\rho_a(X) \equiv -a^{-1} E[X : X \le F_X^{-1}(a)] = -E[X | X \le F_X^{-1}(a)] = -a^{-1} \int_0^a F_X^{-1}(x) \, dx \tag{4.3}$$

for $a > 0$, and $\rho_0(X) \equiv -\operatorname{essinf}(X)$. Notice that from this definition $\rho_1(X) = EX$. For technical reasons, we shall only consider risk measures ρ^μ for μ in *the set $\mathcal{P}_0$ of probability measures on $[0, 1]$ which do not charge $\{0, 1\}$*. The reasons for this are as follows. A measure which charges 0 puts a hard lower bound on the wealth values which may be attained. We prefer to reinterpret such a hard lower bound by saying that this forces the principal to set aside in a bank account sufficient wealth at time 0 so as to guarantee the hard lower bound at time T; but in this case, we may as well look at utility for wealth in excess of the lower bound, and suppose that the lower bound is zero. Excluding mass on 1 also has an interpretation, since this would require the expectation of X to be bigger than some chosen value. By readjusting the terminal wealth to be very large on the event that ζ is very small, we can arrange for an arbitrarily large expectation at arbitrarily small cost; so such a constraint is not a constraint at all.

In the assumed complete market setting, it is well known that the possible terminal wealths w_T are exactly the $\mathcal{F}_T$-measurable random variables X which satisfy the budget constraint

$$w_0 = E[\zeta_T X]; \tag{4.4}$$

see, for example, [9]. The task therefore is to maximise[12] $EU(X)$ subject to this constraint, and the risk-measure constraint

$$\rho(X) \le -b \tag{4.5}$$

for some constant b. Since our objective only depends on the law of X, as does the risk-measure constraint, by the result of Dybvig we know that it is enough to look only at terminal

---

[12]We are considering the problem of the principal; nonetheless, we shall use the notation U in place of U_P for brevity of exposition.

wealths of the form $X = \psi(\zeta_T)$ for decreasing functions ψ. However, in this case the quantiles of X are simply related to the quantiles of ζ_T;

$$F_X^{-1}(a) = \psi(F_\zeta^{-1}(1-a)).\qquad(4.6)$$

where we write $F_\zeta \equiv F_{\zeta_T}$ as an obvious abbreviation. Hence the shortfall risk measure

$$
\begin{aligned}
\rho_a &= -a^{-1}\int_0^a F_X^{-1}(x)\,dx\\
&= -a^{-1}\int_{1-a}^1 \psi(F_\zeta^{-1}(y))\,dy\\
&= -a^{-1}\int_{F_\zeta^{-1}(1-a)}^\infty \psi(z)\,F_\zeta(dz)\qquad(4.7)
\end{aligned}
$$

is expressed in terms of the quantiles of ζ_T, and then the risk measure ρ^μ can be expressed[13] in terms of the distribution of ζ_T also:

$$
\begin{aligned}
\rho^\mu(X) &= -\int_0^1\left\{a^{-1}\int_{F_\zeta^{-1}(1-a)}^\infty \psi(z)\,F_\zeta(dz)\right\}\mu(da)\\
&= -\int \psi(z)\left\{\int_{1-F_\zeta(z)}^1 a^{-1}\,\mu(da)\right\}F_\zeta(dz)\\
&= -E[\psi(\zeta_T)g_\mu(\zeta_T)],\qquad(4.8)
\end{aligned}
$$

where we see that g_μ is non-negative increasing. The optimisation problem therefore becomes

$$\max_\psi EU(\psi(\zeta_T))\quad\text{subject to}\quad w_0 = E[\zeta_T\psi(\zeta_T)],\quad E[\psi(\zeta_T)g_\mu(\zeta_T)]\geq b\quad \forall\mu\in\mathcal{M}\qquad(4.9)$$

where the function ψ is understood to be decreasing and non-negative.

We shall explore this problem under the simplifying assumption that

$$\mathcal{M} = \{\mu_1,\ldots,\mu_n\}\qquad(4.10)$$

is a finite set. Writing $g_i \equiv g_{\mu_i}$, we could (and shall) allow a slightly more general form of the problem (4.9), by taking $E[\psi(\zeta_T)g_i(\zeta_T)]\geq b_i$ for each i, where the b_i may be different. Recall that we assume (2.3), $\lim_{x\to\infty}U'(x) = 0$, and we have $g_i(-\infty) = 0$ for all i since we suppose that $\mu\in\mathcal{P}_0$. For further technical reasons, we shall insist that

$$\lim_{z\to\infty}\frac{g_i(z)}{z} = 0\qquad(4.11)$$

for each i.

---

[13]Recall that μ does not charge $\{0,1\}$.

The Lagrangian form of the optimisation problem is to maximise over non-increasing ψ and non-negative slack variables z_i the Lagrangian[14]

$$
\begin{aligned}
L(\psi, z) &\equiv E\left[U(\psi(\zeta)) + \lambda(w_0 - \zeta\psi(\zeta)) + \sum_{i=1}^{n} \alpha_i\{\psi(\zeta)g_i(\zeta) - b_i - z_i\} \right] \\
&= E\left[U(\psi(\zeta)) - \psi(\zeta)\{\lambda\zeta - \sum_{i=1}^{n} \alpha_i \; g_i(\zeta)\} - \alpha \cdot (z + b) \right] + \lambda w_0. \quad (4.12)
\end{aligned}
$$

Dual-feasibility requires that $\alpha \geq 0$, and complementary slackness gives $\alpha \cdot z = 0$ at optimality. Moreover, dual-feasibility also requires that

$$
\lambda \geq \sup_{x>0} \frac{\sum_{i=1}^{n} \alpha_i \; g_i(x)}{x}, \quad (4.13)
$$

otherwise for some $x > 0$ we could make the objective unbounded by taking $\psi(x)$ very large.

The optimisation of L over non-increasing ψ is straightforward if the function

$$
h(z) \equiv \lambda z - \sum_{i=1}^{n} \alpha_i \; g_i(z) \quad (4.14)
$$

is monotone increasing, for then we simply use the pointwise maximisation $U'(\psi(z)) = h(z)$, which defines the value $\psi(z)$ uniquely. However, if h is not monotone increasing, the story is more subtle. To explain what happens, define

$$
\begin{aligned}
\tilde{h}(x) &\equiv h(F_\zeta^{-1}(x)) \\
&= \lambda F_\zeta^{-1}(x) - \sum_{i=1}^{n} \alpha_i \int_{1-x}^{1} a^{-1} \; \mu_i(da), \\
\tilde{\psi}(x) &\equiv \psi(F_\zeta^{-1}(x)),
\end{aligned}
$$

mapping $[0, 1]$ to $\mathbb{R}$, where we require that $\tilde{\psi}$ is non-increasing. The interesting part of the Lagrangian can be expressed as

$$
E\left[U(\psi(\zeta)) - \psi(\zeta)h(\zeta) \right] = \int_0^1 \{U(\tilde{\psi}(x)) - \tilde{\psi}(x)\tilde{h}(x)\} \; dx, \quad (4.15)
$$

to be optimised over non-increasing $\tilde{\psi}$ which decrease to 0. Now set

$$
H(x) \equiv \int_0^x \tilde{h}(y) \; dy,
$$

and let $\underline{H}$ be the *greatest convex minorant* of H, which we may express as

$$
\underline{H}(x) = H(x) + \eta(x)
$$

---

[14]We abbreviate ζ_T to ζ for this part of the discussion.

for a non-positive function η which is differentiable almost everywhere, and is equal to the integral of its derivative. In addition, $\eta(0) = \eta(1) = 0$. Since $\underline{H}$ is convex, its derivative $\tilde{h}(x) + \eta'(x)$ is non-decreasing, and so we may estimate (4.15)

$$
\begin{aligned}
\int_0^1 \{U(\tilde{\psi}(x)) - \tilde{\psi}(x)\tilde{h}(x)\} \, dx &= \int_0^1 \{U(\tilde{\psi}(x)) - \tilde{\psi}(x)(\tilde{h}(x) + \eta'(x))\} \, dx + \int_0^1 \tilde{\psi}(x)\eta'(x) \, dx \\
&\leq \int_0^1 \tilde{U}(\tilde{h}(x) + \eta'(x)) \, dx + [\tilde{\psi}(x)\eta(x)]_0^1 - \int_0^1 \eta(x) \, d\tilde{\psi}(x) \quad (4.16) \\
&\leq \int_0^1 \tilde{U}(\tilde{h}(x) + \eta'(x)) \, dx, \qquad\qquad\qquad\qquad (4.17)
\end{aligned}
$$

where in (4.16) we have integrated by parts, and used the notation $\tilde{U}$ for the convex dual of U, and to reach (4.17) we have used the fact that η is non-positive and $\tilde{\psi}$ is decreasing. Moreover, the bound (4.17) is achieved when we use[15]

$$
\tilde{\psi}(x) = I(\tilde{h}(x) + \eta'(x)), \qquad\qquad\qquad\qquad (4.18)
$$

because at any x where $\eta(x) < 0$, the greatest convex minorant $\underline{H}$ is strictly less than H, and so its slope is not changing; thus $d\tilde{\psi}$ does not charge the set $\{x : \eta(x) < 0\}$, and the second integral in (4.16) vanishes.

Before finishing this Section, we record a simple fact that is needed to justify the assumptions (3.2) imposed on the derived utility u_P in Section 3.

Proposition 3 *The left derivative of $\underline{H}$ at 1 is infinite.*

PROOF. Notice that since we have insisted (4.11) that $g_i(z)/z \to 0$ as $z \to \infty$, it follows from (4.14) that there exists some z_1 such that $h(z) \geq \lambda z/2$ for all $z \geq z_1$. Hence for x close enough to 1

$$
\begin{aligned}
\frac{\underline{H}(1) - \underline{H}(x)}{1 - x} &\geq \frac{H(1) - H(x)}{1 - x} \\
&\geq \frac{\frac{1}{2}\lambda \int_x^1 F_\zeta^{-1}(y) \, dy}{1 - x} \\
&= \frac{\frac{1}{2}\lambda \int_{F_\zeta^{-1}(x)}^\infty z \, F_\zeta(dz)}{1 - x} \\
&= \frac{1}{2}\lambda E[\zeta \,|\, \zeta > F_\zeta^{-1}(x)] \\
&\geq \frac{1}{2}\lambda F_\zeta^{-1}(x) \\
&\to \infty.
\end{aligned}
$$

$\square$

---

[15]The function I is the inverse marginal utility $(U')^{-1}$.

13

In view of the representation (4.18) of the solution, and the fact (3.1) that $U_P'(0) < \infty$, we learn that ψ satisfies Assumption A.

Proposition 3 ensures the property (3.2) of u_P defined by $u_P' = (\psi)^{-1}$, because we have seen that the optimal $\tilde{\psi}$ maximises $U(y) - y\underline{H}'(x)$. Since we have assumed that $U = U_P$ has finite derivative at 0, it follows from Proposition 3 that for all x close enough to 1, the maximising value is $y = \tilde{\psi}(x) = 0$, and so $\psi(z) = 0$ for all $z \geq z^*$, for some z^*. Hence $u_P' \equiv \psi^{-1}$ is bounded above, so $u_P'(0)$ is finite. Moreover, since u_P' is bounded, it is integrable near zero, so we can without loss of generality assume that $u_P(0) = 0$.

4.1 Numerical examples.

We present here and discuss a few numerical examples which illustrate the ideas. Throughout, we shall take $\sigma = 0.35$, $\mu = 0.2$, $r = 0.05$, $T = 1$ and $w_0 = 1$. The utilities of the principal and agent are of the form

$$U(x) = f_R(x + a) - f_R(a), \tag{4.19}$$

where $f_R(x) = x^{1-R}/(1 - R)$, and $R = 0.5$ for the principal, $R = 0.8$ for the agent, with $a = 0.05$ for both. Thus the utilities satisfy the conditions (2.3), (3.1), (3.4), as well as Assumption B. The reservation utility level of the agent is set at $\underline{u} = U_A(0.05)$.

For the risk constraints, specified by the finite set $\mathcal{M}$, we consider five examples:

1. A single measure $\mu_1 = \delta_a$, $b_1 = 0.8$. Thus this is a constraint on the expected shortfall below the a-percentile, and we have

$$\tilde{g}_1(x) = a^{-1} I_{\{x \geq 1-a\}}. \tag{4.20}$$

   For the numerical solution, we picked $a = 0.2$.

2. Again, a single measure $\mu_1(dx) = I_{[\varepsilon,1]}(x) dx/(1 - \varepsilon)$, $b_1 = 1.03$. In this example,

$$\tilde{g}_1(x) = -\frac{\log(\max\{\varepsilon, 1 - x\})}{1 - \varepsilon}. \tag{4.21}$$

3. Here we took two measures of the expected-shortfall form (4.20) used for the first example, taking $a_1 = 0.65$, $a_2 = 0.05$, and $b_1 = 1$, $b_2 = 0.3$.

4. Here we took just one measure, $\mu_1(dx) = \frac{3}{2}\sqrt{x}\,dx$ and $b_1 = 0.525$. This gives

$$\tilde{g}(x) = \frac{3}{2}\left(1 - \sqrt{1 - x}\right). \tag{4.22}$$

5. The final example had three measures, μ_1 and μ_2 both of the expected-shortfall form (4.20), with parameters $a_1 = 0.25$, $b_1 = 0.85$, $a_2 = 0.05$, $b_2 = 0.75$; and μ_3 the same as in Example 4, taking $b_3 = 0.465$.

For each of the examples, we show plots of the function $\tilde{h}$, and of ψ, Figures 1, 3, 5, 7, and 9. These characterise the principal's optimal solution. Notice that in all cases we see the function ψ has one or more flat sections. This has the interesting implication that the optimal terminal

14

wealth $\psi(\zeta_T)$ will actually take a constant value with positive probability. Nevertheless, the portfolio which implements this will not actually stop at that level at any time prior to T, in view of the smoothing effect of the Brownian transition density. Notice also that in all cases the function $\tilde{h}$ is non-monotone; this is in effect why the plots of ψ must have flat sections.

We also present plots of the three contracts studied in Section 3, Figures 2, 4, 6, 8, and 10. In each case, we see that while the second-best and robust contracts are very nearly the same[16], the first best contract is substantially different. We find that the first best contract pays the agent very little when the wealth generated is low, but offers large payment when a large amount of wealth is generated. This qualitative feature is hardly surprising; low levels of generated wealth correspond to high values of the state-price density, and so the principal will not wish to pay much to the agent in such states, since the contribution to the cost $E[\zeta_T Y]$ will be larger. Nevertheless, we see that the impractical first-best contract is very different from the second best.

5 Conclusions.

This paper makes contributions at various levels of generality. At the most general level, it proposes a solution concept for contracting problems rather different from the traditional first best and second best solutions. The essence of a *robust* contract is that it perfectly aligns the objectives of principal and agent, whereas the second best contract reconciles their objectives. In both cases, the principal is happy to let the agent act without supervision, but for different reasons: under second best contracting, the principal knows that the agent's choice will be optimal for himself too; but under robust contracting, the principal knows that the agent is actually optimising the principal's objective.

At the next level of generality, in the context of a complete Brownian market, we find that the solution to an expected utility problem with constraints on the law of the terminal wealth is actually the solution to a different expected utility problem with *no* constraint on the terminal wealth, and this is technically easier to deal with.

When we specialise further to the case where the constraint on the law of the terminal wealth comes from a law-invariant coherent risk measure, we are able to characterise the constrained solution sufficiently explicitly that it is possible in simple examples to compute the solution numerically. We find that the second best and robust contracts are very similar, and that the fee to be paid to the agent is virtually the same in all cases. In the light of this, we might suppose that the robust contract is to be preferred, because it is simpler, and because it is robust; if the agent has superior knowledge of the asset dynamics than the principal, the principal can rest assured that the agent cannot exploit that superior knowledge

---

[16]The agent's fee is reported in the caption to each of the figures. Notice that we should always have that the fee for the robust contract is at least as big as the fee for the second best contract. In some instances, due to numerical inaccuracy, this theoretical inequality is violated by a little. The main source of inaccuracy is probably the numerical integration required; this was done by the trapezium rule for speed, but better schemes could be considered. Increasing the number of grid points is a possible solution, but we found that using too many led to instabilities in the numerical solution to the ODE.

to the principal's disadvantage. However, the construction of the robust contract was based on assumptions about the asset dynamics, so the contract is still sensitive to those assumptions.

Appendix A : some calculations.

We gather here a few calculations required in the main part of the paper.

1. Participation constraint, CRRA utility. Here we suppose that the agent's utility (omitting the A-subscript) takes the form

$$U(x) = \frac{(x+a)^{1-R} - a^{1-R}}{1-R}$$

for some $a > 0$. Using the form of the agent's optimal wealth (3.5), we obtain that for some $\lambda > 0$

$$Y = I(\lambda\zeta),$$

where the agent's inverse marginal utility I is

$$I(y) = (-a + y^{-1/R})^+.$$

Hence

$$U(Y) = U(I(\lambda\zeta)) = E\left[\frac{((\lambda\zeta) \wedge a^{-R})^{1-1/R} - a^{1-R}}{1-R}\right].$$

We know that $\log\zeta = -(r + \frac{1}{2}\kappa^2)T - \kappa W_T$, where for the purposes of this calculation we may (by replacing κ by $|\kappa|$ if need be) suppose $\kappa > 0$. Thus $\lambda\zeta \geq a^{-R}$ if and only if $W_T \leq b \equiv (\log\lambda + R\log a - (r + \frac{1}{2}\kappa^2)T)/\kappa$. We can therefore develop (abbreviating $\beta \equiv 1 - R^{-1}$)

$$
\begin{aligned}
EU(Y) &= E\left[\frac{(\lambda\zeta)^{1-1/R} - a^{1-R}}{1-R} : W_T > b\right] \\
&= \frac{\lambda^\beta}{1-R}\int_b^\infty \exp(-(r + \tfrac{1}{2}\kappa^2)\beta T - \kappa\beta x - x^2/2T)\frac{dx}{\sqrt{2\pi T}} - \frac{a^{1-R}}{1-R}\bar\Phi\left(\frac{b}{\sqrt{T}}\right) \\
&= \frac{\lambda^\beta}{1-R}e^{-\beta T(r+\kappa^2/2R)}\bar\Phi\left(\frac{b + \kappa\beta T}{\sqrt{T}}\right) - \frac{a^{1-R}}{1-R}\bar\Phi\left(\frac{b}{\sqrt{T}}\right), \qquad (\text{A.1})
\end{aligned}
$$

where as usual Φ denotes the standard Normal distribution function, $\bar\Phi$ its tail.

2. Participation constraint, CARA utility. This time, $U(x) = -e^{-\gamma x}/\gamma$, and we constrain the wealth to be non-negative. Then the optimal terminal wealth is (again ignoring the additive constant)

$$Y = (-\gamma^{-1}\log(\lambda\zeta))^+$$

and hence

$$U(Y) = \begin{cases} -\frac{1}{\gamma} & (\lambda\zeta > 1) \\ \\ -\frac{\lambda\zeta}{\gamma} & (\lambda\zeta \leq 1). \end{cases}$$

Now $\lambda\zeta \geq 1$ if and only if $W_T \leq b \equiv (\log\lambda - rT)/\kappa$, and so after a few calculations we derive

$$EU(Y) = -\gamma^{-1}\Phi\left(\frac{b}{\sqrt{T}}\right) - \gamma^{-1}\lambda e^{-rT + \frac{1}{2}\kappa^2 T}\bar\Phi\left(\frac{b + \kappa T}{\sqrt{T}}\right).$$

17

References

[1] C. Acerbi and P. Simonetti. Portfolio optimization with spectral measures of risk. Technical report, Axabank, 2002.

[2] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent risk measures. *Mathematical Finance*, 9:203–228, 1999.

[3] S. Basak and A. Shapiro. Value-at-risk-based management: optimal policies and asset prices. *Review of Financial Studies*, 14:371–405, 2001.

[4] A.S. Cherny and P.G. Grigoriev. Dilatation monotone risk measures are law invariant. *Finance and Stochastics*, 11:291–298, 2007.

[5] P.H. Dybvig. Distributional analysis of portfolio choice. *Journal of Business*, 61:369–393, 1988.

[6] P.H. Dybvig. Inefficient dynamic portfolio strategies or how to throw away a million dollars in the stock market. *Review of Financial Studies*, 1:67–88, 1988.

[7] S. Emmer, C. Klüppelberg, and R. Korn. Optimal portfolios with bounded capital at risk. *Mathematical Finance*, 11:365–384, 2002.

[8] H. Föllmer and A. Schied. *Stochastic Finance.* de Gruyter, Berlin, 2002.

[9] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance.* Springer, New York, 1998.

[10] C. Klüppelberg and R. Korn. Optimale portfolios mit beschränktem risiko. *Solutions*, 3:23–32, 1999.

[11] S. Kusuoka. On law invariant coherent risk measures. *Advances in Mathematical Economics*, 3:83–95, 2001.

[12] F. Palomino and A. Prat. Risk taking and optimal contracts for money managers. *RAND Journal of Economics*, 34:113–137, 2003.

[13] S.T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Vol 1 Theory.* Springer, Berlin, 1998.

[14] Lars P. Stole. Lectures on the theory of contracts and organizations. Unpublished monograph, 2001.

[15] A. Takahashi and K. Yamamoto. Generating a target payoff distribution with the cheapest dynamic portfolio: an application to hedge fund replication. Technical report, University of Tokyo, 2009.

[16] S. Wang, V. Young, and H. Panjer. Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, 21:173–183, 1997.
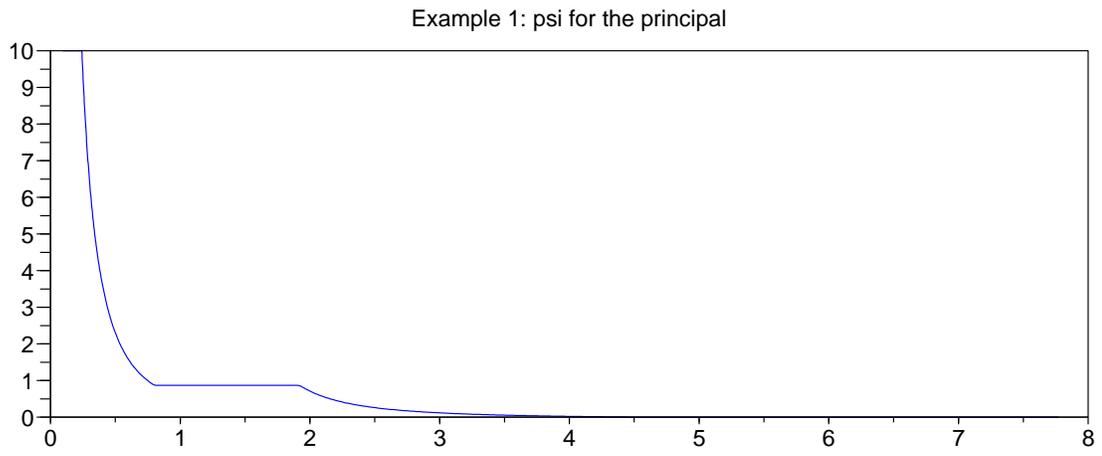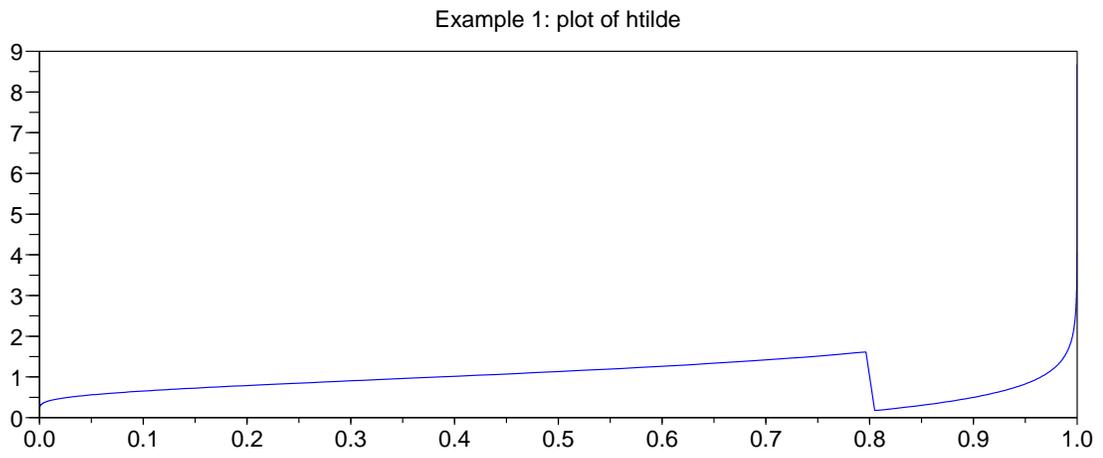
Figure 1: Principal's solution, Example 1.

EXAMPLE 1: First best (black), agent fee = 0.0376854
Second best (blue), agent fee = 0.0435922
Robust (green), agent fee = 0.0437189
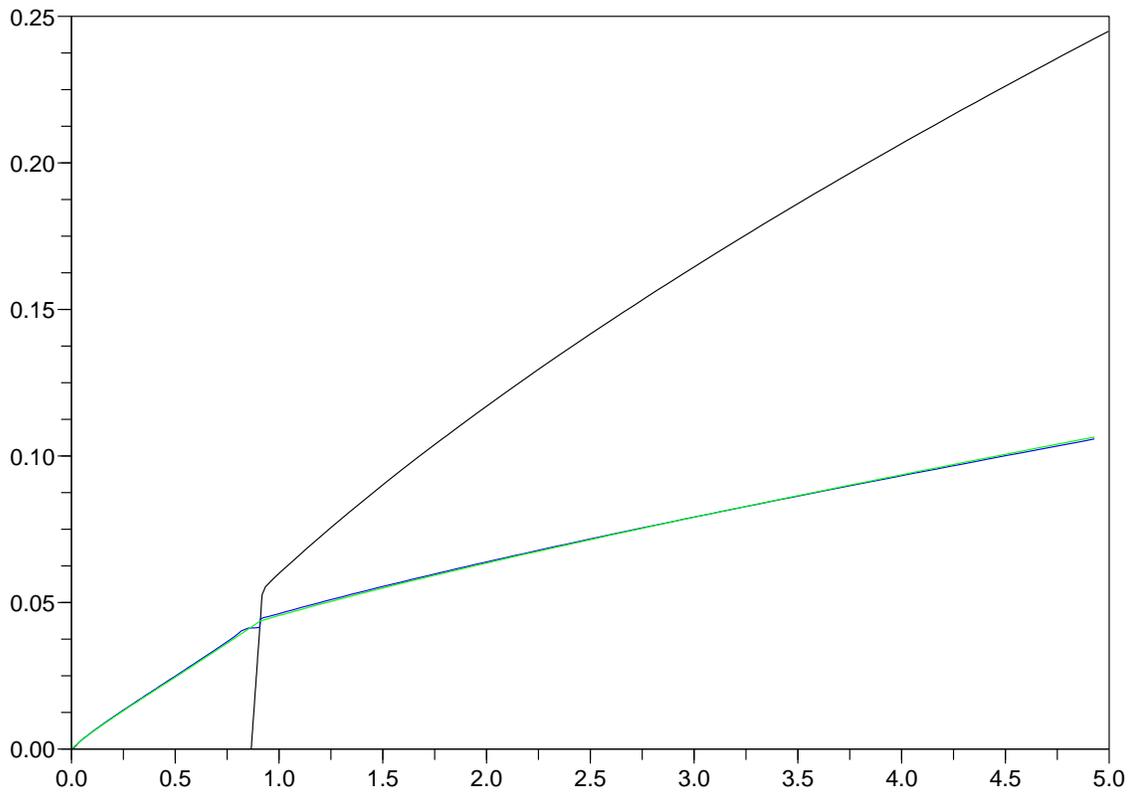

Figure 2: Contracts for Example 1.
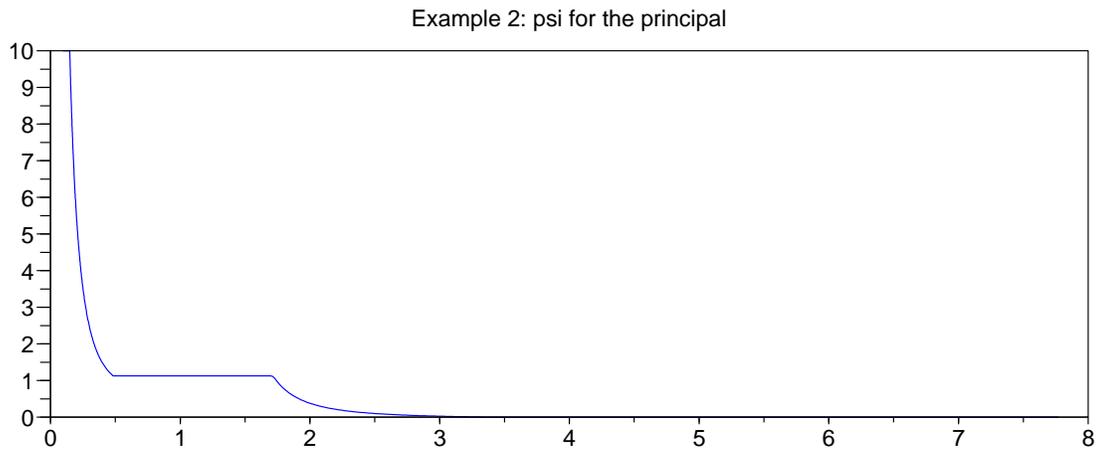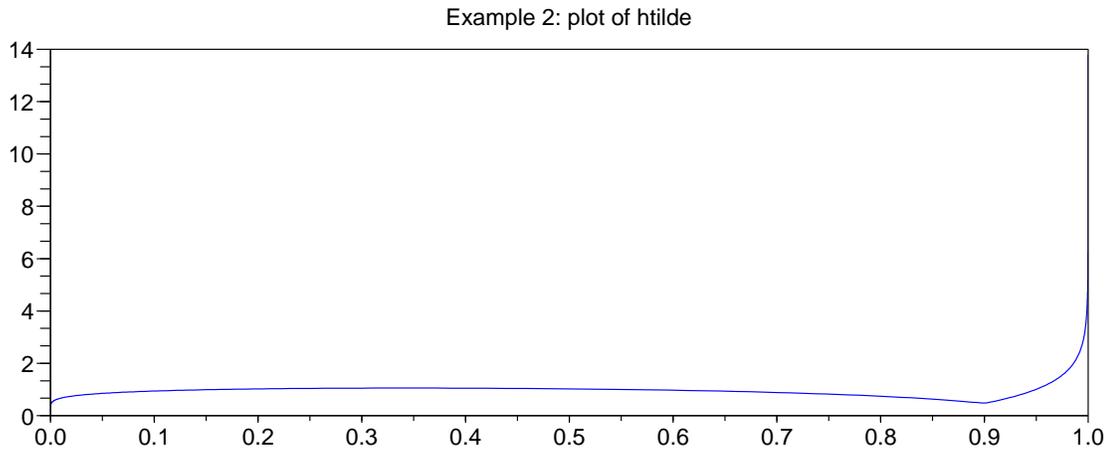
Figure 3: Principal's solution, Example 2.

EXAMPLE 2: First best (black), agent fee = 0.0376854
Second best (blue), agent fee = 0.0454062
Robust (green), agent fee = 0.0454048
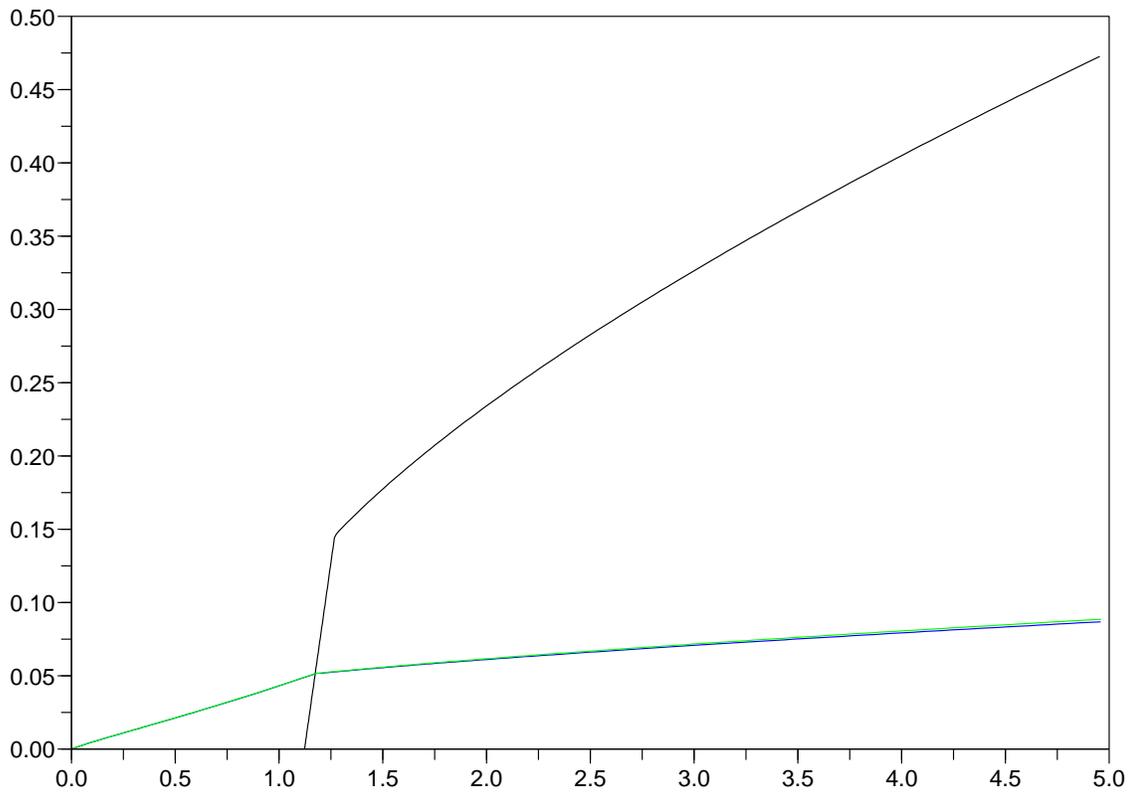
Figure 4: Contracts for Example 2.
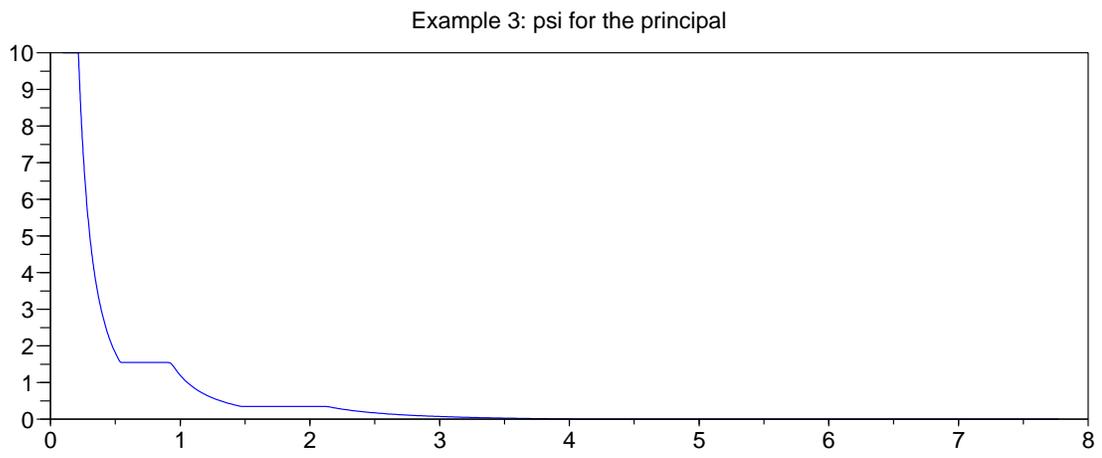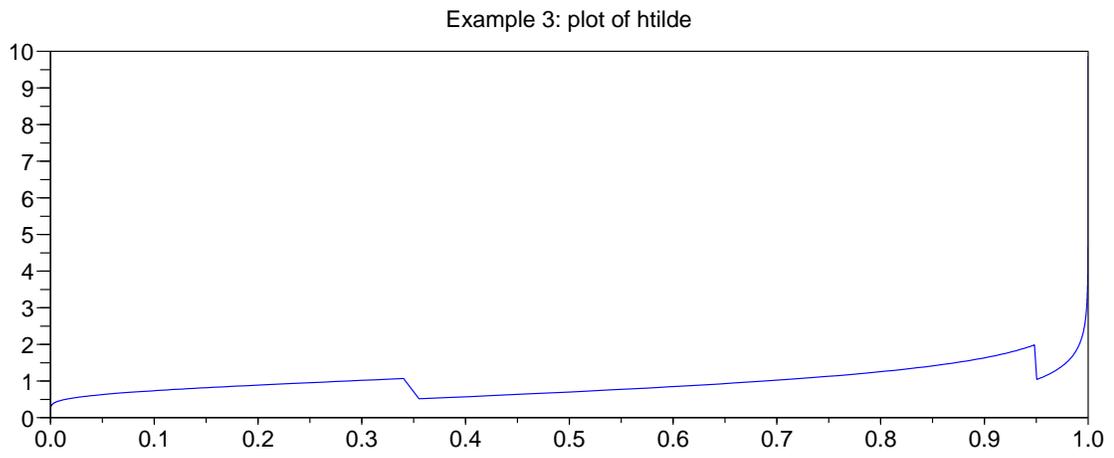
Figure 5: Principal's solution, Example 3.

EXAMPLE 3: First best (black), agent fee = 0.0376854
Second best (blue), agent fee = 0.0414965
Robust (green), agent fee = 0.0414110
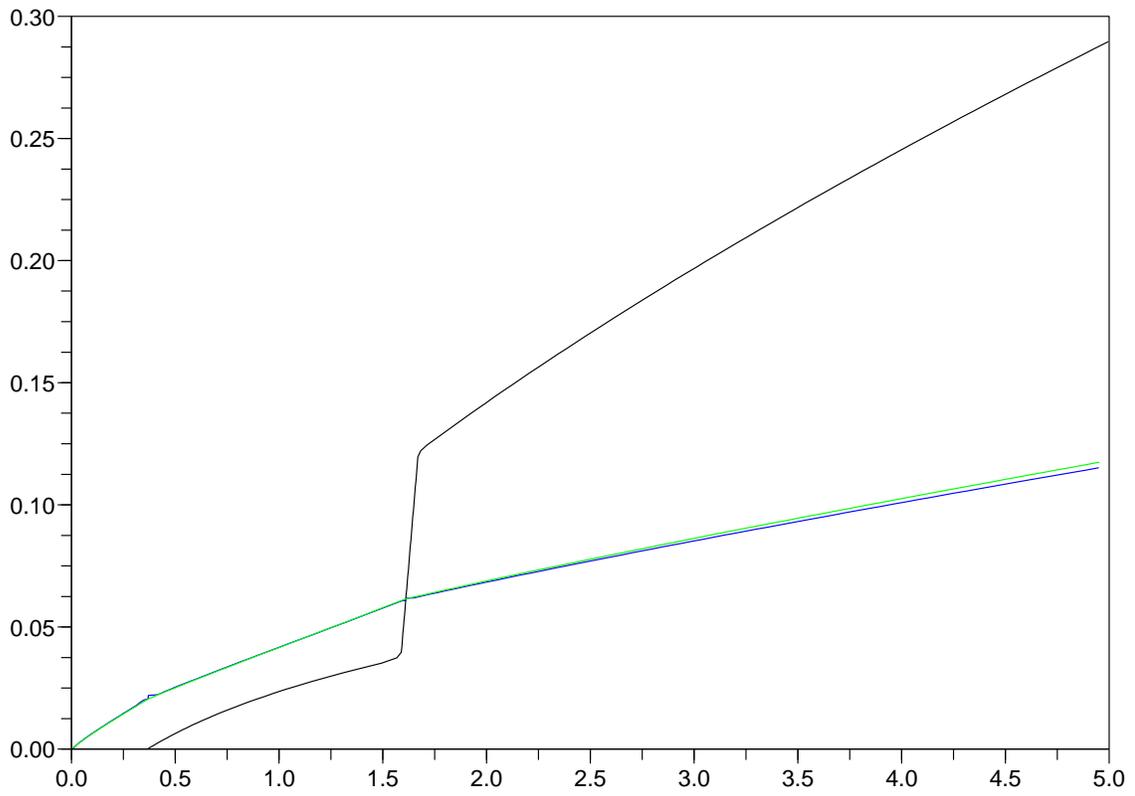
Figure 6: Contracts for Example 3.
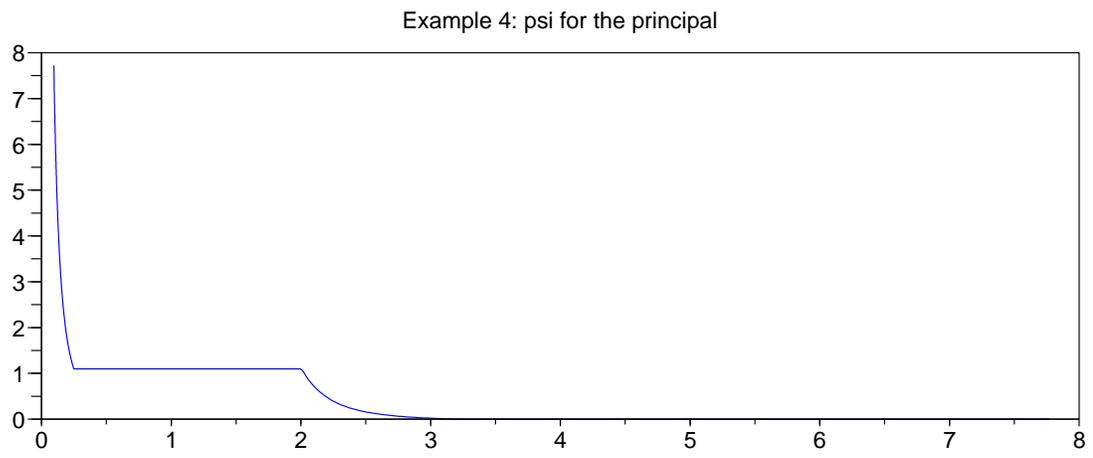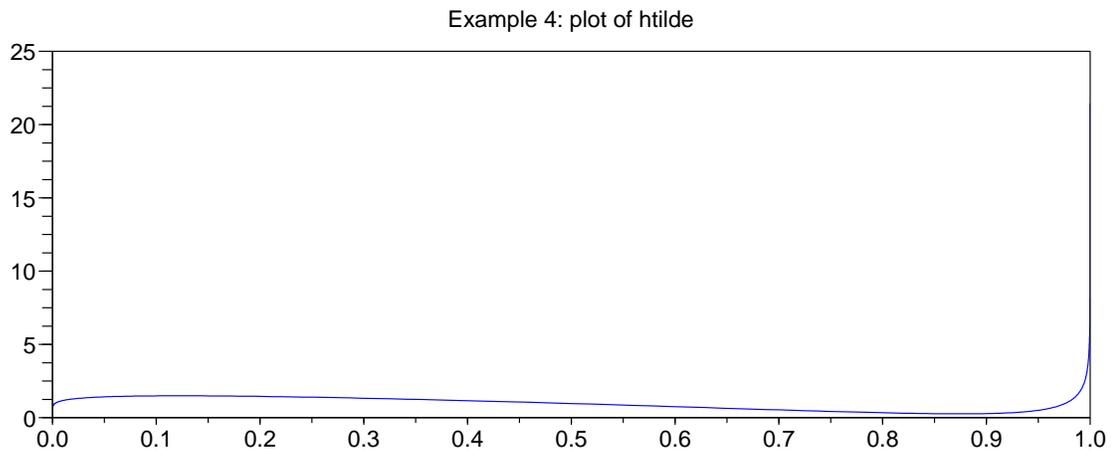
Figure 7: Principal's solution, Example 4.

EXAMPLE 4: First best (black), agent fee = 0.0376854
Second best (blue), agent fee = 0.0464361
Robust (green), agent fee = 0.0464447
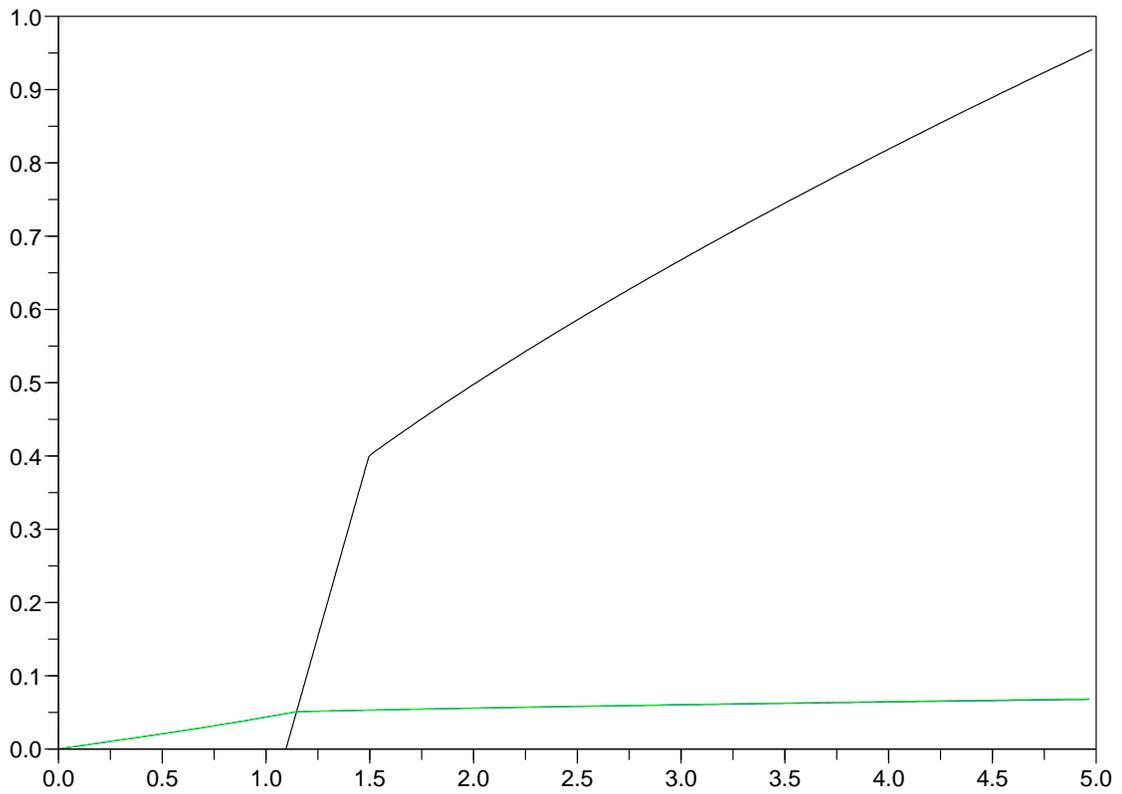
Figure 8: Contracts for Example 4.
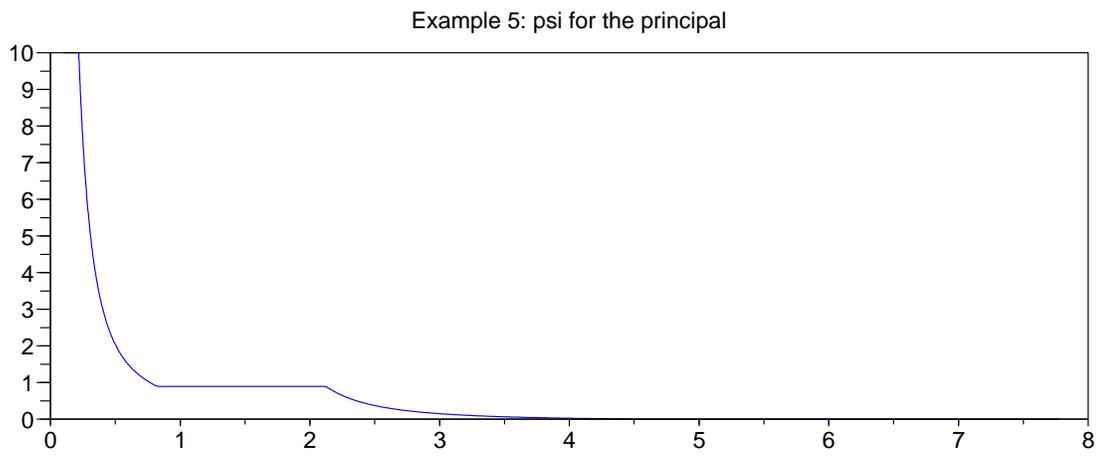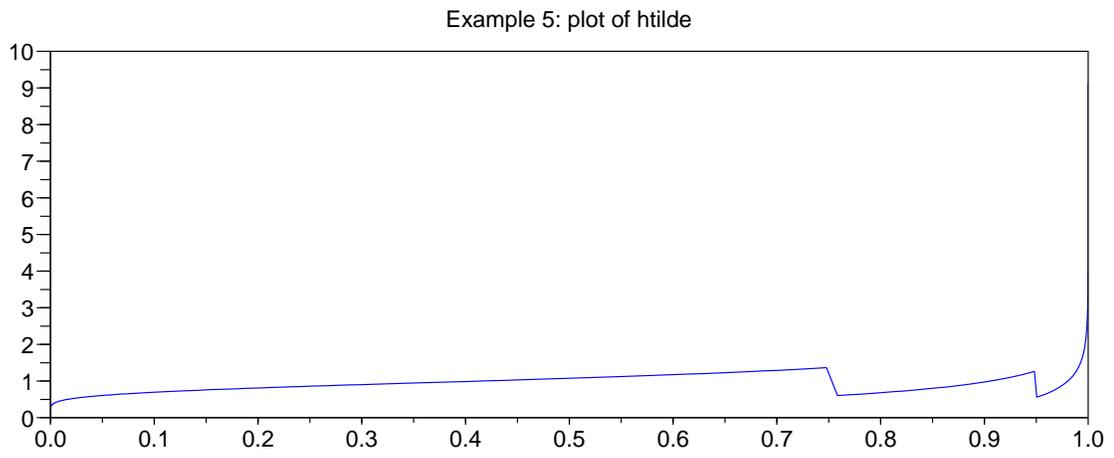
Figure 9: Principal's solution, Example 5.

EXAMPLE 5: First best (black), agent fee = 0.0376854
Second best (blue), agent fee = 0.0445436
Robust (green), agent fee = 0.0446055
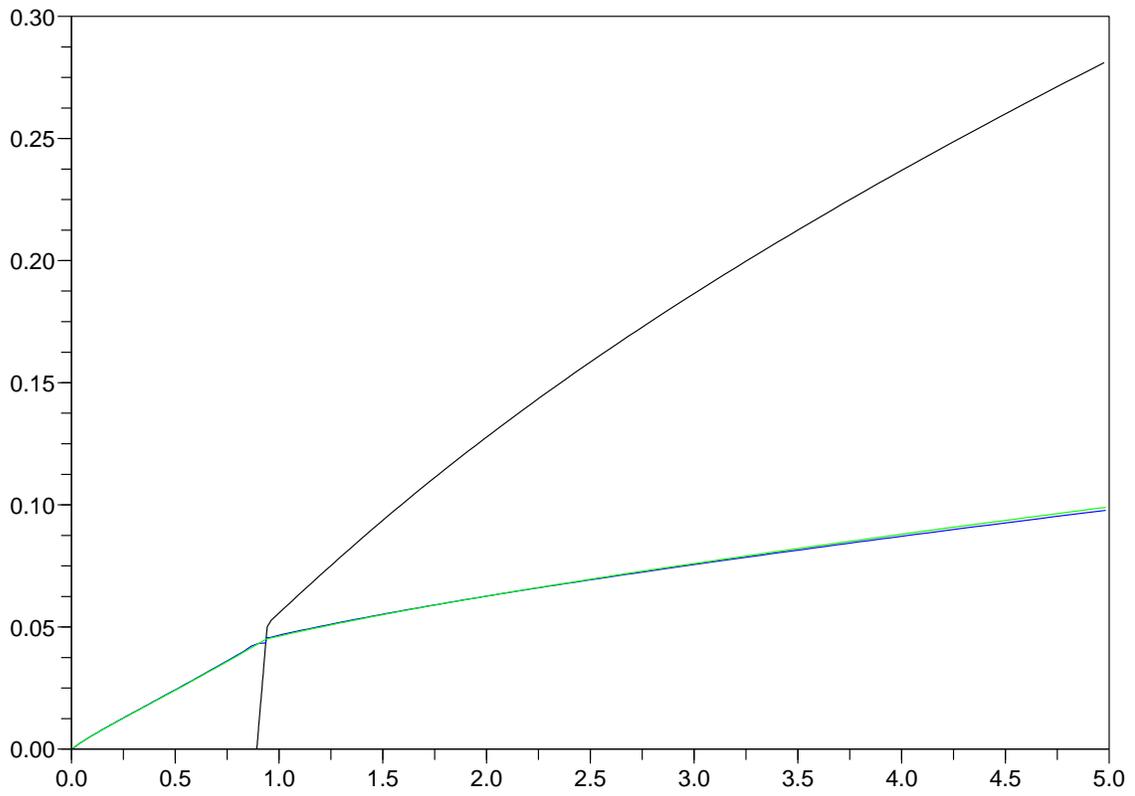
Figure 10: Contracts for Example 5.