Characterization of a Risk Sharing Contract with One-Sided Commitment

Zhang, Yuzhe

Texas AM University

November 2012

Characterization of a Risk Sharing Contract with One-Sided Commitment

Yuzhe Zhang[1]

Department of Economics, Texas A&M University, College Station, TX, 77843, United States

Abstract

In this paper I provide a stopping-time-based solution to a long-term contracting problem between a risk-neutral principal and a risk-averse agent. The agent faces a stochastic income stream and cannot commit to the long-term contracting relationship. To compute the optimal contract, I also design an algorithm that is more efficient than value-function iteration.

Keywords: Limited commitment, Risk sharing, Stopping time, Value-function iteration
JEL: C63, D82, D86

1. Introduction

The theory of contracting with limited commitment has been applied to study a wide variety of economic issues, including asset pricing (cf. Kehoe and Levine (1993), Alvarez and Jermann (2000)), consumption inequality (cf. Krueger and Perri (2006)), and the welfare effects of a progressive tax (cf. Krueger and Perri (2011)). The standard approach to solving these contracting problems is to iterate on the principal's value function.[2] However, value-function iteration (VFI) provides little general analytical characterization; further, when the discount factor is close to one, the value function converges slowly, making it computationally inefficient. The main contribution of this paper is to provide a constructive stopping-time-based procedure for solving the optimal contract with one-sided commitment. This method fully reveals the risk sharing dynamics in the contract. Moreover, I design a stopping-time-based algorithm that is two orders of magnitude faster than value-function iteration.

[1]E-mail address: yuzhe-zhang@econmail.tamu.edu.

[2]Relevant aspects of the agent's history are first summarized in a single variable, which is the promised utility to the agent. Then the contracting problem is transformed into a dynamic programming problem, and recursive techniques are applied to solve the problem (cf. Spear and Srivastava (1987), Abreu et al. (1990)).

My model features a risk-neutral fully committed principal and a risk-averse noncommitted agent, and generalizes Ljungqvist and Sargent (2004, Chapter 19) along three dimensions. While they assume that the agent's income is independently and identically distributed (i.i.d.), the outside option is autarky, and the principal and the agent discount the future at a common rate, I allow for a Markov-chain income process, an arbitrary outside option, and different discount rates. The three generalizations in my model are motivated by the following observations. First, it is well documented that people experience large and persistent income shocks over the life cycle. The quantitative features of the income process are poorly approximated by i.i.d. shocks. Second, agents in a number of long-term relationships have outside options better than autarky. For instance, in wage contracting between a firm and a worker, the worker has the option to quit the current job and find a new one. Last, I allow for different discount rates because when the principal in the model takes the interpretation of a financial intermediary, his discount factor should be determined by the interest rate. In a general equilibrium model, the endogenously determined interest rate is typically lower than the reciprocal of the agent's discount factor.

In the optimal contract, the agent's consumption follows a simple recursive rule: consumption deviates each period from the first-best level by the smallest amount necessary to bring it above some (state-dependent) minimum level. Because the recursive rule is relatively easy to verify, my paper focuses on finding the minimum levels. I first solve a stopping-time optimization problem: the moment when the participation constraint binds is a stopping time, and the stopping time is chosen to minimize the agent's consumption flow before it arrives. Then I guess and verify that the minimum level is the minimized consumption flow in the above problem.

My characterization of the contract is related to the solutions in Ljungqvist and Sargent (2004, Section 19.3.3), Krueger and Uhlig (2006, Section 3.5), Thomas and Worrall (2007, Section 3.2) and Krueger and Perri (2011, Section 4). Similar to my stopping-time approach, their solutions do not rely on value-function iteration. However, they assume i.i.d. incomes and rely on the monotonic mapping between incomes and minimum consumption levels. Broer (2009, 2011) extends the methods of Krueger and Uhlig (2006) to the Markov case and provides a sufficient condition under which the mapping between incomes and minimum consumption levels is monotone. However, his sufficient condition is violated for empirically relevant income processes such as the one calibrated by Krueger and Perri (2006). By contrast, my stopping-time approach does not depend on any particulars of the income process or the ordering of minimum consumption levels.

Stopping-time approaches have been used in continuous-time models. For instance, in liquidity constraint models in finance, Detemple and Serrat (2003) show that the optimal consumption portfolio problem of an individual is equivalent to a stopping-time problem in which wealth is optimally allocated over a random time period, during which the individual is not constrained. Grochulski and Zhang (2011) study a contracting problem in which the agent's income follows a geometric Brownian motion. Their analysis relies heavily on the fact that the stopping time is generated by a Brownian motion, hence is of limited value in other contexts. This paper allows for any Markov-chain income process and is, to the best of my knowledge, the first that applies stopping-time techniques to a discrete-time limited-commitment problem.

The recursive rule in this paper is related to a similar rule in two-sided limited-commitment models (cf. Thomas and Worrall (1988) and Ljungqvist and Sargent (2004, Chapter 20)). Because neither the principal nor the agent can commit in their models, both a minimum and a maximum level exist for the agent's consumption in each state. Their proof, however, is not constructive; therefore, to obtain these endogenous minimum and maximum consumption levels, they still need to iterate on value functions. By contrast, I study only one-sided limited commitment, but I am able to analytically construct the minimum consumption levels. Thus this paper solves completely, in the context of one-sided commitment, the fundamental problem concerning risk sharing dynamics.

This paper is also consistent with the findings in Ray (2002) and Krueger and Uhlig (2006). When the principal and the agent are equally patient, the agent's continuation utility in the long run will be sufficiently high so that the participation constraint no longer binds. When the principal is more patient, the agent's consumption has a downward drift, and the agent's participation constraint binds even in the long run. The model then predicts a nontrivial stationary distribution of consumption.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 uses an example to motivate the general result. Section 4 presents the stopping-time characterization of the optimal contract. In Section 5, I design an efficient algorithm to compute the minimum consumption levels. This algorithm does not involve VFI and terminates in finite steps. Section 6 discusses extensions and limitations of the model. The proofs of all the results in the paper are provided in an appendix.

2. A Risk Sharing Problem

Consider a risk-neutral principal and a risk-averse agent who engage in long-term contracting at time 0. Time is discrete and infinite. Preferences of the agent are represented by the expected utility function

$$E\left[\sum\nolimits_{t=0}^{\infty}\beta^t u(c_t)\right],$$

where c_t is the agent's consumption at time t, $\beta \in (0,1)$ is his discount factor and E is the expectation operator. I make the following assumption on the utility function:

Assumption 1. (i) *$u:\mathbb{R}_{++}\to\mathbb{R}$ is twice continuously differentiable, $u'>0$, and $u''<0$.*

(ii) *u satisfies the Inada conditions, i.e.,* $\lim_{c\to 0} u'(c) = \infty$ *and* $\lim_{c\to\infty} u'(c) = 0$.
(iii) *u is unbounded below, i.e.,* $\lim_{c\to 0} u(c) = -\infty$.

Part (iii) in Assumption 1 is only for the simplicity of the exposition; without it, the analysis in this paper can proceed with minor modifications. Note that I do not restrict $\bar{u} \equiv \lim_{c\to\infty} u(c)$, which can be either finite or infinite. In each period t, the agent's income y_t is in one of N states, i.e., $y_t \in Y \equiv \{\bar{y}_1, \bar{y}_2, ..., \bar{y}_N\}$, where $\bar{y}_1 < \bar{y}_2 < ... < \bar{y}_N$. The income stream $\{y_t; 0 \le t < \infty\}$ is a Markov chain on a probability space $(\Omega \equiv Y^\infty, \mathscr{F}, P)$ with transitional probability $\pi(y'|y)$. The sample space $(\Omega, \mathscr{F})$ is equipped with a filtration, i.e., an increasing family of σ-fields $\{\mathscr{F}_t; t \ge 0\}$, where $\mathscr{F}_t \equiv \sigma(y_s; 0 \le s \le t)$. Assume $\pi(y'|y) > 0$ for all $y, y' \in Y$ so that every finite-length path occurs with positive probability.

A contract specifies that in each period, the agent contributes his income y_t to the principal, who then returns c_t to the agent. The principal has a discount factor $\delta \in (0,1)$ and minimizes the expected discounted cost

$$E\left[\sum\nolimits_{t=0}^{\infty} \delta^t (c_t - y_t)\right].$$

The principal is committed to the contract, while the agent is not. In particular, after the realization of y_t in period t, the agent is always free to walk away from the principal and obtains outside utility $\underline{U}(y_t) \in \mathbb{R}$. One commonly used specification of $\underline{U}(y_t)$ arises for the case of autarky, in which the agent would consume his own income after leaving the contract and thus

$$\underline{U}(y_t) = E\left[\sum\nolimits_{s=t}^{\infty} \beta^{s-t} u(y_s) | y_t\right].$$

I assume that the outside option is not extremely high, so that long-term contracting is always profitable at any point in time, and thus the principal would design contracts in which the agent never leaves. In particular, the assumption implies:

Assumption 2. *If $\bar{u} < \infty$, then $\underline{U}(\bar{y}_i) < \frac{\bar{u}}{1-\beta}$, for all $i = 1, ..., N$.*

For an initial state y_0 and a given level of promised utility $U_0 \in \left[\underline{U}(y_0), \frac{\bar{u}}{1-\beta}\right)$, the principal's problem is to find a consumption plan $\{c_t; 0 \le t < \infty\}$ that minimizes the cost and satisfies the participation constraints, i.e.,

$$\min_{\{c_t; 0 \le t < \infty\}} \quad E\left[\sum\nolimits_{t=0}^{\infty} \delta^t (c_t - y_t)\right] \tag{1}$$

$$\text{subject to} \quad E\left[\sum\nolimits_{s=t}^{\infty} \beta^{s-t} u(c_s) | \mathscr{F}_t\right] - \underline{U}(y_t) \ge 0, \text{ for } t \ge 1, \tag{2}$$

$$E\left[\sum\nolimits_{t=0}^{\infty} \beta^t u(c_t)\right] = U_0. \tag{3}$$

Before characterizing the optimal solution to problem (1), I study as a benchmark the first-best allocation with full commitment (i.e., the optimal solution to (1) where the participation constraints (2) are absent). If the first-best allocation exists, then the first-order condition at the associated consumption levels is $u'(c_t) = \beta\delta^{-1}u'(c_{t+1})$. Let

$$f(c) \equiv (u')^{-1}(\beta^{-1}\delta u'(c))$$

represent what the next period's consumption would be in the first best if the current consumption is c. Note that $c < f(c)$ (= or >) if and only if $\delta < \beta$ (= or >). Moreover $\lim_{t\to\infty} f^t(c) = \infty$ when $\delta < \beta$ and $\lim_{t\to\infty} f^t(c) = 0$ when $\delta > \beta$.

The following assumption guarantees the existence of the first-best allocation.

ASSUMPTION 3. *If $\delta < \beta$ and $\bar{u} = \infty$, then $\sum_{t=0}^{\infty} \beta^t u(f^t(c_0))$ is finite for all $c_0 > 0$, where f^t is the composition of f with itself for t times.*

If $u = \log(\cdot)$, then this assumption is satisfied, because $f^t(c_0) = \beta^t\delta^{-t}c_0$, and $\sum_{t=0}^{\infty} \beta^t \log(\beta^t\delta^{-t}c_0)$ is always finite. In general, this assumption requires some curvature of the utility function, since otherwise, the consumption path $f^t(c_0)$ may grow too fast and $\sum_{t=0}^{\infty} \beta^t u(f^t(c_0))$ may fail to be finite.

LEMMA 1. *Under Assumption 3, the first-best allocation exists for any promised utility $U_0 \in \left(-\infty, \frac{\bar{u}}{1-\beta}\right)$.*

3. A Motivating Example

To provide economic intuition, I first discuss the properties of the optimal contract in the simplest case, namely, when $\delta = \beta$, the outside option is autarky, and the income stream $\{y_t; t \geq 0\}$ is a time-varying but deterministic sequence. Such a sequence can be generated by a deterministic difference equation, hence may be viewed as a special Markov chain whose transitional probability distribution $\pi(\cdot|y_t)$ is degenerate. Consider a principal's problem in which $y_0 = \bar{y}_i$ and initial promised utility $U_0 = \underline{U}(\bar{y}_i)$, that is,

$$\min_{\{c_t; 0 \leq t < \infty\}} \quad \sum_{t=0}^{\infty} \beta^t (c_t - y_t) \tag{4}$$

$$\text{subject to} \quad \sum_{s=t}^{\infty} \beta^{s-t} u(c_s) - \underline{U}(y_t) \geq 0, \text{ for } t \geq 1, \tag{5}$$

$$\sum_{t=0}^{\infty} \beta^t u(c_t) = \underline{U}(\bar{y}_i), \tag{6}$$

where $\underline{U}(y_t) = \sum_{s=t}^{\infty} \beta^{s-t} u(y_s)$.

3.1. When Income Is Monotonic

It would be illuminating to first examine two extreme cases, namely, when income $\{y_t; t \geq 0\}$ is monotonically decreasing or increasing with t. In the former case, the first-best allocation $u(c_t) \equiv (1-\beta)\underline{U}(y_0), \forall t \geq 0$ satisfies (5) (because $\underline{U}(y_0) \geq \underline{U}(y_t)$), and thus is the optimal solution to (4). In the latter case, the autarkic allocation $c_t \equiv y_t, \forall t \geq 0$ is the solution to (4). Consumption is not smooth, so for efficiency, the principal wants to raise $u(c_t)$ and lower $u(c_{t+1})$ until they are equal. However, any such attempt would violate the participation constraint at $t+1$.

The zero insurance in the latter stands in sharp contrast to the complete insurance in the former. In the latter case, the agent in any period would like to borrow against the higher income in future periods, but the principal refuses to lend because such a loan would trigger a default in the future. In the former case, the principal is essentially a saving technology with which the agent saves resources from earlier to later periods to smooth consumption. Note that whenever the agent saves, the Euler equation is not distorted. This turns out to be a useful observation later on.

3.2. When Income Is Nonmonotonic

In this subsection I heuristically derive a characterization of the optimal allocation, which sheds light on the general case in Section 4. Assume the existence of an optimal solution and denote the initial consumption c_0 in the optimal solution by $\bar{c}(\bar{y}_i)$. In problem (4), consumption is nondecreasing, and $c_{t-1} < c_t$ if and only if the participation constraint in t binds.[3] If $\tau_i^* \geq 1$ is the first time when the participation constraint binds, then consumption is perfectly smoothed before τ_i^*, i.e., $c_t = c_0 = \bar{c}(\bar{y}_i)$ for all $t \leq \tau_i^* - 1$. I can pin down $\bar{c}(\bar{y}_i)$ by the promise-keeping constraint $\sum_{t=0}^{\infty} \beta^t u(y_t) = \sum_{t=0}^{\tau_i^*-1} \beta^t u(\bar{c}(\bar{y}_i)) + \beta^{\tau_i^*} \underline{U}(y_{\tau_i^*})$,

$$\bar{c}(\bar{y}_i) = u^{-1}\left(\frac{\sum_{t=0}^{\tau_i^*-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau_i^*-1} \beta^t} \right) \geq u^{-1}\left(\min_{\tau \geq 1} \frac{\sum_{t=0}^{\tau-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau-1} \beta^t} \right).$$

I then argue that

$$\bar{c}(\bar{y}_i) = u^{-1}\left(\min_{\tau \geq 1} \frac{\sum_{t=0}^{\tau-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau-1} \beta^t} \right). \tag{7}$$

This is certainly correct for the two extreme cases discussed above. When income is monotonically decreasing, the participation constraint never binds (i.e., $\tau_i^* = \infty$), which is consistent with (7) because $\frac{\sum_{t=0}^{\tau-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau-1} \beta^t}$ is monotonically decreasing in τ. When income is increasing, the participation constraint binds in every period (i.e., $\tau_i^* = 1$), which is consistent with (7) because $\frac{\sum_{t=0}^{\tau-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau-1} \beta^t}$ is increasing in τ.

---

[3] Let $\beta^t \alpha_t$ be the nonnegative Lagrange multiplier on (5) and ϕ be the multiplier on (6). The first-order condition with respect to c_t is

$$1 = u'(c_t)(\phi + \alpha_1 + \alpha_2 + ... + \alpha_t).$$

Hence $c_{t-1} < c_t$ if and only if $\alpha_t > 0$.

To see (7) more generally, it remains to show that

$$\bar{c}(\bar{y}_i) \leq \min_{\tau \geq 1} u^{-1} \left(\frac{\sum_{t=0}^{\tau-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau-1} \beta^t} \right). \tag{8}$$

Pick τ' and τ'' such that $\tau' < \tau_i^* < \tau''$. The participation constraints at τ' and τ'' are, respectively,

$$\begin{aligned} \sum\nolimits_{t=\tau'}^{\infty} \beta^t u(c_t) &\geq \sum\nolimits_{t=\tau'}^{\infty} \beta^t u(y_t), \\ \sum\nolimits_{t=\tau''}^{\infty} \beta^t u(c_t) &\geq \sum\nolimits_{t=\tau''}^{\infty} \beta^t u(y_t). \end{aligned}$$

It then follows from equation (6) that

$$\begin{aligned} \sum\nolimits_{t=0}^{\tau'-1} \beta^t u(c_t) &\leq \sum\nolimits_{t=0}^{\tau'-1} \beta^t u(y_t), \\ \sum\nolimits_{t=0}^{\tau''-1} \beta^t u(c_t) &\leq \sum\nolimits_{t=0}^{\tau''-1} \beta^t u(y_t). \end{aligned}$$

Since $c_t = \bar{c}(\bar{y}_i)$ for all $t \leq \tau_i^* - 1$ and $c_t > \bar{c}(\bar{y}_i)$ for all $t \geq \tau_i^*$, then

$$\begin{aligned} \sum\nolimits_{t=0}^{\tau'-1} \beta^t u(\bar{c}(\bar{y}_i)) &\leq \sum\nolimits_{t=0}^{\tau'-1} \beta^t u(y_t), \\ \sum\nolimits_{t=0}^{\tau''-1} \beta^t u(\bar{c}(\bar{y}_i)) &\leq \sum\nolimits_{t=0}^{\tau''-1} \beta^t u(y_t), \end{aligned}$$

which imply inequality (8).

In equation (7), it is relatively easy to understand the perfect smoothing of consumption from 0 to $\tau - 1$, but it may be less clear why we need to find a τ to *minimize* $\frac{\sum_{t=0}^{\tau-1} \beta^t u(y_t)}{\sum_{t=0}^{\tau-1} \beta^t}$. We can obtain some intuition from the following example. Suppose $y_0 > y_2 > y_1$, and $y_s = y_2, \forall s \geq 2$ (see Figure 1). Starting from autarky, the principal certainly has incentive to smooth consumption in the first two periods by moving utility from period 0 to 1, since doing this does not violate any participation constraint and reduces the principal's cost. To maintain the same promised utility, the new consumption in the first two periods $\tilde{c}$ satisfies $u(\tilde{c}) + \beta u(\tilde{c}) = u(y_0) + \beta u(y_1)$, thus $\tilde{c} = u^{-1}\left(\frac{u(y_0)+\beta u(y_1)}{1+\beta}\right)$. The next question is: Can the principal provide more consumption smoothing than this? The answer critically depends on whether $\tilde{c} > y_2$ holds or not. If $\tilde{c} > y_2$ (see Figure 2), then the principal can move utility from the first two periods to later periods, for the same reason as before. Doing this lowers the consumption in the first two periods. However, if $\tilde{c} \leq y_2$ (see Figure 3), the principal wants to move utility from later periods to the first two periods but cannot achieve it because of the agent's default risk. In contrast to Figure 2, consumption smoothing in Figure 3 raises the consumption in the first two periods. To summarize, whether more consumption smoothing is feasible is indicated by whether

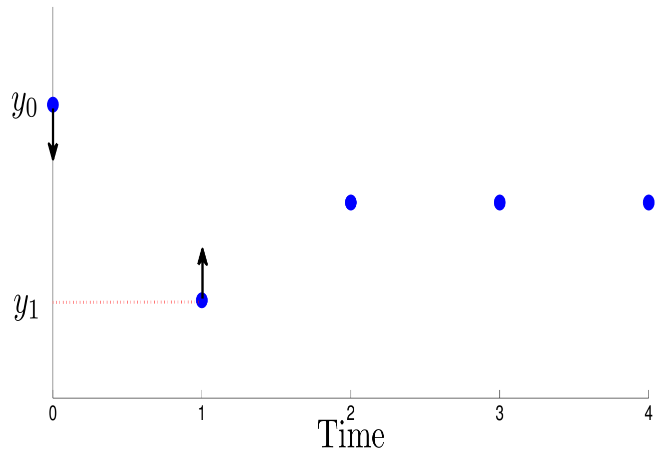

Figure 1: Endowments.

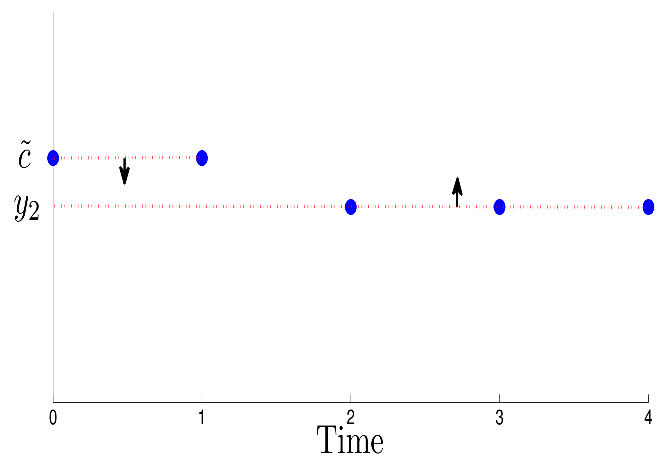

Figure 2: More consumption smoothing is feasible.

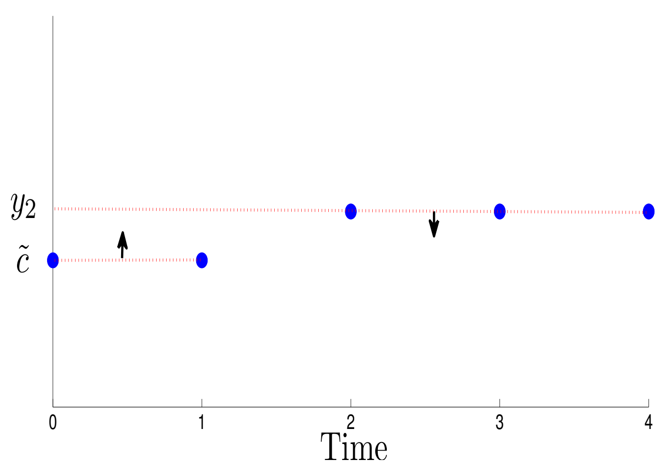

Figure 3: More consumption smoothing is infeasible.

consumption in the first two periods can be further reduced. By searching for a τ that minimizes the average consumption from 0 to $\tau-1$, the principal tries to achieve the most consumption smoothing that is feasible.

After computing $\bar{c}(\bar{y}_i)$ for all $\bar{y}_i$, consumption follows the recursive rule:

$$c_t = \max\{c_{t-1}, \bar{c}(y_t)\}, \quad t \geq 1. \tag{9}$$

To see (9), recall that $c_{t-1} < c_t$ if and only if the participation constraint in t binds. If it binds, then the continuation utility at t is $\underline{U}(y_t)$ and $c_t = \bar{c}(y_t)$.

4. A Characterization of the Optimal Contract

Based on the example in Section 3, I characterize the optimal contract as follows. First, the minimum consumption level $\bar{c}(\bar{y}_i)$ is defined by a generalization of equation (7). The generalization allows τ to be a stopping time, since when income is stochastic, the first time when the participation constraint binds is no longer deterministic. Then, I show that $\bar{c}(\bar{y}_i)$ is the initial consumption when $y_0 = \bar{y}_i$ and the promised utility is $\underline{U}(\bar{y}_i)$. Meanwhile, I guess and verify (a generalization of) the recursive rule in (9).

4.1. The Minimum Consumption Levels

This subsection defines a minimum level of consumption $\bar{c}(y)$ for each income state y. To do this, I first express the outside option as a discounted sum of flows (denoted as g). Specifically, I introduce a function $g : Y \to \mathbb{R}$ by

$$g(\bar{y}_i) = \underline{U}(\bar{y}_i) - \beta \sum_{j=1}^{N} \underline{U}(\bar{y}_j)\pi(\bar{y}_j|\bar{y}_i),$$

which leads to a natural identity

$$\underline{U}(\bar{y}_i) = E\left[\sum_{t=0}^{\infty} \beta^t g(y_t)|y_0 = \bar{y}_i\right].$$

If today's income is y and the agent receives $g(y)$, then he is indifferent between defaulting today and tomorrow. If the outside option is autarky, then $g(\cdot) = u(\cdot)$.

A random time τ is a *stopping time* of the filtration $\{\mathscr{F}_t; t \geq 0\}$, if the event $\{\tau \leq t\}$ belongs to $\mathscr{F}_t$ for every $t \geq 0$. For a stopping time $\tau \geq 1$, let $d_i(\tau)$ be the solution to

$$\begin{aligned} E\left[\sum_{t=0}^{\tau-1} \beta^t u(f^t(d_i(\tau)))|y_0 = \bar{y}_i\right] &= \underline{U}(\bar{y}_i) - E\left[\beta^\tau \underline{U}(y_\tau)|y_0 = \bar{y}_i\right] \\ &= E\left[\sum_{t=0}^{\tau-1} \beta^t g(y_t)|y_0 = \bar{y}_i\right]. \end{aligned} \tag{10}$$

Note that Assumption 3 implies that $E\left[\sum_{t=0}^{\tau-1} \beta^t u(f^t(c_0))|y_0 = \bar{y}_i\right]$ is finite for any c_0 and any $\tau \geq 1$. If $E\left[\sum_{t=0}^{\tau-1} \beta^t \bar{u}|y_0 = \bar{y}_i\right] > \underline{U}(\bar{y}_i) - E\left[\beta^\tau \underline{U}(y_\tau)|y_0 = \bar{y}_i\right]$, it follows from

Assumption 1 (iii) that $d_i(\tau)$ exists, is unique, and is finite; otherwise the solution does not exist and I simply let $d_i(\tau) = \infty$. In particular, for the stopping time $\tau \equiv \infty$, because $\frac{\bar{u}}{1-\beta} > \underline{U}(\bar{y}_i)$, $d_i(\infty)$ is finite. Note that if the first-best allocation is delivered from 0 to $\tau - 1$ and the agent's initial consumption is $d_i(\tau)$, then he is indifferent between defaulting at 0 and at τ, conditional on $y_0 = \bar{y}_i$. The minimum consumption $\bar{c}(\bar{y}_i)$ is defined as

$$\bar{c}(\bar{y}_i) = \inf_{\tau \geq 1} d_i(\tau), \tag{11}$$

where the infimum is over all stopping times $\tau \geq 1$. For notational simplicity, I write $\bar{c}(\bar{y}_i)$, $g(\bar{y}_i)$ as $\bar{c}_i$, g_i.

Lemma 2. *For all i, $\bar{c}_i$ is well-defined and finite. If $\delta \geq \beta$, then $u^{-1}(\min_j g_j) \leq \bar{c}_i \leq d_i(\infty)$. If $\delta < \beta$, then $e \leq \bar{c}_i \leq d_i(\infty)$, where e satisfies*

$$\sum_{t=0}^{\infty} \beta^t u(f^t(e)) = \frac{\min_j g_j}{1-\beta}.$$

4.2. Guess and Verify

I guess that the consumption in the optimal contract is updated recursively by

$$c_t = \max\{f(c_{t-1}), \bar{c}(y_t)\}, t \geq 1. \tag{12}$$

That is, consumption deviates each period from the first-best level by the smallest amount necessary to bring it above some minimum level. Equation (12) reveals the dynamics of consumption without using value functions.[4]

In (12), c_0 is set at a value to satisfy equation (3). Since the left side of (3) is continuous and strictly increasing in c_0 (when $\{c_t; t \geq 1\}$ are all interpreted as functions of c_0), such a value uniquely exists, which I denote by $c_0(U_0, y_0)$. The next lemma shows that $c_0(\underline{U}(\bar{y}_i), \bar{y}_i) = \bar{c}_i$ (thus $c_0(U_0, \bar{y}_i) > \bar{c}_i$ for $U_0 > \underline{U}(\bar{y}_i)$). This and (12) imply that consumption c_t stays weakly above $\bar{c}(y_t)$ for all t (including $t = 0$), which justifies the name "minimum consumption levels" for $\bar{c}(\cdot)$.

Lemma 3. *If $y_0 = \bar{y}_i$ and $c_0 = \bar{c}_i$, then $\{c_t; t \geq 0\}$ defined in (12) delivers $\underline{U}(\bar{y}_i)$ to the agent.*

It remains to verify that $\{c_t; t \geq 0\}$ is indeed the optimal contract.

Theorem 1. *If $y_0 = \bar{y}_i$, $U_0 \geq \underline{U}(\bar{y}_i)$ and $c_0 = c_0(U_0, \bar{y}_i)$, then $\{c_t; t \geq 0\}$ defined in (12) satisfies participation constraints (2) for all $t \geq 1$, and is the optimal solution to (1).*

[4]It is derived by a number of papers in the literature (cf. Thomas and Worrall (1988, 2007), Ljungqvist and Sargent (2004), Krueger and Perri (2006, 2011), Broer (2009)).

4.3. Stationary Distribution of Consumption

In this subsection, I briefly discuss the asymptotic behavior of consumption. If the agent is more patient than the principal, then the agent's consumption has a upward drift and diverges to infinity in the long run. If they are equally patient, then the first-best consumption path has no drift; consumption stays at $\max_i \bar{c}_i$ forever after reaching it the first time. This is consistent with the findings in Ray (2002), who shows that in environments where the principal and the agent are equally patient and the agent lacks commitment, the optimal allocation eventually exhibits a continuation that maximizes the agent's payoff over all self-enforcing sequences. In these cases, the long-run consumption either diverges or its distribution is degenerate.

If the principal is more patient, then the agent's consumption has a downward drift, and the results in Ray (2002) no longer hold, i.e., the agent's participation constraint binds even in the long run. Thus, the model with a more patient principal predicts a nontrivial stationary distribution of consumption. The agent's consumption in the long run is always between $\min_i \bar{c}_i$ and $\max_i \bar{c}_i$. Moreover, since $c_t = f^{t-s}(\bar{c}(y_s))$, where $s \le t$ is the last period with a binding participation constraint, we have

$$c_t \in \{f^n(\bar{c}_i) : 0 \le n \le T, 1 \le i \le N\}.$$

where T satisfies $f^T(\max_i \bar{c}_i) < \min_i \bar{c}_i$. That is, T is the maximal number of periods between two binding participation constraints. Since T is finite, the stationary distribution of consumption has a finite support. Similar finite-support results are derived in Krueger and Uhlig (2006, Proposition 17) and Broer (2009, Proposition 1).

5. Computation

Because searching over all stopping times in (11) is computationally prohibitive, I need an alternative method to calculate the vector $\bar{c} : Y \to \mathbb{R}_{++}$. In this section, I develop such an algorithm. In the following, I will sketch the main idea, leaving a detailed explanation to the Appendix.

For a sequence $A \equiv \{A_t \subseteq Y; t \ge 1\}$ of subsets of Y, let $\Gamma(A)$ denote the exit time of A, i.e.,

$$\Gamma(A) = \min_t \{t \ge 1 : y_t \notin A_t\}. \tag{13}$$

The following lemma states that the optimal stopping time in (11) takes the form of an exit time.

LEMMA 4. *For all i, $\bar{c}_i = d_i(\tau_i^*)$ where $\tau_i^* = \Gamma(A(\bar{c}_i))$, $A(\bar{c}_i) \equiv \{A_t(\bar{c}_i); t \ge 1\}$, $A_t(\bar{c}_i) = \{y \in Y : \bar{c}(y) \le f^t(\bar{c}_i)\}, t \ge 1$.*

To compute $\bar{c}(\cdot)$, I first consider the case of $\delta = \beta$. Suppose further that the ordering of $\bar{c}_i$ concurs with the ordering of the states. The equation $\bar{c}_i = d_i(\Gamma(A(\bar{c}_i)))$ in Lemma 4 is $V_i^{(i)} = 0$, where

$$V_k^{(i)} = E\left[\sum\nolimits_{t=0}^{\Gamma(A(\bar{c}_i))-1} \beta^t (u(\bar{c}_i) - g(y_t))|y_0 = \bar{y}_k\right], 1 \le k \le i.$$

The vector $\{V_k^{(i)}; 1 \le k \le i\}$ satisfies a linear system of equations

$$V_k^{(i)} = u(\bar{c}_i) - g(\bar{y}_k) + \beta \sum\nolimits_{j=1}^{i} \pi_{kj} V_j^{(i)}, 1 \le k \le i, \tag{14}$$

from which $u(\bar{c}_i)$ and $V_k^{(i)}, 1 \le k \le i$ can be easily obtained.[5]

Second, consider the case of $\delta > \beta$. Because $f^t(c)$ decreases with t, Lemma 4 states that for any income state $\bar{y}_i$, I could obtain $\bar{c}_i$ through a fixed-point iteration, as long as the collection of states with values of $\bar{c}(\cdot)$ below $\bar{c}_i$ is known. This means that, if I know the ordering of $\bar{c}(\cdot)$, then I can compute $\bar{c}(\cdot)$ from smaller to larger values. Although the algorithm requires a separate fixed-point iteration for each income state, it is still faster than value-function iteration because all fixed-point iterations terminate in finite steps.

Complications arise when the ordering of $\bar{c}(\cdot)$ is unknown a priori, and hence needs to be obtained along with the computation of $\bar{c}(\cdot)$ from smaller to larger values. More specifically, suppose I have obtained the values of $\bar{c}(\cdot)$ on B, where B is a collection of income states with smaller values of $\bar{c}(\cdot)$. Lemma 4 allows me to compute $\bar{c}_i$, where $\bar{c}_i$ is the next higher value of $\bar{c}(\cdot)$ (i.e., the minimum of $\bar{c}(\cdot)$ on B^c). The challenge here is that, because I have not found the values of $\bar{c}(\cdot)$ on B^c yet, I do not know at which state the next higher value of $\bar{c}(\cdot)$ is achieved. To solve the problem, I compute an auxiliary vector ($\tilde{c}(\cdot)$ in step 2 below) that has the same minimizer as $\bar{c}(\cdot)$ on B^c. The computation of the auxiliary vector is feasible as it depends only on the known values of $\bar{c}(\cdot)$ on B.

The following algorithm implements the above idea. Let B denote the set of states for which $\bar{c}$ has been found, while B^c denotes the set of states for which $\bar{c}$ is still unknown; $\max_{\bar{y}_i \in B} \bar{c}_i < \min_{\bar{y}_i \in B^c} \bar{c}_i$ is satisfied throughout the computation. The algorithm expands B point by point until it terminates at $B = Y$.

---

[5] Using (14), I can easily reproduce the results in Ljungqvist and Sargent (2004, Chapter 19). If the outside option is autarky and income is i.i.d., then $g(\cdot) = u(\cdot)$ and

$$0 = \quad V_i^{(i)} = u(\bar{c}_i) - u(\bar{y}_i) + \beta \sum\nolimits_{j=1}^{i} \pi_j V_j^{(i)}, \tag{15}$$

$$V_k^{(i)} = u(\bar{c}_i) - u(\bar{y}_k) + \beta \sum\nolimits_{j=1}^{i} \pi_j V_j^{(i)}, 1 \le k \le i-1. \tag{16}$$

Subtracting (15) from (16) yields $V_k^{(i)} = u(\bar{y}_i) - u(\bar{y}_k)$. Substituting this into (15) yields $u(\bar{c}_i) = u(\bar{y}_i) - \beta \sum_{j=1}^{i} \pi_j (u(\bar{y}_i) - u(\bar{y}_j))$, which is equation (19.3.25) in Ljungqvist and Sargent (2004).

Algorithm for the case of $\delta \geq \beta$:

- Step 1. This step initializes B as the collection of states at which $\bar{c}(\cdot)$ is minimized. To do so, set $B = \{y \in Y : g(y) = \min_i g_i\}$ and $\bar{c}(y) = u^{-1}(g(y))$ for $y \in B$. The Appendix provides a detailed explanation for why $\{y : \bar{c}(y) = \min_i \bar{c}_i\} = \{y : g(y) = \min_i g_i\}$, as well as explanations for the rest of this algorithm.

- Step 2. This step computes an auxiliary vector $\tilde{c} : B^c \to \mathbb{R}$, such that it has the same minimizer as $\bar{c}(\cdot)$ on B^c, as follows. For each $\bar{y}_i \in B^c$, choose an initial sequence of sets $C^0 = \{C_t^0; t \geq 1\}$ where $C_t^0 = B \cup \{\bar{y}_i\}, \forall t$. Compute $x^0 = d_i(\Gamma(C^0))$ and set $k = 0$.

  - Step 2.1 Compute the next iterate: $x^{k+1} = d_i(\Gamma(C^{k+1}))$, where $C^{k+1} = \{C_t^{k+1}; t \geq 1\}$ and $C_t^{k+1} = \{y \in B : \bar{c}(y) \leq f^t(x^k)\}$ for $t \geq 1$.
  - Step 2.2 Check the stopping criterion: If $x^{k+1} = x^k$, let $\tilde{c}_i = x^k$ and go to step 2.3; else go to step 2.1.
  - Step 2.3 If $\tilde{c}_i$ has been obtained for all $\bar{y}_i \in B^c$, go to step 3; else compute $\tilde{c}$ for the next point in B^c.

- Step 3. This step expands B. To do so, replace B by $B \cup D$, where $D \equiv \{y \in B^c : \tilde{c}(y) = \min_{\bar{y}_i \in B^c} \tilde{c}_i\}$. Let $\bar{c}(y) = \tilde{c}(y)$ for $y \in D$.

- Step 4. This step decides whether to terminate the computation. If $B = Y$, stop; else go to step 2 with the updated B.

Last, consider the case of $\delta < \beta$. The algorithm is analogous, except that here I compute $\bar{c}(\cdot)$ from larger to smaller values. That is, the states for which $\bar{c}$ has been found, B, are states with larger values of $\bar{c}(\cdot)$ and $\min_{\bar{y}_i \in B} \bar{c}_i > \max_{\bar{y}_i \in B^c} \bar{c}_i$ is satisfied throughout the computation.

Algorithm for the case of $\delta < \beta$:

- Step 1. This step initializes B as the collection of states at which $\bar{c}(\cdot)$ is maximized. To do so, set $B = \{y \in Y : \underline{U}(y) = \max_i \underline{U}(\bar{y}_i)\}$ and $\bar{c}_i = d_i(\infty)$ for $\bar{y}_i \in B$.

- Step 2. This step computes an auxiliary vector $\tilde{c} : B^c \to \mathbb{R}$, such that it has the same maximizer as $\bar{c}(\cdot)$ on B^c, as follows. For each $\bar{y}_i \in B^c$, choose an initial sequence of sets $C^0 = \{C_t^0; t \geq 1\}$ where $C_t^0 = Y, \forall t$. Compute $x^0 = d_i(\Gamma(C^0)) = d_i(\infty)$ and set $k = 0$.

  - Step 2.1 Compute the next iterate: $x^{k+1} = d_i(\Gamma(C^{k+1}))$, where $C^{k+1} = \{C_t^{k+1}; t \geq 1\}$ and $C_t^{k+1} = \{y \in B : \bar{c}(y) \leq f^t(x^k)\} \cup B^c$ for $t \geq 1$.
  - Step 2.2 Check the stopping criterion: If $x^{k+1} = x^k$, let $\tilde{c}_i = x^k$ and go to step 2.3; else go to step 2.1.

– Step 2.3 If $\tilde{c}_i$ has been obtained for all $\bar{y}_i \in B^c$, go to step 3; else compute $\tilde{c}$ for the next point in B^c.

- Step 3. This step expands B. To do so, replace B by $B \cup D$, where $D \equiv \{y \in B^c : y = \max_{\bar{y}_i \in B^c} \tilde{c}_i\}$. Let $\bar{c}(y) = \tilde{c}(y)$ for $y \in D$.

- Step 4. This step decides whether to terminate the computation. If $B = Y$, stop; else go to step 2 with the updated B.

I test the algorithm in a numerical example below.

5.1. A Numerical Example

I assume that the agent's utility is $\log(\cdot)$, the outside option is autarky, and set $\beta = 0.9 < \delta = 0.94$. The income process y_t is the sum of a persistent AR(1) process m_t, with persistence ρ and variance σ_m^2, and a completely transitory component ϵ_t which has mean zero and variance σ_ϵ^2. That is,

$$\begin{aligned} y_t &= m_t + \epsilon_t, \\ m_t &= \rho m_{t-1} + v_t. \end{aligned}$$

I employ the Tauchen and Hussey (1991) procedure to discretize the income process. In the benchmark, I use a two-state Markov chain for m_t and a two-state Markov chain for ϵ_t. Hence, there are four income levels. The parameter values are $\rho = 0.999$, $\sigma_m^2 = 0.2$, and $\sigma_\epsilon^2 = 0.1$. I compute the principal's cost using both the algorithm in this paper and value-function iteration (VFI).[6] I use a Fortran program on a PC with a 1.83GHz Intel CPU. In the rest of this subsection, I compare the speed of my algorithm with that of VFI. I then explain why the ordering of $\bar{c}(\cdot)$ is nontrivial. Finally, I compute the invariant distribution of consumption.

[6] In VFI, $C(\bar{y}_i, U)$ denotes the principal's cost function given promise U. For each $\bar{y}_i$, I choose a grid with 20 points as the domain of $C(\bar{y}_i, \cdot)$. The lower bound of the domain is $\underline{U}(\bar{y}_i)$, while the choice of upper bound is somewhat arbitrary (the results reported here are not sensitive to this choice). The Bellman equation is

$$\begin{aligned} C(\bar{y}_i, U) = & \min_{\{c, U_j'; j=1,\ldots,N\}} \left\{ c - \bar{y}_i + \delta \sum_{j=1}^{N} \pi(\bar{y}_j | \bar{y}_i) C(\bar{y}_j, U_j') \right\} \qquad (17) \\ & \text{subject to} \quad U = \log(c) + \beta \sum_{j=1}^{N} \pi(\bar{y}_j | \bar{y}_i) U_j', \\ & \qquad\qquad U_j' \geq \underline{U}(\bar{y}_j), \forall j. \end{aligned}$$

To solve the right side of the Bellman equation, I use the numerical minimizer BCPOL in the IMSL Math Library. Iteration terminates when $||TC(y, U) - C(y, U)||_\infty \leq 10^{-3}$. In contrast, the algorithm in this paper terminates in finite steps, and does not need a convergence criterion.

Table 1: Running times of the two algorithms in seconds.

Parameters	Value-function iteration	My algorithm
$\delta = 0.94, \lvert Y \rvert = 4$ (Benchmark)	3.9 s	1.0×10^{-2} s
$\delta = 0.98, \lvert Y \rvert = 4$	17.3 s	1.0×10^{-2} s
$\delta = 0.94, \lvert Y \rvert = 10$	82.3 s	3.0×10^{-2} s

5.1.1. Comparison of Algorithms

The first row in Table 1 reports the running times of the two algorithms. It shows that the algorithm in this paper is two orders of magnitude faster than VFI. There are three reasons why my algorithm is faster. First, for each state $\bar{y}_i$ VFI approximates and computes the cost function on the entire domain $[\underline{U}(\bar{y}_i), \frac{\bar{u}}{1-\beta})$, while I only compute a single value $\bar{c}_i$. Second, VFI involves an indefinite number of iterations, with the speed of convergence deteriorating as the discount factor approaches one; in contrast, the iterations in Steps 2.1-2.3 of my algorithm terminate in finite steps. In this example, it takes 71 iterations for VFI to converge, while the number of iterations is less than 10 in my algorithm. Third, my algorithm avoids any numerical minimizer, which is expensive in computation time. In contrast, VFI calls a minimizer at every node within one iteration.

The difference between the two algorithms becomes more dramatic in two experiments (see the last two rows of Table 1). In the first experiment, I increase δ from 0.94 to 0.98. It takes more iterations (and more time) for VFI to converge, while the running time of my algorithm is essentially unchanged. In the second experiment, I approximate m_t by a five-state Markov chain, i.e., the number of income states increases from four to ten. Both algorithms slow down, but VFI is affected more. For VFI, not only are there more grid points in the domain, computation at each point also slows down significantly. With a larger Y, the minimization problem on the right side of the Bellman equation has more choice variables (see equation (17)). Consequently, it takes more time for the numerical minimizer to find solutions.

VFI may be fast enough in a simple setting where the optimal contract is computed once. However, when researchers need to compute the optimal contract repeatedly, a longer running time is problematic.[7] In instances such as this, a faster algorithm such as the one I develop provides a distinct and important advantage.

[7]In Krueger and Perri (2006), Krueger and Uhlig (2006), and Broer (2009), the principal's discount factor δ (or the market-clearing interest rate) must be endogenously determined so that aggregate consumption equals aggregate income. In searching for the market-clearing interest rate, researchers need to compute the optimal contract for a guessed δ and then update guesses. Even if the computation of one optimal contract takes only a few seconds, when it is repeated many times, it can slow down the whole computation significantly.

5.1.2. Ordering of $\bar{c}(\cdot)$

Because the outside option is autarky, $g(y)$ is equal to $u(y)$ and hence is monotonic in y. However, the minimum consumption $\bar{c}(\cdot)$ is not, due to the nonmonotonicity of the autarky value function. (For $g(\cdot)$ and $\bar{c}(\cdot)$ in the case of $|Y| = 10$, see Figure 4.) To understand why $\underline{U}(\cdot)$ is nonmonotonic, consider $\underline{U}(\bar{y}_3)$ and $\underline{U}(\bar{y}_4)$: $\underline{U}(\bar{y}_3) > \underline{U}(\bar{y}_4)$ because the income $\bar{y}_3$ has a larger persistent component but a smaller transitory component than $\bar{y}_4$, and the larger persistent component implies a higher autarky value. There is comovement between $\bar{c}(\cdot)$ and $\underline{U}(\cdot)$ because a higher outside option generally requires that more consumption be delivered to the agent. This comovement, however, is not perfect: Figure 5 shows that the ordering of $\bar{c}(\cdot)$ does not coincide with that of $\underline{U}(\cdot)$. This makes the ordering of $\bar{c}(\cdot)$ hard to obtain.

5.1.3. Stationary Distribution of Consumption

Because the principal is more patient than the agent in this example, there is a stationary distribution of long-run consumption, as discussed in subsection 4.3. Recall that the support of this distribution is $\{f^n(\bar{c}_i) : 0 \leq n \leq T, 1 \leq i \leq N\}$. That is, if the last period with a binding participation constraint is in income $\bar{y}_i$, then $c_t = f^n(\bar{c}_i)$ for some n. Figure 6 shows the stationary distribution of consumption in the case of $N = |Y| = 10$. For each i $(1 \leq i \leq N)$, consumption levels in $\{f^n(\bar{c}_i) : 0 \leq n \leq T\}$ share the same color.

6. Extension and Limitation

The characterization in this paper can be easily extended to allow for a risk-averse principal and non-Markov income process (or outside option process). For instance, if outside options depend on calendar time, then we may still use (11) to define the minimum consumption level, acknowledging that now it may depend on calendar time.

One limitation of this paper is that my results cannot easily generalize to models with endogenous outside options. In Kehoe and Perri (2002) and Albuquerque and Hopenhayn (2004), the agent in the event of default would retain his capital stock, and thus the autarky value is a function of the endogenous capital. Problems with endogenous capital (and outside options) should be split into two stages: In the first stage the principal allocates capital in each period, and in the second stage the principal delivers consumption subject to the agent's outside options determined in the first stage. While my paper addresses only the second stage, the problem in the first stage is more challenging and deserves further study in the future.

This paper focuses on one-sided commitment, which may be another limitation. Extending the stopping-time approach to two-sided limited-commitment problems is difficult, but not impossible. To see the difficulties, suppose τ^* is the first time when either the principal's or the agent's participation constraint binds. In a one-sided limited-commitment problem, the agent's continuation utility at τ^* equals his outside

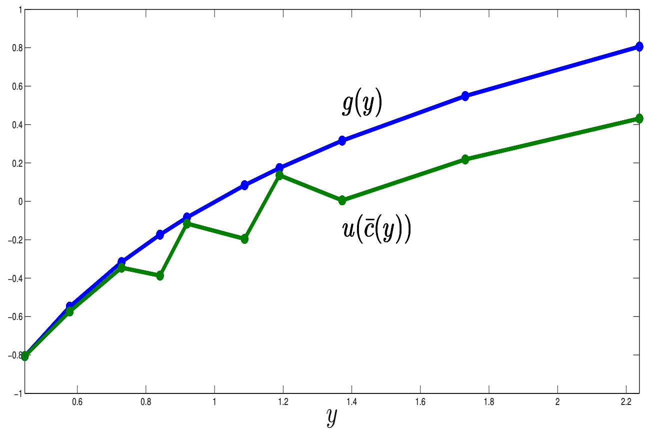

Figure 4: Ordering of $g(y)$ and $u(\bar{c}(y))$.

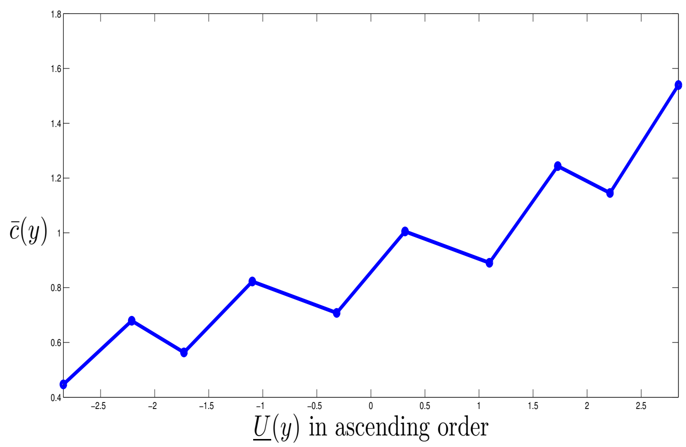

Figure 5: Ordering of $\bar{c}(\cdot)$ does not concur with that of $\underline{U}(\cdot)$.

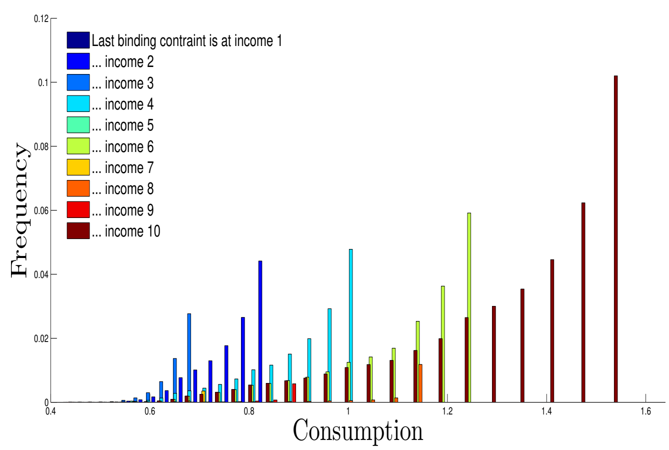

Figure 6: Distribution of consumption.

option. This continuation utility is used to infer the agent's endogenous utility flow before τ^*. In a two-sided limited-commitment problem, however, the agent's utility flow before τ^* cannot be easily inferred: the agent's continuation utility at τ^* is unknown a priori and must be determined together with the optimal stopping time, if the principal's participation constraint binds but the agent's constraint is slack at τ^*. In Grochulski and Zhang (2011, Appendix C), we make significant progress toward applying the stopping-time approach to this more challenging problem. Our assumptions of symmetric agents and Brownian motion income processes greatly simplify the computation of stopping times.

Acknowledgements

I am grateful to the Editor Paul Klein and two anonymous referees, whose comments greatly improve the paper. I also thank Hari Govindan, Borys Grochulski, Dirk Krueger, B Ravikumar for their comments and encouragement. All remaining errors are my own.

Appendix

Proof of Lemma 1: First, the first-best allocation exists if and only if $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0))$ is finite for all $c_0 > 0$. If the first-best allocation exists, then there is a $c^* > 0$, such that $\sum_{t=0}^{\infty}\beta^t u(f^t(c^*)) = U_0$. The finiteness of $\sum_{t=0}^{\infty}\beta^t u(f^t(c^*))$ is equivalent to the finiteness of $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0))$ for all $c_0 > 0$. On the converse, if $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0))$ is finite for all $c_0 > 0$, then $\lim_{c_0\to\infty}\sum_{t=0}^{\infty}\beta^t u(f^t(c_0)) = \frac{\bar{u}}{1-\beta}$ and $\lim_{c_0\to 0}\sum_{t=0}^{\infty}\beta^t u(f^t(c_0)) = -\infty$. For any $U_0 < \frac{\bar{u}}{1-\beta}$, the intermediate value theorem implies the existence of a $c^* > 0$ such that $\sum_{t=0}^{\infty}\beta^t u(f^t(c^*)) = U_0$.

In the rest of the proof, I verify that $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0))$ is always finite. If $\delta \geq \beta$, then it follows from $u(f^t(c_0)) \geq u(c_0) + u'(f^t(c_0))(f^t(c_0) - c_0)$ that

$$\begin{aligned}\frac{u(c_0)}{1-\beta} \geq \sum_{t=0}^{\infty}\beta^t u(f^t(c_0)) \ &\geq\ \sum_{t=0}^{\infty}\beta^t u(c_0) + \sum_{t=0}^{\infty}\beta^t\left(\frac{\delta}{\beta}\right)^t u'(c_0)(f^t(c_0) - c_0)\\ &=\ \frac{u(c_0)}{1-\beta} - \frac{u'(c_0)c_0}{1-\delta} + u'(c_0)\sum_{t=0}^{\infty}\delta^t f^t(c_0) > \frac{u(c_0)}{1-\beta} - \frac{u'(c_0)c_0}{1-\delta}.\end{aligned}$$

Thus both $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0))$ and $\sum_{t=0}^{\infty}\delta^t f^t(c_0)$ are finite. If $\delta < \beta$ and $\bar{u} < \infty$, then $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0)) \in [0, \frac{\bar{u}}{1-\beta}]$ is finite. If $\delta < \beta$ and $\bar{u} = \infty$, we assume that $\sum_{t=0}^{\infty}\beta^t u(f^t(c_0))$ is finite in Assumption 3. *Q.E.D.*

Proof of Lemma 2: The definition of $\bar{c}_i$ in (11) implies that $\bar{c}_i \leq d_i(\infty)$. To prove the lower bound when $\delta \geq \beta$, first note that $u^{-1}(\min_j g_j)$ is well defined because $\min_j g_j < \bar{u}$. It suffices to prove that $d_i(\tau) \geq u^{-1}(\min_j g_j)$ for all τ. Suppose $d_i(\tau) < u^{-1}(\min_j g_j)$ for some τ, then it follows from $c \geq f^t(c), \forall t$ that the left side of

(10) is strictly below the right side, which is a contradiction. Thus $u^{-1}(\min_j g_j) \leq \bar{c}_i$. It remains to prove the lower bound when $\delta < \beta$. Letting $\tilde{d}_i(\tau)$ be the solution to

$$E\left[\sum\nolimits_{t=0}^{\tau-1} \beta^t u(f^t(\tilde{d}_i(\tau)))|y_0 = \bar{y}_i\right] \quad = \quad E\left[\sum\nolimits_{t=0}^{\tau-1} \beta^t \min_j g_j|y_0 = \bar{y}_i\right],$$

it follows that $d_i(\tau) \geq \tilde{d}_i(\tau)$ and

$$\sum\nolimits_{t=0}^{\infty} P_t u(f^t(\tilde{d}_i(\tau))) = \min_j g_j,$$

where $P_t = \frac{\beta^t \Pr(t \leq \tau-1)}{\sum_{s=0}^{\infty} \beta^s \Pr(s \leq \tau-1)}$. By Abel's partial summation formula,

$$\begin{aligned}
&\sum\nolimits_{t=0}^{\infty} P_t u(f^t(\tilde{d}_i(\tau))) - \sum\nolimits_{t=0}^{\infty}(1-\beta)\beta^t u(f^t(\tilde{d}_i(\tau))) \\
= \;&\sum\nolimits_{t=0}^{\infty}\left(\sum\nolimits_{s=0}^{t} P_s - \sum\nolimits_{s=0}^{t}(1-\beta)\beta^s\right)\left(u(f^t(\tilde{d}_i(\tau))) - u(f^{t+1}(\tilde{d}_i(\tau)))\right) \\
= \;&\sum\nolimits_{t=0}^{\infty}\left(\frac{\sum_{s=0}^{t}\beta^s \Pr(s \leq \tau-1)}{\sum_{s=0}^{\infty}\beta^s \Pr(s \leq \tau-1)} - \frac{\sum_{s=0}^{t}\beta^s}{\sum_{s=0}^{\infty}\beta^s}\right)\left(u(f^t(\tilde{d}_i(\tau))) - u(f^{t+1}(\tilde{d}_i(\tau)))\right) \leq 0,
\end{aligned}$$

where the inequality holds because $\Pr(s \leq \tau-1)$ is nonincreasing in s, $\frac{\sum_{s=0}^{t}\beta^s \Pr(s \leq \tau-1)}{\sum_{s=0}^{\infty}\beta^s \Pr(s \leq \tau-1)} \geq \frac{\sum_{s=0}^{t}\beta^s}{\sum_{s=0}^{\infty}\beta^s}$ and $u(f^t(\tilde{d}_i(\tau))) < u(f^{t+1}(\tilde{d}_i(\tau)))$. It follows that

$$\min_j g_j = \sum\nolimits_{t=0}^{\infty} P_t u(f^t(\tilde{d}_i(\tau))) \leq \sum\nolimits_{t=0}^{\infty}(1-\beta)\beta^t u(f^t(\tilde{d}_i(\tau))).$$

Therefore, $e \leq \tilde{d}_i(\tau) \leq d_i(\tau)$ and since τ can be any stopping time, $e \leq \bar{c}_i$. *Q.E.D.*

Before proving Lemma 3, I first prove Lemma A.1 and Lemma 4 below as intermediate steps. Lemma A.1 and Lemma 4 discuss the properties of the optimal stopping time in (11). Let $\tau \geq 1$ be a stopping time and visualize the collection of nodes before the stopping time τ as a tree. Let $y^s \equiv (y_0, y_1, ..., y_s)$ be a finite-length history such that $s \geq 1$ and $y_0 = \bar{y}_i$. If $\tau(y^s) > s$, then the collection of nodes that start with y^s and end before τ represents a subtree. Suppose for the moment $\beta = \delta$ and $f(c) = c$. Then $u(d_i(\tau))$ is the average of the utility flow $g(\cdot)$ in the whole tree (see (10)); while $u(\bar{c}(y_s))$ is a lower bound for the average of the utility flow $g(\cdot)$ within the subtree, because $\bar{c}(y_s)$ is the infimum by definition. If $\bar{c}(y_s)$ is above $d_i(\tau)$, it implies that the average flow in the subtree exceeds the average flow in the whole tree; thus the subtree is *consumption-enhancing* in the tree, and modifying τ so that it stops at s (conditional on y^s) rather than proceeds further will lower $d_i(\tau)$. If $\beta \neq \delta$, then the condition for a *consumption-enhancing* subtree shall be modified to be $\bar{c}(y_s) > f^s(d_i(\tau))$. Lemma A.1 proves this intuition and Lemma 4 extends the intuition to show that the optimal stopping time is achieved when all *consumption-enhancing* subtrees are removed. To formally present Lemma A.1, recall $\Omega = Y^\infty$ and let $X(y^s) \equiv \{\omega \in \Omega : \omega_t = y_t, 0 \leq t \leq s\}$ denote the set of paths starting with

y^s. Note that $\tau(y^s) > s$ (or $= s$) is equivalent to $\tau(\omega) > s$ (or $= s$) for all $\omega \in X(y^s)$ because τ is a stopping time. Let τ' denote the modified τ that stops earlier on $X(y^s)$, that is

$$\tau'(\omega) = \begin{cases} \tau, & \text{if } \omega \notin X(y^s), \\ s, & \text{if } \omega \in X(y^s). \end{cases}$$

Lemma A.1. (i) *If $d_i(\tau) < \infty$, $\tau(\omega) > s$ for $\omega \in X(y^s)$, and $\bar{c}(y_s) > f^s(d_i(\tau))$, then $d_i(\tau') < d_i(\tau)$.*

(ii) *Suppose $\tau(\omega) = s$ for $\omega \in X(y^s)$ and $d_i(\tau) \leq M$, where M is a scalar. If $\bar{c}(y_s) < f^s(M)$, $y_s = \bar{y}_j$, $\chi \geq 1$ is a stopping time such that $d_j(\chi) < f^s(M)$, and τ'' is*

$$\tau''(\omega) = \begin{cases} \tau, & \text{if } \omega \notin X(y^s), \\ s + \chi(\omega_s, \omega_{s+1}, ...), & \text{if } \omega \in X(y^s), \end{cases}$$

where $(\omega_s, \omega_{s+1}, ...)$ is the tail of ω starting from period s, then $d_i(\tau'') < M$. In particular, if $d_i(\tau) = M$ and $\bar{c}_j < f^s(d_i(\tau))$, then $d_i(\tau'') < d_i(\tau)$.

Proof of Lemma A.1:

(i) Note that

$$\begin{aligned}
& E\left[\sum_{t=0}^{\tau-1} \beta^t g(y_t)\right] \\
= \ & E\left[\sum_{t=0}^{\tau-1} \beta^t u(f^t(d_i(\tau)))\right] \\
= \ & E\left[\sum_{t=0}^{\tau-1} \beta^t u(f^t(d_i(\tau)))1_{[X(y^s)]^c}\right] + E\left[\sum_{t=0}^{s-1} \beta^t u(f^t(d_i(\tau)))1_{X(y^s)}\right] \\
& + E\left[\sum_{t=s}^{\tau-1} \beta^t u(f^t(d_i(\tau)))1_{X(y^s)}\right] \\
= \ & E\left[\sum_{t=0}^{\tau'-1} \beta^t u(f^t(d_i(\tau)))\right] + E\left[\sum_{t=s}^{\tau-1} \beta^t u(f^t(d_i(\tau)))1_{X(y^s)}\right]. \qquad (18)
\end{aligned}$$

It follows from $\bar{c}(y_s) > f^s(d_i(\tau))$ that

$$\begin{aligned}
E\left[\sum_{t=s}^{\tau-1} \beta^t u(f^t(d_i(\tau)))1_{X(y^s)}\right] & < E\left[\sum_{t=s}^{\tau-1} \beta^t u(f^{t-s}(\bar{c}(y_s)))1_{X(y^s)}\right] \\
& \leq E\left[\sum_{t=s}^{\tau-1} \beta^t g(y_t)1_{X(y^s)}\right]. \qquad (19)
\end{aligned}$$

Subtracting (19) from (18) yields

$$E\left[\sum_{t=0}^{\tau'-1} \beta^t u(f^t(d_i(\tau)))\right] > E\left[\sum_{t=0}^{\tau'-1} \beta^t g(y_t)\right],$$

which implies that $d_i(\tau) > d_i(\tau')$.

(ii) The proof is analogous to that of part (i) and thus omitted.

Q.E.D.

PROOF OF LEMMA 4: If $\delta < \beta$, then $A_t(c)$ is nondecreasing in t and $A_t(c) = Y$ for sufficiently large t; while if $\delta > \beta$, then $A_t(c)$ is nonincreasing in t and $A_t(c) = \emptyset$ for sufficiently large t.

First it is possible to pick a sufficiently small $\epsilon > 0$ such that $A_t(\bar{c}_i) = A_t(\bar{c}_i + \epsilon)$ for all t. To see this, note that for each t, $A_t(\bar{c}_i) = A_t(\bar{c}_i + \epsilon)$ for sufficiently small $\epsilon > 0$. The uniformity in t when $\delta = \beta$ is because $A_t(\bar{c}_i)$ is independent of t. The uniformity in t when $\delta \neq \beta$ follows from the fact that $A_t(\bar{c}_i)$ differs from $A_t(\bar{c}_i + \epsilon)$ only finitely many times, because either both $A_t(\bar{c}_i)$ and $A_t(\bar{c}_i + \epsilon)$ are $\emptyset$ or both are Y when t is large.

Second I find a stopping time S such that $d_i(S) \leq \bar{c}_i + \epsilon$ and $S \geq \tau_i^*$. Pick a stopping time $S_1 \geq 1$ such that $d_i(S_1) < \bar{c}_i + \epsilon$. If $S_1 \geq \tau_i^*$, then the process is done; otherwise, $\{S_1 < \tau_i^*\} = \cup_{t=1}^{\infty}\{S_1 = t, \tau_i^* > t\} \neq \emptyset$. Letting t_1 be the smallest t such that $\{S_1 = t, \tau_i^* > t\} \neq \emptyset$, it follows from $t_1 < \tau_i^*$ that $\omega_{t_1} < f^{t_1}(\bar{c}_i + \epsilon)$ for all $\omega \in \{S_1 = t_1, \tau_i^* > t_1\}$. Since the number of distinct finite-length histories $(y_0, y_1, ..., y_{t_1})$ in $\{S_1 = t_1, \tau_i^* > t_1\}$ is finite, apply Lemma A.1 (ii) and append S_1 (history by history) finitely many times on the set $\{S_1 = t_1, \tau_i^* > t_1\}$ to obtain a stopping time S_2, such that

$$d_i(S_2) < \bar{c}_i + \epsilon.$$

$S_1 < S_2$ on $\{S_1 = t_1, \tau_i^* > t_1\}$ implies that $\{S_2 = t_1, \tau_i^* > t_1\} = \emptyset$, i.e., if t_2 is the smallest t such that $\{S_2 = t, \tau_i^* > t\} \neq \emptyset$, then $t_2 > t_1$. In general, as long as $\{S_n < \tau_i^*\}$ is not empty, S_n can be appended (history by history) on the set $\{S_n = t_n, \tau_i^* > t_n\}$ to obtain a stopping time S_{n+1}, such that

$$d_i(S_{n+1}) < \bar{c}_i + \epsilon, \tag{20}$$

where t_n is the smallest t such that $\{S_n = t, \tau_i^* > t\} \neq \emptyset$. If $S = \lim_n S_n$, then it follows from (20) and $\lim_n t_n = \infty$ that $d_i(S) \leq \bar{c}_i + \epsilon$ and $S \geq \tau_i^*$.

Third, since $\bar{c}(y_{\tau_i^*}) > f^{\tau_i^*}(\bar{c}_i + \epsilon) \geq f^{\tau_i^*}(d_i(S))$, Lemma A.1 (i) states that removing all consumption-enhancing subtrees from τ_i^* to S lowers $d_i(S)$, that is

$$d_i(\tau_i^*) < d_i(S) \leq \bar{c}_i + \epsilon.$$

This implies that $\bar{c}_i = d_i(\tau_i^*)$, because ϵ can be arbitrarily small. *Q.E.D.*

PROOF OF LEMMA 3: Because Lemma 4 states that $f^t(\bar{c}_i) \geq \bar{c}(y_t)$ for $0 \leq t \leq \tau_i^* - 1$, $\{c_t; t \geq 0\}$ defined in (12) with $c_0 = \bar{c}_i$ satisfies $c_t = f^t(\bar{c}_i)$, for $0 \leq t \leq \tau_i^* - 1$.

Then construct a nondecreasing sequence of stopping times $\{\chi_n; 1 \leq n < \infty\}$ as follows. Let $\chi_1 = \tau_i^*$ and it follows from $c_t = f^t(\bar{c}_i)$, $0 \leq t \leq \tau_i^* - 1$ and Lemma 4 that

$$\begin{aligned} & E\left[\sum_{t=0}^{\chi_1 - 1} \beta^t u(c_t) | y_0 = \bar{y}_i\right] = E\left[\sum_{t=0}^{\chi_1 - 1} \beta^t u(f^t(\bar{c}_i)) | y_0 = \bar{y}_i\right] \\ = \; & E\left[\sum_{t=0}^{\chi_1 - 1} \beta^t u(f^t(d_i(\chi_1))) | y_0 = \bar{y}_i\right] = E\left[\sum_{t=0}^{\chi_1 - 1} \beta^t g(y_t) | y_0 = \bar{y}_i\right]. \end{aligned} \tag{21}$$

Let χ_2 be χ_1 appended by a stopping time $\tau^*_{y_{\tau_i^*}}(y_{\tau_i^*}, y_{\tau_i^*+1}, ...)$, that is, $\chi_2(y) = \tau_i^*(y) + \tau^*_{y_{\tau_i^*}}(y_{\tau_i^*}, y_{\tau_i^*+1}, ...)$, for $y \in \Omega$.[8] Since $f^{\tau_i^*}(\bar{c}_i) < \bar{c}(y_{\tau_i^*})$, $c_{\tau_i^*} = \bar{c}(y_{\tau_i^*})$. It follows again from Lemma 4 that $c_t = f^{t-\tau_i^*}(\bar{c}(y_{\tau_i^*}))$ for $\chi_1 \le t \le \chi_2 - 1$. Hence

$$E\left[\sum_{t=\chi_1}^{\chi_2-1} \beta^t u(c_t)|y_{\chi_1}\right] = E\left[\sum_{t=\chi_1}^{\chi_2-1} \beta^t u(f^{t-\chi_1}(\bar{c}(y_{\tau_i^*})))|y_{\chi_1}\right] = E\left[\sum_{t=\chi_1}^{\chi_2-1} \beta^t g(y_t)|y_{\chi_1}\right]. \quad (22)$$

Combining (21) and (22) yields

$$E\left[\sum\nolimits_{t=0}^{\chi_2-1} \beta^t u(c_t)|y_0 = \bar{y}_i\right] = E\left[\sum\nolimits_{t=0}^{\chi_2-1} \beta^t g(y_t)|y_0 = \bar{y}_i\right].$$

Inductively, letting χ_{n+1} be χ_n appended by a stopping time $\tau^*_{y_{\chi_n}}(y_{\chi_n}, y_{\chi_n+1}, ...)$, i.e., $\chi_{n+1}(y) = \chi_n(y) + \tau^*_{y_{\chi_n}}(y_{\chi_n}, y_{\chi_n+1}, ...)$, I have that

$$E\left[\sum\nolimits_{t=0}^{\chi_{n+1}-1} \beta^t u(c_t)|y_0 = \bar{y}_i\right] = E\left[\sum\nolimits_{t=0}^{\chi_{n+1}-1} \beta^t g(y_t)|y_0 = \bar{y}_i\right].$$

Since $\chi_n < \chi_{n+1}$ whenever χ_n is finite, this implies that $\lim_{n\to\infty} \chi_n = \infty$ almost surely. Note that $\{\sum_{t=0}^{\chi_n-1} \beta^t u(c_t); 1 \le n < \infty\}$ is a sequence of uniformly bounded random variables, because when $\delta \ge \beta$, $\min_i \bar{c}_i \le c_t \le \max_i \bar{c}_i$ is bounded and when $\delta < \beta$,

$$\begin{aligned}\sum\nolimits_{t=0}^{\chi_n-1} \beta^t |u(c_t)| &\le \sum\nolimits_{t=0}^{\chi_n-1} \beta^t |u(c_0)| + \sum\nolimits_{t=0}^{\chi_n-1} \beta^t (u(c_t) - u(c_0)) \\ &\le \frac{|u(c_0)|}{1-\beta} + \sum\nolimits_{t=0}^{\infty} \beta^t (u(c_t) - u(c_0)) < \infty,\end{aligned}$$

where the last inequality follows from Assumption 3.

Taking limit $n \to \infty$ and applying bounded convergence theorem yield

$$E\left[\sum_{t=0}^{\infty} \beta^t u(c_t)|y_0 = \bar{y}_i\right] = \lim_{n\to\infty} E\left[\sum_{t=0}^{\chi_n-1} \beta^t u(c_t)|y_0 = \bar{y}_i\right] = E\left[\sum_{t=0}^{\infty} \beta^t g(y_t)|y_0 = \bar{y}_i\right] = \underline{U}(\bar{y}_i).$$

Q.E.D.

Proof of Theorem 1: To verify the participation constraint (2) at any period t, note that the continuation utility is delivered by consumption $\{c_s; s \ge t\}$ which follows the recursive formula (12), while $\underline{U}(y_t)$ can be delivered by a consumption plan which starts with $\bar{c}(y_t)$ and follows the same recursive formula, according to Lemma 3. Hence it follows from $c_t \ge \bar{c}(y_t)$ that the participation constraint (2) is satisfied.

[8] If $\chi_1 = \infty$, simply let $\chi_2 = \chi_1 = \infty$.

As to the optimality of $\{c_t; t \geq 0\}$, note that the problem has a linear objective function and a convex feasibility set. For this convex optimization problem, Kuhn-Tucker conditions are sufficient for the optimality of $\{c_t; t \geq 0\}$. The principal's Lagrangian is

$$E\sum_{t=0}^{\infty}\delta^t\left[-(c_t-y_t)+\alpha_t\left[E\left[\sum_{s=t}^{\infty}\beta^{s-t}u(c_s)|\mathscr{F}_t\right]-\underline{U}(y_t)\right]\right]+\phi\left[E\left[\sum_{t=0}^{\infty}\beta^t u(c_t)\right]-U_0\right],$$

where $\{\alpha_t; t \geq 1\}$ is a stochastic process of nonnegative Lagrangian multipliers on the participation constraint (2) and ϕ is the strictly positive multiplier on the initial promise-keeping constraint (3). I shall construct Lagrange multipliers and then verify the sufficient Kuhn-Tucker conditions. Construct $\alpha_t = \frac{1}{u'(c_t)} - \frac{1}{u'(f(c_{t-1}))}$ and $\phi = \frac{1}{u'(c_0)}$. From $c_t \geq f(c_{t-1})$, it follows that all the multipliers are nonnegative. The first-order condition with respect to c_t is

$$\delta^t = \left[\beta^t\phi + \beta^{t-1}\delta\alpha_1 + \beta^{t-2}\delta^2\alpha_2 + ... + \delta^t\alpha_t\right]u'(c_t),$$

which holds because $\frac{\beta^{t-s}\delta^s}{u'(c_s)} = \frac{\beta^{t-s-1}\delta^{s+1}}{u'(f(c_s))}$ for $s \leq t$. The complementary slackness condition for α_t is

$$\alpha_t\left[E\left[\sum\nolimits_{s=t}^{\infty}\beta^{s-t}u(c_s)|\mathscr{F}_t\right]-\underline{U}(y_t)\right]=0,$$

which holds because Lemma 3 implies that $E\left[\sum_{s=t}^{\infty}\beta^{s-t}u(c_s)|\mathscr{F}_t\right]-\underline{U}(y_t)=0$, when $\alpha_t > 0$ (i.e., $c_t = \bar{c}(y_t) > f(c_{t-1})$). *Q.E.D.*

To show the validity of the algorithm, the following Lemma A.2 is needed.

LEMMA A.2. (i) *If $\delta < \beta$ and $c > \bar{c}_i$, then $c > d_i(\Gamma(A(c)))$, where $A(c) = \{A_t(c); t \geq 1\}$, $A_t(c) = \{y \in Y : \bar{c}(y) \leq f^t(c)\}$ and Γ is the exit time in* (13).

(ii) *If $\delta > \beta$ and $c > \bar{c}_i$, then $c > d_i(\Gamma(F(c)))$, where $F(c) = \{F_t(c); t \geq 1\}$, $F_t(c) = \{y \in Y : \bar{c}(y) \leq f^t(c), \bar{c}(y) < \bar{c}_i\}$ and Γ is the exit time in* (13).

PROOF OF LEMMA A.2:

(i) First note that $c > d_i(\Gamma(A(c)))$ is equivalent to

$$\begin{aligned} h(c) &\equiv E\left[\sum\nolimits_{t=0}^{\Gamma(A(c))-1}\beta^t u(f^t(c))|y_0=\bar{y}_i\right]-E\left[\sum\nolimits_{t=0}^{\Gamma(A(c))-1}\beta^t g(y_t)|y_0=\bar{y}_i\right] \\ &> 0. \end{aligned} \tag{23}$$

To prove (23), it is sufficient to show that $h(c)$ is strictly increasing in c, since Lemma 4 implies $h(\bar{c}_i) = 0$. For each t, $A_t(c)$ is a nondecreasing upper hemicontinuous correspondence in c, and there exists a T, such that $A_t(c) = Y$ for all $t \geq T$ and $c \geq \bar{c}_i$. If $c < c'$ and $A(c) = A(c')$, then $\Gamma(A(c)) = \Gamma(A(c'))$

and $h(c) < h(c')$ because $u(f^t(c)) < u(f^t(c'))$. In other words, $h(c)$ is strictly increasing and continuous when $A(c)$ is flat in c. To finish the proof, I show that $h(c)$ is continuous, even if $A(c)$ is discontinuous at some c^* (i.e., there exists a t, such that $\tilde{A}_t \equiv \lim_{c \uparrow c^*} A_t(c) \subsetneq A_t(c^*)$). To prove this, pick $\bar{y}_j \in A_t(c^*) \backslash \tilde{A}_t$. Because $\bar{c}_j = f^t(c^*)$, it follows from Lemma 4 that

$$\begin{aligned} & E\left[\sum\nolimits_{s=t}^{\Gamma(A(c^*))-1} \beta^s u(f^s(c^*))|y_t = \bar{y}_j\right] \\ = \; & E\left[\sum\nolimits_{s=t}^{\Gamma(A(c^*))-1} \beta^s u(f^{s-t}(\bar{c}_j))|y_t = \bar{y}_j\right] = E\left[\sum\nolimits_{s=t}^{\Gamma(A(c^*))-1} \beta^s g(y_s)|y_t = \bar{y}_j\right]. \end{aligned}$$

This implies that

$$\begin{aligned} \lim_{c \uparrow c^*} h(c) &= E\left[\sum\nolimits_{t=0}^{\Gamma(\tilde{A})-1} \beta^t u(f^t(c^*))|y_0 = \bar{y}_i\right] - E\left[\sum\nolimits_{t=0}^{\Gamma(\tilde{A})-1} \beta^t g(y_t)|y_0 = \bar{y}_i\right] \\ &= E\left[\sum\nolimits_{t=0}^{\Gamma(A(c^*))-1} \beta^t u(f^t(c^*))|y_0 = \bar{y}_i\right] - E\left[\sum\nolimits_{t=0}^{\Gamma(A(c^*))-1} \beta^t g(y_t)|y_0 = \bar{y}_i\right] \\ &= h(c^*), \end{aligned}$$

where $\tilde{A} = \{\tilde{A}_t; t \geq 1\}$. Therefore, $h(c)$ is continuous at any c. Because $h(c)$ is strictly increasing at continuous points of $A(c)$, and the number of discontinuous points of $A(c)$ is finite, $h(c)$ is strictly increasing in c.

(ii) The proof is analogous to that of part (i) and thus omitted.

Q.E.D.

Explanation of algorithm ($\delta \geq \beta$):

- **Step 1.**
  - I show that $\{y : \bar{c}(y) = \min_i \bar{c}_i\} = \{y : g(y) = \min_i g_i\}$ and $\min_i \bar{c}_i = u^{-1}(\min_i g_i)$.
  - To prove $\{y : \bar{c}(y) = \min_i \bar{c}_i\} \supseteq \{y : g(y) = \min_i g_i\}$, suppose $g(\bar{y}_j) = \min_i g_i$ for some j. The definition of $\bar{c}_j$ implies that $\bar{c}_j \leq d_i(1) = u^{-1}(g_j)$. It follows from $u^{-1}(g_j) = u^{-1}(\min_i g_i) \leq \bar{c}_i, \forall i$ (see Lemma 2) that $\bar{c}_j = u^{-1}(g_j)$ and $\bar{c}_j \leq \bar{c}_i, \forall i$. Hence $\bar{c}_j = \min_i \bar{c}_i = u^{-1}(\min_i g_i)$.
  - To prove $\{y : \bar{c}(y) = \min_i \bar{c}_i\} \subseteq \{y : g(y) = \min_i g_i\}$, suppose $\bar{c}_n = \min_i \bar{c}_i = u^{-1}(\min_i g_i)$ for some n. Since $\bar{c}_n = d_n(\tau_n^*)$ and $f^t(\bar{c}_n) \leq \bar{c}_n = u^{-1}(\min_i g_i)$, equation (10) implies

$$\begin{aligned} E\left[\sum\nolimits_{t=0}^{\tau_n^*-1} \beta^t g(y_t)|y_0 = \bar{y}_n\right] &= E\left[\sum\nolimits_{t=0}^{\tau_n^*-1} \beta^t u(f^t(\bar{c}_n))|y_0 = \bar{y}_n\right] \\ &\leq E\left[\sum\nolimits_{t=0}^{\tau_n^*-1} \beta^t \min_i g_i|y_0 = \bar{y}_n\right], \end{aligned}$$

    which implies that $g_n = g(y_0) = \min_i g_i$. Hence $\bar{y}_n \in \{y : g(y) = \min_i g_i\}$.

- **Step 2.** I discuss only the case of $\delta > \beta$, as the case of $\delta = \beta$ is analogous.
  - **Goal:** Compute a vector $\tilde{c} : B^c \to \mathbb{R}$, such that $\tilde{c}_i = d_i(\Gamma(C))$ for $\bar{y}_i \in B^c$, where $C = \{C_t; t \geq 1\}$ and $C_t = \{y \in B : \bar{c}(y) \leq f^t(\tilde{c}_i)\}$ for $t \geq 1$.
  - As described in the algorithm, I compute $\tilde{c}$ by iteration. If $\delta > \beta$, the definition of C^{k+1} and Lemma A.1 (i) imply that $x^k \geq x^{k+1}$, which then implies that $C_t^{k+1} \supseteq C_t^{k+2}, \forall t$. It follows from induction that x^k decreases and C^k shrinks in k.
  - Here I show that the iteration of x^k finishes in finite steps. If T is such that $f^T(d_i(\infty)) < u^{-1}(\min_i g_i)$, then $C_t(x) = \{y \in B : \bar{c}(y) \leq f^t(x)\} = \emptyset$ for all $t \geq T$ and $x \leq d_i(\infty)$, which implies that there are only finitely many sequences of subsets C^k. Hence there exists a finite k^* at which $C^{k^*+1} = C^{k^*}$, and $x^{k^*+1} = d_i(\Gamma(C^{k^*+1})) = d_i(\Gamma(C^{k^*})) = x^{k^*}$.

- **Step 3.**
  - I show that $\tilde{c}(\cdot)$ and $\bar{c}(\cdot)$ have the same minimizer on B^c, i.e., $D \equiv \{y \in B^c : \tilde{c}(y) = \max_{\bar{y}_i \in B^c} \tilde{c}_i\} = \{y \in B^c : \bar{c}(y) = \min_{\bar{y}_i \in B^c} \bar{c}_i\}$, and the two minimums are identical.
  - $\bar{c}_i \leq \tilde{c}_i$ follows from the definition of $\tilde{c}_i$ in step 2 and $\bar{c}_i \leq d_i(\tau)$ for all τ. Therefore $\min_{\bar{y}_i \in B^c} \bar{c}_i \leq \min_{\bar{y}_i \in B^c} \tilde{c}_i$.
  - To show $\min_{\bar{y}_i \in B^c} \bar{c}_i \geq \min_{\bar{y}_i \in B^c} \tilde{c}_i$, suppose $\bar{c}_j = \min_{\bar{y}_i \in B^c} \bar{c}_i$ for some j. Since $C_t(\tilde{c}_j) = \{y \in B : \bar{c}(y) \leq f^t(\tilde{c}_j)\} = \{y \in Y : \bar{c}(y) \leq f^t(\tilde{c}_j), \bar{c}(y) < \bar{c}_j\} \equiv F_t(\tilde{c}_j)$ for all t, it follows from $\tilde{c}_j = d_j(\Gamma(C(\tilde{c}_j))) = d_j(\Gamma(F(\tilde{c}_j)))$ and Lemma A.2 (ii) that $\tilde{c}_j > \bar{c}_j$ is impossible. Therefore, $\tilde{c}_j = \bar{c}_j$ and $\min_{\bar{y}_i \in B^c} \bar{c}_i \geq \min_{\bar{y}_i \in B^c} \tilde{c}_i$.
  - To prove that $\{y \in B^c : \bar{c}(y) = \min_{\bar{y}_i \in B^c} \bar{c}_i\} = \{y \in B^c : \tilde{c}(y) = \min_{\bar{y}_i \in B^c} \tilde{c}_i\}$, note that if $\bar{c}_m > \min_{\bar{y}_i \in B^c} \bar{c}_i = \min_{\bar{y}_i \in B^c} \tilde{c}_i$ for some m, then $\tilde{c}_m \geq \bar{c}_m > \min_{\bar{y}_i \in B^c} \tilde{c}_i$.
  - Note that $\max_{\bar{y}_i \in B} \bar{c}_i < \min_{\bar{y}_i \in B^c} \bar{c}_i$ is maintained after B is replaced by $B \cup D$.

Explanation of algorithm ($\delta < \beta$):

- **Step 1.**
  - I show that $\{y : \bar{c}(y) = \max_i \bar{c}_i\} = \{y : \underline{U}(y) = \max_i \underline{U}(\bar{y}_i)\}$.
  - To prove $\{y : \bar{c}(y) = \max_i \bar{c}_i\} \supseteq \{y : \underline{U}(y) = \max_i \underline{U}(\bar{y}_i)\}$, suppose $\underline{U}(\bar{y}_j) = \max_i \underline{U}(\bar{y}_i)$ for some j. Hence $d_j(\infty) = \max_i d_i(\infty)$ and $\bar{c}_i \leq d_i(\infty) \leq d_j(\infty)$ for all i. It follows from $d_j(\infty) < f^t(d_j(\infty))$ that $A_t(d_j(\infty)) = Y$ for all $t \geq 1$, where $A_t(c) = \{y \in Y : \bar{c}(y) \leq f^t(c)\}$. Therefore $\Gamma(A(d_j(\infty))) \equiv \infty$ and $d_j(\infty) = d_j(\Gamma(A(d_j(\infty))))$. Then Lemma A.2 (i)

implies that $\bar{c}_j = d_j(\infty)$ and thus $\bar{c}_j = \max_i \bar{c}_i$. Therefore $\{y : \bar{c}(y) = \max_i \bar{c}_i\} \supseteq \{y : \underline{U}(y) = \max_i \underline{U}(\bar{y}_i)\}$.

- To prove $\{y : \bar{c}(y) = \max_i \bar{c}_i\} \subseteq \{y : \underline{U}(y) = \max_i \underline{U}(\bar{y}_i)\}$, suppose $\bar{c}_m = \max_i \bar{c}_i = \bar{c}_j$ for some m. Then Lemma 2 implies that $d_m(\infty) \geq \bar{c}_m = d_j(\infty) = \max_i d_i(\infty)$. Therefore, $\underline{U}(\bar{y}_m) = \max_i \underline{U}(\bar{y}_i)$.

- **Step 2.**
  - **Goal:** Compute a vector $\tilde{c} : B^c \to \mathbb{R}$, such that $\tilde{c}_i = d_i(\Gamma(C))$ for $\bar{y}_i \in B^c$, where $C = \{C_t; t \geq 1\}$ and $C_t = \{y \in B : \bar{c}(y) \leq f^t(\tilde{c}_i)\} \cup B^c$ for $t \geq 1$.
  - As described in the algorithm, I compute $\tilde{c}$ by iteration. The definition of C^{k+1} and Lemma A.1 (i) imply that $x^k \geq x^{k+1}$, which then implies that C^{k+1} contains C^{k+2}. It follows from induction that x^k decreases and C^k shrinks in k.
  - The iteration of x^k finishes in finite steps. Recall that e is a lower bound for $\bar{c}$ in Lemma 2. If T is such that $f^T(e) > \max_i d_i(\infty)$, then $C_t(x) \equiv \{y \in B : \bar{c}(y) \leq f^t(x)\} \cup B^c = Y$ for all $t \geq T$ and $x \geq e$, which implies that there are only finitely many sequences of subsets C^k. Hence there exists a finite k^* at which $C^{k^*+1} = C^{k^*}$, and $x^{k^*+1} = d_i(\Gamma(C^{k^*+1})) = d_i(\Gamma(C^{k^*})) = x^{k^*}$.

- **Step 3.**
  - I show that $\tilde{c}(\cdot)$ and $\bar{c}(\cdot)$ have the same maximizer on B^c, i.e., $D \equiv \{y \in B^c : \tilde{c}(y) = \max_{\bar{y}_i \in B^c} \tilde{c}_i\} = \{y \in B^c : \bar{c}(y) = \max_{\bar{y}_i \in B^c} \bar{c}_i\}$, and the two maximums are identical.
  - $\bar{c}_i \leq \tilde{c}_i$ follows from the definition of $\tilde{c}_i$ in step 2 and $\bar{c}_i \leq d_i(\tau)$ for all τ. Therefore $\min_{\bar{y}_i \in B^c} \bar{c}_i \leq \min_{\bar{y}_i \in B^c} \tilde{c}_i$.
  - To prove $\max_{\bar{y}_i \in B^c} \bar{c}_i \geq \max_{\bar{y}_i \in B^c} \tilde{c}_i$, I prove two things. First, if $\bar{c}_j = \max_{\bar{y}_i \in B^c} \bar{c}_i$ for some j, then $\tilde{c}_j = \bar{c}_j$. Since $C_t(\tilde{c}_j) = \{y \in B : \bar{c}(y) \leq f^t(\tilde{c}_j)\} \cup B^c = \{y \in Y : \bar{c}(y) \leq f^t(\tilde{c}_j)\} \equiv A_t(\tilde{c}_j)$, it follows from $\tilde{c}_j = d_j(\Gamma(C(\tilde{c}_j)))$ and Lemma A.2 (i) that $\tilde{c}_j > \bar{c}_j$ is impossible, therefore, $\tilde{c}_j = \bar{c}_j$. Second, if $\bar{c}_m < \bar{c}_j = \max_{\bar{y}_i \in B^c} \bar{c}_i$ for some m, then $\tilde{c}_m < \bar{c}_j$. By contradiction, suppose $\tilde{c}_m \geq \bar{c}_j = \max_{\bar{y}_i \in B^c} \bar{c}_i$, then $C_t(\tilde{c}_m) = \{y \in B : \bar{c}(y) \leq f^t(\tilde{c}_m)\} \cup B^c = \{y \in Y : \bar{c}(y) \leq f^t(\tilde{c}_m)\} \equiv A_t(\tilde{c}_m)$. Therefore, $\tilde{c}_m = d_m(\Gamma(C(\tilde{c}_m))) = d_m(\Gamma(A(\tilde{c}_m)))$ and Lemma A.2 (i) imply that $\tilde{c}_m = \bar{c}_m$, which contradicts $\tilde{c}_m \geq \bar{c}_j > \bar{c}_m$.
  - It follows from above that $\{y \in B^c : \bar{c}(y) = \max_{\bar{y}_i \in B^c} \bar{c}_i\} = \{y \in B^c : \tilde{c}(y) = \max_{\bar{y}_i \in B^c} \tilde{c}_i\}$.
  - Note that $\min_{\bar{y}_i \in B} \bar{c}_i > \max_{\bar{y}_i \in B^c} \bar{c}_i$ is maintained after B is replaced by $B \cup D$.

References

Abreu, D., Pearce, D., Staccheti, E., 1990. Toward a theory of discounted repeated games with imperfect monitoring. Econometrica 58 (5), 1041–1063.

Albuquerque, R., Hopenhayn, H., 2004. Optimal lending contracts and firm dynamics. Review of Economic Studies 71, 285–315.

Alvarez, F., Jermann, U., 2000. Efficiency, equilibrium, and asset pricing with risk of default. Econometrica 68 (4), 775–797.

Broer, T., 2009. Stationary equilibrium distributions in economies with limited commitment. Economics Working Papers, European University Institute.

Broer, T., 2011. The wrong shape of insurance? what cross-sectional distributions tell us about models of consumption-smoothing. CEPR Discussion Papers 8701, C.E.P.R. Discussion Papers.

Detemple, J., Serrat, A., 2003. Dynamic equilibrium with liquidity constraints. Review of Financial Studies 16 (2), 597–629.

Grochulski, B., Zhang, Y., 2011. Optimal risk sharing and borrowing constraints in a continuous-time model with limited commitment. Journal of Economic Theory 146, 2356–2388.

Kehoe, P., Perri, F., 2002. International business cycles with endogenous incomplete markets. Econometrica 70 (3), 907–928.

Kehoe, T., Levine, D., 1993. Debt-constrained asset markets. Review of Economic Studies 60 (4), 865–888.

Krueger, D., Perri, F., 2006. Does income inequality lead to consumption inequality? evidence and theory. Review of Economic Studies 73, 163–193.

Krueger, D., Perri, F., 2011. Public versus private risk sharing. Journal of Economic Theory 146, 920–956.

Krueger, D., Uhlig, H., 2006. Competitive risk sharing contracts with one-sided commitment. Journal of Monetary Economics 53, 1661–1691.

Ljungqvist, L., Sargent, T., 2004. Recursive Macroeconomic Theory, Second Edition. The MIT Press, Cambridge, Massachusetts, USA.

Ray, D., 2002. The time structure of self-enforcing agreements. Econometrica 70 (2), 547–582.

Spear, S., Srivastava, S., 1987. On repeated moral hazard with discounting. Review of Economic Studies 54 (4), 599–617.

Tauchen, G., Hussey, R., 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. Econometrica 59 (2), 371–96.

Thomas, J., Worrall, T., 1988. Self-enforcing wage contracts. Review of Economic Studies 55, 541–554.

Thomas, J., Worrall, T., 2007. Unemployment insurance under moral hazard and limited commitment: Public versus private provision. Journal of Public Economic Theory 9 (1), 151–181.