Federal Reserve Bank of Minneapolis
Research Department

Dynamic Contracting, Persistent Shocks and Optimal Taxation*

Yuzhe Zhang

Working Paper 640

Revised December 2005

ABSTRACT

In this paper I develop continuous-time methods for solving dynamic principal-agent problems in which the agent's privately observed productivity shocks are persistent over time. I characterize the optimal contract as the solution to a system of ordinary differential equations, and show that, under this contract, the agent's utility converges to its lower bound—immiseration occurs. I also show that, unlike in environments with i.i.d. shocks, the principal would like to renegotiate with the agent when the agent's productivity is low—it is not renegotiation-proof. I apply the theoretical methods I have developed and numerically solve this (Mirrleesian) dynamic taxation model. I find that it is optimal to allow a wedge between the marginal rate of transformation and individuals' marginal rate of substitution between consumption and leisure. This wedge is significantly higher than what is found in the i.i.d. case. Thus, using the i.i.d. assumption is not a good approximation quantitatively when there is persistence in productivity shocks.

*Zhang, University of Minnesota and Federal Reserve Bank of Minneapolis. I thank Javi Fernandez Blanco, Narayana Kocherlakota, Erzo Luttmer, Andrew Mclennan, Christopher Phelan for their helpful comments during this work. I thank V.V. Chari and Larry Jones for their continuous encouragement and support. All remaining errors are mine. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

A common assumption in the dynamic mechanism design literature is that the agent's privately observed shocks are i.i.d. As pointed out by Fernandes and Phelan[7], this assumption is merely for the sake of tractability. It implies that, at the beginning of a given date, an agent's forward looking utility of following a given strategy when facing a given contract is independent of past histories[1].

However, in a lot of economic environments with hidden information, the agent's shocks are highly persistent. For example, in (Mirrleesian)dynamic optimal taxation with hidden productivities shocks, Kocherlakota[12] comments:

> (The i.i.d. assumption is not)... particularly good approximation to what we know about individual skills from the empirical literature on individual wages. This literature documents that individuals experience large and persistent shocks to their wages(and presumably to their skills as well) throughout their lives(see Storesletten, Telmer and Yaron[15] and Meghir and Pistaferri[10]).

In the design of optimal health insurance, it is well known that a customer's health condition today is strongly correlated with her previous conditions. And in unemployment insurance where an unemployed worker's searching effort is hidden, it is typical that the worker's chance of finding a new job depends not only on her current effort, but also on her searching effort in the past.

Fernandes and Phelan[7] developed a recursive formulation of the contracting problem in which private types are serially correlated. In these situations, different types of agents derive different continuation utilities from the same continuation contract. When the agent

[1] Fernandes and Phelan called this property *common knowledge of preferences* over continuation contracts.

chooses between truth-telling and lying, she compares the continuation utility as a truth-teller and the continuation utility as a liar. Thus it is necessary for the principal to enforce a vector of utilities for all the potentially different types. They showed that this vector of continuation utilities is the state variable in their recursive formulation.

This paper is based on Fernandes and Phelan's recursive formulation. Different from Fernandes and Phelan, who solved the optimal contract by numerical iteration following the idea of Abreu, Pearce and Staccheti[1], I use continuous-time methods and characterize the optimal contract as the solution to a system of ordinary differential equations. The advantage of the continuous-time method is twofold. First, it simplifies the mechanism design problem and thus allows us to discover new qualitative properties of the optimal contract. For example, I find that the cost of delivering a utility vector is increasing in the promised utility but decreasing in the threat utility. The contract is not renegotiation-proof when the agent reports to be the low-productivity type. I am also able to show that asymptotically the agent's utility converges to its lower-bound almost surely. Neither of these properties can be easily derived from a numerical iteration. Second, the ordinary differential equations allow different numerical methods which are faster and more accurate than the commonly used iteration methods. Thus the continuous-time methods provide a new approach to the mechanism design problem both qualitatively and numerically.

Our method of solving dynamic contract has immediate applications in numerous economic problems with hidden information. Besides the benchmark model with taxation, I give two more examples here. In health insurance, a typical situation between an insurance company and a customer is that the customer's privately observed health condition is varying but persistent over time. For simplicity, let us assume that her condition θ_t at time t can be either serious($\theta_t = \theta_H$), or non-serious($\theta_t = \theta_L < \theta_H$). Let her consumption of medicare be c_t and her instantaneous utility function be $\theta_t u(c_t)$, implying that she desires more medicare while condition being serious than non-serious. The design of the optimal consumption plan[2] $\mathscr{C} = \{c_t\}_{t=0}^{\infty}$ would be the solution to

$$\begin{aligned} \min_{\mathscr{C}} \quad & E\left[\int_0^\infty e^{-rt} c_t dt\right] \\ \text{s.t.} \quad & E\left[\int_0^\infty e^{-rt} \theta_t u(c_t) dt\right] \geq \bar{U}, \end{aligned}$$

where $\bar{U}$ is the outside option that the customer has.

In executive compensation, it is often argued that the manager knows more about the profitability of the firm than shareholders. Let θ_t be the profit of the firm that is only observable to the manager and let d_t be the dividend payment. I assume that the managers can consume the rest of the profit not taken by the shareholders. Then the efficient design of contract $\mathscr{D} = \{d_t\}_{t=0}^{\infty}$ would be to maximize the dividend payment given the promised utility of the manager.

$$\begin{aligned} \max_{\mathscr{D}} \quad & E\left[\int_0^\infty e^{-rt} d_t dt\right] \\ \text{s.t.} \quad & E\left[\int_0^\infty e^{-rt} u(\theta_t - d_t) dt\right] \geq \bar{U}. \end{aligned}$$

If the utility function $u(c)$ takes the form of $-e^{-\sigma c}$, the problem can be transformed into the previous model with taste shocks.

This paper is motivated by Sannikov[13, 14], who recently introduced a new continuous-time framework to study dynamic models with hidden actions. In his environments with imperfect monitoring, the public signal is a Brownian Motion with a drift term driven by the agent's actions. Using continuous-time methods, he characterized the equilibrium set([13]) and the optimal contract([14]) both by differential equations. This paper has the same spirit as his in the sense that we provide similar characterizations of the optimal contract. One thing that is different between his methods and mine is that we use

[2] See Section 2 for a detailed description of the contracting problem.

different stochastic processes in the construction of the model. He uses drifted Brownian motion(with continuous sample paths) while I choose a jumping process.[3] This seemly technical consideration actually explains the drastic difference between the two continuous-time methods.

Kapicka[11] also attacked the optimal taxation problem with persistent shocks by using a first-order approach. He studied an environment in which the agent can potentially have a continuum of types and claimed that the state variable of the original problem can be simplified into two numbers. However, his paper did not show the equivalence between the original problem and the simplified one, thus it is hard to tell, among his characterizations of the simplified problem, which can carry over to the true problem, and which can not.

The rest of the paper is organized as follows. Section 2 lays out the economic environment and sets up the social planner's contracting problem. In section 3, I derive the continuous-time law of motion of the state variable as one differential equation for the promised utility and one differential inequality for the threat utility. The resulting differential equations are put to use in Section 4 to characterize the set of implementable utility pairs and in Section 5 to study the long-run dynamics of the optimal contract. Here I show that the agents utility converges to its lower bound almost surely. In Section 6, through a numerical example, I show that models with persistent shocks implies significantly higher wedges than the models with i.i.d. shocks. The last section contains concluding comments. All proofs omitted in the main text are in the Appendix.

[3]Sannikov's framework can not be readily adopted to study hidden information models, because Brownian motion(or more generally diffusion process) has a continuum of states, making the state variable in Fernandes and Phelan's recursive formulation too large to handle.

2. A Dynamic Taxation Problem

I shall consider an environment similar to that in Golosov, Kocherlakota and Tsyvinski [8] with privately observed productivity shocks. Time is continuous and $t \in [0, \infty)$. At time 0, a risk-neutral principal and a risk-averse agent can enter into binding contracts(put it differently, they can commit). The principal is able to borrow or lend from the outside at a constant interest rate r. Thus given a utility level she promises to the agent, her objective is to design a contract that minimizes the expected cost of the consumption-output plan she is committed to deliver. The preferences of the agent are

$$E\left[\int_0^\infty e^{-rt}\left[u(c_t) - \theta_t v(y_t)\right] dt\right], \tag{1}$$

where c_t and y_t are the agent's consumption and output at time t, and E is the expectation operator. I assume that the instantaneous utility function u has bounded domain $[0, \bar{c}]$[4]. In $(0, \bar{c})$, u is twice continuously differentiable, u' is positive, u'' is negative and $u(0) = 0$, $\lim_{c\to 0} u'(c) = \infty$, $\lim_{c\to\bar{c}} u'(c) = 0$. Disutility function v has bounded domain $[0, \bar{y}]$. v', v'' are both positive and $v(0) = 0$, $\lim_{y\to 0} v'(y) = 0$, $\lim_{y\to\bar{y}} v'(y) = \infty$. For convenience, I use $\bar{u}, \bar{v}$ to denote $u(\bar{c}), v(\bar{y})$, respectively.

The agent's privately observed θ_t is her taste shock. It can be re-interpreted as productivity shock if $v(y) = y^\gamma, \gamma > 1$. In this case $\theta v(y) = v(l)$, where $l = y/\phi, \phi = \theta^{-1/\gamma}$, the agent is able to transform 1 unit of labor into ϕ units of output and her disutility depends on the amount of labor l she spends to produce y. To keep matters simple, I shall consider the case where θ_t may assume only two values, θ_L, θ_H with $\theta_H > \theta_L > 0$. Formally, $\{\theta_t\}_{t=0}^\infty \in \{\theta_L, \theta_H\}$ is a two-state continuous-time Markov process that spends an exponential time with rate λ in one state before going to the other. Notice that in

[4]The purpose of this assumption is to make the range of utility function a compact set, which helps to generate a compact set of promised values. Having a merely bounded utility function without bounded domain can not guarantee the compactness of the set.

my formulation, the H type agent has lower productivity, in the sense that her marginal disutility is higher than that of the L type when producing the same amount of output. I assume that initially the agent knows her type, while the principal holds a priori belief $(p_L, 1-p_L)$, with p_L being the probability of L type. At time 0, the principal offers a contract, which the agent may accept or reject. If the agent accepts, she sequentially reports the newly-observed shocks to the principal, and the principal implements the contract based on the reported history. Before I go into the details of the agent's strategy space and the planner's contracting problem, it would be crucial to first understand the sequence of actions after the agent accepts the contract. I adopt an approach commonly used in game theory, that is, I first describe a discrete-time analogue of the model, and then think of my continuous-time model as the limit of a sequence of discrete-time models when I let the length of each period converge to 0.

Let $dt > 0$ be the length of one period. Period n represents the time interval $[ndt, (n+1)dt)$. Productivity shocks, reports and consumption-output decisions all happen at the beginning of each period. That is, at $t = ndt$, after the agent observes her new type, she immediately reports it to the principal and based on the history of report, the principal implements consumption-output plan (c_t, y_t). For the rest of the period, (θ_t, c_t, y_t) remain unchanged, and the agent and the principal wait for the beginning of the next period. To be consistent with the continuous-time Markov process, I also require $\Pr(\theta_{ndt+dt} = \theta_j | \theta_{ndt} = \theta_i) = \lambda dt$, for $i \neq j$.

Now I continue the discussion of the continuous-time contracting problem. I use the following notation. I denote the set of all possible histories up to time t(but not including t) by Θ^{t-}, and the set of all possible histories up to and including t by Θ^t.

$$\Theta^{t-} = \{f : f \text{ is a right continuous function } [0,t) \mapsto \{L, H\}, \text{ and has finite jumps.}\}$$

$$\Theta^{t} = \{f : f \text{ is a right continuous function } [0,t] \mapsto \{L, H\}, \text{ and has finite jumps.}\}$$

I use θ^{t-} and h^{t-} to denote generic elements of Θ^{t-} and θ^t, h^t to denote those of Θ^t. A strategy for the principal is to offer a contract $\mathscr{C} = \{c_t, y_t\}_{t=0}^{\infty}$ at time 0, where $c_t : \Theta^t \mapsto [0, \bar{c}]$ specifies the consumption, and $y_t : \Theta^t \mapsto [0, \bar{y}]$ specifies the output at t after the agent's report. A strategy for the agent is a collection of functions $\sigma = \{\sigma_t\}_{t=0}^{\infty}$, where each $\sigma_t : \Theta^{t-} \times \{L, H\} \mapsto \{L, H\}$ maps a history of shocks before t together with the newly observed shock at the beginning of t into a report. Given a type realization $\theta^t \in \Theta^t$, $\sigma^t(\theta^t) = \{\sigma_s(\theta^s)\}_{s=0}^t$ is the reported history under strategy σ up to t. A strategy σ is *truth-telling* if for all t and $\theta^{t-} \in \Theta^{t-}$, $\sigma_t(\theta^{t-}, L) = L$ and $\sigma_t(\theta^{t-}, H) = H$. With a bit of abuse of notation, I use θ to denote the truth-telling strategy.

Given a contract $\mathscr{C}$, the agent with initial type $i (i = L, H)$ receives ex-ante utility $w_i(\sigma; \mathscr{C})$ if she follows strategy σ, where

$$w_i(\sigma; \mathscr{C}) = E\left[\int_0^{\infty} e^{-rt}\left[u(c_t(\sigma^t(\theta^t))) - \theta_t v(y_t(\sigma^t(\theta^t)))\right] dt | \theta_0 = \theta_i\right].$$

A contract $\mathscr{C}$ is said to be *incentive-compatible*(I.C.) if

$$w_i(\theta; \mathscr{C}) = \max_{\sigma} w_i(\sigma; \mathscr{C}), \text{ for } i = L, H.$$

In our environment with commitment, the revelation principle is applicable. Therefore we restrict attention to I.C. contracts. Notice that the optimal contract minimizes the principal's expected cost, thus optimality implicitly depends on the principal's belief $(p_L, 1 - p_L)$. Formally, given the outside option $(\bar{U}_L, \bar{U}_H)$ for the two types at time 0, the principal's problem is

$$\begin{aligned}
&\min_{\mathscr{C}} p_L E\left[\int_0^{\infty} e^{-rt}\left[c_t(\theta^t) - y_t(\theta^t)\right] dt | \theta_0 = \theta_L\right] + \\
&(1 - p_L) E\left[\int_0^{\infty} e^{-rt}\left[c_t(\theta^t) - y_t(\theta^t)\right] dt | \theta_0 = \theta_H\right] \\
&s.t. \quad w_L(\theta; \mathscr{C}) = \max_{\sigma} w_L(\sigma; \mathscr{C}) \geq \bar{U}_L, \\
&\qquad w_H(\theta; \mathscr{C}) = \max_{\sigma} w_H(\sigma; \mathscr{C}) \geq \bar{U}_H.
\end{aligned}$$

The above problem can be easily solved once we know the solutions for special cases where $p_L = 0$ or 1. This can be seen as follows. For (w_L, w_H), let $V_L(w_L, w_H)$ be the cost of optimally implementing (w_L, w_H) under belief $p_L = 1$,

$$
\begin{aligned}
V_L(w_L, w_H) &= \min_{\mathscr{C}} E\left[\int_0^\infty e^{-rt}\left[c_t(\theta^t) - y_t(\theta^t)\right] dt|\theta_0 = \theta_L\right] \\
&\quad s.t. \quad w_L = w_L(\theta; \mathscr{C}) = \max_\sigma w_L(\sigma; \mathscr{C}), \\
&\qquad\qquad w_H = w_H(\theta; \mathscr{C}) = \max_\sigma w_H(\sigma; \mathscr{C}).
\end{aligned}
$$

and $V_H(w_L, w_H)$ the cost under belief $p_L = 0$. Then the original problem would be

$$
\begin{aligned}
\min_{(w_L, w'_L, w_H, w'_H)} \quad & p_L V_L(w_L, w'_H) + (1 - p_L) V_H(w'_L, w_H) \\
s.t. \quad & w_L \geq w'_L, w_L \geq \bar{U}_L, \\
& w_H \geq w'_H, w_H \geq \bar{U}_H.
\end{aligned}
$$

In the rest of the paper, I shall focus on the optimal contract when the principal holds the belief of either 0 or 1.

3. Incentive Constraints and Law of Motion of (w_L, w_H)

Given a contract $\mathscr{C}$ and a report $h^{t-} \in \Theta^{t-}$, let $(w_L(h, t), w_H(h, t))$ be the continuation utility of the agent with newly observed type L and H respectively, before the agent makes the report at time t. Because the utility pair $(w_L(h, t), w_H(h, t))$ serves as the state variable in a recursive formulation, I derive the law of motion of the state variable in this section.

Let $w_i(h, t, j)$ be the continuation utility of the type i agent, after she makes a type j report at time t. For example, $w_L(h, t, H)$ is the utility of the agent after she observes that her true type is L, and cheats the principal. For $s > t$, (h, t, j, s) denotes the history in which the report is h^{-t} before t and j from t to s(but not including s). $w_i(h, t, j, s)$

denotes the continuation utility of the type i agent at time s, before she makes the report, and given that her reported history before s is (h,t,j,s). Obviously,

$$w_L(h,t) = \max\left\{w_L(h,t,L), w_L(h,t,H)\right\}, w_H(h,t) = \max\left\{w_H(h,t,L), w_H(h,t,H)\right\}.$$

Right continuity of the report implies that, after the report at t, the agent needs to wait a small but positive amount of time before she reports a switch in her type, and this implies

$$w_i(h,t,j) = \lim_{s\downarrow t} w_i(h,t,j,s), i = L,H, j = L,H. \tag{2}$$

THEOREM 1 *Let $\mathscr{C}$ be a contract and (w_L, w_H) be an arbitrary stochastic process. $\mathscr{C}$ is I.C. and $(w_L(h,t), w_H(h,t))$ is the promised utility pair under $\mathscr{C}$ if and only if the following statements hold.*

(1) There is a uniform bound $B > 0$, such that for any time t and history h^{t-},

$$-B \le w_i(h,t) \le B, i = L,H. \tag{3}$$

(2) The evolution of $(w_L(h,t), w_H(h,t))$ satisfies

$$\begin{aligned} w_L(h,t) &= \lim_{s\downarrow t} w_L(h,t,L,s) \ge \lim_{s\downarrow t} w_L(h,t,H,s) && (4)\\ w_H(h,t) &= \lim_{s\downarrow t} w_H(h,t,H,s) \ge \lim_{s\downarrow t} w_H(h,t,L,s) && (5)\end{aligned}$$

$$\begin{aligned} \lim_{s\downarrow t}\frac{w_L(h,t,L,s)-w_L(h,t)}{s-t} &= (\lambda+r)w_L(h,t) - \lambda w_H(h,t,L) - \\ &\quad u(c_t(h,t,L)) + \theta_L v(y_t(h,t,L)) && (6)\\ \lim_{s\downarrow t}\frac{w_H(h,t,L,s)-w_H(h,t)}{s-t} &\le (\lambda+r)w_H(h,t,L) - \lambda w_L(h,t) - \\ &\quad u(c_t(h,t,L)) + \theta_H v(y_t(h,t,L)) && (7)\\ \lim_{s\downarrow t}\frac{w_L(h,t,H,s)-w_L(h,t)}{s-t} &\le (\lambda+r)w_L(h,t,H) - \lambda w_H(h,t) - \\ &\quad u(c_t(h,t,H)) + \theta_L v(y_t(h,t,H)) && (8)\\ \lim_{s\downarrow t}\frac{w_H(h,t,H,s)-w_H(h,t)}{s-t} &= (\lambda+r)w_H(h,t) - \lambda w_L(h,t,H) - \\ &\quad u(c_t(h,t,H)) + \theta_H v(y_t(h,t,H)). && (9)\end{aligned}$$

Proof. (Necessity.) If $\mathscr{C}$ is I.C., truth-telling is an optimal strategy. I obtain

$$\begin{aligned} w_L(h,t) &= w_L(h,t,L) \geq w_L(h,t,H), \\ w_H(h,t) &= w_H(h,t,H) \geq w_H(h,t,L). \end{aligned}$$

Together with equation (2), I obtain equations (4),(5). Next I will prove equations (6),(7). The proofs for equations (8),(9) are analogous and thus omitted.

For $s > t$, I discretize the time interval $[t,s)$ into $[t,t+dt),[t+dt,t+2dt)\ldots$, where dt is a small number. Recall that θ remains constant in each subinterval and

$$\begin{aligned} \Pr(\theta_{t+dt} = \theta_L | \theta_t = \theta_L) &= 1 - \lambda dt \\ \Pr(\theta_{t+dt} = \theta_H | \theta_t = \theta_L) &= \lambda dt. \end{aligned}$$

I have

$$\begin{aligned} w_L(h,t) &= [u(c_t(h,t,L)) - \theta_L v(y_t(h,t,L))]\, dt + e^{-rdt}[\\ &\quad (1-\lambda dt) w_L(h,t,L,t+dt) + \lambda dt w_H(h,t,L,t+dt)] \\ &= w_L(h,t,L,t+dt) + [u(c_t(h,t,L)) - \theta_L v(y_t(h,t,L)) + \\ &\quad \lambda w_H(h,t,L,t+dt) - (\lambda + r) w_L(h,t,L,t+dt)] dt. \end{aligned}$$

Notice that in the second equality, I ignore all the terms with order two or above. Taking limit in the above yields

$$\begin{aligned} &\lim_{dt \to 0} \frac{w_L(h,t,L,t+dt) - w_L(h,t)}{dt} \\ = \;& \lim_{dt \to 0} ((\lambda + r) w_L(h,t,L,t+dt) - \lambda w_H(h,t,L,t+dt) - u(c_t(h,t,L)) + \theta_L v(y_t(h,t,L))) \\ = \;& (\lambda + r) w_L(h,t) - \lambda w_H(h,t,L) - u(c_t(h,t,L)) + \theta_L v(y_t(h,t,L)). \end{aligned}$$

Last I prove equation (7). Since under $\mathscr{C}$, the H type agent can not obtain higher utility than $w_H(h,t)$ by pretending to be L type for period $[t, t+dt)$,

$$\begin{aligned} w_H(h,t) &\geq [u(c_t(h,t,L)) - \theta_H v(y_t(h,t,L))]dt + e^{-rdt}[\\ &\quad (1-\lambda dt) w_H(h,t,L,t+dt) + \lambda dt w_L(h,t,L,t+dt)] \\ &= w_H(h,t,L,t+dt) + [u(c_t(h,t,L)) - \theta_H v(y_t(h,t,L)) + \\ &\quad \lambda w_L(h,t,L,t+dt) - (\lambda + r) w_H(h,t,L,t+dt)]dt. \end{aligned}$$

I obtain

$$\begin{aligned} & \lim_{dt \to 0} \frac{w_H(h,t,L,t+dt) - w_H(h,t)}{dt} \\ \leq \; & \lim_{dt \to 0} ((\lambda + r) w_H(h,t,L,t+dt) - \lambda w_L(h,t,L,t+dt) - u(c_t(h,t,L)) + \theta_H v(y_t(h,t,L))) \\ = \; & (\lambda + r) w_H(h,t,L) - \lambda w_L(h,t) - u(c_t(h,t,L)) + \theta_H v(y_t(h,t,L)). \end{aligned}$$

(Sufficiency.) I verify two things. First truth-telling obtains the promised utility $(w_L(h,t), w_H(h,t))$. Second, truth-telling is optimal. From equation (6), I have

$$w_L(h,t) = (u(c_t) - \theta_L v(y_t))dt + e^{-rdt}\left[(1-\lambda dt) w_L(h,t,L,t+dt) + \lambda dt w_H(h,t,L,t+dt)\right].$$

This implies that truth-telling delivers $w_L(h,t)$ at time t as long as it delivers utilities $(w_L(h,t,L,t+dt), w_H(h,t,L,t+dt))$ from $t+dt$ on. Recursively, this generates

$$\begin{aligned} w_L(h,t) &= \sum_{k=0}^{n} e^{-rkdt} E_t[u(c_{t+kdt}) - \theta_{t+kdt} v(y_{t+kdt})]dt + \\ &\quad e^{-r(n+1)dt} E_t[w_j(h^{t+(n+1)dt})], \end{aligned}$$

where $h^{t+(n+1)dt}$ denotes a possible history up to $t+(n+1)dt$, with the history before t identical to h^t and E_t denotes the conditional expectation based on the history h^t. Since $w_j(h^{t+(n+1)dt})$ is a bounded random variable, taking limit $n \to \infty$ in the above equation gives

$$w_L(h,t) = \sum_{k=0}^{\infty} e^{-rkdt} E_t[u(c_{t+kdt}) - \theta_{t+kdt} v(y_{t+kdt})]dt.$$

The utility $w_L(h,t)$ is delivered under the truth-telling strategy.

Cheating can not obtain a utility higher than $w_L(h,t)$. This can be seen from the following. From equation (8),

$$\begin{aligned} w_L(h,t) &\geq w_L(h,t,H,t+dt) + [u(c_t(h,t,H)) - \theta_L v(y_t(h,t,H)) + \\ &\quad \lambda w_H(h,t,H,t+dt) - (\lambda+r)w_L(h,t,H,t+dt)]dt \\ &= [u(c_t(h,t,H)) - \theta_L v(y_t(h,t,H))]dt + e^{-rdt}[\\ &\quad (1-\lambda dt)w_L(h,t,H,t+dt) + \lambda dt w_H(h,t,H,t+dt)] \\ &= w_L(h,t,H). \end{aligned}$$

An analogous argument shows $w_H(h,t) \geq w_H(h,t,L)$. Therefore, $(w_L(h,t), w_H(h,t))$ is the promised utility pair under $\mathscr{C}$ and $\mathscr{C}$ is I.C. ■

Remark 1 *If we think of $w_L(h,t,L,s)$ as a function of s, then it is a continuous function, while $w_H(h,t,L,s)$ may be a discontinuous function with downward jumps. The number of jumps is at most countable.*

Remark 2 *If $w_H(h,t) = w_H(h,t,L) = lim_{s\downarrow t} w_H(h,t,L,s)$, then equations (6),(7) can be simplified to*

$$\begin{aligned} \lim_{s\downarrow t} \frac{w_L(h,t,L,s) - w_L(h,t)}{s-t} &= (\lambda+r)w_L(h,t) - \lambda w_H(h,t) - \\ &\quad u(c_t(h,t,L)) + \theta_L v(y_t(h,t,L)), \\ \lim_{s\downarrow t} \frac{w_H(h,t,L,s) - w_H(h,t)}{s-t} &\leq (\lambda+r)w_H(h,t) - \lambda w_L(h,t) - \\ &\quad u(c_t(h,t,L)) + \theta_H v(y_t(h,t,L)). \end{aligned}$$

And given that jumping points are countable, the above equations hold almost everywhere.

Remark 3 *Given that the precise meanings of the symbols are understood, I rewrite the conditions in Theorem 1 in a more concise and intuitive way. There are three stages in*

the report, the stage when the agent reports to be L, the stage to be H and the time point when a switch occurs. Conditions in $(4-9)$ *are equivalent to:*
There is a non-negative stochastic process μ*, such that, if the report is L,*

$$\begin{aligned}\frac{dw_L}{dt} &= (\lambda+r)w_L - \lambda w_H - u(c_t) + \theta_L v(y_t),\\ \frac{dw_H}{dt} &= (\lambda+r)w_H - \lambda w_L - u(c_t) + \theta_H v(y_t) - \mu_t.\end{aligned}$$

If the report is H,

$$\begin{aligned}\frac{dw_L}{dt} &= (\lambda+r)w_L - \lambda w_H - u(c_t) + \theta_L v(y_t) - \mu_t,\\ \frac{dw_H}{dt} &= (\lambda+r)w_H - \lambda w_L - u(c_t) + \theta_H v(y_t).\end{aligned}$$

If there is a switch from L to H,

$$\begin{aligned}w_L(t) &\geq \lim_{s\downarrow t} w_L(s),\\ w_H(t) &= \lim_{s\downarrow t} w_H(s).\end{aligned}$$

If there is a switch from H to L,

$$\begin{aligned}w_L(t) &= \lim_{s\downarrow t} w_L(s),\\ w_H(t) &\geq \lim_{s\downarrow t} w_H(s).\end{aligned}$$

We interpret possible jumps in the differential inequality as $\mu_t = \infty$*.*

4. The Set of Implementable Utilities

Before we try to find the optimal contract to implement any particular utility pair, it will be helpful to consider the set of all utility pairs that are implementable by at least one contract. Conceptually we could obtain the set by the following procedure. Pick any

I.C. contract and calculate two types' utilities under the truth-telling strategy. This will generate a point for us. Then move to another contract, and obtain a different point. After we go over all possible I.C. contracts and collect all points, we get the set. This set is similar to the set of equilibrium payoffs in the context of repeated games. For any vector in the set, there is an equilibrium strategy to achieve it, while any vector not in the set can not be achieved. Formally, define

$$W = \{(w_L(\theta; \mathscr{C}), w_H(\theta; \mathscr{C})) : \mathscr{C} \text{ is I.C. } \}.$$

The common approach in the literature is to compute this set by iteration. Following the idea of Abreu, Pearce and Staccheti[1], we may start with an initial guess which contains W. We then iterate until the sequence of sets converges to W, which is the largest fixed point of the operator. However, using continuous-time methods, I shall show that this set can be obtained directly. In fact the boundary of W can be characterized by differential equations. The rest of this section will be devoted to this characterization.

I first study some simple contracts. If the contract always specifies maximal consumption $\bar{c}$ and minimal output 0, regardless of reports(i.e. $c_t(h^t) = \bar{c}, y_t(h^t) = 0, \forall h^t \in \Theta^t$), then the contract can implement the pair $(\bar{u}/r, \bar{u}/r)$, which is the upper-right corner of W. If consumption 0 and output $\bar{y}$ are always specified , the lower-left corner is implemented. I denote it by (x_L, x_H), where $x_L = -\bar{v}((\lambda + r)\theta_L + \lambda\theta_H)/(r(2\lambda + r)), x_H = -\bar{v}((\lambda + r)\theta_H + \lambda\theta_L)/(r(2\lambda + r))$. It is easy to see that the "consumption 0, output 0" contract implements the utility pair $(0, 0)$, while the "consumption $\bar{c}$, output $\bar{y}$" contract implements $(x_L + \bar{u}/r, x_H + \bar{u}/r)$.

Next I look at four families of contracts. The first two families are indexed by $c^* \in (0, \bar{c})$. Any contract $\mathscr{C}^{1c^*}$ in the first family is

$$(c_t^{1c^*}(h^t), y_t^{1c^*}(h^t)) = (c^*, 0), \forall h^t \in \Theta^t,$$

which implements the utility pair

$$(u(c^*)/r, u(c^*)/r), c^* \in (0, \bar{c}). \tag{10}$$

Any contract $\mathscr{C}^{2c^*}$ in the second family

$$(c_t^{2c^*}(h^t), y_t^{2c^*}(h^t)) = (c^*, \bar{y}), \forall h^t \in \Theta^t.$$

implements the utility pair

$$(x_L + u(c^*)/r, x_H + u(c^*)/r), c^* \in (0, \bar{c}). \tag{11}$$

The third and fourth families of contracts are indexed by $t^* \in (0, \infty)$. Any contract $\mathscr{C}^{3t^*}$ in the third family is

$$(c_t^{3t^*}(h^t), y_t^{3t^*}(h^t)) = \begin{cases} (0, 0), t \leq t^* \\ (0, \bar{y}), t > t^*, \end{cases}$$

while in the fourth family, a contract $\mathscr{C}^{4t^*}$ is

$$(c_t^{4t^*}(h^t), y_t^{4t^*}(h^t)) = \begin{cases} (\bar{c}, \bar{y}), t \leq t^* \\ (\bar{c}, 0), t > t^*. \end{cases}$$

The utility pair $(w_L^{3t^*}, w_H^{3t^*})$ implemented by $\mathscr{C}^{3t^*}$ can be solved in the following way. Under contract $\mathscr{C}^{3t^*}$ and when $t \leq t^*$, the promised utility evolves according to the differential equation system

$$\begin{aligned} \frac{dw_L}{dt} &= (\lambda + r) w_L - \lambda w_H \\ \frac{dw_H}{dt} &= (\lambda + r) w_H - \lambda w_L, \end{aligned}$$

and (w_L, w_H) will hit (x_L, x_H) at time t^*. Therefore I solve the differential equations together with the boundary condition and obtain

$$w_L^{3t^*} = -\bar{v} \left[\frac{\theta_L - \theta_H}{2(2\lambda + r)} e^{-(2\lambda+r)t^*} + \frac{\theta_L + \theta_H}{2r} e^{-rt^*} \right], \tag{12}$$

$$w_H^{3t^*} = -\bar{v} \left[\frac{\theta_H - \theta_L}{2(2\lambda + r)} e^{-(2\lambda+r)t^*} + \frac{\theta_L + \theta_H}{2r} e^{-rt^*} \right]. \tag{13}$$

Similarly, I obtain

$$w_L^{4t^*} = \bar{v}\left[\frac{\theta_L - \theta_H}{2(2\lambda + r)} e^{-(2\lambda+r)t^*} + \frac{\theta_L + \theta_H}{2r} e^{-rt^*}\right] + x_L + \bar{u}/r, \tag{14}$$

$$w_H^{4t^*} = \bar{v}\left[\frac{\theta_H - \theta_L}{2(2\lambda + r)} e^{-(2\lambda+r)t^*} + \frac{\theta_L + \theta_H}{2r} e^{-rt^*}\right] + x_H + \bar{u}/r. \tag{15}$$

It turns out that the utility pairs delivered by $(\mathscr{C}^{1c^*}, \mathscr{C}^{2c^*}, \mathscr{C}^{3t^*}, \mathscr{C}^{4t^*})$ form the boundary of W (see Figure 1).

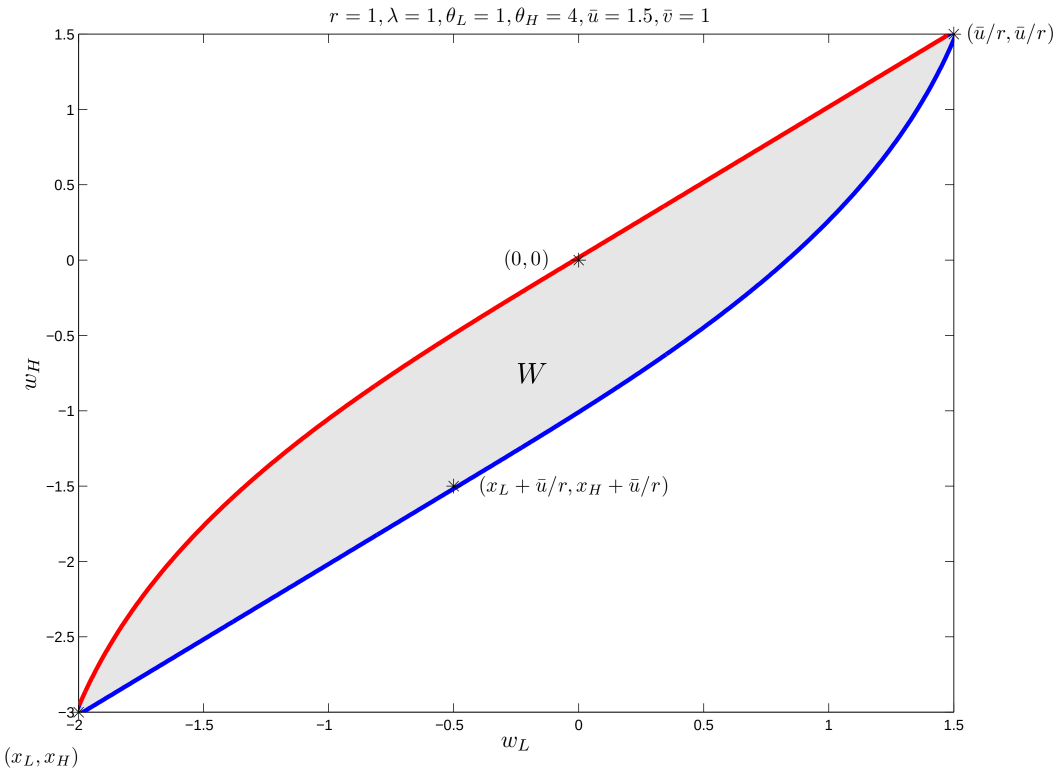

Figure 1: Set of implementable utility pairs.

Theorem 2 *The boundaries of W consist of four vertices $((\bar{u}/r, \bar{u}/r)$, (x_L, x_H), $(0,0)$, $(x_L + \bar{u}/r, x_H + \bar{u}/r))$ and four pieces of curves that connect these vertices. The upper boundary is specified in (10,12,13), while the lower boundary is specified in (11,14,15).*

Proof. I verify two things. First any point between the two boundaries can be implemented by some contract. Second, any point either above the upper boundary or below the lower boundary can not be implemented by any contract.

From the above argument, any point on the upper or lower boundary can be implemented by definition. To implement a point (w_L, w_H) between the two boundaries, the principal may start with the policy $(c_t, y_t) = (0, 0)$ and let the promised utilities evolve according to the following law of motion until time s^*,

$$\frac{dw_L}{dt} = (\lambda + r)w_L - \lambda w_H$$
$$\frac{dw_H}{dt} = (\lambda + r)w_H - \lambda w_L,$$

where s^* is the time at which the path hits some point (w_L^*, w_H^*) on the boundary. Then starting from s^*, the principal implements (w_L^*, w_H^*) using the contracts discussed previously.

Second, I will show that any point below the lower boundary can not be implemented by any contract. The proof for points above the upper boundary is analogous.

Let function $g : [x_L, \bar{u}/r] \mapsto [x_H, \bar{u}/r]$ be the lower boundary of W. Pick a point (w_L, w_H) with $w_H < g(w_L)$, and a contract $\mathscr{C}$. I will prove that the continuation utility will eventually be impossible to implement under the history of reporting type L for a long time. To see this, let us calculate the distance between the lower boundary and the continuation utility $(w_L(h,t), w_H(h,t))$ under contract $\mathscr{C}$, where $h^{t-}(s) = L, \forall s \in [0,t)$(see Figure 2). Recall

$$\begin{aligned} \frac{dw_L(h,t)}{dt} &= (\lambda + r)w_L(h,t) - \lambda w_H(h,t) - u(c_t(h,t,L)) + \theta_L v(y_t(h,t,L)), \\ \frac{dw_H(h,t)}{dt} &\leq (\lambda + r)w_H(h,t) - \lambda w_L(h,t) - u(c_t(h,t,L)) + \theta_H v(y_t(h,t,L)). \end{aligned}$$

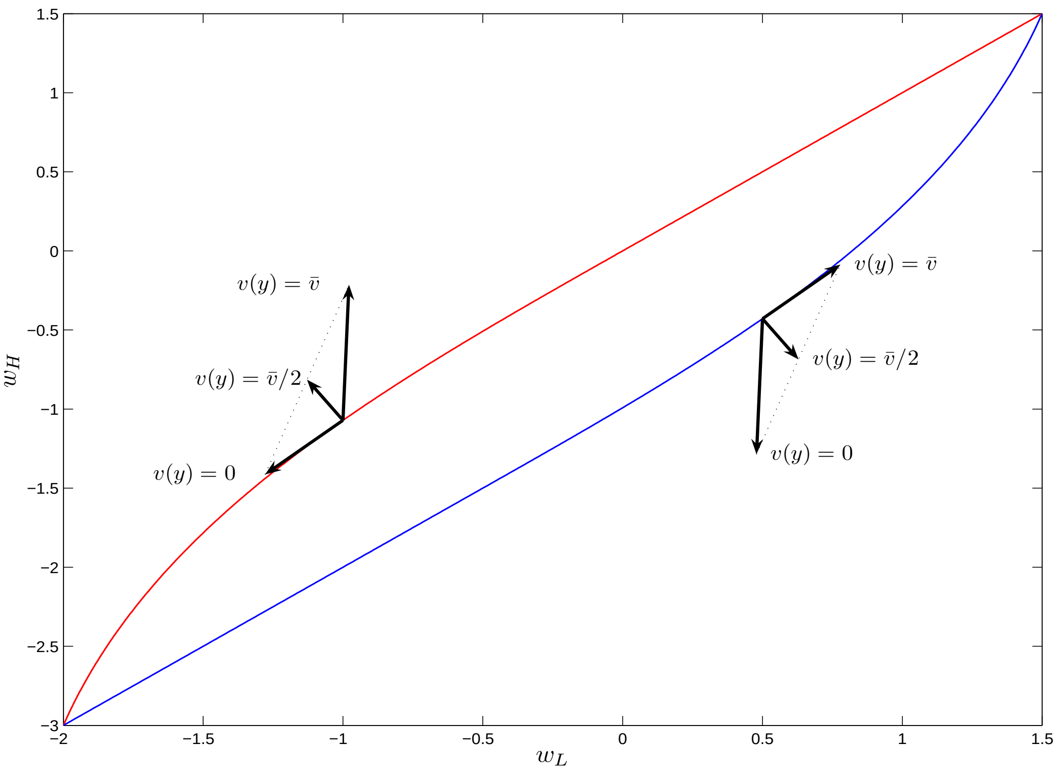

Figure 2: Law of motion on the lower and upper boundaries.

The distance between $w_H(h,t)$ and $g(w_L(h,t))$ satisfies

$$
\begin{aligned}
\frac{d(g(w_L(h,t)) - w_H(h,t))}{dt} \geq\; & \frac{dg(w_L(h,t))}{dw_L}((\lambda + r)w_L(h,t) - \lambda w_H(h,t) - \\
& u(c_t(h,t,L)) + \theta_L v(y_t(h,t,L))) - \\
& ((\lambda + r)w_H(h,t) - \lambda w_L(h,t) - \\
& u(c_t(h,t,L)) + \theta_H v(y_t(h,t,L))) \\
=\; & \frac{dg(w_L(h,t))}{dw_L}((\lambda + r)w_L(h,t) - \lambda w_H(h,t) - \bar{u} + \theta_L \bar{v}) - \\
& ((\lambda + r)w_H(h,t) - \lambda w_L(h,t) - \bar{u} + \theta_H \bar{v}) + \\
& (\theta_H - \frac{dg(w_L(h,t))}{dw_L}\theta_L)(\bar{v} - v(y_t)) +
\end{aligned}
$$

$$\frac{dg(w_L(h,t))}{dw_L}(\bar{u} - u(c_t(h,t,L))).$$

Since $1 \leq \frac{dg(w_L(h,t))}{dw_L} \leq \frac{\theta_H}{\theta_L}$ and $w_H(h,t) \leq g(w_L(h,t))$,

$$\begin{aligned}
\frac{d(g(w_L(h,t)) - w_H(h,t))}{dt} &\geq \frac{dg(w_L(h,t))}{dw_L}((\lambda+r)w_L(h,t) - \lambda g(w_L(h,t)) - \bar{u} + \theta_L \bar{v}) - \\
&\quad ((\lambda+r)w_H(h,t) - \lambda w_L(h,t) - \bar{u} + \theta_H \bar{v}) \\
&= ((\lambda+r)g(w_L(h,t)) - \lambda w_L(h,t) - \bar{u} + \theta_H \bar{v}) - \\
&\quad ((\lambda+r)w_H(h,t) - \lambda w_L(h,t) - \bar{u} + \theta_H \bar{v}) \\
&= (\lambda+r)(g(w_L(h,t)) - w_H(h,t)).
\end{aligned}$$

Therefore, the distance is increasing exponentially and in finite time, $w_H(h,t)$ will be less than x_H. This is a contradiction because the worst scenario for the H type agent is "consumption 0 and maximal output $\bar{y}$", which provides utility x_H. ■

5. Dynamics of the Optimal Contract

In the last section, I characterized the set of all implementable utility pairs. Now I study the optimal I.C. contract to implement each $(w_L, w_H) \in W$. Lemma 1 gives some elementary properties of the value functions.

Lemma 1 *Value functions V_L, V_H have the following properties.*

(1) $V_L(x_L, x_H) = V_H(x_L, x_H) = -\bar{y}/r, V_L(\bar{u}/r, \bar{u}/r) = V_H(\bar{u}/r, \bar{u}/r) = \bar{c}/r.$

(2) V_L, V_H *are convex.*

(3) If $w_L > x_L, w_H > x_H$, then[5]

$$V_{Lw_L}(w_L - x_L) + V_{Lw_H}(w_H - x_H) > 0,$$
$$V_{Hw_L}(w_L - x_L) + V_{Hw_H}(w_H - x_H) > 0.$$

(4) $V_{Lw_H} \leq 0, V_{Lw_L} > 0, V_{Hw_L} \leq 0, V_{Hw_H} > 0$.

(5) $V_{Lw_Lw_H} \leq 0, V_{Hw_Lw_H} \leq 0$.

Part (4) in the above lemma says that value functions are monotonic, increasing in the promised utility but decreasing in the threat utility. The monotonicity in the promised utility is straightforward because the principal needs to give more consumption(and less output) if she promises more utility to the agent. The intuition for $V_{Lw_H} \leq 0$ is as follows. Think of w_H as the threat utility that the principal imposes on the cheater. Imposing harsher punishment necessarily puts more restrictions on the contracting problem and thus increases the cost.

For any $w_L \in [x_L, \bar{u}/r]$, let $g(w_L), h(w_L)$ denote the lower and upper boundary of W, respectively.

Lemma 2 *For any $w_L \in (x_L, \bar{u}/r)$,*

(1) $V_{Lw_H}(w_L, g(w_L)) < 0$.

(2) $V_{Hw_L}(w_L, h(w_L)) = -\infty$.

(3) $V_{Lw_H}(w_L, h(w_L)) = 0, V_{Hw_L}(w_L, g(w_L)) = 0$.

[5] In this paper, I use $V_{Lw_L}, V_{Lw_H}, V_{Lw_Lw_L}, V_{Lw_Lw_H}, V_{Lw_Hw_H}$ to denote $\frac{\partial V_L}{\partial w_L}$, $\frac{\partial V_L}{\partial w_H}$, $\frac{\partial^2 V_L}{\partial w_L \partial w_L}$, $\frac{\partial^2 V_L}{\partial w_L \partial w_H}$, $\frac{\partial^2 V_L}{\partial w_H \partial w_H}$, respectively. Similar notations are used for V_H.

Based on LEMMA 2, I can define two curves f_L, f_H,

$$f_L(w_L) = \min\{w_H : V_{Lw_H}(w_L, w_H) = 0, (w_L, w_H) \in W\},$$
$$f_H(w_H) = \min\{w_L : V_{Hw_L}(w_L, w_H) = 0, (w_L, w_H) \in W\}.$$

Since $V_{Lw_Hw_H} \geq 0, V_{Lw_Lw_H} \leq 0, V_{Hw_Lw_L} \geq 0, V_{Hw_Lw_H} \leq 0$, it is easy to see that these curves are increasing. I will call these curves the *efficiency* curves, because for each level of the promised utility, they indicate the optimal level of the threat utility to minimize the cost. For example, if initially the principal holds belief $p_L = 1$ and wants to deliver utility w_L to the agent, the optimal contract for the principal is to start with $(w_L, f_L(w_L))$. These curves are also critical for our study of the dynamics, because starting from certain initial conditions, the state variable will jump onto the efficiency curves under the optimal contract.

LEMMA 3 *Starting from $(w_L, w_H) \in W$ with $w_H > f_L(w_L)$, the optimal contract satisfies*

$$\lim_{t\downarrow 0} w_H(0, L, t) = f_L(w_L).$$

Similarly, starting from $(w_L, w_H) \in W$ with $w_L > f_H(w_H)$,

$$\lim_{t\downarrow 0} w_L(0, H, t) = f_H(w_H).$$

Proof. By contradiction, if $\lim_{t\downarrow 0} w_H(0, L, t) < f_L(w_L)$, then

$$V_L(w_L, w_H) = \lim_{t\downarrow 0} V_L(w_L(0, L, t), w_H(0, L, t)) > V_L(w_L, f_L(w_L)).$$

If $\lim_{t\downarrow 0} w_H(0, L, t) > f_L(w_L)$, then the contracts starting from (w_L, w_H) and $(w_L, f_L(w_L))$ are not equal to each other almost surely. Convex combination can be used to lower the cost. Thus $\lim_{t\downarrow 0} w_H(0, L, t) = f_L(w_L)$. ■

Now it is ready to lay out the *Hamilton-Jacobi-Bellman(HJB)* equations that value functions satisfy. For any (w_L, w_H) with $w_H \leq f_L(w_L)$, V_L satisfies

$$\begin{aligned}(\lambda+r)V_L(w_L,w_H) &= \min_c\{c-(V_{Lw_L}+V_{Lw_H})u(c)\}+\min_y\{-y+(\theta_L V_{Lw_L}+\theta_H V_{Lw_H})v(y)\}\\ &\quad +\lambda V_H(w_L,w_H)+V_{Lw_L}((\lambda+r)w_L-\lambda w_H)\\ &\quad +\min_{\mu\geq 0}\{V_{Lw_H}((\lambda+r)w_H-\lambda w_L-\mu)\}.\end{aligned}$$

Similarly, for (w_L, w_H) with $w_L \leq f_H(w_H)$,

$$\begin{aligned}(\lambda+r)V_H(w_L,w_H) &= \min_c\{c-(V_{Hw_L}+V_{Hw_H})u(c)\}+\min_y\{-y+(\theta_L V_{Hw_L}+\theta_H V_{Hw_H})v(y)\}\\ &\quad +\lambda V_L(w_L,w_H)+\min_{\mu\geq 0}\{V_{Hw_L}((\lambda+r)w_L-\lambda w_H-\mu)\}\\ &\quad +V_{Hw_H}((\lambda+r)w_H-\lambda w_L).\end{aligned}$$

In addition, notice that μ can be non-zero only if $V_{Lw_H}=0$ or $V_{Hw_L}=0$. Therefore, if $w_H < f_L(w_L)$, I can rewrite the HJB equation as

$$\begin{aligned}(\lambda+r)V_L(w_L,w_H) &= \min_c\{c-(V_{Lw_L}+V_{Lw_H})u(c)\}+\min_y\{-y+(\theta_L V_{Lw_L}+\theta_H V_{Lw_H})v(y)\}\\ &\quad +\lambda V_H(w_L,w_H)+V_{Lw_L}((\lambda+r)w_L-\lambda w_H)\\ &\quad +V_{Lw_H}((\lambda+r)w_H-\lambda w_L).\end{aligned}\tag{16}$$

Totally differentiating (16) with respect to w_L, w_H, I obtain

$$\begin{aligned}0 &= V_{Lw_Lw_L}((\lambda+r)w_L-\lambda w_H-u(c)+\theta_L v(y))+\\ &\quad V_{Lw_Lw_H}((\lambda+r)w_H-\lambda w_L-u(c)+\theta_H v(y))+\lambda V_{Hw_L}-\lambda V_{Lw_H},\\ 0 &= V_{Lw_Hw_L}((\lambda+r)w_L-\lambda w_H-u(c)+\theta_L v(y))+\\ &\quad V_{Lw_Hw_H}((\lambda+r)w_H-\lambda w_L-u(c)+\theta_H v(y))+\lambda V_{Hw_H}-\lambda V_{Lw_L}.\end{aligned}$$

The above equations together with $dw_L = ((\lambda+r)w_L-\lambda w_H-u(c)+\theta_L v(y))dt$, $dw_H = ((\lambda+r)w_H-\lambda w_L-u(c)+\theta_H v(y))dt$ constitute an ODE system to describe the dynamics under report L.

$$\begin{aligned}
\frac{dw_L}{dt} &= (\lambda + r)w_L - \lambda w_H - u(c) + \theta_L v(y) \\
\frac{dw_H}{dt} &= (\lambda + r)w_H - \lambda w_L - u(c) + \theta_H v(y) \\
\frac{dV_{Lw_L}}{dt} &= \lambda V_{Lw_H} - \lambda V_{Hw_L} \\
\frac{dV_{Lw_H}}{dt} &= \lambda V_{Lw_L} - \lambda V_{Hw_H}.
\end{aligned}$$

Similarly, the ODE system under report H is

$$\begin{aligned}
\frac{dw_L}{dt} &= (\lambda + r)w_L - \lambda w_H - u(c) + \theta_L v(y) \\
\frac{dw_H}{dt} &= (\lambda + r)w_H - \lambda w_L - u(c) + \theta_H v(y) \\
\frac{dV_{Hw_L}}{dt} &= \lambda V_{Hw_H} - \lambda V_{Lw_L} \\
\frac{dV_{Hw_H}}{dt} &= \lambda V_{Hw_L} - \lambda V_{Lw_H}.
\end{aligned}$$

LEMMA 4 *For* $w_L \in (x_L, \bar{u}/r), w_H \in (x_H, \bar{u}/r)$,

$$V_{Lw_L}(w_L, f_L(w_L)) \leq V_{Hw_H}(w_L, f_L(w_L)),$$

$$V_{Lw_L}(f_H(w_H), w_H) \geq V_{Hw_H}(f_H(w_H), w_H).$$

LEMMA 5 *Curve* f_L *is strictly above* f_H*. That is, for each* $w_L \in (x_L, \bar{u}/r)$,

$$f_L(w_L) > f_H^{-1}(w_L).$$

LEMMA 6

$$V_{Lw_L}(w_L, f_L(w_L)) = V_{Hw_H}(w_L, f_L(w_L)),$$

$$V_{Lw_L}(f_H(w_H), w_H) > V_{Hw_H}(f_H(w_H), w_H).$$

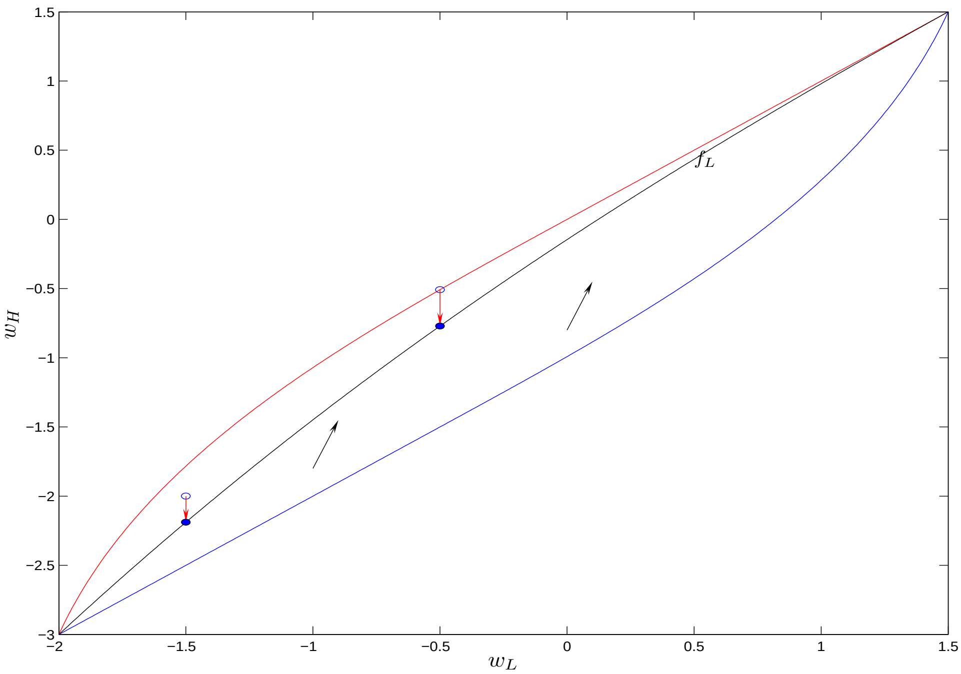

Figure 3: Dynamics under report L.

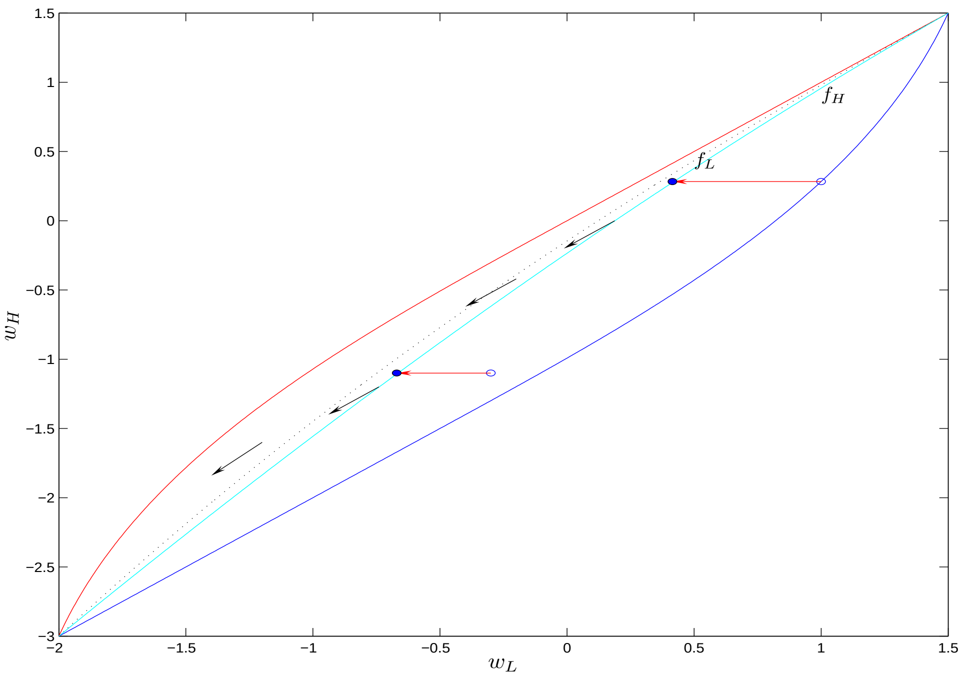

Figure 4: Dynamics under report H.

LEMMA 7 *Under report L, time paths starting from f_L will remain on f_L(see Figure 3) and move toward $(\bar{u}/r, \bar{u}/r)$. Under report H, time paths starting from f_H will move above f_H(see Figure 4).*

Proof. $V_{Lw_L} = V_{Hw_H}$ on f_L implies that $V_{Lw_L} > V_{Hw_H}$ below f_L. Under report L and starting from f_L, the time path will not leave f_L, because if it moves below f_L, V_{Lw_H} will be positive along the time path, which is a contradiction to the fact that $V_{Lw_H} \leq 0$. Furthermore, we know that

$$\frac{dw_L}{dt} = (\lambda + r)w_L - \lambda f_L(w_L) - u(c) + \theta_L v(y) > 0.$$

Otherwise $\frac{dw_L}{dt} \leq 0$ implies $\frac{dV_{Lw_L}}{dt} \leq 0$, which is a contradiction to

$$\frac{dV_{Lw_L}}{dt} = \lambda(V_{Lw_H} - V_{Hw_L}) > 0.$$

Under report H and starting from the curve f_H, since

$$\frac{dV_{Hw_L}}{dt} = \lambda(V_{Hw_H} - V_{Lw_L}) < 0,$$

we know that the time path will leave f_H and move above f_H. ■

Intuitively, as long as the agent claims that her type is H and productivity is low, the contract specifies low level of output, but in order to prevent a high productivity agent from lying, it necessarily lowers the promised utility of a potential liar. On one hand, this keeps incentive-compatibility, on the other hand, maintaining a low threat incurs additional cost besides that of providing promised utility to the truth-teller.

The dynamics of the time path under high report also implies that the optimal contract under commitment is no longer renegotiation-proof in the environment with persistent shocks. To see this, suppose the contract starts from the curve f_H, and the agent experiences a period of high shocks, then the time path moves to the left of f_H, which is

$w_L < f_H(w_H)$ and $V_H(w_L, w_H) > V_H(f_H(w_H), w_H)$. Should the principal had the chance to renegotiate with the agent, she would be willing to move the state from (w_L, w_H) to $(f_H(w_H), w_H)$, and doing this makes the agent indifferent, and lowers the principal's cost. However it violates the ex-ante incentive constraints that prevent the type L agent from lying.

In order to further investigate the dynamics of the optimal contract, I need to make certain assumptions about patterns of the dynamics. These patterns are what I observe in my numerical approximations, but at this moment I am not able to provide mathematical proofs for them.

Assumption 1 Under report H, time paths move southwest.

Assumption 2 On f_L, V_{Hw_L} is monotonic.

LEMMA 8 *Under report H, each time path will converge to (x_L, x_H). On each time path,*

$$\lim_{(w_L,w_H)\to(x_L,x_H)} [V_{Hw_L} + V_{Hw_H}] \leq 0,$$
$$\lim_{(w_L,w_H)\to(x_L,x_H)} [\theta_L V_{Hw_L} + \theta_H V_{Hw_H}] = 0.$$

LEMMA 9 *The region between the upper boundary and f_L is absorbing. More precisely, starting from f_L, under report H, (w_L, w_H) enters the interior of the region.*

The dynamics under the L report is relatively simple. All the points (w_L, w_H) approach f_L, and remain on the curve once they reach f_L. So our focus is on the dynamics under the H report.

THEOREM 3 *(1) The dynamics under the H report is described by a function f_L and an*

ODE system.

$$\begin{aligned}
\frac{dw_L}{dt} &= (\lambda + r)w_L - \lambda w_H - u(c) + \theta_L v(y) \\
\frac{dw_H}{dt} &= (\lambda + r)w_H - \lambda w_L - u(c) + \theta_H v(y) \\
\frac{dV_{Hw_L}}{dt} &= \lambda V_{Hw_H} - \lambda V_{Lw_L}(w_L, f_L(w_L)) \\
\frac{dV_{Hw_H}}{dt} &= \lambda V_{Hw_L},
\end{aligned}$$

where $c \in argmin_c\{c-(V_{Hw_L}+V_{Hw_H})u(c)\}$, $y \in argmin_y\{(\theta_L V_{Hw_L}+\theta_H V_{Hw_H})v(y)-y\}$.

(2) *There exist time paths for which*

$$\lim_{(w_L,w_H)\to(x_L,x_H)} V_{Hw_L} < 0.$$

(3) *(Immiserization) Each time path in Part (2) will hit* (x_L, x_H) *in finite time. However, although the time path under report L moves toward* $(\bar{u}/r, \bar{u}/r)$ *along the curve* f_L*, it never reaches* $(\bar{u}/r, \bar{u}/r)$ *in finite time. Asymptotically, all the time paths will be absorbed into* (x_L, x_H) *and remain there forever.*

Proof.

(1) Notice that in the region between upper boundary and f_L,

$$V_{Lw_H}(w_L, w_H) = 0, V_{Lw_L}(w_L, w_H) = V_{Lw_L}(w_L, f_L(w_L)).$$

(2) By contradiction, assume that for all time paths,

$$\lim_{(w_L,w_H)\to(x_L,x_H)} V_{Hw_L} = \lim_{(w_L,w_H)\to(x_L,x_H)} V_{Hw_H} = 0.$$

Then we can locally linearize the dynamic system. Since the system converges to the point $(x_L, x_H, 0, 0)$ in $\mathbb{R}^4$, any point $(w_L, w_H, V_{Hw_L}, V_{Hw_H})$ on the time path is on the stable manifold of the linear system. Continuity implies that (V_{Hw_L}, V_{Hw_H}) is close to $(0,0)$, when (w_L, w_H) is close to (x_L, x_H). However, we know that at points close to the upper boundary, (V_{Hw_L}, V_{Hw_H}) is close to $-\infty$ and $+\infty$ respectively. Therefore, it is a contradiction.

(3) First recall $\frac{dV_{Hw_H}}{dt} = \lambda V_{Hw_L}$. If $\lim_{(w_L,w_H)\to(x_L,x_H)} V_{Hw_L} < 0$, the time path has to meet (x_L, x_H) in finite time, otherwise, V_{Hw_H} will decrease below 0 in finite time, which is contradiction to the non-negativity of V_{Hw_H}. Now I show that under the L report, (w_L, w_H) can never hit $(\bar{u}/r, \bar{u}/r)$ in finite time. Recall

$$\frac{dV_{Hw_H}}{dt} = \frac{dV_{Lw_L}}{dt} = -\lambda V_{Hw_L} \leq \lambda \frac{\theta_H}{\theta_L} V_{Hw_H}.$$

Therefore, by Gronwall's inequality, in finite time, V_{Hw_H} remains finite. This means that (w_L, w_H) can not be $(\bar{u}/r, \bar{u}/r)$ in finite time.

■

6. Parameterization

In this section, I numerically solve the model with hidden productivity shocks. First I choose the parameters so that they match observed empirical facts. Then I artificially change the productivity process to the one with i.i.d. shocks, and keep all the other elements of the model fixed. I shall make comparisons between the implications of the persistent-shock model and the i.i.d.-shock model.

I assume that the agent's preferences are

$$E\left[\int_0^\infty e^{-rt}\left(\frac{c_t^{1-\sigma}}{1-\sigma} - \kappa\frac{l_t^{1+\gamma}}{1+\gamma}\right)dt\right], \tag{17}$$

where c_t and l_t are consumption and labor supply respectively. When the agent with productivity ϕ_t works l_t units of time, she produces $y_t = \phi_t l_t$ units of output. The output

Parameter	Value
r	0.0408
σ	1.4610
γ	2.0
κ	1.1840
$\underline{U}$	-79.6110
ϕ_L	0.6425
ϕ_H	1.9495
λ	0.0249

Table 1: Parameters

y_t is observable, while the productivity ϕ_t and labor l_t are the agent's private information. I may rewrite the agent's preference as

$$E\left[\int_0^\infty e^{-rt}\left(\frac{c_t^{1-\sigma}}{1-\sigma}-\kappa\phi_t^{-(1+\gamma)}\frac{y_t^{1+\gamma}}{1+\gamma}\right)dt\right]. \tag{18}$$

I set r to 0.0408 to match an annual discount factor of 0.96. I follow Albanesi and Sleet[2] in setting σ, κ to be 1.461 and 1.1840 respectively, and follow Chari, Kehoe and Mcgrattan[5] in setting γ to be 2. This implies that the elasticity of labor supply is 0.5. The lower bound $\underline{U} = -79.6110$ is set in a way similar to that in Albanesi and Sleet[2]. I choose parameter values for ϕ_L, ϕ_H, λ to match the unconditional mean, unconditional variance and the covariance of the skill process described in Golosov and Tsyvinski[9]. This implies a value for λ to be 0.0249. The productivity process is highly persistent, which is the driving force of the pattern of wedges shown below. All parameter values are listed in Table 1.

6.1 The Wedges

I first define several wedges discussed in Albanesi and Sleet[2].

(1) THE INSURANCE WEDGE: For a type $i(i = L, H)$ agent with history h^{t-} at t, it is defined as

$$\frac{u'(c(h,t,i,t+1,L))}{u'(c(h,t,i,t+1,H))} - 1.$$

This measures the consumption smoothing implied by the optimal contract. The larger the wedge, the worse the insurance.

(2) THE CONSUMPTION-LEISURE WEDGE: For a type $i(i = L, H)$ agent with history h^{t-} at t, it is defined as

$$\frac{u'(c(h,t,i))}{v'(y(h,t,i)/\phi_i)/\phi_i} - 1.$$

This measures the ratio of marginal utility to marginal disutility for the type i agent at time t.

(3) THE INTERTEMPORAL WEDGE: For a type $i(i = L, H)$ agent with history h^{t-} at t, it is defined as

$$\frac{E_t[u'(c(h,t,i,t+1))]}{u'(c(h,t,i))} - 1.$$

The expectation is over the uncertain types of the agent at $t + 1$. Golosov, Kocherlakota and Tsyvinski[8] showed that this wedge is positive when the state variable $(w_L(h,t), w_H(h,t))$ is above the lower bound $\underline{U}$.

The wedges defined above measure the degree of insurance from different dimensions. It is easy to see that in the optimal allocation without information frictions, all the wedges should be 0.

	Persistent	I.I.D.
Insurance wedge	(0.5851, 0.3497)	$(0.0171, 0.0151)$
C-L wedge	(0.2689, -0.00015)	$(0.00744, 0)$
Intertemporal wedge	(0.000107, 0.0032)	$(-0.0000723, 0.000356)$

Table 2: Wedges

6.2 Numerical Results

I follow the same procedure for the persistent-shock model and the i.i.d.-shock model. I first solve the contracting problem and derive the optimal policy function $c_{it}(w_L, w_H)$, $y_{it}(w_L, w_H)$, $i = L, H$. Then I simulate the model and compute the stationary distribution of promised utilities for each type. I calculate the three wedges for each type with different levels of promised utilities, and report the average of the wedges in the stationary distribution in Table 2.

A striking feature of the results from the i.i.d. shock model is that all the wedges are close to 0. This implies that the allocation under the i.i.d. shocks is close to the first-best allocation. The intuition for the results is as follows. Given that the discount factor e^{-r} is close to 1, the agent cares about her utility as the long-run average. If the shocks are i.i.d. and thus transitory, the effect of any productivity shock at t is small and will be smoothed into many periods in the future. If the agent has a bad shock, the principal will still provide the consumption level close to that of the high-productivity agent, but will lower the discounted utility from $t + 1$ on. In the long run(by law of large numbers), the effects of high and low productivity shocks cancel out, and the agent does not experience large deviations from the first-best allocation. The intertemporal taxation and subsidy play an essential role in the optimal contract to smooth consumption.

The patterns of wedges with persistent shocks are significantly different from the i.i.d. model. We see that the insurance wedge is more than 10 times bigger than the wedge in i.i.d case, implying that the consumption smoothing is far from being perfect. The consumption-leisure wedge is also quantitatively large, meaning that the low-productivity agent is distorted in her labor-consumption decision. Despite this, the persistent shock model does not imply a large intertemporal wedge. In order to understand these patterns, it would be helpful to consider the permanent-shock model, which is the opposite extreme of the i.i.d. shocks. Suppose that initially the agent has a permanent productivity shock that is only privately observed. Then the optimal allocation is to specify a type-specific but constant stream of consumption and output for each type. It is well known that the optimal allocation implies distortion for the low productivity agent. And intertemporal wedge is 0 simply because that the consumption process is constant. Our results show that the pattern of wedges with persistent shocks is similar to a permanent-shock model. Quantitatively, this is driven by the low value of λ. The productivity process is so persistent that it is almost permanent.

7. Conclusion

In this paper, I develop continuous-time methods for solving repeated principal-agent problem with persistent shocks. I first simplify the incentive constraints in the discrete time model into a system of differential equations and then use the differential equations to infer the qualitative and quantitative properties of the optimal contract. My main result is that the optimal contract still implies immiserization, but different from the i.i.d. case, the contract is no longer renegotiation-proof. I demonstrate my method in the context of dynamic taxation, but the same technique is applicable to a large variety of problems with asymmetric information.

My numerical results from a calibrated version of the model imply that the high persistence of the hidden information has significant effects on the properties of the optimal contract. In the context of taxation, high persistence of productivity shocks implies much larger distortion than i.i.d. shocks.

I conclude by briefly mentioning one important direction for future research. I have assumed that the agent's type can take only two values. I would like to extend the method to any finite number of types. Making this extension will leave us more freedom when we approximate a stochastic process with finite states, thus it would be necessary before we do a more careful quantitative work. With more than two types, the differential equations I derived before are still available, but the exact dynamics of the optimal contract are hard to analyze. Future research should explore these questions.

Appendix

Proof of Lemma 1

I only give a proof for V_L, since the proof for V_H is the same.

(1) In order to implement (x_L, x_H), the principal uses the contract in which the agent does not consume anything and produces $\bar{y}$ under all possible reports. This is clearly the contract that minimizes cost. In order to implement $(\bar{u}/r, \bar{u}/r)$, the principal has to specify consumption $\bar{c}$ and output 0 almost surely, therefore the cost is $\bar{c}/r$.

(2) This follows from the fact that the contracting problem has a linear objective function and a convex constraint set. For $\lambda \in [0, 1]$, pick two implementable pairs $(w_L^1, w_H^1), (w_L^2, w_H^2)$, suppose $\mathscr{C}^1, \mathscr{C}^2$ implement (w_L^1, w_H^1) and (w_L^2, w_H^2) respectively.

Design a new contract $\mathscr{C}$ by

$$\begin{aligned} c_t(h^t) &= u^{-1}(\lambda u(c_t^1(h^t)) + (1-\lambda)u(c_t^2(h^t))), \\ y_t(h^t) &= v^{-1}(\lambda v(y_t^1(h^t)) + (1-\lambda)v(y_t^2(h^t))), \forall t > 0, h^t \in \Theta^t. \end{aligned}$$

This contract will be I.C.

(3) Since V_L is a convex function,

$$V_{Lw_L}(w_L - x_L) + V_{Lw_H}(w_H - x_H) \geq V_L(w_L, w_H) - V_L(x_L, x_H) > 0.$$

(4) I show $V_{Lw_H} \leq 0$ first. Pick $(w_L, w_H), (w_L, w'_H)$ with $w_H < w'_H$. Let $\mathscr{C}$ be the optimal contract implementing (w_L, w_H), and let $w_i(0, j)$ be the continuation utility of the type i agent after she reports to be type j at time 0.

$$\begin{aligned} w_L = w_L(0, L) \geq w_L(0, H), \\ w_H = w_H(0, H) \geq w_H(0, L). \end{aligned}$$

I shall design a new contract $\mathscr{C}'$ to implement (w_L, w'_H), but with the same cost as $\mathscr{C}$. First recall that I can always implement (w_L, w'_H) by some simple contract in which I use $(c_t = 0, y_t = 0)$ initially, and follow the contract on boundaries starting from the time s^* when the continuation utility hits the boundary. Notice that this contract does not depend on the report, but only on time. I denote this simple contract by $\hat{\mathscr{C}}$. Now define $\mathscr{C}'$ by

$$(c'_t(h^t), y'_t(h^t)) = \begin{cases} (c_t(h^t), y_t(h^t)), & \text{if } h^t(0) = L, \\ (\hat{c}_t(h^t), \hat{y}_t(h^t)), & \text{if } h^t(0) = H. \end{cases}$$

Under contract $\mathscr{C}'$,

$$\begin{aligned} w_L = w_L(0, L) = w_L(0, H), \\ w'_H = w_H(0, H) \geq w_H(0, L). \end{aligned}$$

Thus $\mathscr{C}'$ is I.C. Notice that with truth-telling, and under belief $p_L = 1$, $\mathscr{C}'$ equals $\mathscr{C}$ almost surely, therefore they have the same cost.

To see that $V_{Lw_L} > 0$, using (3), I obtain

$$V_{Lw_L} > -\frac{w_H - x_H}{w_L - x_L} V_{Lw_H} \geq 0.$$

(5) It is equivalent to show that, for (w_L, w_H), and $h > 0$,

$$V_L(w_L, w_H) + V_L(w_L + h, w_H + h) \leq V_L(w_L, w_H + h) + V_L(w_L + h, w_H).$$

We only need to prove these facts for the value functions in the discrete-time model, since the value functions in discrete-time will converge to V_L, V_H and these properties still hold after we take the limit. In discrete-time model, the value function is the unique fixed point of the operator T, which maps the cost function (V_L^{t+dt}, V_H^{t+dt}) into (V_L^t, V_H^t).

We proceed as follows. First assume that (V_L^{t+dt}, V_H^{t+dt}) has the above property, we will show that (V_L^t, V_H^t) also has the property. To see this, let (c_t^{12}, y_t^{12}) and (c_t^{21}, y_t^{21}) be the optimal policy at $(w_L, w_H + h)$ and $(w_L + h, w_H)$ respectively. Then

$$\begin{aligned}
V_L^t(w_L, w_H + h) &= (c_t^{12} - y_t^{12})dt + e^{-rdt}[(1 - \lambda dt)V_L^{t+dt}(w'_{L12}, w'_{H12}) + \\
&\quad \lambda dt V_H^{t+dt}(w'_{L12}, w'_{H12})], \\
V_L^t(w_L + h, w_H) &= (c_t^{21} - y_t^{21})dt + e^{-rdt}[(1 - \lambda dt)V_L^{t+dt}(w'_{L21}, w'_{H21}) + \\
&\quad \lambda dt V_H^{t+dt}(w'_{L21}, w'_{H21})], \\
\text{where } w'_{L12} &= w_L + dt((\lambda + r)w_L - \lambda(w_H + h) - u(c_t^{12}) + \theta_L v(y_t^{12})), \\
w'_{H12} &= (w_H + h) + dt((\lambda + r)(w_H + h) - \lambda w_L - u(c_t^{12}) + \theta_H v(y_t^{12})), \\
w'_{L21} &= (w_L + h) + dt((\lambda + r)(w_L + h) - \lambda w_H - u(c_t^{21}) + \theta_L v(y_t^{21})), \\
w'_{H21} &= w_H + dt((\lambda + r)w_H - \lambda(w_L + h) - u(c_t^{21}) + \theta_H v(y_t^{21})).
\end{aligned}$$

Since V_L^{t+dt}, V_H^{t+dt} both have the property, if $-u(c_t^{21})+\theta_L v(y_t^{21}) \geq -u(c_t^{12})+\theta_L v(y_t^{12})$,

$$
\begin{aligned}
& V_L^{t+dt}(w'_{L12}, w'_{H12}) + V_L^{t+dt}(w'_{L21}, w'_{H21}) \\
\geq\ & V_L^{t+dt}(w'_{L12} + \lambda hdt, w'_{H12} - \lambda hdt) + V_L^{t+dt}(w'_{L21} - \lambda hdt, w'_{H21} + \lambda hdt) \\
\geq\ & V_L^{t+dt}(w'_{L12} + \lambda hdt + (-u(c_t^{21}) + \theta_L v(y_t^{21}) + u(c_t^{12}) - \theta_L v(y_t^{12}))dt, w'_{H12} - \lambda hdt) + \\
& V_L^{t+dt}(w'_{L21} - \lambda hdt - (-u(c_t^{21}) + \theta_L v(y_t^{21}) + u(c_t^{12}) - \theta_L v(y_t^{12}))dt, w'_{H21} + \lambda hdt) \\
\geq\ & V_L^{t+dt}(w'_{L12} + \lambda hdt + (-u(c_t^{21}) + \theta_L v(y_t^{21}) + u(c_t^{12}) - \theta_L v(y_t^{12}))dt, w'_{H21} + \lambda hdt) + \\
& V_L^{t+dt}(w'_{L21} - \lambda hdt - (-u(c_t^{21}) + \theta_L v(y_t^{21}) + u(c_t^{12}) - \theta_L v(y_t^{12}))dt, w'_{H12} - \lambda hdt) \\
=\ & V_L^{t+dt}(w_L + dt((\lambda + r)w_L - \lambda w_H - u(c_t^{21}) + \theta_L v(y_t^{21})), \\
& w_H + dt((\lambda + r)w_H - \lambda w_L - u(c_t^{21}) + \theta_H v(y_t^{21}))) + \\
& V_L^{t+dt}((w_L + h) + dt((\lambda + r)(w_L + h) - \lambda(w_H + h) - u(c_t^{12}) + \theta_L v(y_t^{12})), \\
& (w_H + h) + dt((\lambda + r)(w_H + h) - \lambda(w_L + h) - u(c_t^{12}) + \theta_H v(y_t^{12}))).
\end{aligned}
$$

Else if $-u(c_t^{12}) + \theta_H v(y_t^{12}) \geq -u(c_t^{21}) + \theta_H v(y_t^{21})$,

$$
\begin{aligned}
& V_L^{t+dt}(w'_{L12}, w'_{H12}) + V_L^{t+dt}(w'_{L21}, w'_{H21}) \\
\geq\ & V_L^{t+dt}(w'_{L12} + \lambda hdt, w'_{H12} - \lambda hdt) + V_L^{t+dt}(w'_{L21} - \lambda hdt, w'_{H21} + \lambda hdt) \\
\geq\ & V_L^{t+dt}(w'_{L12} + \lambda hdt, w'_{H12} - \lambda hdt - (-u(c_t^{12}) + \theta_H v(y_t^{12}) + u(c_t^{21}) - \theta_H v(y_t^{21}))dt) + \\
& V_L^{t+dt}(w'_{L21} - \lambda hdt, w'_{H21} + \lambda hdt + (-u(c_t^{12}) + \theta_H v(y_t^{12}) + u(c_t^{21}) - \theta_H v(y_t^{21}))dt) \\
\geq\ & V_L^{t+dt}(w'_{L12} + \lambda hdt, w'_{H21} + \lambda hdt + (-u(c_t^{12}) + \theta_H v(y_t^{12}) + u(c_t^{21}) - \theta_H v(y_t^{21}))dt) + \\
& V_L^{t+dt}(w'_{L21} - \lambda hdt, w'_{H12} - \lambda hdt - (-u(c_t^{12}) + \theta_H v(y_t^{12}) + u(c_t^{21}) - \theta_H v(y_t^{21}))dt) \\
=\ & V_L^{t+dt}(w_L + dt((\lambda + r)w_L - \lambda w_H - u(c_t^{12}) + \theta_L v(y_t^{12})), \\
& w_H + dt((\lambda + r)w_H - \lambda w_L - u(c_t^{12}) + \theta_H v(y_t^{12}))) + \\
& V_L^{t+dt}((w_L + h) + dt((\lambda + r)(w_L + h) - \lambda(w_H + h) - u(c_t^{21}) + \theta_L v(y_t^{21})), \\
& (w_H + h) + dt((\lambda + r)(w_H + h) - \lambda(w_L + h) - u(c_t^{21}) + \theta_H v(y_t^{21}))).
\end{aligned}
$$

Else

$$-u(c_t^{21}) + \theta_L v(y_t^{21}) < -u(c_t^{12}) + \theta_L v(y_t^{12}),$$
$$-u(c_t^{12}) + \theta_H v(y_t^{12}) < -u(c_t^{21}) + \theta_H v(y_t^{21}).$$

Then

$$u(c_t^{21}) > u(c_t^{12}), v(y_t^{21}) > v(y_t^{12}).$$

$$\begin{aligned}
& V_L^{t+dt}(w'_{L12}, w'_{H12}) + V_L^{t+dt}(w'_{L21}, w'_{H21}) \\
\geq\ & V_L^{t+dt}(w'_{L12} + (\lambda h + \theta_L v(y_t^{21}) - \theta_L v(y_t^{12}))dt, w'_{H12} - (\lambda h + u(c_t^{21}) - u(c_t^{12}))dt) + \\
& V_L^{t+dt}(w'_{L21} - (\lambda h + \theta_L v(y_t^{21}) - \theta_L v(y_t^{12}))dt, w'_{H21} + (\lambda h + u(c_t^{21}) - u(c_t^{12}))dt) \\
\geq\ & V_L^{t+dt}(w'_{L12} + (\lambda h + \theta_L v(y_t^{21}) - \theta_L v(y_t^{12}))dt, w'_{H21} + (\lambda h + u(c_t^{21}) - u(c_t^{12}))dt) + \\
& V_L^{t+dt}(w'_{L21} - (\lambda h + \theta_L v(y_t^{21}) - \theta_L v(y_t^{12}))dt, w'_{H12} - (\lambda h + u(c_t^{21}) - u(c_t^{12}))dt) \\
=\ & V_L^{t+dt}(w_L + dt((\lambda + r)w_L - \lambda w_H - u(c_t^{12}) + \theta_L v(y_t^{21})), \\
& w_H + dt((\lambda + r)w_H - \lambda w_L - u(c_t^{12}) + \theta_H v(y_t^{21}))) + \\
& V_L^{t+dt}((w_L + h) + dt((\lambda + r)(w_L + h) - \lambda(w_H + h) - u(c_t^{21}) + \theta_L v(y_t^{12})), \\
& (w_H + h) + dt((\lambda + r)(w_H + h) - \lambda(w_L + h) - u(c_t^{21}) + \theta_H v(y_t^{12}))).
\end{aligned}$$

Therefore, at (w_L, w_H) we can use policy (c_t^{12}, y_t^{12}), or (c_t^{21}, y_t^{21}), or (c_t^{12}, y_t^{21}), depending on different cases. In every case, we have

$$V_L^t(w_L, w_H) + V_L^t(w_L + h, w_H + h) \leq V_L^t(w_L, w_H + h) + V_L^t(w_L + h, w_H).$$

Proof of Lemma 2

(1) Since at $(w_L, g(w_L))$, the only feasible policy function following report L is $y_t = \bar{y}$, therefore

$$V_L(w_L, g(w_L))$$

$$
\begin{aligned}
&= (c_t - \bar{y})dt + e^{-rdt}[(1-\lambda dt)V_L(w_L + ((\lambda+r)w_L - \lambda g(w_L) - u(c_t) + \theta_L \bar{v})dt, \\
&\quad g(w_L) + ((\lambda+r)g(w_L) - \lambda w_L - u(c_t) + \theta_H \bar{v})dt) + \\
&\quad \lambda dt V_H(w_L, g(w_L))] \\
&= V_L(w_L, g(w_L)) + [-(\lambda+r)V_L(w_L, w_H) + \\
&\quad V_{Lw_L}((\lambda+r)w_L - \lambda g(w_L) - u(c_t) + \theta_L \bar{v}) + \\
&\quad V_{Lw_H}((\lambda+r)g(w_L) - \lambda w_L - u(c_t) + \theta_H \bar{v}) + \lambda V_H(w_L, g(w_L))]dt.
\end{aligned}
$$

$$
\begin{aligned}
&V_L(w_L, g(w_L) + \theta_H \epsilon dt) \\
\leq\ & [c_t - v^{-1}(\bar{v} - \epsilon)]dt + e^{-rdt}[\\
&(1-\lambda dt)V_L(w_L + ((\lambda+r)w_L - \lambda g(w_L) - u(c_t) + \theta_L \bar{v})dt - \theta_L \epsilon dt, \\
&g(w_L) + ((\lambda+r)g(w_L) - \lambda w_L - u(c_t) + \theta_H \bar{v})dt) \\
&+\lambda dt V_H(w_L, g(w_L))] \\
=\ & V_L(w_L, g(w_L)) + [c_t - v^{-1}(\bar{v} - \epsilon) - (\lambda+r)V_L(w_L, w_H) + \\
&V_{Lw_L}((\lambda+r)w_L - \lambda g(w_L) - u(c_t) + \theta_L \bar{v} - \theta_L \epsilon) + \\
&V_{Lw_H}((\lambda+r)g(w_L) - \lambda w_L - u(c_t) + \theta_H \bar{v}) + \lambda V_H(w_L, g(w_L))]dt.
\end{aligned}
$$

I have,

$$
\begin{aligned}
V_{Lw_H}(w_L, g(w_L)) &= \lim_{\epsilon \to 0} \frac{V_L(w_L, g(w_L) + \theta_H \epsilon dt) - V_L(w_L, g(w_L))}{\theta_H \epsilon dt} \\
&\leq \frac{\bar{y} - v^{-1}(\bar{v} - \epsilon) - V_{Lw_L}\theta_L \epsilon}{\theta_H \epsilon} \\
&< 0.
\end{aligned}
$$

(2) I first show that $V_{Hw_H}(w_L, h(w_L)) = \infty$. At $(w_L, h(w_L))$, the only feasible policy

function following report H is $y_t = 0$, therefore

$$\begin{aligned}
& V_H(w_L, h(w_L)) \\
= \;& c_t dt + e^{-rdt}[(1-\lambda dt)V_H(w_L + ((\lambda+r)w_L - \lambda h(w_L) - u(c_t))dt, \\
& h(w_L) + ((\lambda+r)h(w_L) - \lambda w_L - u(c_t))dt) \\
& +\lambda dt V_L(w_L, h(w_L))] \\
= \;& V_H(w_L, h(w_L)) + [c_t - (\lambda+r)V_H(w_L, h(w_L)) + \\
& V_{Hw_L}((\lambda+r)w_L - \lambda h(w_L) - u(c_t)) + V_{Hw_H}((\lambda+r)h(w_L) - \\
& \lambda w_L - u(c_t)) + \lambda V_L(w_L, h(w_L))]dt.
\end{aligned}$$

$$\begin{aligned}
& V_H(w_L, h(w_L) - \theta_H \epsilon dt) \\
\leq \;& [c_t - v^{-1}(\epsilon)]dt + e^{-rdt}[(1-\lambda dt)V_H(w_L + ((\lambda+r)w_L - \lambda h(w_L) - u(c_t) + \theta_L \epsilon)dt, \\
& h(w_L) + ((\lambda+r)h(w_L) - \lambda w_L - u(c_t))dt) + \\
& \lambda dt V_L(w_L, h(w_L)) \\
= \;& V_H(w_L, h(w_L)) + [c_t - v^{-1}(\epsilon) - (\lambda+r)V_H(w_L, h(w_L)) + \\
& V_{Hw_L}((\lambda+r)w_L - \lambda h(w_L) - u(c_t) + \theta_L \epsilon) + V_{Hw_H}((\lambda+r)h(w_L) - \lambda w_L - u(c_t)) + \\
& \lambda V_L(w_L, h(w_L))]dt.
\end{aligned}$$

I have,

$$\begin{aligned}
V_{Hw_H}(w_L, h(w_L)) &= \lim_{\epsilon \to 0} \frac{V_H(w_L, h(w_L)) - V_H(w_L, h(w_L) - \theta_H \epsilon dt)}{\theta_H \epsilon dt} \\
&\geq \frac{v^{-1}(\epsilon) - V_{Hw_L}\theta_L \epsilon}{\theta_H \epsilon} \\
&= \infty.
\end{aligned}$$

Denote $dw_L = ((\lambda+r)w_L - \lambda h(w_L) - u(c_t))dt, dw_H = ((\lambda+r)h(w_L) - \lambda w_L - u(c_t))dt,$

I also have

$$
\begin{aligned}
& \lim_{dt\to 0} \frac{V_H(w_L+dw_L, h(w_L)+dw_H) - V_H(w_L, h(w_L))}{dw_L} \\
= & \lim_{dt\to 0} \frac{c_t dt + e^{-rdt}[(1-\lambda dt)V_H(w_L, h(w_L)) + \lambda dt V_L(w_L, h(w_L))] - V_H(w_L, h(w_L))}{dw_L} \\
= & \text{ finite number.}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& V_{Hw_L}(w_L, h(w_L)) \\
= & \lim_{dt\to 0} \frac{V_H(w_L+dw_L, h(w_L)) - V_H(w_L, h(w_L))}{dw_L} \\
= & \lim_{dt\to 0} \frac{V_H(w_L+dw_L, h(w_L)) - V_H(w_L+dw_L, h(w_L)+dw_H)}{dw_L} + \\
& \lim_{dt\to 0} \frac{V_H(w_L+dw_L, h(w_L)+dw_H) - V_H(w_L, h(w_L))}{dw_L} \\
= & -\lim_{dt\to 0} \frac{V_H(w_L+dw_L, h(w_L)+dw_H) - V_H(w_L+dw_L, h(w_L))}{dw_H} \frac{dw_H}{dw_L} + \\
& \lim_{dt\to 0} \frac{V_H(w_L+dw_L, h(w_L)+dw_H) - V_H(w_L, h(w_L))}{dw_L} \\
= & -\infty.
\end{aligned}
$$

(3) I only prove $V_{Lw_H}(w_L, h(w_L)) = 0$, since the other is analogous. Denote

$$B = \lim_{t\downarrow 0} w_H(0, L, t).$$

Starting from the initial state $(w_L, h(w_L))$, if $B < h(w_L)$, $V_L(w_L, h(w_L)) = V_L(w_L, B)$, thus $V_{Lw_H}(w_L, h(w_L)) = 0$. If $B = h(w_L)$, then y_t maybe 0 or bigger than 0. If it is 0, I can get a contradiction similar to the argument used in part (2). Therefore, consider the case where $y_t > 0$. Let μ be the control with which the continuation remains inside W.

$$\begin{aligned}
& V_L(w_L, h(w_L)) \\
= \; & [c_t - y_t]dt + e^{-rdt}[(1-\lambda dt)V_L(w_L + ((\lambda+r)w_L - \lambda h(w_L) - u(c_t) + \theta_L v(y_t))dt, \\
& h(w_L) + ((\lambda+r)h(w_L) - \lambda w_L - u(c_t) + \theta_H v(y_t) - \mu)dt) + \\
& \lambda dt V_H(w_L, h(w_L))] \\
= \; & V_L(w_L, h(w_L)) + [c_t - y_t - (\lambda+r)V_L(w_L, h(w_L)) + \\
& V_{Lw_L}((\lambda+r)w_L - \lambda h(w_L) - u(c_t) + \theta_L v(y_t)) + \\
& V_{Lw_H}((\lambda+r)h(w_L) - \lambda w_L - u(c_t) + \theta_H v(y_t) - \mu) + \lambda V_H(w_L, h(w_L))]dt.
\end{aligned}$$

$$\begin{aligned}
& V_L(w_L, h(w_L) - \mu dt) \\
\leq \; & [c_t - y_t]dt + e^{-rdt}[(1-\lambda dt)V_L(w_L + ((\lambda+r)w_L - \lambda h(w_L) - u(c_t) + \theta_L v(y_t))dt, \\
& h(w_L) + ((\lambda+r)h(w_L) - \lambda w_L - u(c_t) + \theta_H v(y_t) - \mu)dt) + \lambda dt V_H(w_L, h(w_L))] \\
= \; & V_L(w_L, h(w_L)) + [c_t - y_t - (\lambda+r)V_L(w_L, h(w_L)) + \\
& V_{Lw_L}((\lambda+r)w_L - \lambda h(w_L) - u(c_t) + \theta_L v(y_t)) + \\
& V_{Lw_H}((\lambda+r)h(w_L) - \lambda w_L - u(c_t) + \theta_H v(y_t) - \mu) + \lambda V_H(w_L, h(w_L))]dt.
\end{aligned}$$

Therefore, $V_{Lw_H}(w_L, h(w_L)) = \lim_{dt\to 0} \frac{V_L(w_L,h(w_L)) - V_L(w_L,h(w_L)-\mu dt)}{\mu dt} \geq 0$.

Proof of Lemma 4

I only prove the first inequality. First I show that for $w_H > f_L(w_L)$,

$$\begin{aligned}
(\lambda+r)V_L(w_L, w_H) \; \leq \; & \min_c\{c - (V_{Lw_L} + V_{Lw_H})u(c)\} + \min_y\{-y + (\theta_L V_{Lw_L} + \theta_H V_{Lw_H})v(y)\} \\
& + \lambda V_H(w_L, w_H) + V_{Lw_L}((\lambda+r)w_L - \lambda w_H) + V_{Lw_H}((\lambda+r)w_H - \lambda w_L).
\end{aligned}$$

By contradiction, suppose that the left-hand side is bigger than the right-hand side, then

$$\begin{aligned}
V_L(w_L, w_H) \quad \leq \quad & (c-y)dt + e^{-rdt}((1-\lambda dt)V_L(w_L + ((\lambda+r)w_L - \lambda w_H - u(c) + \theta_L v(y))dt, \\
& w_H + ((\lambda+r)w_H - \lambda w_L - u(c) + \theta_H v(y))dt) + \\
& \lambda dt V_L(w_L + (w_L + ((\lambda+r)w_L - \lambda w_H - u(c) + \theta_L v(y))dt, \\
& w_H + ((\lambda+r)w_H - \lambda w_L - u(c) + \theta_H v(y))dt) \\
= \quad & V_L(w_L, w_H) + (\min_c\{c - (V_{Lw_L} + V_{Lw_H})u(c)\} + \\
& \min_y\{-y + (\theta_L V_{Lw_L} + \theta_H V_{Lw_H})v(y)\} + \lambda V_H(w_L, w_H) + \\
& V_{Lw_L}((\lambda+r)w_L - \lambda w_H) + V_{Lw_H}((\lambda+r)w_H - \lambda w_L) - \\
& (\lambda+r)V_L(w_L, w_H))dt \\
< \quad & V_L(w_L, w_H).
\end{aligned}$$

Now I know that

$$\begin{aligned}
0 \quad \leq \quad & \partial[\min_c\{c - V_{Lw_L}u(c) - V_{Lw_H}u(c)\} \\
& + \min_y\{-y + \theta_L V_{Lw_L}v(y) + \theta_H V_{Lw_H}v(y)\} + \lambda V_H(w_L, w_H) + V_{Lw_L}((\lambda+r)w_L - \lambda w_H) \\
& + V_{Lw_H}((\lambda+r)w_H - \lambda w_L) - (\lambda+r)V_L(w_L, w_H)]/\partial w_H \\
= \quad & \lambda(V_{Hw_H}(w_L, w_H)|_{w_H = f_L(w_L)} - V_{Lw_L}(w_L, w_H)|_{w_H = f_L(w_L)}).
\end{aligned}$$

Proof of Lemma 5

We first notice that $f_L(w_L) \geq f_H^{-1}(w_L)$. Fix a value of w_L, we know that $V_{Lw_L}(w_L, w_H)$ is an decreasing function of w_H, while $V_{Hw_H}(w_L, w_H)$ is an increasing function. Since

$$V_{Hw_H}(w_L, f_L(w_L)) - V_{Lw_L}(w_L, f_L(w_L)) \geq 0,$$
$$V_{Hw_H}(w_L, f_H^{-1}(w_L)) - V_{Lw_L}(w_L, f_H^{-1}(w_L)) \leq 0.$$

I obtain $f_L(w_L) \geq f_H^{-1}(w_L)$.

To show the strict inequality, we assume by contradiction that for some $w^* \in (x_L, \bar{u}/r)$, $f_L(w_L^*) = f_H^{-1}(w_L^*)$. Then since f_H is below f_L by the previous arguments, f_L and f_H has to be tangent at point $(w_L^*, f_L(w_L^*))$. On curve f_L, HJB equations are

$$\begin{aligned}(\lambda+r)V_L(w_L, f_L(w_L)) &= \min_c\{c - V_{Lw_L}u(c)\} + \min_y\{-y + \theta_L V_{Lw_L}v(y)\} + \\ &\quad \lambda V_H(w_L, f_L(w_L)) + V_{Lw_L}((\lambda+r)w_L - \lambda f_L(w_L)). \end{aligned} \tag{19}$$

$$\begin{aligned}(\lambda+r)V_H(w_L, f_L(w_L)) &= \min_c\{c - V_{Hw_L}u(c) - V_{Hw_H}u(c)\} + \\ &\quad \min_y\{-y + \theta_L V_{Hw_L}v(y) + \theta_H V_{Hw_H}v(y)\} + \\ &\quad \lambda V_L(w_L, f_L(w_L)) + V_{Hw_L}((\lambda+r)w_L - \lambda f_L(w_L)) + \\ &\quad V_{Hw_H}((\lambda+r)f_L(w_L) - \lambda w_L). \end{aligned} \tag{20}$$

Substituting equation (20) into (19) yields

$$\begin{aligned}\frac{2\lambda r + r^2}{\lambda+r}V_L(w_L, f_L(w_L)) &= \min_c\{c - V_{Lw_L}u(c)\} + \min_y\{-y + \theta_L V_{Lw_L}v(y)\} + \\ &\quad V_{Lw_L}((\lambda+r)w_L - \lambda f_L(w_L)) + \\ &\quad \frac{\lambda}{\lambda+r}\min_c\{c - V_{Hw_L}u(c) - V_{Hw_H}u(c)\} + \\ &\quad \frac{\lambda}{\lambda+r}\min_y\{-y + \theta_L V_{Hw_L}v(y) + \theta_H V_{Hw_H}v(y)\} + \\ &\quad \frac{\lambda}{\lambda+r}V_{Hw_L}((\lambda+r)w_L - \lambda f_L(w_L)) + \\ &\quad \frac{\lambda}{\lambda+r}V_{Hw_H}((\lambda+r)f_L(w_L) - \lambda w_L). \end{aligned}$$

Notice that

$$\begin{aligned}\frac{dV_{Hw_L}(w_L, f_L(w_L))}{dw_L}\Big|_{w_L=w_L^*} &= 0, \\ \frac{dV_{Hw_H}(w_L, f_L(w_L))}{dw_L}\Big|_{w_L=w_L^*} &= \frac{dV_{Lw_L}(w_L, f_L(w_L))}{dw_L}\Big|_{w_L=w_L^*}. \end{aligned}$$

I differentiate the HJB equation at $(w_L^*, f_L(w_L^*))$ and obtain

$$\begin{aligned}
\frac{2\lambda r + r^2}{\lambda + r} V_{Lw_L} &= \frac{dV_{Lw_L}}{dw_L}(-u(c_L) + \theta_L v(y_L)) + \\
&\quad \frac{dV_{Lw_L}}{dw_L}((\lambda + r)w_L - \lambda f_L(w_L)) + \\
&\quad V_{Lw_L}((\lambda + r) - \lambda \frac{df_L(w_L)}{dw_L}) + \\
&\quad \frac{\lambda}{\lambda + r} \frac{dV_{Hw_H}}{dw_L}(-u(c_H) + \theta_H v(y_H)) + \\
&\quad \frac{\lambda}{\lambda + r} \frac{dV_{Hw_H}}{dw_L}((\lambda + r) f_L(w_L) - \lambda w_L) + \\
&\quad \frac{\lambda}{\lambda + r} V_{Hw_H}((\lambda + r)\frac{df_L(w_L)}{dw_L} - \lambda),
\end{aligned}$$

which can be simplified to

$$w_L^* = \tfrac{\lambda + r}{(2\lambda + r)r}(u(c_L) - \theta_L v(y_L)) + \tfrac{\lambda}{(2\lambda + r)r}(u(c_H) - \theta_H v(y_H)).$$

Similarly, I use the HJB equation on f_H and obtain

$$w_H^* = \tfrac{\lambda}{(2\lambda + r)r}(u(c_L) - \theta_L v(y_L)) + \tfrac{\lambda + r}{(2\lambda + r)r}(u(c_H) - \theta_H v(y_H)).$$

This implies that

$$(\lambda + r)w_L^* - \lambda w_H^* = u(c_L) - \theta_L v(y_L),$$

$$(\lambda + r)w_H^* - \lambda w_L^* = u(c_H) - \theta_H v(y_H).$$

Substituting back into equation (19,20) yields

$$(\lambda + r)V_L(w_L^*, w_H^*) = c_L - y_L + \lambda V_H(w_L^*, w_H^*),$$

$$(\lambda + r)V_H(w_L^*, w_H^*) = c_H - y_H + \lambda V_L(w_L^*, w_H^*).$$

Implementing (w_L^*, w_H^*) uses full information, which is impossible given our informational constraints.

Proof of Lemma 6

I first show that $V_{Lw_L} = V_{Hw_H}$ on f_L. By contradiction, I assume that $V_{Lw_L} < V_{Hw_H}$ on f_L. This would imply that $V_L(w_L, w_H) > V_H(w_L, w_H)$ for all $w_H \leq f_L(w_L), (w_L, w_H) \in W$.

This can be seen as follows. First starting from $(w_L, f_L(w_L))$, the time path under report L moves below f_L because

$$\frac{dV_{Lw_H}}{dt} = \lambda(V_{Lw_L} - V_{Hw_H}) < 0.$$

This means that on the time path, instead of being an inequality,

$$\frac{dw_H}{dt} = (\lambda + r)w_H - \lambda w_L - u(c) + \theta_H v(y).$$

The type H agent is indifferent between telling the truth and cheating, because reporting L also delivers the promised utility. Now I show that when the principal's belief is $p_L = 0$ and wants to deliver a utility pair (w_L, w_H), she can utilize the consumption-output plan implied by the value function $V_L(w_L, w_H)$ and incurs the expected cost $V_L(w_L, w_H)$(but under the belief $p_L = 0$, not under $p_L = 1$!). The only thing different is that instead of proposing truth-telling as the strategy for the agent, the contract require randomization of the reporting strategy. Let θ_t^1 be a Markov process with initial state L, which is independent with the agent's privately observed types θ_t. Let t^{1*} be the time of θ_t^1's first transition from L to H, and let t^{2*} be the time of θ_t's first transition from H to L. The strategy is

$$\sigma_t(\theta^{1t}, \theta^t) = \begin{cases} L, \text{if } t < \min\{t^{1*}, t^*\}, \\ \theta_t, \text{if } t \geq \min\{t^{1*}, t^*\}. \end{cases}$$

Up to the time of the first transition, the agent cheats by telling her type to be L. At the time of switch, if $t^{1*} < t^*$, the agent report that her type switches from L to H, although she does not have one. If $t^{1*} \geq t^*$, she continues to claims to be the type L type although she has a switch from H to L in the past. After t^*, the contract requires truth-telling. This strategy is optimal at $t < \min\{t^{1*}, t^*\}$ because when the agent's type is H, she could still obtain the promised utility by cheating, as we argued before. It is easy to verify that if the agent uses this strategy, the principal incurs the expected cost $V_L(w_L, w_H)$. Since the above mechanism and strategy is one option for the principal when she wants to deliver (w_L, w_H) under belief $p_L = 0$,

$$V_L(w_L, w_H) > V_H(w_L, w_H), \forall w_H \leq f_L(w_L), (w_L, w_H) \in W.$$

Under the assumption that $V_{Lw_L} < V_{Hw_H}$ on f_L, it is not possible that $V_{Lw_L} > V_{Hw_H}$ on f_H, because by the same argument used before, it would imply

$$V_L(w_L, w_H) < V_H(w_L, w_H), \forall w_L \leq f_H(w_H), (w_L, w_H) \in W.$$

which is a contradiction. Therefore, we need to get a contradiction in the case

$$\begin{aligned} V_{Lw_L} &< V_{Hw_H}, \text{on } f_L, \\ V_{Lw_L} &= V_{Hw_H}, \text{on } f_H. \end{aligned}$$

Since f_H is located between the lower and upper boundary of W,

$$\lim_{w_L \to x_L} \frac{d(f_H^{-1})(w_L)}{dw_L} \geq 1.$$

Thus

$$\begin{aligned} V_L(w_L, f_H^{-1}(w_L)) &> V_H(w_L, f_H^{-1}(w_L)) \\ &= V_H(x_L, x_H) + \int_{x_L}^{w_L} \left(V_{Hw_L} + V_{Hw_H} \frac{df_H^{-1}(s)}{ds} \right) ds \\ &= V_H(x_L, x_H) + \int_{x_L}^{w_L} \left(V_{Hw_H} \frac{df_H^{-1}(s)}{ds} \right) ds \\ &> V_L(x_L, x_H) + \int_{x_L}^{w_L} \left(V_{Lw_L} + V_{Lw_H} \frac{df_H^{-1}(s)}{ds} \right) ds \\ &= V_L(w_L, f_H^{-1}(w_L)). \end{aligned}$$

From this contradiction, we know that $V_{Lw_L} = V_{Hw_H}$ on f_L. This immediately implies that $V_{Lw_L} > V_{Hw_H}$ on f_H, because f_H is below f_L, and $(V_{Lw_L} - V_{Hw_H})$ is decreasing in w_H.

Proof of Lemma 8

We show that a time path under report H can not converge to any interior point. By contradiction, suppose a point (w_L, w_H) is stationary, then

$$\frac{dV_{Hw_L}}{dt} = \lambda(V_{Hw_H} - V_{Lw_L}),$$
$$\frac{dV_{Hw_H}}{dt} = \lambda(V_{Hw_L} - V_{Lw_H}).$$

$V_{Hw_H} - V_{Lw_L} = 0$ implies that (w_L, w_H) is on f_L, but then $V_{Hw_L} - V_{Lw_H} < 0$.

Take a time path $(w_L(t), w_H(t))$ that converges to (x_L, x_H). The HJB equation is

$$\begin{aligned}(\lambda + r)V_H(w_L, w_H) &= \min_c\{c - (V_{Hw_L} + V_{Hw_H})u(c)\} + \min_y\{-y + (\theta_L V_{Hw_L} + \theta_H V_{Hw_H})v(y)\} \\ &\quad +\lambda V_L(w_L, w_H) + V_{Hw_L}((\lambda + r)w_L - \lambda w_H) \\ &\quad +V_{Hw_H}((\lambda + r)w_H - \lambda w_L).\end{aligned}$$

Taking limit,

$$\begin{aligned}(\lambda + r)(-\bar{y}/r) &= c^* - y^* + \lambda(-\bar{y}/r), \\ \text{where } c^* &\in \text{argmin}_c\{c - Bu(c)\}, y^* \in \text{argmin}_y\{-y + Cv(y)\}, \\ B &= \lim_{(w_L,w_H)\to(x_L,x_H)} V_{Hw_L} + V_{Hw_H}, \\ C &= \lim_{(w_L,w_H)\to(x_L,x_H)} \theta_L V_{Hw_L} + \theta_H V_{Hw_H}.\end{aligned}$$

which implies $c^* = 0, y^* = \bar{y}, B \leq 0, C = 0$.

Proof of Lemma 9

Under *Assumption* 3, I first show that

$$\frac{dV_{Hw_L}(w_L, f_L(w_L))}{dw_L} < 0.$$

This is implied by the following two facts:

$$\lim_{w_L \downarrow x_L} V_{Hw_L}(w_L, f_L(w_L)) = 0,$$
$$\lim_{w_L \uparrow \bar{u}/r} V_{Hw_L}(w_L, f_L(w_L)) = -\infty.$$

To prove the first one, notice that

$$\lim_{w_L \downarrow x_L} \frac{V_L(w_L, f_L(w_L)) - V_L(x_L, x_H)}{w_L - x_L} = 0.$$

Then I have $\lim_{w_L \downarrow x_L} V_{Lw_L}(w_L, f_L(w_L)) = 0$. Since on f_L, $V_{Lw_L} = V_{Hw_H}$, I obtain

$$\begin{aligned}
0 &\geq \lim_{w_L \downarrow x_L} V_{Hw_L}(w_L, f_L(w_L)) \\
&\geq \lim_{w_L \downarrow x_L} \left[V_{Hw_L}(w_L, f_L(w_L)) + \frac{df_L(w_L)}{dw_L} V_{Hw_H}(w_L, f_L(w_L)) \right] \\
&= \lim_{w_L \downarrow x_L} \frac{V_H(w_L, f_L(w_L))}{dw_L} \\
&\geq 0.
\end{aligned}$$

Now I show

$$\lim_{w_L \uparrow \bar{u}/r} V_{Hw_L}(w_L, f_L(w_L)) = -\infty.$$

Assume by contradiction that $\lim_{w_L \uparrow \bar{u}/r} V_{Hw_L}(w_L, f_L(w_L))$ is finite,

$$\begin{aligned}
& \frac{V_H(\bar{u}/r, \bar{u}/r) - V_H(\bar{u}/r - dw_L, f_L(\bar{u}/r - dw_L))}{dw_L} \\
= \; & \frac{dw_H}{dw_L} V_{Hw_H}(w_L, f_L(w_L)) + V_{Hw_L}(w_L, f_L(w_L)) \\
\geq \; & V_{Lw_L}(w_L, f_L(w_L)) \\
= \; & \frac{V_L(\bar{u}/r, \bar{u}/r) - V_L(\bar{u}/r - dw_L, f_L(\bar{u}/r - dw_L))}{dw_L}.
\end{aligned}$$

This implies $V_H(w_L, f_L(w_L)) \leq V_L(w_L, f_L(w_L))$, a contradiction to $V_H(w_L, f_L(w_L)) > V_L(w_L, f_L(w_L))$.

Now recall the ODE system describing the dynamics under report H,

$$\frac{dV_{Hw_L}}{dt} = \lambda V_{Hw_H} - \lambda V_{Lw_L}.$$

Therefore, on f_L, the time path of (w_L, w_H) will be tangent to the isoquants of V_{Hw_L}. Since the isoquants go above f_L, the time path of (w_L, w_H) must enter the interior of the region.

References

[1] Dilip Abreu, David Pearce and Ennio Staccheti: Toward a Theory of Discounted Repeated Games with Imperfect Monitoring, Econometrica **58(5)**, 1041-1063(1990).

[2] Stefania Albanesi and Christopher Sleet: Dynamic Optimal Taxation with Private Information, forthcoming, Review of Economic Studies.

[3] Andrew Atkeson and Robert E. Lucas, Jr: On Efficient Distributions with Private Informaton, Review of Economic Studies, **59(3)**, 427-453(1992).

[4] Marco Battaglini and Stephen Coate: Pareto Efficient Income Taxations with Stochastic Abilities, NBER Working Paper 10119, 2003.

[5] V. V. Chari, Patrick J. Kehoe and Ellen R. McGrattan: Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?, Staff Report 217, Federal Reserve Bank of Minneapolis.

[6] Emmanuel Farhi and Iván Werning: Inequality, Social Discounting and Estate Taxation, Working Paper, MIT, 2005.

[7] Ana Fernandes and Christopher Phelan: A Recursive Formulation for Repeated Agency with History Dependence, Journal of Economic Theory **91**, 223–247(2000).

[8] Mikhail Golosov, Narayana Kocherlakota and Aleh Tsyvinski: Optimal Indirect and Capital Taxation, Review of Economic studies **70**, 569-587(2003).

[9] Mikhail Golosov and Aleh Tsyvinski: Designing Optimal Disability Insurance: A Case for Asset Testing, Working Paper, MIT and Harvard University, 2004.

[10] Costas Meghir and Luigi Pistaferri: Income Variance and Heterogeneity, Econometrica **72**, 1-32(2004).

[11] Marek Kapicka: Optimal Taxation with Persistent Shocks, Working Paper, University of California, Santa Barbara.

[12] Narayana, Kocherlakota: Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, Econometrica **73**, 1587-1621(2005).

[13] Yuliy Sannikov: Games with Imperfectly Observable Actions in Continuous Time, Working Paper, University of California, Berkeley.

[14] Yuliy Sannikov: A Continuous-Time Version of the Principal-Agent Problem, Working Paper, University of California, Berkeley.

[15] Storesletten K., Telmer C. and Yaron A: How Important are Idiosyncratic Shocks? Evidence from Labor Supply, American Economic Review Papers and Proceedings, **91**,413–417(2001).