Optimal asset management with hidden savings

Sebastian Di Tella  Yuliy Sannikov

April 20, 2015

PRELIMINARY

Abstract

We characterize the optimal contract for an asset manager. The agent can continuously invest in capital and has CRRA preferences over consumption, but faces a moral hazard problem: he can secretly divert funds and has access to hidden savings. We show that the presence of hidden savings creates a dynamic hedging motive. After good outcomes the agent is not only rewarded with more consumption and more capital, but the contract also features a higher growth rate, higher volatility, and a more back loaded consumption path. We prove the validity of the first order approach, and show that the dynamic hedging behavior ensures incentive compatibility of the optimal contract.
1 Introduction

Delegated asset management plays an important role in modern economies, from financial intermediaries such as fund managers, to CEOs or entrepreneurs who manage real capital assets. However, financial frictions can limit the efficient allocation of capital to its most productive users, and have important effects on both asset prices and macroeconomic outcomes. In particular, in the presence of moral hazard it will be necessary to expose the asset manager to risk in order to provide incentives, and this will make it costly to delegate capital. Providing incentives is particularly difficult when the agent has access to hidden savings, which he can use to undo the incentive scheme. This paper studies the impact of hidden savings on the optimal dynamic contract for an agent who manages capital.

We consider a classical investment setting. An agent with CRRA preferences over consumption can continuously invest in risky capital. He would like to raise funds and share risk with a complete financial market, but he faces a moral hazard problem: he can secretly divert funds and has access to hidden savings. The full commitment contract specifies compensation and capital under management contingent on the agent’s returns.

The setting has several features that make it well suited to asset pricing and macro applications. First, the ability to continuously invest in capital and adjust the scale of the project is important in these settings where a key focus is on the allocation and pricing of capital assets, rather than the optimal effort level or compensation scheme. Second, CRRA preferences are widely used both in macro and asset pricing where the relative risk aversion and the elasticity of intertemporal substitution play important roles. Third, hidden savings place realistic constraints on incentive compatible contracts, but can be technically challenging due to the double-deviation problem, especially if we move away from the CARA or linear preferences case. One of the methodological contributions of this paper is to establish the general validity of the first order approach in this environment.

The optimal contract takes a very tractable form with rich dynamics. After good outcomes the contract not only rewards the agent with more consumption and capital, but it also features higher growth, higher volatility, and more back loaded consumption. Conversely, after bad outcomes the contract not only scales down by giving the agent less consumption and capital, but it also features lower growth, lower volatility, and more front loaded consumption. This dynamic behavior is the result of the hidden savings assumption. Without hidden savings the optimal contract would have a constant growth rate, volatility, and intertemporal consumption profile.

To see how the optimal contract works, consider the decision of how much capital to give to the agent. The principal would like to give capital to the agent because it pays an excess return. In the first best without moral hazard, this would create an arbitrage opportunity. In our setting, however, there are two costs associated with capital. The first is that since the agent can secretly divert funds for his own consumption, he must be given some “skin in the game” to deter him from stealing. Since stealing makes bad outcomes more likely, he must be punished after bad outcomes with lower consumption and capital. This exposes the agent to risk proportionally to the capital he manages, which is costly because he is risk averse and the market is risk neutral with respect to
the agent’s idiosyncratic risk (the first best has full insurance). This cost would appear even if the agent didn’t have access to hidden savings, as in Di Tella (2014).

With hidden savings, however, there is a second cost associated with capital. Since giving capital to the agent requires exposing him to risk, it induces a precautionary motive for saving. The agent will want to postpone consumption to insure against the higher risk. This is costly because it distorts the optimal intertemporal consumption smoothing conditional on risk exposure. The optimal contract must therefore trade off the excess return of capital against the required distortions in both risk sharing and intertemporal consumption smoothing. In fact, these two tradeoffs interact. The more the agent postpones consumption, the higher his marginal utility from consumption will be at a given time. This raises the private benefit from diverting funds and immediately consuming them, and so forces the contract to expose the agent to more risk to provide incentives for good behavior. At the same time, the more risk the agent is exposed to the stronger his precautionary motive for postponing consumption will be.

So at one extreme, if the principal wants to give the agent the first best consumption path without any risk, he can’t give him any capital to manage. However, since capital pays an excess return it will be optimal to give him some capital. With more exposure to risk and a more back loaded consumption, the agent’s continuation utility will then have a higher growth rate, and therefore so will the scale of the contract. These “high growth/high volatility” contracts therefore give a large amount of capital to the agent, have high growth rates and volatility, and as a result a very back loaded consumption path. Up to a point, they reduce the cost of delivering utility to the agent, because the principal benefits from the excess return of capital.

We can now understand the contract dynamics. After good outcomes, the agent’s continuation utility will rise because of the “skin in the game” constraint. This could be achieved with a proportional increase in capital and consumption, keeping the growth rate, volatility, and intertemporal consumption profile unchanged. In fact, this would be the optimal response in the case without hidden savings.\(^1\) However, because the agent has access to hidden savings, after good outcomes the contract will shift towards “high growth/high volatility”. Capital will increase more than proportionally to the agent’s continuation utility, the expected growth rate and volatility will go up, and consumption will become more back loaded. We can understand this in terms of intertemporal hedging. Since “high growth/high volatility” contracts are a cheaper way of delivering utility to the agent, the principal prefers to use them when he must deliver more utility (after good outcomes), and use the more costly “low growth/low volatility” contracts when he must deliver less utility (after bad outcomes).

This intertemporal hedging behavior helps ensure that the optimal contract is incentive compatible. Contractual environments where the agent has hidden savings often suffer from the problem of double deviations.\(^2\) Dealing with single deviations is relatively straightforward. Giving the agent some skin in the game can deter him from stealing and immediately consuming the proceeds. Like-

---

1. See Di Tella (2014).
2. See for example Kocherlakota (2004).
wise, incorporating the agent’s Euler equation as a constraint in the contract design problem ensures that he won’t secretly save his recommended consumption for later. But what if the agent both steals and saves the proceeds for later? Since stealing makes bad outcomes more likely, the agent expects to be punished with lower consumption in the future. In other words, he expects a high marginal utility in the future, so it seems like stealing and secretly saving the proceeds for later is an attractive double deviation. We prove this is not the case by establishing an upper bound on the agent’s continuation utility from any valid deviation, after any history, thus establishing the validity of the first order approach in this environment.

There are two important insights regarding incentive compatibility. The first is that the marginal utility of the last dollar in hidden savings is lower than the marginal utility of the first dollar (which is equal to the marginal utility of legitimate consumption). It is true that if the agent steals he expects the marginal utility of legitimate consumption to be high in the future (because he expects to be punished with lower consumption), but he also expects the marginal utility of his hidden savings to be lower than that. This is a result of the assumption that the hidden action takes the form of diverting funds, instead of hidden effort or shirking as in, for example, Kocherlakota (2004). However, this is not enough to ensure incentive compatibility. The second insight is that the agent faces an incomplete market with in his hidden savings problem. If after some history the agent expects his legitimate consumption (which is his income in his hidden savings problem) to be highly volatile looking forward, he will place a large marginal value on hidden savings that can help him self insure. As a result, if the contract increased the volatility of the agents’ consumption after bad outcomes, stealing and saving the proceeds for later would be an attractive double deviation, since the agent would expect to have hidden savings precisely when they are most valuable to him. It is here that the intertemporal hedging behavior helps ensure incentive compatibility. After bad outcomes, the contract shifts to “low growth/low volatility”, so hidden savings becomes less valuable. In fact, it is enough to ensure incentive compatibility that the contract maintains the same risk and intertemporal consumption profile. It is a remarkable feature of the contractual environment that the principal’s cost minimizing intertemporal hedging, created by the introduction of hidden savings, also ensures incentive compatibility with respect to double deviations.

Turning to the long-run behavior of the contract, there are several features that make the contract tractable and easy to embed in general equilibrium asset pricing or macro problems. First, there is no endogenous termination. Because capital can be continuously adjusted, after a very good history the agent will be managing a large amount of capital, but he will not be retired nor outgrow the moral hazard problem as in Sannikov (2008) or Hopenhayn and Clementi (2006) respectively. Since there is no outside option or lower bound on utility, neither will he retire after sufficiently bad outcomes as in DeMarzo and Sannikov (2006): the contract just gives him a very small amount of capital and consumption, but there’s always the chance that he will recover.4

In fact, the long-run behavior of the contract features a non-degenerate stationary distribution.

3See Werning (2001).
4While endogenous termination can be an interesting feature of optimal contracts, in many general equilibrium applications we may want to abstract from them to gain tractability.
In the absence of shocks, contract dynamics would lead to a “steady state” with constant growth rate and volatility. However, the “steady state” can be a very misleading guide to the contract’s long run behavior. The contract spends most of the time in the “low growth/low volatility” region. The reason for this is that with low growth and low volatility the only way for the contract to get out of this region is to slowly move towards the steady state. On the other hand, when the contract finally gets to a “high growth/high volatility” region, even small shocks will quickly move the contract away and it will quickly fall back into the “low growth/low volatility” region.

We also allow for an exogenous retirement with independent Poisson arrival rate, since in many macro applications this is required to ensure a stationary wealth distribution. When the agent retires he cannot manage capital any longer, but he can still consume. We show the in the optimal contract the agent’s utility drops on retirement, and in exchange it grows faster while the agent is still managing capital. The intuition is also related to intertemporal hedging. Once the agent cannot manage capital, it is more costly to deliver utility to him. It is therefore cheaper to give him less utility when he retires, and more utility while he can still manage capital.\(^5\)

Finally, the contract is easy to characterize with a second order ODE. While contractual environments with hidden savings typically require working with two state variables, in our setting the homothetic CRRA preferences and the linear technology with continuous trading yield a scale invariance property that significantly reduce the complexity of the problem. This also makes it easy to embed into general equilibrium models.

**Literature Review.** This paper fits within the literature on dynamic agency problems, such as DeMarzo and Fishman (2007), Sannikov (2008), Ihe (2011), Biais et al. (2007), and Hopenhayn and Clementi (2006). It is most closely related to DeMarzo and Sannikov (2006) and Di Tella (2014). The two main differences with DeMarzo and Sannikov (2006) are the CRRA preferences (vs. linear preferences) and the ability to continuously adjust the capital given to the agent (vs. a fixed project size). Permitting the contract to trade capital and vary the scale of the project allows us to focus on the demand and pricing of capital. CRRA preferences are important for this purpose as well, since we need to incorporate risk aversion and elasticity of intertemporal substitution. In addition, with linear preferences, the optimal contracts with and without hidden savings are the same. Once concave preferences are introduced, the principal has incentives to front load consumption in order to reduce the private benefit of stealing and relax the risk sharing problem. In fact, this is precisely what happens in Di Tella (2014) where the same contractual environment is used to study the optimal financial regulation policy in a general equilibrium setting. The only difference is that the agent doesn’t have access to hidden savings. As a result, in a stationary environment such as the one considered here, the optimal contract would be characterized by a constant growth rate and volatility, and front loaded consumption.\(^6\)

\(^5\)The same mechanism appears in Di Tella (2014).

\(^6\)When the agent doesn’t have access to hidden savings, the principal has a lot of power over him and might be able to eliminate the moral hazard problem. Indeed, if the agent has a low elasticity of intertemporal substitution, by front loading his consumption the agent can achieve an arbitrage, and as a result the optimal contract won’t exist.
Several papers deal with hidden savings. Werning (2001) introduces the first order approach in the context of unemployment insurance. He (2011) characterizes the optimal contract when the agent must put in unobservable effort and has access to hidden savings. He obtains incentive compatibility in the context of a fixed project size and CARA preferences. Williams (2013) studies a similar environment with fixed project size, CARA preferences, hidden effort and hidden savings. CARA preferences play important roles in ensuring incentive compatibility in both papers because they eliminate the wealth effect. Edmans et al. (2011) obtains a tractable incentive scheme in a setting where the agent can take a hidden action and has access to hidden savings. They allow for CRRA preferences, but assume the hidden action enters the agent’s utility function in a multiplicative way with consumption. In addition, the size of the project is also fixed. In contrast to these papers, here we allow CRRA preferences, continuous trading of capital, and the hidden action consists of diverting funds.

TO BE COMPLETED

2 The model

Let $(\Omega, P, F)$ be a probability space with a filtration $F$ generated by a brownian motion $W$ and a Poisson process $N$. Throughout, all stochastic processes are adapted to $F$. There is a complete financial market with equivalent martingale measure $Q$. The risk-free interest rate is $r > 0$ and both $W$ and $N$ are idiosyncratic risks and therefore not priced by the market. In this setting there is no aggregate risk, so $Q = P$, but in general $P$ and $Q$ could differ.

The agent can manage capital and obtain a risky return per dollar invested

$$\frac{dG_t}{G_t} = (r + \alpha)dt + \beta dW_t$$

where $\alpha > 0$ is an excess return and $\beta > 0$ is the volatility. Notice $W$ is agent-specific idiosyncratic risk. If we think of the agent as a fund manager, it represents the outcome of his particular trading activity.\textsuperscript{7} If we take the agent to be an entrepreneur it represents the outcome of his particular project.

The agent wants to raise funds and share risk with the market, but he faces a moral hazard problem. He signs a contract $C = (c, U, k)$ with full commitment on both sides, that specifies a consumption stream $c = \{c_t; t \leq \tau\}$, a terminal utility $U = \{U_t; t \leq \infty\}$ when the agent retires, and an investment strategy $k = \{k_t \geq 0; t \leq \tau\}$, where $k_t$ is the dollar value of the capital he manages at time $t$. The agent retires at a stopping time $\tau$, which arrives with exogenous Poisson intensity $\theta > 0$. After retirement the agent can’t manage capital any longer, so $k_{\tau+u} = 0$ for $u \geq 0$, and the principal delivers utility $U_\tau$. Without any loss of generality, the best way to deliver this utility is with a deterministic consumption stream $c_{\tau+u} = \hat{c}_h \left( (1 - \gamma)U_\tau \right)^{\frac{1}{1 - \gamma}} e^{\frac{\gamma}{1 - \gamma} u}$, where $\hat{c}_h = \left( \frac{\rho - r(1 - \gamma)}{\gamma} \right)^{\frac{1}{1 - \gamma}} > 0$.

\textsuperscript{7}If we give $1 to invest to two fund managers, they will obtain different returns depending on exactly which assets they buy or sell, and the exact timing and price of their trades.
This is not a restriction on consumption after retirement. As we will see below, this retirement contract both minimizes the cost of delivering utility after retirement and also relaxes the problem before retirement.

After signing the contract \( C \) the agent can choose a hidden action \( a = (\tilde{c}, s) \). We interpret 
\[
s = \{ s_t; \ t \leq \tau \}
\]
as stealing, and 
\[
\tilde{c} = \{ \tilde{c}_t; \ t \leq \tau \}
\]
as hidden consumption. The stealing activity changes the distribution of observed outcomes from \( P \) to \( P^s \) so that the return of the agent’s investment is
\[
\frac{dG_t}{G_t} = (r + \alpha - s_t)dt + \beta dW^s_t
\]
where \( W^s_t = W_t + \int_0^t s_u du \) is a Brownian Motion under \( P^s \). For each stolen dollar, the agent keeps only \( \phi \in (0, 1) \) which he can either consume or add to his hidden savings \( h_t \). His hidden savings then evolve
\[
dh_t = (h_tr + c_t - \tilde{c}_t + \phi k_t s_t)\ dt
\]
with \( h_0 = 0 \). Notice the agent gets the risk free interest rate for his hidden savings, but he cannot invest them in capital which is contractible.\(^8\)

The agent has CRRA preferences. Given \( C \), the utility process under action \( a = (\tilde{c}, s) \) is
\[
U^a = \{ U^a_t; \ t \leq \tau \}
\]
defined as the unique square integrable solution to
\[
U^a_t = \mathbb{E}^{s}_{t}[\int_t^\tau e^{-\rho(u-t)}(\frac{\tilde{c}_u^{1-\gamma}}{1-\gamma} - \tilde{c}_u) du + e^{-\rho(\tau-t)} B(\bar{U}_\tau, h_\tau)]
\]
where the expectation is taken under probability distribution \( P^s \) generated by the stealing activity \( s \). The function \( B(U, h) \) is the continuation utility that an agent can achieve if he has been promised utility \( U \) and has accumulated hidden savings \( h \) at the time of retirement. Since utility \( \bar{U}_\tau \) is delivered with a constant consumption stream, the best the agent can do is to “scale up” his consumption using his hidden savings. As a result, his continuation utility takes the form
\[
B(U, h) = \left(1 + h((1 - \gamma)U \frac{\tilde{c}_0^{1-\gamma}}{1-\gamma})\right)^{1-\gamma} U
\]
In particular, if \( h = 0 \) we get \( B(U, 0) = U \).

We say an action \( a = (\tilde{c}, s) \) is valid if 1) \( s \) is bounded,\(^9\) 2) \( U^a \) exists, and 3) \( h_t \geq 0 \) always.\(^10\)

Let \( \mathbb{A}(C) \) be the set of valid hidden actions \( a \) given contract \( C \). It is always optimal to implement \( (\tilde{c}, s) = (c, 0) \) in this environment.\(^11\)

---

\(^8\)If the agent could enter into secret contracts whose payoffs are contingent on his return, he would be able to undo any incentives.

\(^9\)Bounded \( s \) guarantees that \( Z_t = \exp(\int_0^t s_u dW_u - \frac{1}{2} \int_0^t s_u^2 du) \) is a \( P \)-martingale, so we can use it to define the equivalent measure \( P^s \) via Girsanov’s theorem.

\(^10\)The agent can have hidden savings but not hidden debt. This is without loss of generality if the contract can exhaust the agent’s credit capacity.

\(^11\)This is a standard feature of this type of problem. If a contract implements \( (\tilde{c}, s) \neq (c, 0) \) in equilibrium, the principal can do better by giving \( c = \tilde{c} \) to the agent legitimately, saving for him, and asking him to not steal instead. See DeMarzo and Sannikov (2006).
The principal tries to minimize the cost of delivering utility $u_0$ to the agent, where the continuation cost is $J = \{J_t; t \leq \tau\}$ given by

$$J_t = \mathbb{E}_t^Q \left[ \int_t^\tau e^{-r(u-t)} \left( c_u dt - k_u [dG_u/G_u - r dt] \right) + e^{-r(\tau-t)} F(\bar{U}_\tau) \right]$$

where $F(U) > 0$ is the cost of delivering utility $U$ to an agent who has retired and can’t manage capital anymore. The cheapest way to deliver utility $U$ at retirement is through a deterministic consumption stream, and the cost is given by

$$F(U) = \hat{v}_h ((1 - \gamma)U)^{\frac{1}{1-\gamma}} > 0$$

where $\hat{v}_h = \left( \frac{\rho - r(1-\gamma)}{\gamma} \right)^{\frac{\gamma}{1-\gamma}} = \hat{c}_h > 0$.

We say the consumption $(c, U)$ is admissible if $U^{c, 0}$ exists. In order to rule out arbitrage from doubling strategies and the like, we say the trading strategy $k$ is admissible if it is martingale-generating.\footnote{A trading strategy $k$ is martingale-generating if $\int_0^t e^{-r(t-s)} k_s \beta dW_t$ is a Q-martingale. See Karatzas and Shreve (1998), Definition 5.9.}

A contract $C = (c, U, k)$ is admissible if 1) $(c, U)$ and $k$ are both admissible, and 2)

$$\mathbb{E}^Q \left[ \int_0^\tau e^{-rt} |c_t - k_t| \alpha dt + \sup_{u \leq \tau} e^{-ru} F(U^{c, 0}_u) \right] < \infty \quad (1)$$

We say an admissible contract $C$ is incentive compatible if\footnote{Notice $\mathcal{A}(C) \neq \emptyset$ because $(c, 0) \in \mathcal{A}(C)$ for an admissible contract.}

$$(c, 0) \in \arg \max_{a \in \mathcal{A}(C)} U^{a, 0}_0$$

Let $\mathcal{I}C$ be the set of incentive compatible contracts. For an initial utility $u_0$ for the agent, an incentive compatible contract is optimal if it minimizes the cost of delivering initial utility $u_0$ to the agent, that is

$$v_0 = \min_{(c, U, k)} J_0$$

$$st : \quad U^{c, 0}_0 \geq u_0$$

$$(c, \bar{U}, k) \in \mathcal{I}C$$

By changing $u_0$ we can trace the Pareto frontier for this problem. Let $v_t$ be the continuation cost of the optimal contract at any time. This is how much it would cost the agent to “buy into” the contract at that time. In particular, at $t = 0$ the agent needs $v_0$ dollars to buy a contract that delivers utility $u_0$ to him.

\textit{Note 1.} About admissibility condition 2): this assumption makes sure the principal’s objective function is well defined, ruling out exploding strategies where the cost of delivering the retirement utility diverges to infinity in present expected value. But it’s slightly stronger than that, since it...
requires $\mathbb{E}^Q \left[ \sup_{u \leq \tau} e^{-ru} F(U_{u}^{c,0}) \right] < \infty$ instead of $\mathbb{E}^Q \left[ e^{-r\tau} F(\bar{U}_\tau) \right] < \infty$. This technical integrability constraint disciplines admissible contracts, and it won’t be binding for the optimal contract. Lemma 7 in the Appendix provides sufficient conditions.

**Incentive compatibility**

As usual, we use the continuation utility of the agent as a state variable for the contract, which we use to provide incentives for not stealing. However, we also need the discounted marginal utility of the agent as a state variable in order to provide incentives to consume the right amount. In fact, both problems are related because the incentives to steal depend on the marginal utility of consumption.

**Lemma 1.** For any admissible contract $C = (c, \bar{U}, k)$, the agent’s continuation utility $U_{c,0}^{c,0}$ satisfies

$$dU_{t}^{c,0} = \left( \rho U_{t}^{c,0} - \frac{c_{t}^{1-\gamma}}{1-\gamma} \right) dt + \tilde{\sigma}_{U,t} dW_{t} - \lambda_{t} (dN_{t} - \theta dt)$$

(2)

for some $\tilde{\sigma}_{U} \in \mathcal{L}^2$ and $\lambda_{t} = U_{t}^{c,0} - \bar{U}_{t}$.\(^{15}\) In the other direction, if $U$ solves (2) for some $\tilde{\sigma}_{U} \in \mathcal{L}^2$ and $\lambda_{t} = U_{t}^{c,0} - \bar{U}_{t}$, then $U_{c,0}^{c,0} = U$ for admissible contract $(c, \bar{U}, k)$.

Notice that because retirement is contractible, the agent’s utility can in principle “jump” when the agent retires. However, if $U_{c,0}^{c,0}$ jumps down when the agent retires, for example, then while he doesn’t retire it must drift up to compensate the agent.

Faced with this contract, the agent might consider stealing and immediately consuming the proceeds, so that $h_{t} = 0$ always. In other words $a = (c + \phi k s, s)$ for some $s$. We can compare the utility from this hidden action $U^{a}$ with the utility from good behavior $a = (c,0)$:

$$U_{t}^{a} - U_{t}^{c,0} = \mathbb{E}_{t}^{s} \left[ \int_{t}^{\tau_{t}} e^{\rho(u-t)} \left( \frac{(c_{u} + \phi k_{u} s_{u})^{1-\gamma}}{1-\gamma} - \frac{c_{u}^{1-\gamma}}{1-\gamma} - \tilde{\sigma}_{U,u} s_{u} \right) du \right]$$

(3)

where we have used $W_{t}^{s} = W_{t} + \int_{0}^{t} \frac{s_{u}}{\beta} du$ to express $U_{c,0}^{c,0}$ as an expected integral under $P^{s}$, and also the fact that $U_{t}^{a} = U_{t}^{c,0} = \bar{U}_{t}$ because both plans have $h_{t} = 0$.\(^{16}\) Incentive compatibility will then require

$$0 \in \arg \max_{s \geq 0} \left( \frac{(c_{u} + \phi k_{u} s_{u})^{1-\gamma}}{1-\gamma} - \frac{c_{u}^{1-\gamma}}{1-\gamma} - \tilde{\sigma}_{U,u} s_{u} \right)$$

(4)

Taking FOC yields

$$\tilde{\sigma}_{U,t} \geq c_{t}^{-\gamma} \phi k_{t} \beta$$

(5)

which is positive. In words, we need to give the agent some “skin in game”. His continuation utility must depend on the observable outcomes, so he is not tempted to take a hidden action that makes

---

\(^{14}\) $N$ is the counting process associated to the Poisson arrival of retirement.

\(^{15}\) In this context, $\mathcal{L}^2$ is the set of $\mathbb{F}$-adapted processes such that $\mathbb{E} \left[ \int_{0}^{t} \tilde{\sigma}_{U,u}^{2} du \right] < \infty$ for any $t$.

\(^{16}\) We also use the fact that since $s$ is bounded, $\mathbb{E}^{s} \left[ \int_{0}^{t} \tilde{\sigma}_{U,u}^{2} du \right] < \infty$. 

9
bad outcomes more likely. This is costly because the principal is risk-neutral with respect to $W$ and $N$ so he would like to provide full insurance to the agent, but can’t because of the moral hazard problem.

Notice how the marginal utility of the hidden action depends on the marginal utility of consumption, so the principal would like to front load his consumption. By distorting the intertemporal consumption profile he can reduce the private benefit of the hidden action and therefore relax the risk-sharing problem. If the agent couldn’t have hidden savings, this condition would be enough to ensure the contract is incentive compatible, as in Di Tella (2014), but because of hidden savings it will not be enough. If the principal tries to front load the agent’s consumption, the agent will secretly save and consume when his marginal utility is higher, and this will in addition affect his incentives to steal.

The optimal contract must therefore respect the agent’s Euler equation: his discounted marginal utility $e^{-(\rho-r)t}c_t^{\gamma}$ must be a supermartingale. The economic intuition is that if the expected discounted marginal utility in the future is higher than the marginal utility today, the agent would like to save today in order to consume later. If instead the expected discounted marginal utility is lower than the marginal utility today, he would like to consume more today, but he can’t.17 We can then write

$$e^{-(\rho-r)t}c_t^{\gamma} = M_t - A_t,$$

where $A$ is an increasing process and $M$ a local martingale.18

We can then write

$$\frac{dc_t^{\gamma}}{c_t^{\gamma}} = (\rho - r)dt + \sigma_{M,t}dW_t - \lambda_{M,t}(dN_t - \theta dt) - \frac{e^{-(\rho-r)t}}{c_t^{\gamma}}dA_t$$

for some $\sigma_M$ and $\lambda_M$ processes.19 Using Ito’s lemma we get an expression for the law of motion of $c$:

$$\frac{dc_t}{c_t} = \left( \frac{r - \rho}{\gamma} + \frac{1}{2} \tilde{\sigma}_{c,t}^2 - \frac{\theta}{\gamma} \lambda_{M,t} \right) dt + \tilde{\sigma}_{c,t}dW_t + \left( (1 - \lambda_{M,t})^{\frac{1}{\gamma}} - 1 \right) dN_t + dL_t$$

(6)

where $L$ is an increasing process. Hidden savings restrict the ability of the principal to control the agent’s consumption. It is easy to make the agent consume with a high expected growth rate or even have his consumption “jump up” (because the agent can’t borrow against future consumption), but he can’t force him to consume with a low expected growth rate, since the agent would secretly save to achieve a better intertemporal smoothing.20 In fact, as it turns out in the optimal contract $e^{-(\rho-r)t}c_t^{\gamma}$ is a proper martingale, and the agent doesn’t want to borrow.

We summarize all the incentive compatibility conditions into the following lemma.

Lemma 2. If $C = (c, \bar{U}, k)$ is an incentive compatible contract, then (5) and (6) must hold for some $(\tilde{\sigma}_c, \lambda_M, L)$.

These are all necessary conditions for incentive compatibility. Lemma 8 in the Appendix provides

17If the agent could borrow and have hidden debt this would be a martingale, but under the constraint $h_t \geq 0$ we only need to make sure it’s a supermartingale.

18Since $e^{-(\rho-r)t}c_t^{\gamma}$ is a supermartingale then it is locally integrable and therefore has the desired representation.

19Note $\sigma_M$ and $\lambda_M$ are already normalized by $e^{-(\rho-r)t}c_t^{\gamma}$.

20If the agent could have hidden debt then $dL_t = 0$. 

10
a verification of incentive compatibility suitable for the optimal contract.

State variables

In the case without hidden savings, we would only need to work with the agent’s continuation utility $U$. However, in light of (6), we also need to keep $c$ as a state variable, and we need to figure out what is the domain of $(U, c)$ for which incentive compatible contracts exits. Because the drift in the agent’s consumption has a lower bound, it may very well be impossible to give an agent a very high $c$ combined with a very low $U$. To deal with this it will be easier to work with the following transformation of the state variables

$$X_t = \left((1 - \gamma)U_t\right)^{\frac{1}{1-\gamma}} > 0$$

$$\hat{c}_t = \frac{c_t}{X_t} \geq 0$$

While $X_t$ can take any positive value, $\hat{c}_t$ has an upper bound.

**Lemma 3.** For any incentive compatible contract $C$ it holds that

$$0 \leq \hat{c}_t \leq \hat{c}_h = \left(\frac{\rho - r(1 - \gamma)}{\gamma}\right)^{\frac{1}{1-\gamma}} > 0$$

In addition, $\hat{c}_h$ is an absorbing state for $\hat{c}_t$. The continuation contract after $\hat{c}_t = \hat{c}_h$ is the deterministic consumption path of retirement that delivers utility $U_t$, with cost $\hat{v}_h X_t$ and no risky investment, i.e. $c_{t+u} = c_t e^{\frac{r - \rho}{\gamma} u}$ and $k_{t+u} = 0$ for $u \geq 0$.

This upper bound has a close connection with the terminal contract. When the agent retires the principal delivers utility $\bar{U}_\tau$ in the cheapest way, with a deterministic consumption path given by $c_{\tau+s} = \hat{c}_h X_{\tau+s} > 0$ where $X_{\tau+s}$ evolves according to the law of motion (7) below. The continuation cost of this plan is $\hat{v}_h X_{\tau+s}$. Looking at the law of motion for $c_t$ in (6), we can use $\lambda_{M,t} > 0$ to make $c$ “jump up” on retirement. Conditional on $X_\tau$ we would like to make $c_\tau$ as high as possible. This has the benefit of both delivering utility to the agent in the cheapest way and also relaxing the constraint on the drift of $c_t$ before retirement (since it reduces the lower bound on the drift).

As a result, we can take without loss of generality that $\hat{c}_\tau = \hat{c}_h$.21

Using Ito’s lemma and $\hat{c}_\tau = \hat{c}_h$, we can obtain laws of motion for $X_t$ and $\hat{c}_t$. Using the normalization $\sigma_{U,t} = (1 - \gamma)U_t \sigma_{X,t}$ and $\lambda_t = (1 - \gamma)U_t \lambda_t$ we obtain22

$$\frac{dX_t}{X_t} = \left(\frac{\rho - c_t^{1-\gamma}}{1 - \gamma} + \frac{\gamma}{2} \sigma_{X,t}^2 + \theta \lambda_t\right) dt + \sigma_{X,t} dW_t + \left(1 - \hat{\lambda}_t(1 - \gamma)\right)^{\frac{1}{1-\gamma}} - 1 dN_t$$

21 Necessary conditions (5) and (6) for incentive compatibility don’t depend on the form of the contract after retirement, and we will verify at the end that the optimal contract is in fact incentive compatible.

22 Notice that $\hat{c}_\tau = \hat{c}_h$ implies $\lambda_M = 1 - \frac{\rho - \gamma}{\gamma} \left(1 - \hat{\lambda}_t(1 - \gamma)\right)^{\frac{1}{1-\gamma}}$, which is why $\lambda_M$ drops from the formulas.
and
\[
d\hat{c}_t \left( \frac{r - \rho}{\gamma} + \frac{1 + \gamma}{2} (\bar{\sigma}_{X,t} + \hat{\sigma}_{\hat{c},t})^2 - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} - \frac{\gamma}{2} \bar{\sigma}_{X,t}^2 - \bar{\sigma}_{X,t} \hat{\sigma}_{\hat{c},t} \right) \right) + dt + \hat{\sigma}_{\hat{c},t} dW_t + \left( \frac{\hat{c}_h}{\hat{c}_t} - 1 \right) dN_t + dL_t
\]
for some \( \hat{\sigma} \) and \( \hat{\lambda}(1 - \gamma) < 1 \). We call the \( dt \) multiplying terms \( \mu_X \) and \( \mu_\hat{c} \) respectively. Notice that using the normalization \( \hat{k} = \frac{k}{\bar{X}} \) we get from (5) a restriction of \( \bar{\sigma}_X \):
\[
\bar{\sigma}_{X,t} \geq \hat{c}_t^{-\gamma} \hat{k}_t \phi \beta
\]
or equivalently, \( \hat{k}_t = \frac{\bar{\sigma}_{X,t}\hat{c}_t^{\gamma}}{\phi \beta} \) and \( \bar{\sigma}_{X,t} \geq 0 \), since conditional on \( \bar{\sigma}_{X,t} \), giving the agent a larger investment is good for the principal because \( \alpha > 0 \).

The HJB equation

Because preferences are homothetic and the principal’s objective is linear, we know the principal’s value function takes the form \( v(X, \hat{c}) = \hat{v}(\hat{c}_t) \bar{X}_t \) with \( \hat{v}(\hat{c}_h) = \hat{v}_h > 0 \) because of Lemma 3. Notice that since we can always move \( \hat{c}_t \) up using \( dL_t \), we know that \( \hat{v}(\hat{c}) \) must be weakly increasing. In order to avoid unnecessary clutter, I will drop \( dL_t \) from what follows, and look for an HJB equation that characterizes the optimal contract while there are no jumps. As it turns out, this will in fact be the case in the optimal contract. I will also sometimes write \( \hat{v}_t = \hat{v}(\hat{c}_t) \), and \( \hat{v} \) instead of \( \hat{v}(\hat{c}) \).

The HJB equation associated with this problem is therefore
\[
r \hat{v} X = \min_{\bar{\sigma}_X, \bar{\sigma}_\hat{c}, \hat{\lambda}} \left( \hat{c} - \hat{k}_t \right) X + E_t^Q \left[ d(\hat{v}_t X_t) \right]
\]
subject to (7), (8), and (9), and \( \hat{\lambda}(1 - \gamma) < 1 \) and \( \bar{\sigma}_{X,t} \geq 0 \).

Using Ito’s lemma and canceling the \( X \) on both sides, we get
\[
r \hat{v} = \min_{\bar{\sigma}_X, \bar{\sigma}_\hat{c}, \hat{\lambda}} \hat{c} - \bar{\sigma}_X \hat{c}^{\gamma} \frac{\alpha}{\phi \beta}
\]
\[
\begin{align*}
+ \hat{v} \left( \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} + \frac{\gamma}{2} \bar{\sigma}_X^2 + \theta \hat{\lambda} + \frac{\hat{v}'}{\hat{v}} \hat{c} \left( \frac{r - \rho}{\gamma} + \frac{1 + \gamma}{2} (\bar{\sigma}_X + \hat{\sigma}_{\hat{c}})^2 - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} - \frac{\gamma}{2} \bar{\sigma}_X^2 - \bar{\sigma}_X \hat{\sigma}_{\hat{c}} \right) \right) \\
- \frac{\theta}{\gamma} \left( \gamma \hat{\lambda} + 1 - \frac{\hat{c}_h}{\hat{c}_t^{1-\gamma}} (1 - \hat{\lambda}(1 - \gamma)) \right)
\end{align*}
\]
\[
\begin{align*}
- \hat{v}'' \hat{c}^2 + \frac{\hat{v}'}{\hat{v}} \bar{\sigma}_X \hat{\sigma}_{\hat{c}} \hat{c} + \theta \left( (1 - \hat{\lambda}(1 - \gamma)) \frac{1}{1 - \gamma} \frac{\hat{v}_h}{\hat{v}} - 1 \right)
\end{align*}
\]
Notice how even though we have two state variables, \( \hat{c} \) and \( X \), the HJB equation boils down to a second order ODE. This is a feature of homothetic preferences and linear technology which makes the problem more tractable.
First order conditions. We can take FOC in (10). For $\bar{\sigma}_X$ we get (ignoring the inequality constraint)

$$\alpha = \gamma \left( \frac{\theta}{\bar{c}} - \gamma \phi \beta \bar{\sigma}_X \right) + \phi \beta \bar{v} \bar{c}^{1-\gamma} \left( (1 + \gamma) \bar{\sigma}_c + \bar{\sigma}_X \right)$$

(11)

If we embed the contract in a general equilibrium framework, we can interpret this FOC as an asset pricing equation for capital that pins down the equilibrium excess return $\alpha$. The condition says that the principal will give the agent a larger risky investment only if it pays a high excess return $\alpha$. In the first best without moral hazard we need $\alpha \leq 0$ for an optimal contract to exist (otherwise we have arbitrage). With moral hazard, however, there are two costs for the principal from giving a large investment to the agent. The first is that the principal will have to expose the agent to risk to provide incentives, and this is costly because the agent is risk averse. This effect would be present even without hidden savings. With hidden savings, however, the principal also knows that a larger exposure to risk induces the agent to postpone consumption for precautionary reasons, making it more expensive to deliver utility to him. The only way to have the agent consume $\hat{c}$ and manage a large amount of capital at the same time is by promising him a higher $\hat{c}$ in the future, with a correspondingly smaller amount of capital.

For this FOC to characterize a minimum (rather than a maximum) we need to check that $\hat{\gamma} + \hat{\gamma} \hat{c} > 0$, which will always be the case since $\hat{v} > 0$ and $\hat{\gamma} \geq 0$. The optimal $\bar{\sigma}_X$ conditional on $\bar{\sigma}_c$ then is

$$\bar{\sigma}_X = \frac{\alpha \hat{\gamma}}{\hat{\gamma}} - \hat{\gamma} \hat{c} (1 + \gamma) \bar{\sigma}_c$$

For $\hat{\lambda}$ we get

$$1 - \frac{\hat{\gamma} \hat{c}}{\hat{\gamma}} \left( 1 - \frac{\hat{c}_h}{\hat{c}} \left( 1 - \hat{\lambda}(1 - \gamma) \right) \right) - \left( 1 - \hat{\lambda}(1 - \gamma) \right) \frac{1}{1 + \hat{\gamma}} \frac{\hat{c}_h}{\hat{\gamma}} = 0$$

(12)

The left hand side is increasing in $\hat{\lambda}$ and is non-positive for $\hat{\lambda} = 0$, so this always has a solution $\hat{\lambda} \geq 0$ with $\hat{\lambda}(1 - \gamma) < 1$. This means that at retirement the agent’s utility drops. For $\bar{\sigma}_c$ let’s first plug in the optimal $\bar{\sigma}_X$ into the HJB equation and re-write it

$$0 = \min_{\bar{\sigma}_c, \hat{\lambda}} \bar{A} + \bar{B} \bar{\sigma}_c + \frac{1}{2} \bar{C} \bar{\sigma}_c^2$$

with

$$\bar{A} = \hat{c} - r \hat{v} - \frac{1}{2} \left( \frac{\alpha \hat{\gamma}}{\hat{\gamma}} \right)^2 + \hat{v} \left( \frac{\rho - \hat{c}^{1-\gamma}}{1 - \gamma} + \theta \left( (1 - \hat{\lambda}(1 - \gamma)) \frac{1}{1 + \gamma} \frac{\hat{c}_h}{\hat{\gamma}} - 1 + \hat{\lambda} \right) \right)$$

$$- \hat{\gamma} \hat{c} \left( \frac{\rho - r}{\gamma} + \frac{\rho - \hat{c}^{1-\gamma}}{1 - \gamma} - \frac{\theta}{\gamma} \left( \gamma \hat{\lambda} + 1 - \frac{\hat{c}_h}{\hat{c}} \left( 1 - \hat{\lambda}(1 - \gamma) \right) \frac{1}{1 + \gamma} \right) \right)$$

$$\bar{B} = \hat{\gamma} \hat{c} (1 + \gamma) \frac{\alpha \hat{\gamma}}{\hat{\gamma}} \frac{\hat{c}_h}{\hat{\gamma}} \geq 0$$
\[ \tilde{C} = \gamma \hat{v}' \hat{c}(1 + \gamma) \frac{\hat{v} - \hat{v}' \hat{c}}{\nu \gamma + \hat{v}' \hat{c}} + \hat{v}'' \hat{c}^2 \]

For this to have a minimum we need \( C \geq 0 \), and then the optimal \( \tilde{\sigma}_c \) is given by

\[ \tilde{\sigma}_c = -\frac{\tilde{B}}{C} \leq 0 \]

and this implies \( \tilde{\sigma}_X > 0 \) (so the inequality is satisfied). Notice how a shock \( dW_t \) moves \( X_t \) and \( \hat{c}_t \) in opposite directions. There is an economic intuition for this: the principal wants to give the agent some investment \( k_t > 0 \) because it has an excess return \( \alpha > 0 \), but this forces him to expose him to risk \( \tilde{\sigma}_X > 0 \). As a result, after a good shock his continuation utility goes up. Since the cost of delivering utility to the agent is increasing in \( \hat{c} \), he prefers to have \( \hat{c} \) move down when \( X \) goes up.

In other words, use the relatively cheaper contracts when you must deliver more utility. This is the same reason why \( \hat{\lambda} \geq 0 \) so the agent’s utility drops at retirement. It is more costly to deliver utility to the agent when he can’t manage assets, so it’s better to give him less utility when he retires, at the cost of giving him more utility while he is still able to manage assets.

**Domain of the HJB equation.** Although incentive compatible contracts can have \( \hat{c}_t \) ranging from 0 to \( \hat{c}_h \), we will look for an optimal contract with \( \hat{c}_t \geq \hat{c}_l \) for some \( \hat{c}_l \in (0, \hat{c}_h) \). For \( \hat{c}_l < \hat{c}_l \) we would make \( \hat{c}_t \) immediately “jump up” to \( \hat{c}_l \) using \( dL_t \) (but this will never happen in the optimal contract). As a result, the HJB equation will hold only on \([\hat{c}_l, \hat{c}_h]\), and \( \hat{v}(\hat{c}) = \hat{v}(\hat{c}_l) \equiv \hat{v}_l \) for \( \hat{c} < \hat{c}_l \).

For this construction to work we will need the HJB equation to hold as an inequality for \( \hat{c} < \hat{c}_l \) and \( \hat{v}'(\hat{c}_l) = 0 \) so that the two parts “smooth paste”. Since \( \tilde{B} = 0 \) when \( \hat{v}' = 0 \), we have \( \tilde{\sigma}_c = 0 \) for \( \hat{c} \leq \hat{c}_l \) and from (12)

\[ \hat{\lambda} = \hat{\lambda}_b(\hat{v}_l) \equiv 1 - \frac{1 - (\hat{v}_l / \hat{v}_h)^{1 - \gamma}}{1 - \gamma} \]

which is non-negative for all \( \hat{v}_l \leq \hat{v}_h \). As a result, we want \( A(\hat{c}; \hat{v}_l) \geq 0 \) on \( \hat{c} < \hat{c}_l \) where

\[ A(\hat{c}; \hat{v}_l) = \hat{c} - r \hat{v}_l - \frac{1}{2} \left( \frac{\partial^2 \hat{c}}{\partial \hat{v}_l^2} \right) + \hat{v}_l \left( \frac{\rho - \hat{c}^{1-\gamma}}{1 - \gamma} + \theta \gamma \hat{\lambda}_b(\hat{v}_l) \right) \]

Notice that by construction, since the HJB holds on \( \hat{c}_l \) with \( \hat{v}(\hat{c}_l) = \hat{v}_l \) and \( \hat{v}'(\hat{c}_l) = 0 \), we have \( A(\hat{c}_l, \hat{v}_l) = 0 \). So we just need to make sure that as \( \hat{c} \) falls \( A(\hat{c}; \hat{v}_l) \) remains non-negative. We can verify that as \( \hat{c} \to 0 \), \( A(\hat{c}; \hat{v}_l) \) becomes positive for any \( \hat{v}_l \in (0, \hat{v}_h) \), so this condition is satisfied if and only if \( \hat{c}_l \) is the smallest root of \( A(\hat{c}; \hat{v}_l) \).

\[ ^{23} \text{Because } A(\hat{c}; \hat{v}_l) \geq 0 \text{ for } \hat{c} \text{ below the smallest root, we know } A'_c(\hat{c}_l; \hat{v}_l) \leq 0. \text{ We don’t need to worry about } A'_c(\hat{c}_l; \hat{v}_l) = 0 \text{ because in that case there is a unique root. This will be formalized later.} \]

---

14
Stationary Contracts

Before characterizing the optimal contract, it is useful to study a set of admissible but not optimal contracts that set $\mu_c = \tilde{\sigma}_c = 0$ and

$$\hat{\lambda}(\hat{c}) = \hat{\lambda}_r(\hat{c}) = 1 - \frac{1}{1 - \gamma} \geq 0$$

which satisfies $\hat{\lambda}_r(1 - \gamma) < 1$. In order to set $\mu_c = 0$ we set $\tilde{\sigma}_X$ so that

$$\frac{1}{2}\tilde{\sigma}_X^2 = \frac{\rho - r}{\gamma} + \frac{\rho - \hat{c}^{1-\gamma} + \theta \hat{\lambda}_r(\hat{c})}{1 - \gamma}$$

where, using $\hat{c}_h^{1-\gamma} = \frac{\rho - r(1 - \gamma)}{\gamma} > 0$, we see the right hand side is non-negative for $\hat{c} \leq \hat{c}_h$. We can also rewrite it using the definition of $\hat{\lambda}_r$

$$\frac{1}{2}\tilde{\sigma}_X^2 = \frac{\rho - r(1 - \gamma) + \theta \gamma \left(1 - \left(\frac{\hat{c}}{\hat{c}_h}\right)^{1-\gamma}\right)}{(1 - \gamma)\gamma}$$

This equation says that the only way to give the agent a permanently high capital relative to his continuation utility, $\hat{k}$, is to accept that he will postpone his consumption (low $\hat{c}$) for precautionary motives. We can then solve for the value of the stationary contract for a given $\hat{c}$ as follows

$$\hat{v}_r(\hat{c}) = \frac{\hat{c} - \frac{\alpha}{\rho \hat{c}^{1-\gamma}} \tilde{\sigma}^{\gamma} \sqrt{\frac{\rho - r(1 - \gamma) + \theta \gamma}{1 - r \gamma}} \left(1 - \left(\frac{\hat{c}}{\hat{c}_h}\right)^{1-\gamma}\right) + \frac{\theta \gamma \hat{c}}{\rho - r(1 - \gamma)}}{2r - \rho - \frac{1}{1 - \gamma} \rho + \frac{\hat{c}^{1-\gamma}}{1 - \gamma} (1 + \gamma) + \theta \left(1 - \hat{\lambda}_r(\hat{c})(1 + \gamma)\right)}$$

(15)

Notice that $\hat{v}_r(\hat{c}_h) = \hat{v}_h = \hat{c}_h^{\gamma}$, which makes sense because $\hat{v}_h$ corresponds to a stationary contract with $\tilde{\sigma}_X = 0$. As we move back from $\hat{c}_h$, however, the denominator can become negative. This means $X_t$ is growing faster than $r + \theta$, and therefore the contract is not actually admissible (the expected discounted value of the termination contract $\hat{v}_h X_r$ diverges to infinity). Since the denominator is increasing in $\hat{c}$ (notice $\hat{\lambda}_r$ is decreasing in $\hat{c}$) $\hat{v}_r$ is the cost of an admissible stationary contract for $\hat{c} \geq \hat{c}_s$ (for $\hat{c} < \hat{c}_s$ there are no admissible stationary contracts)

$$\hat{c}_s = \left(\frac{2\gamma}{1 + \gamma} \frac{\rho - r(1 - \gamma)}{\gamma}\right)^{1-\gamma} \in (0, \hat{c}_h)$$

For $\hat{c} \geq \hat{c}_s$ the constant drift $\mu_X < \theta + r$ and the volatility $\sigma_X$ and $\hat{\lambda}_r \geq 0$ is constant as well. Therefore in light of Lemma 7 in the Appendix the stationary contract is admissible, and because $\tilde{\sigma}_c = 0$ by construction, Lemma 8 ensures that it is incentive compatible. These stationary contracts are not necessarily optimal, so $\hat{v}(\hat{c}) \leq \hat{v}_r(\hat{c})$. A necessary condition for the existence of an optimal contract therefore is that $\hat{v}_r(\hat{c}) > 0$ for all $\hat{c} \geq \hat{c}_s$. Otherwise it is possible to deliver utility to the
agent and obtain unbounded profits. This implies an upper bound on the excess returns of the risky investment $\alpha$.

**Lemma 4.** If

$$\alpha \geq \bar{\alpha} \equiv \frac{\phi \beta \gamma \sqrt{2}}{\sqrt{1 + \gamma}} \sqrt{\frac{\rho - r(1 - \gamma) + \theta \gamma}{\gamma}} > 0$$

then $\hat{v}_r(\hat{c}) \leq 0$ for some $\hat{c} \geq \hat{c}_*$. 

In light of this result, we need to assume $\alpha < \bar{\alpha}$ to have the problem well defined.\textsuperscript{24}

**Verification theorem**

As explained above, although admissible contracts can have $\hat{c}_t$ ranging from 0 to $\hat{c}_h$, we will look for an optimal contract with $\hat{c}_t \geq \hat{c}_i$ for some $\hat{c}_i \in (0, \hat{c}_h)$. The HJB equation will hold only on $[\hat{c}_i, \hat{c}_h)$, and $\hat{v}(\hat{c}) = \hat{v}(\hat{c}_i) \equiv \hat{v}_l$ for $\hat{c} < \hat{c}_i$. For this construction to work we will need the HJB equation to hold as an inequality for $\hat{c} < \hat{c}_i$ (i.e. $A(\hat{c}; \hat{v}_l) \geq 0$ for all $\hat{c} \leq \hat{c}_i$) and $\hat{v}'(\hat{c}_i) = 0$ so that the two parts “smooth paste”.

From a solution to the HJB equation we can build an optimal contract using the policy functions for $\tilde{\sigma}_X$, $\tilde{\sigma}_\hat{c}$ and $\hat{\lambda}$. Since $\hat{v}(\hat{c})$ must be increasing, we will start the contract at $\hat{c}_i$. For a given solution to the HJB equation, let $C^* = (c^*, \tilde{U}^*, k^*)$ be the contract generated by the policy functions from the HJB, with associated state variables $X^*$ and $\hat{c}^*$, with $X^*_0 = ((1 - \gamma)u_0)^{1/\gamma}$ and $\hat{c}^*_0 = \hat{c}_i$.

Below we will provide a formal verification theorem. Let us summarize the assumptions we need to make the problem well defined.

**Assumption 1.** The following inequalities must hold for an optimal contract to exist:

$$\hat{c}_h = \left(\frac{\rho - r(1 - \gamma)}{\gamma}\right)^{\frac{1}{1 - \gamma}} > 0$$

$$\alpha < \bar{\alpha} \equiv \frac{\phi \beta \gamma \sqrt{2}}{\sqrt{1 + \gamma}} \sqrt{\frac{\rho - r(1 - \gamma) + \theta \gamma}{\gamma}} > 0$$

We are now ready to provide a formal verification theorem.

**Theorem 1** (Verification Theorem). Let $\hat{v}(\hat{c}) : [\hat{c}_i, \hat{c}_h) \to (0, \hat{v}_h]$ be an increasing $C^2$ solution to the HJB equation (10) with $\hat{v}(\hat{c}_h) = \hat{v}_h$ and $\hat{v}'(\hat{c}_i) = 0$, for some $\hat{c}_i \in [0, \hat{c}_h]$, such that $A(\hat{c}; \hat{v}_l) \geq 0$ for all $\hat{c} \leq \hat{c}_i$, where $\hat{v}_l \equiv \hat{v}(\hat{c}_i) > 0$. Then,

1) for any incentive compatible contract $C = (c, \tilde{U}, k)$ that delivers at least utility $u_0$ to the agent, we have $\hat{v}(\hat{c}_i) \left((1 - \gamma)u_0\right)^{\frac{1}{1 - \gamma}} \leq J_0(C)$.

\textsuperscript{24}This is a necessary, but not sufficient condition. It is analogous to upper-bounds on excess returns for portfolio problems to have a solution.
2) If in addition \( C^* \) is an admissible contract with bounded \( \hat{\sigma}_t^* \), then \( C^* \) is an optimal contract, with cost \( J_0(C^*) = \hat{v}(\hat{c}) \left( (1 - \gamma)u_0 \right)^{\frac{1}{1 - \gamma}} \). Furthermore, \( \hat{c}_t^* \in [\hat{c}_l, \hat{c}_h] \) always and \( \hat{\sigma}_X^* \geq 0, \hat{\sigma}_X^* \leq 0, \mu_X^* \), and \( \hat{\lambda}^* \geq 0 \) are all bounded.

Note 2. The condition that \( \hat{\sigma}_t^* \) be bounded can be relaxed. It is used to verify that some local martingales are true martingales.

The HJB equation can be solved as an ODE by plugging in the FOCs and checking at the end that \( \bar{C} \geq 0 \) everywhere (to ensure we have in fact a minimum). Verifying that \( A(\hat{c}; \hat{v}_t) \geq 0 \) for all \( \hat{c} \leq \hat{c}_l \) is then straightforward, but as it turns out if \( \gamma \geq \frac{1}{2} \) the inequality is guaranteed so we can skip this step. If \( \gamma < \frac{1}{2} \) we can’t skip the step, but we can replace it by a local condition

**Lemma 5.** In the setting of Theorem 1, if \( \gamma \geq \frac{1}{2} \) then we don’t need to check that \( A(\hat{c}; \hat{v}_t) \geq 0 \) for all \( \hat{c} \leq \hat{c}_l \) since this condition is always satisfied. If \( \gamma < \frac{1}{2} \) then it is enough to check that

\[
A'_t(\hat{c}_t, \hat{v}_t) \equiv 1 - \hat{v}_t \left( \hat{c}_t^{\gamma} + \hat{c}_t^{2\gamma-1} \left( \frac{\alpha}{\phi \beta} \right)^2 \frac{1}{\hat{v}_t^2} \right) \leq 0
\]

We still need to verify that \( \hat{\sigma}_t^* \) is bounded and the contract generated by the HJB \( C^* \) is indeed admissible. The following sufficiency result is useful here.

**Lemma 6.** Let \( C^* \) be as in part 2) of Theorem 1, except we don’t know if it is admissible. If \( \mu_X^* \) is strictly bounded above by \( \theta + r \), then \( C^* \) is admissible.

Note 3. The upper bound on \( \mu_X^* \) ensures that the integrability constraint (1) holds, but it is not necessary for a contract to be admissible. In practice, however, this sufficiency theorem is often enough.

**Verifying incentive compatibility.** An important step in the proof of verification Theorem 1 is making sure that the candidate optimal contract \( C^* \) is incentive compatible. While (5) and (6) will ensure that neither stealing and immediately consuming, nor secretly saving without stealing are attractive on their own, they leave open the possibility that a double deviation (stealing and saving the proceeds for later) could be attractive to the agent. To see how this can happen, consider that while stealing reduces the continuation utility of the agent under good behavior, it might increase the expected marginal utility of consumption \( E_t^s \left[ e^{(\rho - \rho)u \phi^{-\gamma}} \right] \) to such an extent that saving the stolen funds for consumption later could be very attractive.\(^{25}\)

We show that the candidate optimal contract is indeed incentive compatible by deriving an upper bound for the utility an agent can get after any history which has left him with \( h_t \) hidden savings, for any valid hidden action \( a = (\hat{c}, s) \):

\[
U_t^a \leq \left( 1 + \frac{h_t \hat{c}_t^{\gamma}}{X_t} \right)^{1-\gamma} U_t^{c,0}
\]

\(^{25}\)In other words, even if \( e^{(\rho - \rho)u \phi^{-\gamma}} \) is a martingale under \( P \), it might be a submartingale under \( P^* \).
Notice how hidden savings can allow the agent to achieve more utility than he should had he always behaved (however remember this is just an upper bound on achievable utility). In particular, if the agent has somehow accumulated hidden savings in the past, he will want to deviate from good behavior \((c, 0)\) in the future, at the very least to increase his consumption. However, if he doesn’t have any hidden savings \(h_t = 0\) (regardless of whether he stole in the past), the most utility he could get is \(U^{c,0}_t\).

There are two important insights here. The first is that while the marginal value of the first dollar in hidden savings is equal to the marginal value of legitimate consumption, the marginal value of the last dollar in hidden savings is lower. We can see this clearly if we consider a particular deviation: steal for a while and save all proceeds until retirement. This double deviation has the nice feature that at retirement the upper bound on the value function of the agent is actually binding, equal to \(B(U, h)\), so the marginal value of a dollar in hidden savings is

\[
\left(1 + \frac{h_\tau}{X_\tau} \hat{c}_\tau^{-\gamma}\right)^{-\gamma} c_\tau^{-\gamma}
\]

If \(h_\tau = 0\), the marginal utility is \(c_\tau^{-\gamma}\) as we would expect. The first dollar stolen at \(t < \tau\) and saved until retirement then has an expected marginal value \(E_t \left[ e^{(r-\rho)(\tau-t)} c_\tau^{-\gamma}\right]\). Under stealing, since the agent is punished and his consumption is expected to go down, his marginal utility at retirement is expected to be higher than at \(t\), i.e. \(E_t \left[ e^{(r-\rho)(\tau-t)} c_\tau^{-\gamma}\right] > c_\tau^{-\gamma}\), so it seems that stealing and consuming at retirement is an attractive double deviation for the agent. However, to the extent that the agent is stealing and changing the probability distribution of \(c_\tau\), he is also accumulating hidden savings \(h_\tau\). The marginal utility of the last dollar is lower than the marginal utility of the first dollar in hidden savings, so that \(E_t \left[ \left(1 + h_\tau X_\tau^{-1} \hat{c}_\tau^{-\gamma}\right)^{-\gamma} e^{(r-\rho)(\tau-t)} c_\tau^{-\gamma}\right] \leq c_\tau^{-\gamma}\).

This is a result of the assumption that the hidden action consists of diverting funds, as opposed to shirking as in Kocherlakota (2004). However, it is not enough to ensure incentive compatibility. If the agent waits until retirement to consume his hidden savings, he knows he will be facing from then on a deterministic consumption path (and therefore we know \(\hat{c}_\tau = \hat{c}_h\)). Before retirement, however, his future consumption is volatile. The second insight is that from the point of view of his hidden savings problem, the legitimate consumption is his income, and since he can’t secretly insure against this risk, the agent faces an incomplete markets problem. If after some history his future consumption is very volatile, the agent will place a very large marginal value on hidden savings that can help him self-insure (even if it is still true that the marginal value of the last dollar in hidden savings is lower than the first). This will be captured by a low \(\hat{c}_t\): to the extent that the agent faces a large volatility, he will have a large precautionary motive to postpone consumption which will result in a back loaded consumption path. We can see in (18) that the marginal value of hidden savings in the upper bound is larger when \(\hat{c}_t\) is small.\(^{26}\)

As a result of this, if after bad outcomes the contract increased the volatility of the agent’s consumption, then stealing could be attractive since the agent would expect to have hidden dollars

\(^{26}\)While (17) is only an upper bound, it captures the intuition.
precisely when they are most valuable. An important condition to prevent this from happening is therefore that in the optimal contract \( \tilde{\sigma}_c \leq 0 \). This means that after bad outcomes the contract exposes the agent to less volatility and he is consequently willing to consume more up front. If the agent steals he expects to have hidden dollars when they are less valuable to him.

Fortunately the optimal contract naturally delivers \( \tilde{\sigma}_c \leq 0 \) for intertemporal hedging motives. The intuition for this is that contracts that give a large amount of capital to the agent and therefore expose him to high volatility are cheaper for the principal (in the relevant range), because capital pays an excess return. From the FOC in equation (13) we see that the principal prefers to use contracts with higher volatility (and therefore more back loaded consumption) when he must deliver more utility to the principal, i.e. after good outcomes. This intertemporal hedging motive is induced by the presence of hidden savings (the optimal contract without hidden savings has a constant growth rate, volatility, and intertemporal consumption profile), but follows a completely different logic from the incentive compatibility of double deviations. It is derived using only the first order approach. It is a remarkable feature of his environment that the presence of hidden savings induces an intertemporal hedging motive that eliminates the double deviations made possible by hidden savings.

**Optimal contract**

We can now solve for the optimal contract. We use the following parameter values: \( \rho = r = 0.05, \gamma = 3, \phi = 0.2, \beta = 1, \alpha = 0.03, \) and let \( \theta \to 0 \). We use the FOC to plug into the HJB and solve as an ODE numerically. Appendix B describes the numerical algorithm in detail. Figure 1 shows the value function \( \hat{v}(\hat{c}) \) and the capital \( \hat{k}(\hat{c}) \). Within the relevant range \( [\hat{c}_l, \hat{c}_h] \), if the principal is willing to accept a more backloaded consumption (low \( \hat{c} \)), he can have the agent manage a larger amount of capital (high \( \hat{k} \)), which yields a low cost \( \hat{v} \) of delivering utility to the agent. Distorting the agent’s consumption smoothing beyond \( \hat{c}_l \) is not cost effective, so the contract will never visit that region of the state space. The red dotted line represents the cost of stationary contracts, \( \hat{v}_r(\hat{c}) \).

Since these contracts are incentive compatible, but not optimal, we have \( \hat{v}_r(\hat{c}) \geq \hat{v}(\hat{c}) \), and the two curves touch at \( \hat{c}_h \), where the only incentive compatible contract is the the stationary “retirement” contract.

First let’s make sure we have the optimal contract. The law of motion for the state variables are given by Figure 2. We see that \( \tilde{\sigma}_c \leq 0 \) at all times, which means \( \tilde{C} \geq 0 \) as required for minimization (we know \( \tilde{B} \geq 0 \) because \( \hat{v}(\hat{c}) \) is increasing). In addition, \( \tilde{\sigma}_c \), and therefore the drift and volatility of \( X \), are bounded (which implies \( \hat{k} \) is bounded as well), and \( \mu_{X,t} < \bar{\mu} < r \), so using Lemma 6, the contract is admissible. We can then use Theorem 1 and Lemma 5 to conclude we have the optimal contract.

In the region with low \( \hat{c} \) the agent receives a relatively large amount of capital \( \hat{k} \), so he is exposed to a large amount of risk in his continuation utility, \( \tilde{\sigma}_X \). With high exposure to risk and back loaded consumption, the growth rate of his continuation utility, \( \mu_X \), is high. This is a “high growth/high volatility” region, while the region with high \( \hat{c} \) can be characterize as a “low growth/low volatility”
Figure 1: The value function $\hat{v}(\hat{c})$ solid in blue, and capital $\hat{k}(\hat{c})$ solid in yellow (right axis). Also, the cost of stationary contracts $\hat{v}_r(\hat{c})$ is dotted in red, and capital in the stationary contract $\hat{k}_r(\hat{c})$ is dotted in light blue (right axis). The dashed lines indicates the domain of the optimal contract $[\hat{c}_l, \hat{c}_h]$, and the cost of the retirement contract $\hat{v}_h$ is dotted in black. The black dot indicates the cost of the best stationary contract.

Figure 2: The drift, $\mu_\hat{c}$ and $\mu_X$, and volatility, $\sigma_\hat{c}$ and $\sigma_X$, of the state variables $\hat{c}$ and $X$. 
Turning to the contract dynamics, because of the moral hazard problem, after good outcomes the agent’s continuation utility $X_t$ goes up and so does his consumption $c_t$ and capital $k_t$, i.e. $\hat{\sigma}_X \geq 0$. This could be achieved with a proportional increase in both $c_t$ and $k_t$, as in the stationary contracts. In fact, this would be optimal if the agent didn’t have access to hidden savings, as in the contractual environment in Di Tella (2014). Instead, after good outcomes the optimal contract shifts towards the “high growth/high volatility” region, i.e. $\hat{\sigma}_c \leq 0$. We can understand this in terms of intertemporal hedging. Because “high growth/high volatility” contracts are a cheaper way of delivering utility to the agent, the principal prefers to use them after good outcomes, when he must deliver more continuation utility to the agent.

Now we can study the contract’s long run behavior. The contract has a scale invariance property for $X$. The drift of $\hat{c}$, on the other hand, takes the contract to a “steady state” $\hat{c}^{ss} \in (\hat{c}_l, \hat{c}_h)$. However, the volatility $\hat{\sigma}_\hat{c}$ is relatively high in that area, so the contract actually spends very little time near the “steady state”. Figure 3 shows the stationary distribution of $\hat{c}$ and $\hat{k}$. The contract spends most of the time in the “low growth/low volatility” region, with consumption that is close to intertemporally optimal but very little capital. While the drift $\mu_{\hat{c}} < 0$ pushes $\hat{c}$ down towards the “steady state”, it is very weak, and so is the volatility $\hat{\sigma}_c$. As a result, the contract can become “trapped” in this high cost region for very long times. When it finally approaches the “steady state”, the volatility is higher, and the contract quickly moves away and falls back into the “low growth/low volatility” region. At the other extreme, when $\hat{c}$ is close to the lower bound $\hat{c}_l$, the drift $\mu_{\hat{c}}$ is large and pushes the system back towards the steady state. The reason for this is that in this region the agent manages a large amount of capital and has a large exposure to risk. As a result, his consumption is very back loaded, so $\hat{c}$ is expected to grow fast. In other words, the only way for the principal to give the agent such a large amount of capital without the agent postponing consumption even more is by promising him less capital and risk in the future. For this reason, while the contract starts at $\hat{c}_l$ in the “high growth/high volatility” region, it quickly moves away and spends most of the time in the “low growth/low volatility” region.

**Stationary contracts.** It is useful to compare the optimal contract to the stationary contracts. In particular, we could chose the stationary contract that minimizes the cost, indexed by $\hat{c}^{\text{min}}_c > \hat{c}_l$. The black dot on Figure 1 indicates the cost of the optimal stationary contract. Since stationary contracts give the agent a constant capital relative to his continuation utility, we can see the tradeoff clearly. If the principal wanted to give the agent more capital, he would have to expose the agent to more risk and therefore accept an even more back loaded consumption (see Figure 1). While this is a good tradeoff for the first units of capital, at some point the distortion in intertemporal smoothing becomes too large, so the cost of stationary contracts goes up below $\hat{c}^{\text{min}}_c$.

However, what the principal can do to further reduce the cost is to move away from stationary contracts. Give the agent more capital today, and in order to avoid too large a distortion in intertemporal smoothing, promise the agent less capital in the future in proportion to his continuation
utility, i.e. \( \hat{c}_t \) has a positive drift. Since this improves the tradeoff between capital and intertemporal consumption smoothing, the principal is now willing to accept a lower \( \hat{c} < \hat{c}_{t \min} \) in exchange for more capital \( \hat{k} \). This yields the optimal contract, and Figure 1 shows the capital under the optimal contract is larger than under the stationary contracts. It is remarkable that this reduces the cost of the contract even though it creates long-run dynamics under which the contract spends most of the time in regions where the cost of delivering utility is larger than under the optimal stationary contract, i.e. \( \hat{v}(\hat{c}_t) > \hat{v}_r(\hat{c}_{t \min}) \) most of the time.

**Implementation**

TO BE COMPLETED

### 3 Conclusions

This paper develops a tractable contractual framework for delegated asset management that is well suited to macro and asset pricing applications. In particular, hidden savings are an important realistic feature, but can be technically challenging to deal with. We show that the introduction of hidden savings creates a hedging motive and rich dynamics in an otherwise stationary environment, and that this hedging motive ensures incentive compatibility of the optimal contract. The optimal contract has a scale invariance property and can be characterized by an ODE, which makes it easy to embed in general equilibrium models.
References


Appendix A - Omitted Proofs

Lemma 1

Consider
\[ Y_t = \mathbb{E}_t \left[ \int_0^\tau e^{-\rho u} \frac{c_{u}^{1-\gamma}}{1-\gamma} du + e^{-\rho \tau} \bar{U}_\tau \right] = \int_0^t e^{-\rho u} \frac{c_{u}^{1-\gamma}}{1-\gamma} du + e^{-\rho \tau} \bar{U}_\tau \]
on on \( \{ t \leq \tau \} \). Since \( Y \) is an \( \mathbb{F} \)-adapted \( \mathbb{P} \)-martingale, and \( \mathbb{F} \) is generated by \( W \) and \( \mathbb{N} \), we can apply a martingale representation theorem to obtain
\[ dY_t = e^{-\rho (u - t)} \frac{c_{u}^{1-\gamma}}{1-\gamma} dt + e^{-\rho U_t^{c,0}} \sigma_{U,t} dW_t - e^{-\rho \lambda_t (dN_t - \theta dt)} \]
and dividing by \( e^{-\rho t} \) and rearranging we get (2). Since \( \lambda_t = U^{c,0}_t - \bar{U}_t \), and from the square integrability of \( U^{c,0} \) we get \( \sigma_{U} \in \mathcal{L}^2 \).

In the other direction, if we have a solution to (2) with the desired properties, we can integrate to obtain
\[ U_t = \mathbb{E}_t \left[ \int_t^\tau e^{-\rho (u-t)} \frac{c_{u}^{1-\gamma}}{1-\gamma} du + e^{-\rho (\tau-t)} \bar{U}_\tau \right] \]
which is square integrable.

Lemma 2

First, in light of (3), take the hidden action \( a = (c + \phi ks, s) \) where \( s \) is zero where (4) is satisfied and \( s = s^*_t \wedge \bar{s} \) where \( s^*_t \) achieves the maximum in (4) and \( \bar{s} > 0 \) is an arbitrary bound. Because the objective in (4) is concave and it’s zero for \( s = 0 \), then for \( s^*_t \wedge \bar{s} \) it is strictly positive when (4) fails. The resulting action \( a \) is valid since \( h_t = 0 \) by construction. If (4) fails on a positive measure \( (\omega, t) \) set, then from (3) we get \( U^a_t > U^{c,0}_t \) and \( C \) is not incentive compatible. (5) is simply the FOC necessary condition for (4) (and sufficient because of concavity).

For (6) this is the result of \( e^{-(\rho-r)t} c_{t}^{-\gamma} \) being a supermartingale, which is a standard necessary condition in a savings/consumption problem. Note this doesn’t involve stealing \( s \), just hidden consumption \( c \).

Lemma 3

For the bound, since both \( c_t \geq 0 \) and \( X_t \geq 0 \), we only need to show that \( \hat{c}_t \leq \hat{c}_h \). Marginal utility of consumption is \( m_t = c_t^{-\gamma} \) and the utility flow \( \frac{c_{u}^{1-\gamma}}{1-\gamma} = \frac{1}{1-\gamma} m_t^{\gamma-1} \). This is a convex and decreasing function of \( m_t \). From Lemma 2, we have by (6) that \( \mathbb{E}_t [m_{t+s}] \leq e^{(\rho-r)s} m_t \). Given any \( m_0 \) we have

\footnote{Notice that for \( t \geq \tau \), \( Y_t \) is constant, so \( \sigma_{U,t} \) and \( \lambda_t \) are zero, but this is not relevant for our purposes.}
by Jensen’s inequality

\[ E \left[ c_{t}^{1-\gamma} \right] \geq \frac{1}{1-\gamma} E [m_{t}^{\frac{1-\gamma}{\gamma}}] \geq \frac{1}{1-\gamma} m_{0}^{\frac{1-\gamma}{\gamma}} e^{(\rho-r)\frac{2-1}{\gamma} t} = \frac{c_{0}^{1-\gamma}}{1-\gamma} e^{(\rho-r)\frac{2-1}{\gamma} t} \]

with equality only if \( c \) is deterministic. So

\[ U_{t}^{c,0} = E_{t} \left[ \int_{t}^{\tau} e^{-\rho(s-t)} \frac{c_{s}^{1-\gamma}}{1-\gamma} ds + e^{-\rho(\tau-t)} \overline{U}_{\tau} \right] = E_{t} \left[ \int_{t}^{\infty} e^{-\rho(s-t)} \frac{c_{s}^{1-\gamma}}{1-\gamma} ds \right] \geq \int_{t}^{\infty} e^{-\rho s} \frac{c_{t}^{1-\gamma}}{1-\gamma} e^{(\rho-r)\frac{2-1}{\gamma} s} ds = \frac{c_{t}^{1-\gamma}}{1-\gamma} \frac{\gamma}{\rho - r(1-\gamma)} \]

where the second equality uses the fact that the termination contract must deliver \( \overline{U} \) to the agent. This bound implies

\[ X_{t} = \left( (1-\gamma)U_{t}^{c,0} \right)^{\frac{1}{1-\gamma}} \geq c_{t} \left( \frac{\gamma}{\rho - r(1-\gamma)} \right)^{\frac{1}{1-\gamma}} \]

So \( \hat{c}_{t} = c_{t}/X_{t} \) has an upper bound

\[ \hat{c}_{t} \leq \left( \frac{\rho - r(1-\gamma)}{\gamma} \right)^{\frac{1}{1-\gamma}} = \hat{c}_{h} \]

In addition, since the upper bound can only be achieved with a deterministic consumption, \( U_{t}^{c,0} \) is deterministic too. This implies that both \( c_{t} \) and \( X_{t} \) grow at rate \( \frac{\rho - r}{\gamma} \), so \( \hat{c}_{h} \) is an absorbing state. In light of (5), we must have \( k_{t+u} = 0 \) in the continuation contract, so we have the retirement contract with cost \( \hat{v}_{h}X_{t} \). This completes the proof.

**Lemma 4**

We need to check the numerator in (15), since the denominator is positive for all \( \hat{c} \geq \hat{c}_{s} \):

\[ \hat{c} - \frac{\alpha}{\phi \beta} \hat{c}^{2} \sqrt{2} \left( \frac{\rho - r(1-\gamma) + \theta \gamma}{(1-\gamma) \gamma} \right) \left( \frac{1}{\hat{c}_{h}} \right)^{\frac{1-\gamma}{\gamma}} + \frac{\hat{c} \theta \gamma}{\rho - r(1-\gamma)} \]

The rest of the proof consists of evaluating this expression at \( \hat{c} = \hat{c}_{s} \) and showing it is non-positive if the bound is violated. We get \( \hat{c} \) times

\[ 1 - \frac{\alpha}{\phi \beta} \sqrt{2} \frac{\rho - r(1-\gamma) + \theta \gamma}{\gamma(1-\gamma)^{\frac{1-\gamma}{\gamma}}} + \frac{\theta \gamma}{\rho - r(1-\gamma)} \]

\[ \frac{\rho - r(1-\gamma)}{\gamma} + \theta \frac{\alpha}{\phi \beta} \sqrt{2} \frac{1}{\gamma^{\frac{1-\gamma}{\gamma}}} \sqrt{\frac{\rho - r(1-\gamma) + \theta \gamma}{\gamma(1-\gamma)^{\frac{1-\gamma}{\gamma}}} \frac{\rho - r(1-\gamma)}{\gamma}} \]
\[
\rho - r(1 - \gamma) + \theta \gamma \leq \frac{\alpha}{\phi \beta} \frac{1}{\sqrt{2 - \frac{\gamma}{\sqrt{1 + \gamma}}}} \sqrt{\frac{\rho - r(1 - \gamma) + \theta \gamma}{\gamma}}
\]

So if \( \alpha \geq \bar{\alpha} \) the numerator is non-positive. This completes the proof.

**Theorem 1**

First let’s extend the function \( \hat{v}(\hat{c}) \) as described above, with \( \hat{v}(\hat{c}) = \hat{v}_t \equiv \hat{v}(\hat{c}_t) \) for all \( \hat{c} < \hat{c}_t \) (we always have \( \hat{c} \in [0, \hat{c}_h] \)). The HJB holds as an equality for \( \hat{c} \geq \hat{c}_t \), and because \( A(\hat{c}; \hat{v}_t) \geq 0 \) for \( \hat{c} \leq \hat{c}_t \), the HJB holds as an inequality there. Then consider any incentive compatible contract \( \mathcal{C} = (c, \bar{U}, k) \) that delivers utility \( u_0 \) to the agent, with associated state variables \( X \) and \( \hat{c} \). We only need to consider contracts with \( \hat{c}_r = \hat{c}_h \). Because \( \hat{v}'(\hat{c}_t) = 0 \) we can use Ito’s lemma\(^{28}\) and the HJB equation to obtain

\[
e^{-r(\tau^* \wedge t)} \hat{v}(\hat{c}_r \wedge \tau) X_{\tau^* \wedge t} \geq \hat{v}(\hat{c}_0) X_0 - \int_0^{\tau^* \wedge t} e^{-rt} \left( \hat{c}_t - \hat{k}_t \alpha \right) X_t dt
\]

\[
+ \int_0^{\tau^* \wedge t} e^{-rt} \hat{v}(\hat{c}_t) X_t \left( \frac{\hat{v}'(\hat{c}_t)}{\hat{v}(\hat{c}_t)} \hat{c}_t \hat{\sigma}_{\tau^* \wedge t} + \hat{\sigma}_{X_t} \right) dW_t
\]

\[
+ \int_0^{\tau^* \wedge t} e^{-rt} \hat{v}(\hat{c}_t) X_t \left( \frac{\hat{v}_h}{\hat{v}(\hat{c}_t)} \left( 1 - \hat{\lambda}(1 - \gamma) \right)^{\frac{1}{\gamma}} - 1 \right) (dN_t - \theta dt)
\]

for the localizing sequence of stopping times \( \{\tau^n\}_{n \in \mathbb{N}} \):

\[
\tau^n = \inf \left\{ T \geq 0 : \int_0^T \left| e^{-rt} \hat{v}(\hat{c}_t) X_t \left( \frac{\hat{v}'(\hat{c}_t)}{\hat{v}(\hat{c}_t)} \hat{c}_t \hat{\sigma}_{\tau^* \wedge t} + \hat{\sigma}_{X_t} \right) \right| dt + \int_0^T \left| e^{-rt} \hat{v}(\hat{c}_t) X_t \left( \frac{\hat{v}_h}{\hat{v}(\hat{c}_t)} \left( 1 - \hat{\lambda}(1 - \gamma) \right)^{\frac{1}{\gamma}} - 1 \right) \right| dt \geq n \}
\]

The stopped stochastic integrals are therefore martingales, so take expectations under \( Q \) to obtain

\[
\mathbb{E}_Q^Q \left[ e^{-r(\tau^* \wedge t)} \hat{v}(\hat{c}_r \wedge \tau) X_{\tau^* \wedge t} \right] \geq \hat{v}(\hat{c}_0) X_0 - \mathbb{E}_Q^Q \left[ \int_0^{\tau^* \wedge t} e^{-rt} (c_t - k_t \alpha) dt \right]
\]

Now we would like to take the limit \( n \to \infty \), but we need to use the dominated convergence theorem. First,

\[
\left| \int_0^{\tau^* \wedge t} e^{-rt} (c_t - k_t \alpha) dt \right| \leq \int_0^\tau e^{-rt} |(c_t - k_t \alpha)| dt
\]

which is integrable because the contract is admissible. Second,

\[
e^{-r(\tau^* \wedge t)} \hat{v}(\hat{c}_r \wedge \tau) X_{\tau^* \wedge t} \leq \sup_{t \leq \tau} e^{-rt} \hat{v}_h X_t
\]

\(^{28}\)Notice \( \hat{v}'' \) could be discontinuous at \( \hat{c}_t \), but this doesn’t change Ito’s formula.
which also is integrable because the contract is admissible. Upon taking the limit \( n \to \infty \) we obtain 
\[ \tau^n \wedge \tau \to \tau \text{ a.s.} \] and therefore
\[ E_0^Q \left[ e^{-rt} \hat{v}_h X_r \right] \geq \hat{v}(\hat{c}_0) X_0 - E_0^Q \left[ \int_0^r e^{-rt} (c_t dt - k_t [dG_t / G_t - r_t dt]) \right] \]
where we used the admissibility of \( k \) to replace \( k_t \alpha dt \) by \( E_t^Q [dG_t / G_t - r_t dt] \) and then used the LIE. Re-arranging and using \( \hat{v}(\hat{c}_1) \leq \hat{v}(\hat{c}_0) \) we obtain the first result.

For the second part, first let’s show that \( C^* \) is incentive compatible. We already know that for 
the HJB to have a solution it must be the case that \( \hat{B} \) and \( \hat{C} \) are both non-negative, and therefore 
\( \hat{\sigma}^*_\hat{c} \leq 0 \) and \( \hat{\sigma}_X^* \geq 0 \). Since \( \hat{\sigma}^*_\hat{c} \) is bounded, so is \( \hat{\sigma}_X \). In addition, since \( \hat{v} \) is bounded away from zero, 
we have \( \left( 1 - \hat{\lambda}^*_\hat{c} (1 - \gamma) \right) \) bounded as well, and therefore \( \mu^*_X \) is bounded too. Since we also know 
that \( \hat{c}_r = \hat{c}_h \) we can use Lemma 8 and the fact that \( h_0 = 0 \) to show that 
\[ U_0^a \leq U_0^c,0 = u_0 \]
for any valid \( a \in A(C^*) \), and therefore \( C^* \) is indeed incentive compatible. To show that the cost of 
the contract is \( \hat{v}_t X_0^h \), we can use the HJB. If \( \hat{c}_t \in [\hat{c}_l, \hat{c}_h] \) always, then the same argument as in 
the first part shows the desired result. Lemma 9 shows this is always the case and completes the proof.

**Lemma 5**

For the case with \( \gamma \geq \frac{1}{2} \) we will show that \( A(\hat{c}; \hat{v}) \) has at most one root in [0, \( \hat{c}_h \)] for any \( \hat{v} \in (0, \hat{v}_h) \).
Combined with the fact that \( \hat{c}_l \) is a root of \( A(\hat{c}; \hat{v}_l) \) by construction and that \( A(\hat{c}; \hat{v}_l) \) is positive 
when \( \hat{c} \to 0 \), we get that \( A(\hat{c}; \hat{v}_l) \geq 0 \) for all \( \hat{c} \leq \hat{c}_l \) as desired. To show that \( A(\hat{c}; \hat{v}) \) has at most 
one root in [0, \( \hat{c}_h \)] for any \( \hat{v} \in (0, \hat{v}_h) \), we will show that \( A'(\hat{c}; \hat{v}) = 0 \implies A(\hat{c}; \hat{v}) > 0 \) for all \( \hat{c} < \hat{c}_h \).
First compute the derivative (dropping the arguments to avoid clutter)
\[ A' = 1 - \hat{v} \hat{c}^{-\gamma} - \hat{c}^{2\gamma-1} \left( \frac{\alpha}{\phi \beta} \right)^2 \frac{1}{\hat{v}} \]
So
\[ A'_c = 0 \implies \hat{c} - \hat{v} \hat{c}^{1-\gamma} \Rightarrow \hat{c} = \hat{v} \hat{c}^{1-\gamma} = \left( \frac{\alpha}{\phi \beta} \right) \frac{1}{\hat{v}} \]
Plug this into the formula for \( A \) to get
\[ A = \hat{c} - r \hat{v} + \hat{v} \frac{\theta - \hat{c}^{1-\gamma}}{1 - \gamma} - \hat{c}^{2\gamma} \left( \frac{\alpha}{\phi \beta} \right)^2 + \hat{v} \hat{\lambda}_b(\hat{v}) \gamma \theta \]
\[ A \geq \hat{c} - r \hat{v} + \hat{v} \frac{\theta - \hat{c}^{1-\gamma}}{1 - \gamma} - \frac{1}{2\gamma} (\hat{c} - \hat{v} \hat{c}^{1-\gamma}) \]
\[ B(\hat{c}, \hat{v}) = \frac{2\gamma - 1}{2\gamma} \hat{c} + \frac{1 - 3\gamma}{2\gamma} \hat{v}^{1-\gamma} + \hat{v} \frac{\rho - r(1 - \gamma)}{1 - \gamma} \equiv B(\hat{c}, \hat{v}) \]

\( B(\hat{c}, \hat{v}) \) is convex in \( \hat{c} \) because \( 1 - 3\gamma < 0 \) for \( \gamma \geq \frac{1}{2} \), so it’s minimized in \( \hat{c} \) when \( B'_c = 0 \):

\[ \frac{2\gamma - 1}{3\gamma - 1} = \hat{v}^{\gamma} \tag{19} \]

and it is strictly decreasing before this point. Now we have two possible cases:

**CASE 1:** The minimum of \( B \) is achieved for \( \hat{c} \geq \hat{c}_h \), so in the relevant range, it is minimized at \( \hat{c}_h \). So let’s plug in \( \hat{c}_h \) into \( B(\hat{c}, \hat{v}) \):

\[
2\gamma B(\hat{c}_h, \hat{v}) = (2\gamma - 1) \hat{c}_h + \frac{\hat{v}}{1 - \gamma} \left( (\rho - r(1 - \gamma))2\gamma + (1 - 3\gamma)\hat{c}_h^{1-\gamma} \right)
\]

\[
= (2\gamma - 1) \hat{c}_h + \frac{\hat{v}}{1 - \gamma} (\rho - r(1 - \gamma)) (2\gamma^2 + (1 - 3\gamma))
\]

\[
= (2\gamma - 1) \hat{c}_h + \frac{\hat{v}}{1 - \gamma} (\rho - r(1 - \gamma)) (1 - 2\gamma)
\]

\[
(2\gamma - 1) \left( \hat{c}_h - \hat{v} \frac{\rho - r(1 - \gamma)}{\gamma} \right) \geq 0
\]

and the inequality is strict if \( \hat{v} < \hat{v}_h \). So \( A(\hat{c}, \hat{v}) \geq B(\hat{c}, \hat{v}) \geq B(\hat{c}_h, \hat{v}) \geq 0 \) for any \( \hat{c} < \hat{c}_h \).

**CASE 2:** If the minimum is achieved for \( \hat{c}_m \in [0, \hat{c}_h) \) it must be that \( \gamma > 1/2 \). Then plugging in (19) into \( B \):

\[
B(\hat{c}, \hat{v}) \geq \frac{2\gamma - 1}{2\gamma} \hat{c}_m - \frac{2\gamma - 1}{2\gamma} \hat{c}_m + \hat{v} \frac{\rho - r(1 - \gamma)}{1 - \gamma}
\]

\[
= \frac{1 - 2\gamma}{2} \hat{c}_m + \hat{v} \frac{\rho - r(1 - \gamma)}{1 - \gamma}
\]

\[
= \frac{1 - 2\gamma}{2} \hat{c}_m + \frac{2\gamma - 1}{3\gamma - 1} \hat{c}_m^\gamma \frac{\rho - r(1 - \gamma)}{1 - \gamma}
\]

and dividing throughout by \( 2\gamma - 1 > 0 \)

\[
= \frac{1}{2} \frac{\hat{c}_m}{1 - \gamma} + \frac{\hat{c}_m^\gamma}{3\gamma - 1} \frac{\rho - r(1 - \gamma)}{1 - \gamma}
\]

and multiplying by \( \hat{c}_m^{1-\gamma} > 0 \) and using \( \frac{\hat{c}_m^{1-\gamma}}{1 - \gamma} < \frac{\hat{c}_h^{1-\gamma}}{1 - \gamma} \):

\[
> \frac{1}{2} \frac{\hat{c}_h^{1-\gamma}}{1 - \gamma} + \frac{\gamma}{3\gamma - 1} \frac{\rho - r(1 - \gamma)}{(1 - \gamma)^{\gamma}} = \frac{\rho - r(1 - \gamma)}{\gamma(1 - \gamma)} \left( \frac{\gamma}{3\gamma - 1} - \frac{1}{2} \right)
\]

\[
= \frac{\rho - r(1 - \gamma)}{\gamma(1 - \gamma)} \frac{1 - \gamma}{(3\gamma - 1)^2} > 0
\]

So \( A(\hat{c}, \hat{v}) \geq B(\hat{c}, \hat{v}) > 0 \) for all \( \hat{c} \in [0, \hat{c}_h] \).
For the case with $\gamma < \frac{1}{2}$, the second derivative of $A$ is

$$A''_c = \gamma \hat{c}^{1 - \gamma} - (2\gamma - 1)\hat{c}^{2\gamma - 2} \left( \frac{\alpha}{\phi \beta} \right)^2 \frac{1}{\hat{v}} > 0$$

So $A(\hat{c}; \hat{v})$ is strictly convex and so can have at most two roots. Furthermore the smaller root has negative derivative

$$A'_c = 1 - \hat{v}\hat{c}^{\gamma} - \hat{c}^{2\gamma - 1} \left( \frac{\alpha}{\phi \beta} \right)^2 \frac{1}{\hat{v}} \leq 0$$

And the larger root the opposite sign. Notice that if the root has $A'_c = 0$ it means there is only one root. Therefore checking (16) is sufficient and we obtain the desired result.

**Lemma 6**

Because $\hat{v} \in [\hat{v}_l, \hat{v}_h]$ bounded away from zero, and $\hat{v}' \geq 0$, then $\hat{c}^* \leq 0$ and bounded implies $\hat{c}^*_X \geq 0$ and bounded too, as well as $\hat{\lambda}^* \geq 0$. As a result, $U^{c^*, 0}_t = \frac{(X_t^*)^{1 - \gamma}}{1 - \gamma}$ is square integrable with $U^{c^*, 0}_0 = u_0$, and using $\mu^*_X \leq \bar{\mu}_X < r + \theta$, Lemma 7 guarantees that

$$\mathbb{E} \left[ \sup_{t \leq \tau} \hat{v}_h e^{-rt} X_t \right] < \infty$$

where $\bar{\sigma}_X$ is the upper bound on $\hat{c}^*_X$.

**Lemma 7.** Let $X$ be as in (7), and assume $\mu_X$, $\bar{\sigma}_X$, and $\hat{\lambda} \geq 0$ are bounded. If $r + \theta > \bar{\mu}_X$, where $\bar{\mu}_X$ is the upper bound on $\mu_X$, then $\mathbb{E} \left[ \sup_{t \leq \tau} \hat{v}_h e^{-rt} X_t \right] < \infty$.

**Note 4.** Recall that we have assumed throughout that $r > 0$ and $\theta > 0$. For $\theta = 0$ retirement never comes, but it is still true that $\mathbb{E} \left[ \sup_{t \leq \tau} \hat{v}_h e^{-rt} X_t \right] < \infty$ as long as $r > \bar{\mu}_X$. Of course Lemma 7 gives sufficient conditions to check the integrability constraint on admissible contracts, so even if it fails the contract might still be admissible.

**Proof.** We have $X_t$ dominated by

$$\tilde{X}_t = \exp \left( \int_0^t \mu_{X,s} ds + \int_0^t \sigma_{X,s} dW_s - \frac{1}{2} \int_0^t \sigma_{X,s}^2 ds \right)$$

(it’s equal for $t < \tau$, and $X_\tau \leq \tilde{X}_\tau$ because with $\hat{\lambda} \geq 0$ $X$ “jumps down” at $t = \tau$). As a result
\[
\mathbb{E}[\sup_{s \leq \tau} e^{-rs} \hat{v}_h X_s] \leq \mathbb{E}[\sup_{s \leq \tau} e^{-rs} \hat{v}_h \tilde{X}_s].
\]
Define on \( t \in [0, \infty) \)
\[
\tilde{X}_t = \exp \left( \int_0^t \hat{\mu}_X ds + \int_0^t \hat{\sigma}_X dW_s - \frac{1}{2} \int_0^t \hat{\sigma}_X^2 ds \right)
\]
Let \( \tilde{S}_t = \sup_{s \leq t} e^{-rs} \tilde{X}_s \) be the running supremum of the discounted \( \tilde{X}_t \), and likewise \( \hat{S}_t = \sup_{s \leq t} e^{-rs} \hat{X}_s \).
It holds that \( \mathbb{E}[\tilde{S}_\tau] \leq \mathbb{E}[\hat{S}_\tau] \). Now consider
\[
\mathbb{E}[\tilde{S}_\tau] \leq \mathbb{E}[\hat{S}_\tau] = \theta \mathbb{E}\left[ \int_0^\infty e^{-\theta t} \hat{S}_t dt \right]
\]
\[
= \theta \int_0^\infty e^{-(\theta - \lambda) t} \mathbb{E}[e^{-\lambda t} \hat{S}_t] dt \leq \theta \int_0^\infty e^{-(\theta - \lambda) t} \mathbb{E} \left[ \sup_s e^{-\lambda s} \hat{S}_s \right] dt
\]
If \( \theta > \hat{\mu}_X - \hat{\lambda} \) then we can find \( \lambda \in (\hat{\mu}_X - \hat{\lambda}, \theta) \) and \( \lambda > 0 \) (because \( \theta > 0 \)). Since the geometric drift of \( \hat{S} \) is \( \hat{\mu}_X - \hat{\lambda} \) then we have
\[
\mathbb{E} \left[ \sup_s e^{-\lambda s} \hat{S}_s \right] < \infty
\]
(see Shepp and Shiryaev (1993)), so we get
\[
\mathbb{E} \left[ \sup_{s \leq \tau} e^{-rs} \hat{v}_h X_s \right] \leq \mathbb{E} \left[ \hat{v}_h \tilde{S}_\tau \right] \leq \hat{v}_h \theta \mathbb{E} \left[ \sup_s e^{-\lambda s} \hat{S}_s \right] \int_0^\infty e^{-(\theta - \lambda) t} dt < \infty
\]
because \( \theta - \lambda > 0 \). This completes the proof. \( \square \)

**Lemma 8.** Let \( C = (c, U, k) \) be an admissible contract with associated processes \( X, \hat{c}, \) satisfying (5) and (6) with \( dL_t = 0 \). Assume further that \( \hat{\sigma}_c \leq 0, \mu_X, \hat{\sigma}_X \) and \( \left( 1 - \hat{\lambda}_t (1 - \gamma) \right) \) are bounded, and \( \hat{c}_t = \hat{c}_h \). Then for any \( a = (\hat{c}, s) \) valid hidden action, with associated hidden savings \( h \), we have the following upper bound on the utility of the agent under \( a \):
\[
U_t^a \leq \left( 1 + \frac{h_t \hat{c}_t^{1-\gamma}}{X_t} \right)^{1-\gamma} U_t^{c,0}
\]
In particular, since \( h_0 = 0 \) for any \( a, U_0^a \leq U_0^{c,0} \) and the contract \( C \) is therefore incentive compatible.

**Note 5.** This lemma can be extended to contracts where \( dL_t \neq 0 \) quite naturally, but it’s not necessary for the verification of the candidate optimal contract.

**Proof.** Let \( F(\hat{h}, \hat{c}) = \left( 1 + \hat{h} \hat{c}^{1-\gamma} \right)^{1-\gamma} \) where \( \hat{h} = h/X \), and let \( F_t = F(\hat{h}_t, \hat{c}_t) \), and likewise for derivatives, e.g. \( F_{\hat{c},t} = \partial_{\hat{c}} F(\hat{h}_t, \hat{c}_t) \). Also, define \( \hat{\hat{c}} = \hat{c}/X \). Write
\[
e^{-\rho t} \left( U_t^{c,s} - F(\hat{h}_t, \hat{c}_t) U_t^{c,0} \right) = \mathbb{E}_t^s \left[ \int_t^\tau e^{-\rho u} \frac{\hat{c}_u^{1-\gamma}}{1-\gamma} du + \int_t^\tau d \left( e^{-\rho u} F_u U_u^{c,0} \right) + e^{-\rho t} \left( B(U_t^{c,0}, \hat{h}_t X_t) - F_t U_t^{c,0} \right) \right]
\]
30
and using $B(U^{c,0}_\tau, \hat{h}_\tau X_\tau) = F_r U^{c,0}_\tau$

$$e^{-pt} \left( U^{c,0}_{\tau t} - F(\hat{h}_t, \hat{c}_t) U^{c,0}_{\tau t} \right) = \mathbb{E}^{\gamma}_{t} \left[ \int_{t}^{\tau} e^{-pu} \frac{\zeta_{1}^{1-\gamma}}{1-\gamma} du + \int_{t}^{\tau} d(e^{-pu} F_u U^{c,0}_u) \right]$$

We will show that the rhs is non-positive. First write

$$\mathbb{E}^{\gamma}_{t} \left[ \int_{t}^{\tau} e^{-pu} \frac{\zeta_{1}^{1-\gamma}}{1-\gamma} du + \int_{t}^{\tau} d(e^{-pu} F_u U^{c,0}_u) \right] = \mathbb{E}^{\gamma}_{t} \left[ \int_{t}^{\tau} e^{-pt}(1-\gamma)U^{c,0}_u Y_u du \right]$$

with

$$Y_t = \frac{\zeta_{1}^{1-\gamma}}{1-\gamma} + \frac{-\rho F_t + \rho F_t - \zeta_{1}^{1-\gamma} F_t}{1-\gamma}$$

$$+ \frac{F_{\hat{c},t}}{(1-\gamma)} \hat{c}_t \left( \frac{r - \rho}{\gamma} + \frac{1 + \gamma}{2} (\hat{\sigma}_{X,t} + \hat{\sigma}_c) \right)^2 + \frac{\rho - \zeta_{1}^{1-\gamma} \hat{c}_t}{1-\gamma} - \frac{\gamma \hat{c}_t^2}{2} \hat{X}_{t} + \hat{c}_t \hat{\sigma}_{c,t} \hat{c}_t$$

$$+ \frac{F_{\hat{h},t}}{(1-\gamma)} \hat{h}_t \left( r + \hat{c}_t - \hat{c}_t \hat{\sigma}_{c} - \hat{\sigma}_{X,t} \hat{\sigma}_{c,t} \right) - \frac{\gamma \hat{c}_t^2}{2} \hat{X}_{t} + \hat{c}_t \hat{\sigma}_{c,t}$$

$$+ \frac{\hat{\sigma}_{c,t}^2 \hat{c}_t^2 F_{\hat{c},c,t}}{2(1-\gamma)} + \frac{\hat{\sigma}_{c,t}^2 \hat{c}_t^2 F_{\hat{c},c,t} + \hat{\sigma}_{c,t}^2 \hat{h}_t F_{\hat{h},t}}{2(1-\gamma)}$$

$$+ \frac{1}{\lambda_{t}} \left( -(1-\gamma) \hat{\sigma}_{X,t} F_t - F_{\hat{c},t} \hat{\sigma}_{c,t} \hat{c}_t + (\hat{\sigma}_{X,t} \hat{h}_t + \hat{\sigma}_{X,t} \hat{c}_t) \hat{F}_{h,t} \right) \frac{s_{t}}{\beta}$$

$$+ \frac{F_{t}}{1-\gamma} \frac{\lambda_{t}}{\lambda_{t}} + \frac{1}{\lambda_{t}} F_{\hat{c},t} \hat{c}_t \left( \frac{\theta_{M,t}}{\gamma} \theta_{M,t} + \hat{\theta}_{M,t} \right) - \frac{1}{(1-\gamma)} F_{\hat{h},t} \hat{h}_t \theta_{M,t}$$

$$+ \theta \left[ (1-\lambda_{t})(1-\gamma) \frac{1}{(1-\gamma)} \left( 1 + \hat{h}_t (1 + \lambda_{t}(1-\gamma))^{\frac{1}{1-\gamma}} \hat{c}_t^{1-\gamma} \right)^{1-\gamma} - \frac{1}{(1-\gamma)} (1 + \hat{h}_t \hat{c}_t^{1-\gamma})^{1-\gamma} \right]$$

where

$$F_{\hat{c},t} = \gamma(\gamma - 1) \hat{h}_t \hat{c}_t^{1-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma}$$

$$F_{\hat{h},t} = (1-\gamma) \hat{c}_t^{1-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma}$$

$$F_{\hat{c},c,t} = -\gamma(\gamma - 1) \hat{h}_t \hat{c}_t^{2-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma} - \gamma^2(\gamma - 1) \hat{h}_t \hat{c}_t^{2-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma} \left( 1 - \hat{h}_t (1 - \gamma) \hat{c}_t^{1-\gamma} \right)$$

$$F_{\hat{c},h} = \gamma(\gamma - 1) \hat{c}_t^{1-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma} - \gamma^2(\gamma - 1) \hat{h}_t \hat{c}_t^{2-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma}$$

$$F_{\hat{h},h} = -\gamma \hat{c}_t^{2-\gamma} (1-\gamma) \left( 1 + \hat{h}_t \hat{c}_t^{1-\gamma} \right)^{-\gamma-1}$$

To obtain (21) we have used the fact that $\mathbb{E}^{\gamma}_{t} \left[ \int_{t}^{\tau} e^{-pt}(1-\gamma)A_t dW_t \right] = 0$ which I will prove at the end. We will show that $Y_t \leq 0$. To do this, we will split the expression into 4 parts,
where we have used $\hat{h}_t \geq 0$ and $\hat{\sigma}_{c,t} \leq 0$. Here is the only place where $\hat{\sigma}_{c,t} \leq 0$ is used.

Second, the terms multiplying $\theta$:

$$B_t = \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} \left\{ \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) \hat{\lambda}_t + \gamma \hat{h}_t \hat{c}_t^{-\gamma} \left(\frac{1}{\gamma} \hat{\lambda}_{M,t} + \hat{\lambda}_t\right) - \hat{c}_t^{-\gamma} \hat{h}_t \hat{\lambda}_t \right\}$$

$$+ \left(1 - \hat{\lambda}(1 - \gamma)\right) \frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t(1 - \hat{\lambda}(1 - \gamma))^{-\frac{1}{1-\gamma}} \hat{c}_h^{-\gamma}\right)^{1-\gamma} - \frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{1-\gamma}$$

so

$$B_t = \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} \left\{ \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) \hat{\lambda}_t + \hat{h}_t \hat{c}_t^{-\gamma} \left(1 - \hat{c}_h^{-\gamma} \hat{\lambda}_t\right) - \hat{c}_t^{-\gamma} \hat{h}_t \hat{\lambda}_t \right\}$$

$$+ \left(1 - \hat{\lambda}(1 - \gamma)\right) \frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t(1 - \hat{\lambda}(1 - \gamma))^{-\frac{1}{1-\gamma}} \hat{c}_h^{-\gamma}\right)^{1-\gamma} - \frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{1-\gamma}$$

Use the concavity of $\frac{1}{(1-\gamma)}$ to write

$$B_t \leq \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} \left\{ \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) \hat{\lambda}_t + \hat{h}_t \hat{c}_t^{-\gamma} \left(1 - \hat{c}_h^{-\gamma} \hat{\lambda}_t\right) - \hat{c}_t^{-\gamma} \hat{h}_t \hat{\lambda}_t \right\}$$

$$+ \left(1 - \hat{\lambda}(1 - \gamma)\right) \left(\frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{1-\gamma} + \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}(1 - \hat{\lambda}(1 - \gamma))^{-\frac{1}{1-\gamma}} - 1 - \hat{h}_t \hat{c}_t^{-\gamma}\right)\right)$$

$$- \frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{1-\gamma}$$

$$B_t \leq \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} \left\{ \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) \hat{\lambda}_t + \hat{h}_t \hat{c}_t^{-\gamma} \hat{\lambda}_t - \hat{c}_t^{-\gamma} \hat{h}_t \hat{\lambda}_t \right\}$$

$$- \hat{\lambda}(1 - \gamma) \left(\frac{1}{(1 - \gamma)} \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) - \hat{h}_t \hat{c}_t^{-\gamma}\right) \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} = 0$$
Third, all the remaining terms that have $\bar{\sigma}_{X,t}$ or $\bar{\sigma}_{c,t}$ are

\[
C_t = \frac{F_{\hat{c},t}}{1 - \gamma} \hat{c}_t \left( \frac{1 + \gamma}{2} (\bar{\sigma}_{X,t} + \bar{\sigma}_{c,t})^2 - \frac{\gamma}{2} \bar{\sigma}_{X,t}^2 - \bar{\sigma}_{X,t} \bar{\sigma}_{c,t} \right) + \frac{F_{\hat{c},t}}{1 - \gamma} (1 - \gamma) \bar{\sigma}_{X,t} \bar{\sigma}_{c,t} \hat{c}_t
\]

\[
+ \frac{F_{\hat{h},t}}{1 - \gamma} \hat{h}_t \left( -\frac{\gamma}{2} \bar{\sigma}_{X,t}^2 + \bar{\sigma}_{c,t}^2 \right) - \frac{F_{\hat{h},t}}{1 - \gamma} (1 - \gamma) \bar{\sigma}_{X,t}^2 \hat{h}_t + \frac{\bar{\sigma}_{c,t}^2 \bar{\sigma}_{X,t}^2 F_{\hat{c},t} - 2 \bar{\sigma}_{c,t} \bar{\sigma}_{X,t} \hat{c}_t \hat{h}_t F_{\hat{c},t} + \bar{\sigma}_{X,t}^2 \hat{h}_t^2 F_{\hat{h},t}}{2(1 - \gamma)}
\]

\[
C_t = \frac{F_{\hat{c},t}}{1 - \gamma} \hat{c}_t \left( \frac{1 + \gamma}{2} (\bar{\sigma}_{X,t} + \bar{\sigma}_{c,t})^2 - \frac{\gamma}{2} \bar{\sigma}_{X,t}^2 - \bar{\sigma}_{X,t} \bar{\sigma}_{c,t} \right) + \frac{F_{\hat{h},t}}{1 - \gamma} \hat{h}_t \left( -\frac{\gamma}{2} \bar{\sigma}_{X,t}^2 + \bar{\sigma}_{c,t}^2 \right)
\]

\[
\frac{\bar{\sigma}_{c,t}^2 \bar{\sigma}_{X,t}^2 F_{\hat{c},t} - 2 \bar{\sigma}_{c,t} \bar{\sigma}_{X,t} \hat{c}_t \hat{h}_t F_{\hat{c},t} + \bar{\sigma}_{X,t}^2 \hat{h}_t^2 F_{\hat{h},t}}{2(1 - \gamma)}
\]

\[
C_t = \frac{\bar{\sigma}_{X,t}^2}{2} \left( \frac{F_{\hat{c},t}}{1 - \gamma} \hat{c}_t + \frac{F_{\hat{h},t}}{1 - \gamma} \hat{h}_t \gamma + \frac{F_{\hat{h},t} \hat{h}_t}{1 - \gamma} \hat{c}_t \right)
\]

\[
+ \frac{\bar{\sigma}_{c,t}^2}{2} \left( \frac{F_{\hat{c},t}}{1 - \gamma} \hat{c}_t + \frac{F_{\hat{h},t} \hat{h}_t}{1 - \gamma} \hat{c}_t \right)
\]

\[
\frac{\bar{\sigma}_{c,t}^2}{2} \left( \frac{F_{\hat{c},t}}{1 - \gamma} \hat{c}_t + \frac{F_{\hat{h},t} \hat{h}_t}{1 - \gamma} \hat{c}_t \right)
\]

\[
\frac{\bar{\sigma}_{c,t}^2}{2} \left( \frac{F_{\hat{c},t}}{1 - \gamma} \hat{c}_t + \frac{F_{\hat{h},t} \hat{h}_t}{1 - \gamma} \hat{c}_t \right)
\]

which plugging in the $F$’s gets us

\[
C_t = \frac{\bar{\sigma}_{X,t}^2}{2} \left( \frac{\gamma(\gamma - 1) \hat{h}_t \hat{c}_t - \gamma^2(1 + \hat{h}_t \hat{c}_t)^{-1}}{1 - \gamma} \right)
\]

\[
\frac{\gamma(\gamma - 1) \hat{h}_t \hat{c}_t - \gamma^2(1 + \hat{h}_t \hat{c}_t)^{-1}}{1 - \gamma} \right)
\]

\[
\frac{\gamma(\gamma - 1) \hat{h}_t \hat{c}_t - \gamma^2(1 + \hat{h}_t \hat{c}_t)^{-1}}{1 - \gamma} \right)
\]

\[
\frac{\gamma(\gamma - 1) \hat{h}_t \hat{c}_t - \gamma^2(1 + \hat{h}_t \hat{c}_t)^{-1}}{1 - \gamma} \right)
\]

\[
\frac{\gamma(\gamma - 1) \hat{h}_t \hat{c}_t - \gamma^2(1 + \hat{h}_t \hat{c}_t)^{-1}}{1 - \gamma} \right)
\]

\[
\frac{\gamma(\gamma - 1) \hat{h}_t \hat{c}_t - \gamma^2(1 + \hat{h}_t \hat{c}_t)^{-1}}{1 - \gamma} \right)
\]

33
which simplifies to:

\[
C_t = \frac{\sigma_{X,t}^2}{2} \left( -\gamma \hat{h}_t \hat{c}_t^{-\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \hat{c}_t + \hat{c}_t^{-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \hat{h}_t \gamma - \gamma \hat{c}_t^{-2\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma-1} \hat{h}_t^2 \right)
+ \frac{\sigma_{\hat{c},t}^2}{2} \left( -\gamma \hat{h}_t \hat{c}_t^{-\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \left( 1 + \gamma \right) \hat{c}_t + \gamma \hat{h}_t \hat{c}_t^{-2\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \hat{c}_t^2 \right)
+ \gamma^2 \hat{h}_t \hat{c}_t^{-\gamma-2} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma-1} \left( 1 - \hat{h}_t (\gamma - 1) \hat{c}_t^{-\gamma} \right) \hat{c}_t^2 \right)
- \sigma_{X,t} \sigma_{\hat{c},t} \left( \gamma \hat{h}_t \hat{c}_t^{-\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \hat{c}_t - \gamma \hat{c}_t^{-\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \hat{c}_t \hat{h}_t + \gamma^2 \hat{h}_t \hat{c}_t^{-2\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma-1} \hat{c}_t \hat{h}_t \right)
\]

and

\[
C_t = -\frac{\sigma_{X,t}^2 + \sigma_{\hat{c},t}^2}{2} \gamma \hat{h}_t \hat{c}_t^{-1-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma-1} \hat{c}_t^{-1-\gamma} \leq 0
\]

And finally all the remaining terms that don’t involve \( \sigma_{X,t} \) or \( \sigma_{\hat{c},t} \) are

\[
D_t = \frac{\hat{c}_t^{1-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{1-\gamma}}{1 - \gamma} - \frac{\hat{c}_t^{1-\gamma} F_t}{(1 - \gamma)} + \frac{\gamma (\gamma - 1) \hat{h}_t \hat{c}_t^{-\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma}}{(1 - \gamma)} \hat{c}_t \left( \frac{r - \rho}{\gamma} - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right)
+ \frac{(1 - \gamma) \hat{c}_t^{-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma}}{(1 - \gamma)} \hat{h}_t \left( r + \frac{\hat{c}_t - \hat{\hat{c}}_t}{\hat{h}_t} - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right)
\]

which is maximized for \( \hat{c}_t = \hat{c}_t + \hat{h}_t \hat{c}_t^{1-\gamma} \). Plugging this in yields

\[
D_t \leq \frac{\hat{c}_t^{1-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{1-\gamma}}{1 - \gamma} - \frac{\hat{c}_t^{1-\gamma} F_t}{(1 - \gamma)} + \frac{\gamma (\gamma - 1) \hat{h}_t \hat{c}_t^{-\gamma-1} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma}}{(1 - \gamma)} \hat{c}_t \left( \frac{r - \rho}{\gamma} - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right)
+ \frac{(1 - \gamma) \hat{c}_t^{-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma}}{(1 - \gamma)} \hat{h}_t \left( r + \frac{\hat{c}_t - \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right) \hat{c}_t}{\hat{h}_t} - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right)
\]

\[
D_t \leq -\gamma \hat{h}_t \hat{c}_t^{-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \left( \frac{r - \rho}{\gamma} - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right)
+ \hat{c}_t^{-\gamma} \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right)^{-\gamma} \hat{h}_t \left( r + \frac{\hat{c}_t - \left( 1 + \hat{h}_t \hat{c}_t^{-\gamma} \right) \hat{c}_t}{\hat{h}_t} - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right)
\]
\[
D_t \leq \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) \hat{c}_t^{-\gamma} \hat{h}_t \left\{ \rho - r + \frac{\rho - \hat{c}_t^{1-\gamma} - r - \hat{c}_t^{1-\gamma}}{1 - \gamma} \right\}
\]
\[
D_t \leq \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right) \hat{c}_t^{-\gamma} \hat{h}_t \left\{ \rho - \left(\rho - \hat{c}_t^{1-\gamma}\right) - \hat{c}_t^{1-\gamma} \right\} = 0
\]

Since \(A_u, B_u, C_u, D_u \leq 0\) and \(e^{-\mu_k(1 - \gamma)}U^c_{u} \geq 0\) we conclude that
\[
e^{-\rho t} \left(U^c_{t} - F(\hat{h}_t, \hat{c}_t)U^c_{t}\right) = \mathbb{E}_t^{s} \left[ \int_t^\tau e^{-\mu_k} \hat{c}_u^{1-\gamma} du + \int_t^\tau d \left(e^{-\mu_k} F(U^c_{u})\right) \right]
\]
\[
= \mathbb{E}_t^{s} \left[ \int_t^\tau e^{-\mu_k} (1 - \gamma)U^c_{u} du \left(A_u + B_u + C_u + D_u\right) \right] \leq 0
\]

We only need to show that \(\mathbb{E}_t^{s} \left[ \int_t^\tau e^{-\rho t} (1 - \gamma)U^c_{t} A_t dW_s^{t}\right] = 0\). Recall \(A_t = \left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma} \hat{h}_t \hat{c}_t^{-\gamma} \hat{c}_{\ell,t}\)
where \(\left(1 + \hat{h}_t \hat{c}_t^{-\gamma}\right)^{-\gamma}\) is bounded, as well as \(\hat{c}_{\ell}\) and \(\hat{c}_{s}\). Since \(U^c_{t} = \frac{X_t^{1-\gamma}}{1-\gamma}\), we have \((1 - \gamma)\hat{h}_t X^c_{t}\gamma = h_t X_t^{1-\gamma}\) which is square integrable under \(P\). To see this, write
\[
h_{t\wedge \tau} = \int_0^{t\wedge \tau} e^{r(t-u)} (k_u \phi s_u + c_u - \tilde{c}_u) du \leq e^{r(t\wedge \tau)} \left(\frac{\sigma_X \hat{c}_t^{1-\gamma}}{\phi \tilde{b}} \phi s + \hat{c}_h\right) \int_0^{t\wedge \tau} X_u du
\]
\[
\leq e^{r(t\wedge \tau)} \left(\sigma_X \hat{c}_h^{\gamma} \beta^{-1} s + \hat{c}_h\right) (t \wedge \tau) \times \sup_{u \leq t\wedge \tau} X_u
\]
where \(\sigma_X\) is the upper bound on \(\hat{c}_{\ell}\) and \(s\) the upper bound on \(s\). Then
\[
\left(h_{t\wedge \tau} X_{t\wedge \tau}^{\gamma}\right)^2 \leq e^{2r(t\wedge \tau)} \left(\sigma_X \hat{c}_h^{\gamma} \beta^{-1} s + \hat{c}_h\right)^2 (t \wedge \tau)^2 \times \sup_{u \leq t\wedge \tau} X_u^{2(1-\gamma)}
\]
\[
\implies \int_0^t \left(h_{u\wedge \tau} X_{u\wedge \tau}^{\gamma}\right)^2 du \leq e^{2rt} \left(\sigma_X \hat{c}_h^{\gamma} \beta^{-1} s + \hat{c}_h\right)^2 t^3 \sup_{u \leq t} X_u^{2(1-\gamma)}
\]

Now since \(\mu_X, \sigma_X,\) and \((1 - \hat{\lambda}_t(1 - \gamma))^{1-\gamma}\) are bounded, we get that \(\mathbb{E} \left[\sup_{u \leq t} X_u^{2(1-\gamma)}\right] < \infty\). This proves that \(e^{-\rho t} (1 - \gamma)U^c_{t} A_t\) is square integrable under \(P\). Because \(s\) is bounded this means that \(e^{-\rho t} (1 - \gamma)U^c_{t} A_t\) is square integrable under \(P^s\) as well. This completes the proof. \(\square\)

**Lemma 9.** Let \(C^*\) be as in Theorem 1. Then \(\hat{c}_s^{*} \in [\tilde{c}_t, \hat{c}_h]\) with \(\hat{c}_{\ell} = 0\) and \(\mu_{\hat{c}}^* > 0\) at \(\hat{c}_t\).

**Proof.** Drop the asterisk to simplify notation, throughout the proof we are dealing the \(C^*\). Because \(\tilde{v}'(\hat{c}_t) = 0\) we have \(\tilde{B} = 0\) and therefore by (13) we get \(\tilde{c}_{\ell} = 0\) (because it is derived from the HJB as in Theorem 1 we know that \(\tilde{C} \geq 0\), so (13) characterizes \(\tilde{c}_{\ell}\)). Since \(\hat{c}_0 = \hat{c}_t\), if we show that \(\mu_{\hat{c}} > 0\) when \(\hat{c}_t = \hat{c}_t\) we are done.

Looking at (8), with \(\tilde{c}_{\ell} = 0\) we get for the drift
\[
\mu_{\hat{c}} = \frac{r - \rho}{\gamma} + \frac{1}{2} \sigma_X^2 \left(\lambda_b(\tilde{v}) \gamma + 1 - \hat{c}_h^{\gamma} \hat{c}_t \left(1 - \lambda_b(\tilde{v})(1 - \gamma)\right)^{1-\gamma}\right) - \frac{\rho - \hat{c}_t^{1-\gamma}}{1 - \gamma}
\]
So $\mu_\hat{c} > 0$ implies

$$\frac{1}{2} \sigma_X^2 > \frac{\rho - r}{\gamma} + \frac{\theta}{\gamma} \left( \lambda_h(\hat{\nu}) \gamma + 1 - \hat{c}_h \gamma \left( 1 - \hat{\lambda}_b(\hat{\nu})(1 - \gamma) \right)^{\frac{1 - \gamma}{\gamma}} \right) + \frac{\rho - \hat{c}_1}{1 - \gamma} \gamma$$

Since we also want $A(\hat{c}; \hat{\nu}) = 0$, we get

$$0 = \hat{c} - r \hat{\nu} + \hat{\nu} \left( \frac{\rho - \hat{c}_1}{1 - \gamma} - \frac{\gamma}{2} \frac{\sigma_X^2}{\lambda_b(\hat{\nu})\gamma} + \hat{\lambda}_b(\hat{\nu}) \gamma \theta \right)$$

$$< \hat{c} - \hat{\nu} \frac{1 - \hat{c}_1}{\gamma} - \hat{\nu} \theta \left( 1 - \frac{\hat{c}_h}{\gamma} \left( 1 - \hat{\lambda}_b(\hat{\nu})(1 - \gamma) \right)^{\frac{1 - \gamma}{\gamma}} \right) = M$$

Notice that if $\hat{\nu} = \hat{c}_1$ we have $M = 0$, because $\hat{\nu}_h = \hat{c}_1 \gamma$ and therefore $\left( 1 - \hat{\lambda}_b(\hat{\nu})(1 - \gamma) \right)^{\frac{1 - \gamma}{\gamma}} = \left( \hat{\nu}_h / \hat{\nu} = \hat{c}_1 \gamma \right)$. If $\hat{\nu} > \hat{c}_1$ we have $M < 0$ and if $\hat{\nu} < \hat{c}_1$ we have $M > 0$. So for $A(\hat{c}; \hat{\nu}) = 0$ and $\mu_\hat{c} > 0$ we need $\hat{\nu} < \hat{c}_1$. In fact, if $\hat{\nu} = \hat{c}_1$ and in addition

$$\frac{1}{2} \left( \frac{\alpha}{\phi \beta \gamma} \right)^2 = \frac{\rho - \hat{c}_1}{1 - \gamma} + \frac{\rho - r}{\gamma} + \theta \left( 1 - \frac{\hat{\nu}}{\hat{\nu}_h} \right)^{\frac{1 - \gamma}{\gamma}}$$

then we have $A = 0$ and $\mu_\hat{c} = 0$ (notice that with $\hat{\nu}' = 0$ and $\hat{\nu} = \hat{c}_1$ the FOC for $\sigma_X$ (11) yields $\sigma_X = \frac{\alpha}{\phi \beta \gamma}$). In this case, $\lambda_b(\hat{\nu}) = \lambda_\hat{\nu}(\hat{\nu})$ and because we have $\mu_\hat{c} = 0$ we therefore have the value of a stationary contract, i.e. $\hat{\nu} = \hat{\nu}_r(\hat{\nu})$. This is a tangency point where $\hat{\nu}_r(\hat{\nu})$ touches the locus $\hat{\nu}_h(\hat{\nu})$ define by $A(\hat{c}; \hat{\nu}_h(\hat{\nu})) = 0$. So let $(\hat{c}_T, \hat{\nu}_T)$ be defined by $\hat{\nu}_T = \hat{c}_T$ and (22). Suppose this point exists for $\hat{c}_T \in [\hat{c}_*, \hat{c}_h]$. It must be the case that $\hat{\nu}_T \geq \hat{\nu}_l$ since $\hat{\nu}_T$ is the cost of a stationary contract (because $\hat{c}_T \geq \hat{c}_s$ so it’s an incentive compatible stationary contract) but not optimal.

Now towards contradiction, suppose $\mu_\hat{c} \leq 0$ at $\hat{c}_l$. Then it must be the case that $\hat{\nu}_l \geq \hat{c}_1$ because we have $A(\hat{c}_l, \hat{\nu}_l) = 0$. We will show that $A(\hat{c}_l, \hat{\nu}_l) > 0$ and get a contradiction. First take the derivative of $A$:

$$A'_l(\hat{c}_l, \hat{\nu}_l) = 1 - \hat{\nu}_l \left( \hat{c}_l - \hat{c}_l \gamma + \hat{c}_l^{2\gamma - 1} \left( \frac{\alpha}{\phi \beta} \right)^2 \frac{1}{\hat{\nu}_l^2} \right) < 0$$

where the inequality is because $\hat{\nu}_l \geq \hat{c}_1$. So $A(\hat{c}_l, \hat{\nu}_l) > A(\hat{c}_l^{\frac{1}{2}}, \hat{\nu}_l)$. Letting $\hat{c}_m = \hat{\nu}_l^{\frac{1}{2}}$ we get

$$A(\hat{c}_l, \hat{\nu}_l) > \hat{c}_m - \hat{\nu}_l \left( \frac{\rho - \hat{c}_m}{1 - \gamma} - \frac{1}{2} \left( \frac{\alpha}{\phi \beta} \right)^2 \frac{1}{\hat{\nu}_l} + \gamma \theta \hat{\lambda}_b(\hat{\nu}_l) \right)$$

$$= \hat{c}_m - \hat{\nu}_l \left( \frac{\rho - \hat{c}_m}{1 - \gamma} - \frac{\rho - \hat{c}_m}{1 - \gamma} - \frac{\rho - \hat{c}_T}{1 - \gamma} + \gamma \theta \hat{\lambda}_b(\hat{\nu}_T) + \gamma \theta \hat{\lambda}_b(\hat{\nu}_l) \right)$$

and using $\hat{\lambda}_b(\hat{\nu}_T) \leq \hat{\lambda}_b(\hat{\nu}_l)$ because $\hat{\nu}_T \geq \hat{\nu}_l$:

$$> \hat{c}_m + \hat{\nu}_l \left( \frac{\rho - \hat{c}_T}{1 - \gamma} - \frac{\rho - \hat{c}_m}{1 - \gamma} \right) = \hat{c}_m \gamma \frac{\hat{c}_T^{1 - \gamma} - \hat{c}_m^{1 - \gamma}}{1 - \gamma} \geq 0$$

36
where the last equality uses \( \hat{v}_l = \hat{c}_m^{\gamma} \) and the last inequality uses \( \hat{c}_m = \hat{v}_l^{\frac{1}{\gamma}} \leq \hat{v}_T^{\frac{1}{\gamma}} = \hat{c}_T \). This is a contradiction, and therefore it must be the case that \( \mu_{\hat{c}} > 0 \) at \( \hat{c}_l \).

It only remains to show that \((\hat{c}_T, \hat{v}_T)\) exists. Because \( \alpha < \bar{\alpha} \), the lhs of (22) is positive and is bounded by

\[
0 < \text{lhs} < \frac{\rho - r(1 - \gamma) + \theta \gamma}{\gamma(1 + \gamma)} > 0
\]

The rhs of (22) can be written

\[
\text{rhs} = \frac{\rho - r(1 - \gamma) + \theta \gamma}{(1 - \gamma)\gamma} - \frac{c^{1 - \gamma}}{1 - \gamma} \left(1 + \theta \hat{c}_h^{\gamma - 1}\right) = \left(\rho - r(1 - \gamma) + \theta \gamma\right) \left(\frac{1}{1 - \gamma} - \frac{c^{1 - \gamma}}{1 - \gamma} \frac{1}{\rho - r(1 - \gamma)}\right)
\]

For \( \hat{c} = \hat{c}_h \) the rhs is zero. For \( \hat{c} = \hat{c}_* \) we get

\[
\text{rhs} = \frac{\rho - r(1 - \gamma) + \theta \gamma}{\gamma(1 + \gamma)}
\]

Then there is a \( \hat{c} \in [\hat{c}_*, \hat{c}_h] \) solving (22) as desired. This completes the proof.

\[
\square
\]

**Appendix B - numerical algorithm**

TO BE COMPLETED