

Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending*

Á. Abraham[†] and N. Pavoni[‡]

First Draft: February 2001. This Version: April 2003.

Abstract

In this paper we develop a recursive approach to study efficient allocations in a dynamic moral hazard setting, where agents can borrow and lend and their decisions about effort, consumption and savings are private information. The recursive formulation of the problem is based on a generalized first order approach, whose validity is verified using a parsimonious numerical procedure based on the recursive formulation itself. In contrast with previous findings, we show that the second best allocation is welfare improving with respect to the case where the agents can self insure themselves only through borrowing and lending. Thanks to the recursive formulation, we are able to quantify the efficiency gains in a number of numerical examples. We find that welfare gains are substantial and do not vary monotonically with the credit market return. We also identify the main observational characteristics of the constrained efficient allocation.

*The authors thank Orazio Attanasio, Piero Gottardi, Per Krusell, Guido Lorenzoni, Albert Marcet, Andy Newman, Fabrizio Perri, the participants at the SITE2002 Conference in Stanford, at the SED2002 Conference in New York, at the CEPR Workshop in Gerzensee, at the LACEA2002 Conference in Madrid, at the SUNY Stony Brook Workshop, at the NYU Macro Workshop, at UCL lunch seminar, at the Macro Workshop at Carlos III, and at the conference on Consumptions Behaviour and Welfare in Verona for useful suggestions. Nicola Pavoni thanks the Marie Curie Fellowship MCFI-2000-00689 and the Spanish Ministry of Science and Technology Grant BEC2001-1653. *Correspondence:* Nicola Pavoni, Department of Economics, University College London, Gower Street, London WC1E 6BT. *E-mail:* n.pavoni@ucl.ac.uk, aabraham@econ.duke.edu

[†]Duke University.

[‡]University College London.

1 Introduction

The aim of this paper is to analyze the efficient allocation in an endogenously incomplete market framework. We consider an environment in which risk-averse individuals have random income and can borrow and lend at a given risk-free interest rate. Their asset and consumption decisions are private information (or social insurance contracts cannot be contingent on these variables). Moreover, each individual can affect future income realizations through his effort decision, which is non-monitorable either (moral hazard). The second best allocation is derived by solving the problem of a risk neutral planner who provides optimal incentive-compatible insurance contracts.

There is a large prior literature which studies optimal (long-term) insurance contracts under moral hazard assuming that agents cannot borrow and save or, equivalently, that the planner can perfectly control the agents' asset holdings.¹ Allen (1985) and Cole and Kocherlakota (2001) (ACK) introduce hidden asset accumulation in a pure endowment economy with adverse selection (hidden information moral hazard). The main finding of ACK is that the constrained efficient allocation does not differ from the one agents would obtain in a "bond economy," i.e. the allocation they could obtain by self insuring themselves through borrowing and lending, without the planner's provision of additional insurance.² This result has very strong implications both from a positive and normative point of view. First, it implies that the constrained efficient allocation cannot be (observationally) distinguished from the bond economy one. Second, it suggests that there are no potential gains from several policy interventions with inherent informational problems (e.g. public unemployment insurance schemes³).

¹See, among many others, Townsend (1982), Rogerson (1985a), Spear and Srivastava (1987), Green (1987), Phelan and Townsend (1991). Atkeson (1991) analyses a moral hazard model with borrowing and lending and default. His model differs crucially from the one studied in this paper since Atkeson assumes that asset holdings are observable.

²Cole and Kocherlakota (2001) extends the result of Allen (1985) for the case where the agent can only save, but cannot borrow.

³Hopenhayn and Nicolini (1997) and Pavoni (2002a,b) study the design of an optimal UI scheme using a dynamic moral hazard framework, where the worker's job search effort level is private information but asset holdings are perfectly observable. Independently, Werning (2002) uses techniques similar to the one used in this paper and extends the optimal UI analysis to the case where asset holdings are not observable. Rather than characterizing the efficient allocation of consumption, Werning's objective is to describe the UI *transfer* scheme. He does not study whether unemployment insurance provision actually improves on

Therefore, a natural preliminary step is to see whether the “self-insurance” result of ACK applies to our framework. In fact, the first key finding of our analysis is that in a dynamic (hidden action) moral hazard setting a constrained efficient allocation *does* improve on self insurance. Given that the efficient allocation differs from the one obtainable without insurance contracts, what are the characteristics of this new allocation? In terms of observables, one key difference between the second best and the allocation obtainable in a pure bond economy turns out to be risk sharing. This insurance property of the efficient allocation induces a discrepancy between the present value of income and consumption flows, which is not present in the ACK/self-insurance framework, and therefore can be used to distinguish empirically the two models. We also discuss how the time series of individual consumption can be used to test our second best allocation against the optimal allocations of the observable asset holdings case and the first best.

Our next objective is to quantify welfare gains and to identify the sources of such welfare improvements. As we will see below this task required the reformulation of the planner problem within the recursive contracts framework. We computed welfare gains for a set of infinite-horizon numerical examples, and we found very significant improvements. In a the second best allocation agents could enjoy an increase in consumption between 5.7 and 10.9 percent in all times and states compared to self insurance.

We also performed some comparative statics exercises. One interesting characteristic of our results is the non-monotonic relationship between the exogenous risk free rate r and the efficiency gains. When the interest rate decreases the welfare gains of the second best compared to self-insurance increase. This result is intuitive since the relevant constraint in the optimal contract is the agent’s incentive to save. And a reduction in the return on savings relaxes this incentive constraint. More surprisingly, in the neighborhood of $\frac{1}{1+r} = \beta$ (β is the agent’s discount factor) an increase in r will result in higher welfare gains as well. This holds despite the fact that the set of incentive-feasible allocations shrinks with increasing r . However, in contrast with the bond economy, in the second best allocation each agent has access to a complete set of contingent commodities.⁴ An increase in the interest rate reduces the (discounted) price of such history contingent goods and the agent

self insurance. Kocherlakota (2002) fully characterizes the optimal UI transfer scheme for the case where agents’ preferences are linear in (job-search) effort, and effort affects linearly job finding probabilities.

⁴Even though the trade of these goods is of course endogenously restricted by the presence of the incentive constraints.

can make additional profits (with respect to self insurance) by intertemporally trading assets related to these commodities. It turns out that, locally, these intertemporal trading gains overcome the increased incentive costs.⁵

The constrained efficient allocation improves on self insurance at least in two main directions. First, each agent is required to exert the second best effort level, which typically differs from the level the agent would choose in a bond economy, although the self insurance level of effort is of course still feasible in the second best problem.⁶ We call this aspect *total output efficiency*, since the (aggregate) level of production is chosen optimally along this dimension. Second, the second best allocation provides incentives in an optimal way. In particular, (i) output levels not only modify the total amount of available resources (this role is also present in the case of self insurance) but also signal the agent's level of effort to the planner. This informational role is the main guide for the planner to allocate risk optimally. This property implies that in contrast with self insurance the way consumption varies (cross-sectionally) with income does not depend on the level of each particular output realization per se, but on its informational content on effort.⁷ Moreover, (ii) although these incentives require some variation of lifetime utility, punishments and rewards are spread out across different states in such a way that, for comparable dispersion levels of lifetime utilities, the second best consumption is intertemporally much smoother than self insurance. We call these aspects *incentive-insurance efficiency*.

In our computations, we find that production efficiency is quantitatively as important as incentive-insurance to explain welfare gains. Moreover, their relative importance is not uniform across periods. Since agents are born with no wealth, at the beginning the dominating effect is risk sharing. However, in self insurance the only way to buy insurance is bond accumulation, hence agents tend to overaccumulate assets. As time passes welfare gains coming from risk sharing decrease and total output efficiency tends to dominate. In

⁵Notice that these intertemporal trade gains are due to the mismatch between the agent's patience $\beta^{-1} - 1$ and the interest rate, and are available in the low r case as well. Of course in the latter case agents tend to frontload consumption, while when r is high intertemporal trade gains are obtained by postponing consumption.

⁶The planner can always implement the bond economy allocation by making no transfers to the agent.

⁷Similarly to the standard moral hazard model (with no hidden assets), the second best consumption is positively correlated with the size of the *likelihood ratio*, which measures the relative likelihood that a specific outcome occurs from the implemented level of effort rather than the levels of effort for which the incentive compatibility constraint is binding.

particular, if leisure is a normal good, wealth overaccumulation reduces the self insurance aggregate level of production well below the second best level.

The specificity of the adverse selection framework of ACK makes the two mentioned welfare improving opportunities not available to the planner. First of all, in a pure endowment economy such as the one of ACK, there is no endogenous production (so, no total output efficiency). Moreover, in an adverse selection framework, informational asymmetries are such that each agent has complete and costless control on public outcome histories. In addition, the free access to the credit market makes the agent's preferences dependent on the net present value of the transfers alone (and in particular not dependent on the timing of transfers). Since the planner's transfers are based on public outcomes alone, when the agent can secretly borrow and lend he will generate the history of public realizations that delivers the transfer sequence with the highest net present value. As a consequence, incentive compatibility prevents the planner from making transfer payments contingent on past public history. It follows that no insurance-efficiency - in particular, no risk-sharing - is possible either, and the only incentive compatible consumption allocation is the one available to the agent operating alone in a pure bond economy. It is crucial, that in our model agents' perfect control over public outcomes is prevented by the *full support condition*, perhaps the most commonly adopted assumption for moral hazard models with hidden actions.⁸

In addition to the mentioned ACK only few papers analyze dynamic moral hazard with non-monitorable borrowing and lending. In a two period principal-agent model, Bizer and DeMarzo (1999) show that hidden access to the credit market reduces total welfare with respect to the no asset market case. They focus their analysis on the possibility of increasing welfare by allowing the agent to default on the debt. We analyze the full-commitment case in a general multi-period setting. As mentioned above, we show that the welfare reduction highlighted by Bizer and DeMarzo is not as drastic as the ACK analysis would suggest. Chiappori et al. (1994) analyze a two period discrete effort model, and obtain a different negative result. They find that under some conditions, the optimal contract is in some sense trivial: a renegotiation-proof contract always implements the minimum level of effort. We solve a general multi-period continuous-effort model, where the planner can commit himself

⁸The full support assumption requires that for any given level of effort supplied by the agent, all conceivable public outcomes might occur with positive probability.

to not renegotiate the contract ex post.

Those papers use particular models and focus on specific issues. In fact, the analysis of the general dynamic moral hazard problem with non-monitorable asset holdings is complicated because the problem fails to have a recursive structure, at least in the usual sense. Fudenberg et al. (1990) provide characterizations of efficient allocations in a wide class of dynamic environments where agents' preferences over continuation contracts are common knowledge after any history. The assumption of common knowledge of preferences is not satisfied in our framework.⁹ In spite of that, by using a generalized *first order condition approach*, we are able to formulate the problem recursively. Together with the agents' expected discounted utility, we introduce an additional endogenous state variable: the agent's marginal utility of consumption. The "dual" counterpart of this state variable is the multiplier related to the marginal value of capital which has been used in the Ramsey taxation literature (see Kydland and Prescott, 1980; and Marcet Marimon, 1999).¹⁰ We are able to obtain a recursive formulation because incentive compatibility guarantees that common knowledge of preferences is maintained along the equilibrium path, and the first order approach allows us to write the problem only in terms of equilibrium values.

One way to justify the use of a first order approach is to show that - given the optimal contract - the agent's problem is concave. However, in a complicated intertemporal problem such as the one analyzed in this paper finding concavity conditions for the agent's problem easily becomes a formidable task.¹¹ We propose an alternative methodology. We take ad-

⁹Fernandes and Phelan (2000) and Doepke and Townsend (2002) propose a way to solve recursively an even wider class of problems than the one analyzed by Fudenberg et al. (1990). Unfortunately, their methods are not viable when hidden actions belong to a continuum as it is the case for savings in our model. Cole and Kocherlakota (2001) extend the Abreu et al. (1990) framework to a relatively wide class of dynamic games. However again, none of their extensions is of any use for us, since our law of motion for bond holdings does not satisfy the "full support" assumption needed in those frameworks.

¹⁰In independent works, Chang (1998) and Phelan and Stacchetti (2001) derive recursive formulations similar to ours. Using a game theoretical approach, Chang studies Calvo's monetary model and Phelan et al. study optimal linear capital taxation. They both analyze essentially deterministic economies with anonymous players (Phelan and Stacchetti introduce an exogenous and payoff irrelevant shock to convexify the problem), and focus on the (sequential/subgame perfect) equilibrium allocation under the assumption that the government has limited commitment. We study a different problem (with moral hazard) and use a contractual approach under full-commitment.

¹¹To our best knowledge, this has been done only for static models. See, for example, Rogerson (1985b) and Jewitt (1988).

vantage of the dynamic programming formulation and use a recursive numerical procedure to *verify ex post* whether the optimal contract derived using the first order approach is in fact incentive compatible. The proposed methodology, turns out to be numerically parsimonious since it requires the solution of a simple dynamic optimization problem with only one endogenous state.

The paper is organized as follows. In Section 2, we present our environment and define constrained efficiency (second-best). Section 3 introduces some benchmark frameworks which are useful to disentangle the different forces operating in our model. Section 4 is dedicated to the qualitative study of the efficient allocation: we prove the existence of welfare gains and identify the main observable characteristics of the second best allocation. In Section 5 we present our approach and characterize the optimal scheme in some detail using a two period model. The recursive formulation is presented in Section 6. In Section 7, we perform an extensive quantitative analysis and implement our ex-post verification procedure. Section 8 concludes.

2 Environment and Constrained Efficiency

The model builds on the typical dynamic moral hazard problem. Consider an economy consisting of a large number of agents that are identical ex-ante, and who each live $T \leq \infty$ periods. Each agent is endowed with a *private* stochastic production technology, which takes the following form. There is a finite set $Y = (y^1, \dots, y^N)$ of possible income levels, with $y^i < y^{i+1}$. At each period t , the realization $y_t \in Y$ is publicly observable, however, the probability distribution over Y is affected by the agent's unobservable effort level e , which we assume belonging to a bounded interval $E \subset \mathfrak{R}_+$. We normalize effort levels such that $0 \in E$. The conditional probabilities over Y are defined by the smooth functions:¹² $p_i(e) = \Pr \{y = y^i \mid e\}$. Hence, agents are subject to idiosyncratic risk, and we assume time independent (conditional) distribution of output.¹³ As it is usual in the optimal contracting literature, we assume *full support*, that is $p_i(e) > 0$ for any $i = 1, \dots, N$. The history of public

¹²In particular we require the function $\mathbf{p} : E \rightarrow \Delta^N$ to be continuous and continuously differentiable (where $\Delta^N = \{x \in \mathfrak{R}^N \mid x \geq 0 \text{ and } \sum_i x_i = 1\}$).

¹³Notice that this model, naturally allows for persistence in idiosyncratic shocks. Without additional complications, we could have defined a Markov process for individual output by writing the conditional probabilities as: $p_i(e, y) = \Pr \{y_{t+1} = y^i \mid e, y_t = y\}$.

outcomes up to period t will be denoted by $h^t \equiv (y_0, y_1, \dots, y_t)$.¹⁴

The crucial feature of our model is that agents are allowed to buy and sell a risk-free bond which pays a constant interest rate $r \geq 0$, and their asset holdings are private information as well. Note, that since they can trade only *risk free* bonds, each agent faces incomplete asset markets. Therefore, we can expect that a social planner could increase overall welfare by providing additional insurance. In this setting, a constrained efficient allocation can be computed by solving the problem of a planner who reallocates resources optimally in order to insure agents, subject to the feasibility and incentive constraints which will be specified below.

An *allocation* (or social contract) in this economy is a contingent plan

$$\mathcal{W} = (\boldsymbol{\tau}, \boldsymbol{\sigma}); \quad \boldsymbol{\tau} = \{\tau_t(h^t)\}_{t=0}^{T-1}, \quad \boldsymbol{\sigma} = \{\mathbf{e}_t(h^t), \mathbf{b}_t(h^t), \mathbf{c}_t(h^t)\}_{t=0}^{T-1}$$

where $\tau_t(h^t)$ represents the transfer the individual receives in period t , $\mathbf{e}_t(h^t)$ the implemented effort, $\mathbf{b}_t(h^t)$ the bond holdings and $\mathbf{c}_t(h^t)$ the agent's consumption level as a function of the realized history h^t . Note, that our notion immediately assumes that agents can only be distinguished through their output histories. In this sense, since all individuals are *ex ante* identical, we restrict ourself to symmetric allocations.

To simplify the analysis, we separate the planner's transfer plan $\boldsymbol{\tau}$, from $\boldsymbol{\sigma}$ the components of the allocation under the agent's control. The metaphor used in contract theory is that the planner proposes $\boldsymbol{\sigma}$, and the plan will be implemented by appropriately designing the transfer scheme $\boldsymbol{\tau}$. The timing of the model is as follows. At period $t \geq 0$, each agent receives a transfer payment $\tau_t = \tau_t(h^t)$ from the planner, contingent on the realized history h^t . Given today's income level y_t , transfer τ_t , asset level b_{t-1} (and the continuation plan $\boldsymbol{\tau} \setminus h^t$), the agent chooses consumption $c_t \geq 0$ and bond holdings b_t subject to the following budget constraint

$$c_t + b_t = y_t + \tau_t + (1 + r)b_{t-1}. \quad (1)$$

As usual, we impose the last period condition $b_{T-1} \geq 0$ on asset holdings.¹⁵ At the beginning of each period, the agent also decides the effort level $e_t \geq 0$, which affects the stochastic

¹⁴Since the only other publicly observable variable is the planner's transfer, and the planner has full commitment, we can restrict public histories to histories of income only without loss of generality (see Pavoni (1999) for details).

¹⁵For the infinite horizon version of the model ($T = \infty$) the corresponding (transversality) condition is: $\lim_{T \rightarrow \infty} \left(\frac{1}{1+r}\right)^{T-1} \mathbf{b}_{T-1}(h^{T-1}) \geq 0$ almost surely for any h^{T-1} .

output realization $y_{t+1} \in Y$, leading to next period output history $h^{t+1} = (h^t, y_{t+1})$. This sequence of events continues until the end of the world. We assume that each agent is born with no wealth ($b_{-1} = 0$).¹⁶

Agents have intertemporally additive separable von Neumann-Morgenstern utility function. The continuation plan $\mathcal{W} \setminus h^t$ ($\equiv \{\boldsymbol{\tau}_\tau(h^\tau), \mathbf{e}_\tau(h^\tau), \mathbf{b}_\tau(h^\tau), \mathbf{c}_\tau(h^\tau) \setminus h^t\}_{\tau=t}^{T-1}$) from node h^t generates the following expected discounted utility at time t

$$\mathbf{U}_t^T(\mathcal{W}; h^t) = \mathbf{E} \left[\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathbf{c}_\tau(h^\tau), \mathbf{e}_\tau(h^\tau)) \setminus \alpha_t(h^t) \right],$$

where $\alpha_t(h^t) \equiv \{\mathbf{e}_\tau(h^\tau) \setminus h^t\}_{\tau=t}^{T-1}$ denotes the implemented effort plan from history h^t onward, \mathbf{E} is the expectation operator, and $\beta \in (0, 1)$. We assume that the choice of \mathcal{W} is implicitly restricted in such a way that both the expectation and the (possibly infinite) summation are well defined. The continuous and real valued utility function u is assumed to be such that $u(\cdot, e)$ is strictly concave, strictly increasing, and as smooth as we need. Similarly, $u(c, \cdot)$ is assumed to be strictly decreasing, concave and smooth for any given c .

To be feasible, an allocation \mathcal{W} must be deviation-proof in all components of $\boldsymbol{\sigma}$, that is, in effort e , bond holdings b , and consumption c . Hence, we say that the allocation \mathcal{W} is *sequentially incentive compatible* if, for any history h^t , we have:

$$\boldsymbol{\sigma} \setminus h^t \in \arg \max_{\tilde{\boldsymbol{\sigma}} \setminus h^t} \mathbf{U}_t^T(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\tau}; h^t) \quad (2)$$

s.t. (1) $\forall h^{t+n} \setminus h^t$
and $\tilde{\mathbf{b}}_{T-1}(h^{T-1}) \geq 0$

The expression $\boldsymbol{\sigma} \setminus h^t$ ($\equiv \{\mathbf{e}_\tau(h^\tau), \mathbf{b}_\tau(h^\tau), \mathbf{c}_\tau(h^\tau) \setminus h^t\}_{\tau=t}^{T-1}$) denotes the implemented continuation of the agent's decisions after node h^t , and the "tilde" refers to deviations over all possible alternative continuation plans of effort, consumption and bond holdings, compatible with the budget constraints (1) and the transfer plan $\boldsymbol{\tau} \setminus h^t$. We denote the set of *incentive feasible allocations* as

$$\Omega = \{\mathcal{W} : \text{for any history } h^t, \mathcal{W} \setminus h^t \text{ satisfies (2)}\}.$$

The goal of this paper is to characterize the constrained efficient allocations in this environment. For technical tractability, we define the optimal contract as the one that maximizes planner's net returns (or minimizes costs), subject to incentive feasibility and

¹⁶In the numerical section, we will also consider the case where y_0 differs among agents and $b_{-1} > 0$.

the (social) restriction that each individual must at least receive a given *ex-ante* expected discounted utility level U_0 . We will then choose U_0 so that planner's expected discounted returns equal zero. The planner is represented by a risk neutral principal who faces the same interest rate as the agents, whose net return at node h^t induced by the continuation plan $\mathcal{W} \setminus h^t$ is:

$$\Pi_t^T(\mathcal{W}; h^t) = \mathbf{E} \left[\sum_{\tau=t}^{T-1} \frac{(1+r)^t}{(1+r)^\tau} (-\tau_\tau(h^\tau)) \setminus \alpha_t(h^t) \right].$$

Given an initial history $h^0 = y_0$ and social restriction U_0 , the T -horizon planner's problem can then be formulated as follows¹⁷

$$\begin{aligned} & \sup_{\mathcal{W} \in \Omega} \Pi_0^T(\mathcal{W}; h^0) \\ \text{s.t. } & \mathbf{U}_0^T(\mathcal{W}; h^0) \geq U_0. \end{aligned} \tag{3}$$

3 Bond Economy, First-Best, and Pure Moral Hazard

Before studying the constrained efficient allocation in our model, we should investigate some different set-ups, which will be used later as benchmarks.

We start by describing the autarchic situation, where agents do not have access either to the credit or to the insurance market. After that, we present the bond economy or self insurance model. In pure bond economy, agents do not have access to social insurance, but can partially smooth consumption through borrowing and lending at a constant interest rate r . If the self insurance result of ACK applied to our model, the second best allocation we described in Section 2 would coincide with the one obtained in this bond economy. In Section 3.3 we analyze the first best allocation. Here agents have unrestricted access to a complete set of assets, hence they can perfectly insure themselves. Then, in Section 3.4 we study an endogenously incomplete markets framework which we call "pure moral hazard." In this model, agents still have access to the insurance market, but their participation is restricted by a moral-hazard constraint on effort decisions. The key (and only) difference between the pure moral hazard model and the framework we study in this paper is that in the former there are no informational asymmetries on asset holdings. Finally, in Section

¹⁷Note that we assume that the planner has full-commitment. In particular, we permit the planner to commit not to recontract, even though there would be ex-post incentives to do so.

3.5 we rank these allocations in terms of total welfare, and we start comparing them with the second best.

3.1 Autarchy

First, suppose that each agent owns the above described production technology, but has no access to the asset market. This implies that the agent is forced to consume the realized output $y_t \in Y$ each period, hence he will face the maximum possible uncertainty. When there are $t \leq T$ periods before the end, each agent solves¹⁸

$$U_t^a(y) = \max_{e \in E} u(y, e) + \beta \sum_i p_i(e) U_{t-1}^a(y^i). \quad (4)$$

Notice that $U_t^a(y^i)$ increases with y^i . If p is “concave”¹⁹ this allocation is completely characterized by the first order condition

$$u'_e(y, e) + \beta \sum_i p'_i(e) U_t^a(y^i) = 0.$$

3.2 Bond Economy (Self Insurance)

We know from consumption theory that allowing an agent to participate in a risk free bond market is Pareto improving. We call *bond economy* or *self-insurance* the allocation derived from autarchy by allowing the agent to participate in the credit market. It has already been pointed out that this model is an important benchmark for our economy with moral hazard and hidden asset accumulation, for if the self insurance result of ACK were true in our framework the second best allocation would coincide with the one the agent could obtain alone in the bond economy we describe here.

Within the dynamic programming framework, an agent’s problem in the bond economy can be completely described in terms of the properties of the value functions U_t^{Bond} , $t = 0, \dots, T$. The function $U_t^{Bond}(z)$ represents the maximal value of lifetime utility that an agent

¹⁸In the infinite horizon version of the model ($T = \infty$), the agent solves the obvious time invariant limit problem.

¹⁹The function $\mathbf{p} : E \rightarrow \Delta^N$ satisfy the concavity of the distribution function condition (CDFC) if $F_I''(e)$ is non-negative for every e and $I \leq N$, where $F_I(e) = \sum_{i=1}^I p_i(e) = 1 - \sum_{i=I+1}^N p_i(e)$.

with wealth level z can get when $t \leq T$ periods are left. It is recursively defined as²⁰

$$U_t^{Bond}(z) = \sup_{b,e,c} u(c, e) + \beta \sum_i p_i(e) U_{t-1}^{Bond}(z^i) \quad (5)$$

s.t.

$$\begin{aligned} c + b &\leq z \\ z^i &= y^i + (1+r)b. \end{aligned}$$

The first constraint above is the usual budget restriction. The second equality represents the law of motion for the agent's total wealth z (the only state), where r is the risk-free interest rate. For future reference, notice that this problem with $T = \infty$ can be seen as the endogenous labor supply extension of the (permanent-income) model studied by Aiyagari (1994) and Huggett (1993). When e is not effective - $p_i(e) = \bar{p}_i$ for $\forall i$ - we are back to their case.

One of the main characteristics of the self insurance model can be obtained by considering the following perturbation of the agent's consumption plan (in the event of a particular income history h^t) : reduce consumption infinitesimally at date t , invest the amount for one period at the return r , then consume the proceeds of the investment at date $t + 1$. If the consumption plan is optimal, this perturbation must not affect the agent's utility level. The first-order necessary condition for his utility not to be affected is the Euler equation:

$$u'_c(c_t, e_t) \geq \beta(1+r) \sum_i p_i(e_t) u'_c(c_{t+1}^i, e_{t+1}^i), \quad (6)$$

with equality if $c_{t+1}^i > 0$ for any i .²¹

3.3 First Best

In general, the introduction of a risk free asset market cannot insure the agent perfectly. Therefore there is a role for a social planner who may provide additional insurance. It is

²⁰It is straightforward to see that $U_1^{Bond}(z) = u(z, 0)$, as the one period problem is

$$\begin{aligned} U_1^{Bond}(z) &= \max_{b \geq 0, e \in E, c \geq 0} u(c, e) \\ \text{s.t.} & : c + b \leq z. \end{aligned}$$

In the infinite horizon case, U_∞^{Bond} becomes the time invariant function that solves (5) when the "natural" debt limit $b \geq -\frac{y_1}{r}$ is imposed each period (Aiyagari, 1994).

²¹When h^t is known, we denote $c_{t+1}^i \equiv \mathbf{c}_t(h^t, y_{t+1}^i)$.

easy to see that without monitoring problems the planner will fully insure agents, requiring them to supply the first best effort level. An unconstrained Pareto optimal allocation can be determined by solving

$$\begin{aligned} & \sup_{\mathcal{W}} \Pi_0^T(\mathcal{W}; h^0) & (7) \\ \text{s.t. } & (1); \text{ and } \mathbf{U}_0^T(\mathcal{W}; h^0) \geq U_0. \end{aligned}$$

Problem (7) can be seen as a version of (3) where the incentive restrictions on the contract \mathcal{W} are removed. Note that if we allowed for a complete set of assets, the agent alone could obtain this allocation.²² In other words, if U_0 were the agent's value with complete markets, then the optimal allocation of (7) would be equivalent to the agent's optimal plan and the planner would make zero surplus. In an unconstrained efficient allocation, consumption and effort are either constant (if $\beta = \frac{1}{1+r}$), or vary deterministically with time. In addition, the optimality condition

$$-\frac{u'_e(c_t, e_t)}{u'_c(c_t, e_t)} = \frac{1}{1+r} \sum_i p'_i(e_t) y^i. \quad (8)$$

guarantees that the first best level of effort is chosen so that to equalize social costs with expected returns.

3.4 Pure Moral Hazard (Monitorable Asset Holdings)

Now, consider the case where each agent has private information on the effort level e , but the planner can monitor consumption and asset decisions. In such environment, the optimal allocation of consumption and effort is identical to the one obtained in a model where agents do not have access to the credit market at all.²³ The optimal allocation hence

²²In this economy, the prices of Arrow securities depend on the (observable) effort level the agent commits to supply.

²³See Lemma 4 in Appendix A.

solves

$$\begin{aligned} & \sup_{\mathcal{W}} \Pi_0^T(\mathcal{W}; h^0) & (9) \\ \text{s.t.} \quad & : \quad \mathbf{U}_0^T(\mathcal{W}; h^0) \geq U_0; \\ & \text{and for any } h^t \end{aligned}$$

$$\mathbf{b}_t(h^t) = 0 \quad (10)$$

$$\mathbf{c}_t(h^t) = y_t + \boldsymbol{\tau}_t(h^t) \quad (11)$$

$$\alpha_t(h^t) \in \arg \max_{\tilde{\boldsymbol{\alpha}}} \mathbf{U}_t^T(\tilde{\boldsymbol{\sigma}}, \boldsymbol{\tau}; h^t). \quad (12)$$

The key difference between this problem and the first-best (7) is the incentive constraint (12), which defines the set of contracts for which the agent will be induced to supply the effort level proposed by the planner. At each node h^t , the agent will (privately) take the effort that maximizes his expected discounted life-time utility. Finally, since in this model agents are not allowed to borrow and save, the budget constraint (11) requires that they consume their wealth (their income together with the transfer received from the planner).

In the additive separable case, the key characteristic of an optimal interior contract is summarized by²⁴

$$\frac{1}{1+r} \sum_i p_i(e_t) \frac{1}{u'(c_{t+1}^i)} = \beta \frac{1}{u'(c_t)}, \quad (13)$$

In order to relate (13) to the Euler equation, notice that the inverse ($1/x$) is a strictly convex transformation. As a consequence, Jensen inequality implies

$$u'(c_t) \leq \beta(1+r) \sum_i p_i(e_t) u'(c_{t+1}^i),$$

with strict inequality if c_{t+1}^i is not constant in i . That is, the optimality condition (13) is incompatible with the Euler equation (6). The optimal pure moral hazard contract tends to frontload transfers, and agents would be *willing to save*. This consideration will play an important role in Proposition 1 below.

3.5 Ranking Different Allocations

We conclude this section by ranking the different allocations in terms of total welfare.

²⁴For later use, we reproduce Rogerson (1985a) in Proposition 6, Appendix A.

Proposition 1 *Assume $T < \infty$, additive separability in consumption and effort, and that the optimal pure moral hazard contract is such that $e_{t-1}, c_{t-1}, c_t^i > 0 \forall i$ for some $t < T$ (interiority). Then: (i) first best improves on pure moral hazard; (ii) pure moral hazard improves both on self insurance (iia), and on moral hazard with hidden borrowing and lending (second best) (iib).*

All the proofs are presented in Appendix A. The assumption of finite horizon is imposed to rule out folk theorem like results, and can be dispensed by imposing an appropriate restriction on the agent's discount factor. Additive separability is used only in the second part of the proposition.²⁵ The idea of result (i) is simple: in order to implement a positive effort level in a pure moral hazard model the planner cannot fully insure the agent, and this of course reduces total welfare. Results (ii) are both derived by comparing (6) with (13), which are known to be incompatible. Result (iia) is derived by following a pure revealed preference argument: the self insurance allocation is incentive compatible, nevertheless the planner did not choose to implement in the efficient allocation with observable assets. In a somehow specular way, result (iib) is implied by the fact that the allocation that maximizes planner's returns in the pure moral hazard model is not incentive feasible for the second best problem with hidden borrowing and lending.

4 Efficiency Gains and Positive Implications

4.1 The Welfare Gains Result

We saw in Proposition 1 (iib) that the introduction of hidden asset holdings in a moral hazard model is not innocuous. Compared with the pure moral hazard model, an additional informational asymmetry between the agent and the planner about asset accumulation reduces total welfare. However, we did not rank the second best allocation (with hidden asset holdings) with respect to self insurance. It has already been pointed out that in a pure endowment economy with adverse selection, Allen and Cole and Kocherlakota (ACK) show that the possibility of secret borrowing and lending is in fact detrimental for efficiency. They find that the constrained efficient allocation coincides with the one the agent would

²⁵Renouncing some elegance, Proposition 1 could be relaxed in many dimensions, however this exercise goes beyond the target of this paper.

obtain alone in a pure bond economy (by self insurance). The following proposition shows that in fact in our framework the second best allocation always improves on self insurance.

Proposition 2 *Assume $T < \infty$ and that u is either unbounded below or unbounded above or both. Then self insurance cannot solve (3), i.e. cannot be a constrained optimal allocation.*

The idea of the proof is that the planner can always provide some additional insurance with respect to self insurance, at least in the last period of the program. To provide *risk sharing* the planner makes positive transfers when the public outcome y is low and levies taxes when y is high. Of course, this scheme is implementable only if the agent's control on public outcomes is imperfect so that he cannot avoid paying taxes. As a consequence, the *full support* assumption is crucial for our result. The finite horizon assumption also plays an important role in the proof. In Section 7, we complement this result with a set of infinite horizon numerical examples where the welfare gains are considerable.

We now discuss what are the specificities of the economy analyzed by Allen and Cole and Kocherlakota that induce their self insurance result. One should indeed be aware that with a flexible enough formulation, the hidden information moral hazard model of ACK can be described in terms of a hidden action framework of the kind we analyze in this paper.²⁶ In the adverse selection model, e is a complete mapping between the set of private endowments Θ and the set of public outcomes Y : e is an announcement $y \in Y$ on the private shock $\theta \in \Theta$. In this model, the conditional probability $p_i(e)$ becomes the chance that the realized endowment lies within the set of θ 's for which e specifies the report y^i .

ACK assume a pure endowment economy, because in their model the probability measure over Θ is exogenous. Consequently, the planner cannot raise total welfare by implementing a different production level. In contrast, in our model effort affects total output and, according to our calculations, a substantial part of welfare gains between the second best and the bond economy are the result of the efficient implementation of effort levels (total output efficiency).

We have already pointed out that one of the pivotal elements in the proof of Proposition 2 is the fact that - due to the full support assumption - agents have imperfect control on

²⁶In Appendix B, we show the exact mapping between the two models. In order to replicate the ACK economy within our framework, we should allow e to affect $p_i(e)$ and u in a non-smooth way. However, continuity is maintained.

output realizations. In contrast, notice that in the economy studied by ACK the assumption of full support is violated as each agent could adopt the strategy of announcing one specific level of y^i for any realization of the private endowment θ , i.e. e is such that $e(\theta) = y^i \forall \theta \in \Theta$. In fact, in a pure adverse selection framework, each agent has complete and costless control of public outcome histories. In addition, they do not have “intrinsic” preferences for any specific outcome y or effort level e . The only trade-off they might face is therefore induced by the transfer scheme. In fact, the free access to the credit market makes agents’ preferences depend only on the net present value of transfers (and in particular not on the timing of transfers). On the other hand, the planner’s transfers are based on public outcomes alone. Consequently, the ability of secretly borrow and lend means that agents will, regardless of the true history of their endowments, claim having the history of endowment realizations, that delivers the transfer sequence with the highest net present value. In other terms, incentive compatibility prevents the planner from improving welfare by making transfer payments contingent on past public history. It follows that no risk-sharing is possible either, and the only incentive compatible consumption allocation is the one of the bond economy.

4.2 Observational Characteristics of the Optimal Allocation

Proposition 2 suggests that the second best allocation must differ from self insurance. We now study the positive aspects of this result. To make our analysis more applicable for empirical work, we will focus on the properties of consumption and income processes alone.

We discuss first briefly how the second best allocation can be empirically distinguished from the first best and “pure moral hazard” (moral hazard with observable bond holdings). The comparison with the first-best is a relatively easy task. As long as the constrained efficient level of effort is positive, incentive compatibility requires consumption not to be constant. This is in stark contrast with the full-insurance property of the first best. This simple observation is at the core of many tests of full insurance (complete markets) proposed in the literature.²⁷

The distinction between our second best allocation and the “pure moral-hazard model” with monitorable borrowing and lending, can be directly associated to the intertemporal

²⁷See for example, Mace (1991), Cochrane (1991), Townsend (1994), Attanasio and Davis (1996) and Hayashi et al. (1996).

properties of consumption. One possibility is based upon the fact that in a second best allocation, incentive feasibility on asset accumulation requires that consumption satisfies the agents' Euler equation (see condition (21)). Roughly speaking, in a constrained efficient allocation the marginal utility of consumption follows a martingale. In contrast, according to condition (13) the *inverse* of the marginal utility follows a martingale in the pure moral hazard model. The test of this property has been successfully implemented empirically by the work of Ligon (1998) with rural South India data.

The aim of Ligon was to test pure moral hazard versus self insurance. Indeed one cannot use time series properties of consumption alone to distinguish the second best from self insurance, since both models satisfy the usual Euler equation. However, there is another key distinction between the two allocations, which - consistently with the line of proof used in Proposition 2 - is associated to *risk sharing*. In a bond economy (with $b_{-1} = 0$), budget feasibility implies that the discounted present value of income flows must be equal to the discounted present value of consumption flows for each agent. In other terms, each agent has a net present value (NPV) equal to zero:

$$\sum_{t=0}^T \left(\frac{1}{1+r} \right)^t (y_t - c_t) = 0 \text{ almost for any } h^T. \quad (14)$$

This property is in contrast with a second best allocation, where the planner might provide additional insurance by making history dependent transfers. In a second best allocation, the zero NPV restriction must be violated for some history h^T , equivalently the zero NPV condition is not satisfied at the *individual* level. Of course, the usual technical feasibility condition requires that, in both allocations the zero NPV condition must be satisfied at the *aggregate* level. By the law of large numbers, this simply means that

$$\mathbf{E}_0 \left[\sum_{t=0}^T \left(\frac{1}{1+r} \right)^t (y_t - c_t) \right] = 0,$$

which is obviously a much weaker condition than (14).

5 Two Period Model

For didactical reasons, we first describe our approach in the two period version of the model. For this model, we will also be able to characterize the scheme with some detail. Let us

start by considering the agent's problem in his last period. At the beginning of the last period, output y_1^i is realized and the planner's transfer is τ_1^i . The wealth of the agent is thus equal to $z_1^i = y_1^i + \tau_1^i + (1+r)b_0$, where b_0 is the level of asset holdings chosen at the initial period. In this last period, the agent's optimal choices satisfy

$$(b_1^i, e_1^i) \in \arg \max_{e \in E, b \geq 0} u(y_1^i + \tau_1^i + (1+r)b_0 - b, e). \quad (15)$$

The restrictions induced by (15) represents the last period incentive compatibility constraint. It is easy to see from the properties of u , that the optimal solution to this problem is $b_1^i = 0$ and $e_1^i = 0$, regardless of the level of b_0 and $y_1^i + \tau_1^i$. Therefore, the final period (incentive compatible) level of utility is $U_1(z_1^i) = u(z_1^i, 0) = u(y_1^i + \tau_1^i + (1+r)b_0, 0)$.

A second best allocation solves

$$V_2^*(U_0, y_0) = \max_{\substack{e_0 \\ \tau_0, \tau_1^i \\ b_0}} -\tau_0 + \frac{1}{1+r} \sum_i p_i(e_0)(-\tau_1^i),$$

subject to the participation constraint (recall that $b_{-1} = 0$)

$$u(y_0 + \tau_0 - b_0, e_0) + \beta \sum_i p_i(e_0)U_1(y_1^i + \tau_1^i + (1+r)b_0) \geq U_0,$$

and to the initial period incentive compatibility constraint²⁸

$$(b_0, e_0) \in \arg \max_{e \in E, b \geq -\min_i \left\{ \frac{y_1^i + \tau_1^i}{1+r} \right\}} u(y_0 + \tau_0 - b, e) + \beta \sum_i p_i(e)U_1(y_1^i + \tau_1^i + (1+r)b).$$

The *first order approach* replaces the latter incentive constraint for the corresponding stationary points of the agent's maximization problem with respect to e_0 and b_0

$$\begin{aligned} -u'_e(y_0 + \tau_0 - b_0, e_0) &= \beta \sum_i p'_i(e_0)U_1(z_1^i), \\ &\text{and} \\ u'_c(y_0 + \tau_0 - b_0, e_0) &= \beta(1+r) \sum_i p_i(e_0)U'_1(z_1^i) \\ &= \beta(1+r) \sum_i p_i(e_0)u'_c(y_1^i + \tau_1^i + (1+r)b_0, 0). \end{aligned}$$

²⁸This constraint says that, given the transfer scheme $\tau = (\tau_0, \{\tau_1^i\}_{i=1}^N)$, the agent chooses effort level e_0 and bond holdings b_0 so that to maximize his expected discounted utility.

Notice that both the above constraints depend on the agent's consumption and effort equilibrium choices alone. This consideration enables us to rewrite the planner's problem in a much simpler way²⁹

$$\max_{\substack{e_0 \in E, \\ c_0, c_1^i \geq 0}} y_0 - c_0 + \frac{1}{1+r} \sum_i p_i(e_0) [y_1^i - c_1^i] \quad (16)$$

subject to the *relaxed* incentive constraints

$$-u'_e(c_0, e_0) = \beta \sum_i p'_i(e_0) u(c_1^i, 0), \quad (17)$$

$$u'_c(c_0, e_0) = \beta(1+r) \sum_i p_i(e_0) u'_c(c_1^i, 0); \quad (18)$$

and the participation condition

$$u(c_0, e_0) + \beta \sum_i p_i(e_0) u(c_1^i, 0) \geq U_0. \quad (19)$$

Following the lines of Proposition 2 below one can show under very general conditions (in particular without requiring that the first order approach is valid) that *the optimal scheme cannot be such that both c_1^i and τ_1^i increase with i for any i* . However, our aim here is to characterize the scheme in more detail by making use of the first order approach. Therefore, for the remaining of this section we will only consider the (16)-(19) formulation of the problem, assuming that we are entitled to do so. To simplify the analysis we further assume $\beta(1+r) = 1$ and additive separability. The first result is about the monotonicity of the transfer scheme.

Proposition 3 *If $-u'$ is concave (non negative prudence),³⁰ then c_1^i moves together with the likelihood ratio $\frac{p'_i(e)}{p_i(e)}$. In particular, if \mathbf{p} satisfies MLRC³¹ then c_1^i increases with i for $\forall i$.*

²⁹Using the agent's budget constraint, the problem of the planner can be rewritten as follows

$$\max_{\substack{e_0 \\ c_0, c_1^i \\ b_0}} y_0 - c_0 - b_0 + \frac{1}{1+r} \sum_i p_i(e_0) [y_1^i - c_1^i + (1+r)b_0].$$

Then we use the fact that b_0 cancels out to set w.l.o.g. $b_0 = 0$. Since the planner faces the same return as the agents do, he can implement the same allocation for any b_0 with appropriately chosen transfers; and all such allocations induce the same net returns.

³⁰We allow u' to be linear.

³¹In a differentiable world, the monotone likelihood ratio condition (MLRC) is equivalent to $\frac{p'_i(e)}{p_i(e)}$ being non-decreasing in i for every e .

Because of the saving incentives, if u displays negative prudence, it is easy to find examples where c_1^i increases with i even though the likelihood ratio decreases for some i and vice versa. We now address the issue of progressivity.

Definition 1 We say that the transfer scheme τ is progressive (regressive) if $\frac{c_1^{i+1}-c_1^i}{y_1^{i+1}-y_1^i}$ is decreasing (increasing) in i .

Proposition 4 Assume that the likelihood ratio is monotone (MLRC) and convex³² (concave) and that $\frac{1}{w(c)}$ is concave (convex) in c and that the absolute risk aversion $a(c)$ is decreasing and convex (constant).³³ Moreover assume that $y_1^i - y_1^{i-1} = k > 0$ for any i . Then τ is regressive (progressive).

CARA utilities with concave likelihood ratios lead to progressive schemes. When the likelihood ratio is convex, CRRA utilities with risk aversion $\sigma \geq 1$ induce regressive schemes since $a(c)(= \frac{1}{c})$ is strictly convex. Interestingly, this case includes the logarithmic utility case, which - in the observable assets case with linear likelihood ratios - would lead to *proportional* schemes.

Note, that none of the constraints (17)-(19) depends on the actual level of output realizations. This implies that in contrast with the bond economy, consumption does not have to vary monotonically with output. To see this, consider an example where for two specific levels of output y^j and y^l , the effort decision does not influence their probability (i.e. $p_j'(e) = p_l'(e) = 0$ for all effort levels). Effort may be effective for the other output realizations $i \neq j, l$. Note, that c_1^j and c_1^l do not enter (17) and therefore not constrained by effort incentives. For this reason Proposition 3 implies that $c_1^j = c_1^l$. Therefore for this two levels of output there is no correlation between consumption and output. In most cases, the overall correlation between output and consumption is reduced compared to a value equal to 1 in the bond economy, and one can easily find examples where the overall correlation between y and c is negative.

³²That is, for any e, i $\frac{p_{i+1}'(e)}{p_{i+1}(e)} - \frac{p_i'(e)}{p_i(e)} \geq \frac{p_i'(e)}{p_i(e)} - \frac{p_{i-1}'(e)}{p_{i-1}(e)}$.

³³Notice that a function cannot be positive, decreasing and concave everywhere.

6 Recursive Formulation

6.1 The First-Order Conditions Approach

The discussion about the two period model also applies to the more general situation. The adoption of the first order approach means that the set of constraints described in (2) are replaced by the agent's corresponding first-order conditions *along the optimal path*. Using the budget constraint (1) and assuming interiority,³⁴ the agent's first order conditions are, for any $h^t \neq h^{T-1}$

$$-u'_c(\mathbf{c}_t(h^t), \mathbf{e}_t(h^t)) = \beta \sum_i p'_i(\mathbf{e}_t(h^t)) \mathbf{U}_{t+1}^T(\mathcal{W}; (h^t, y^i)); \quad (20)$$

and

$$u'_c(\mathbf{c}_t(h^t), \mathbf{e}_t(h^t)) = \beta(1+r) \sum_i p_i(\mathbf{e}_t(h^t)) u'_c(\mathbf{c}_{t+1}(h^t, y^i), \mathbf{e}_{t+1}(h^t, y^i)). \quad (21)$$

Now, define the set of social contracts that satisfies these first order conditions as follows:

$$\Omega_{FOC} = \{ \mathcal{W} : \text{for any history } h^t \text{ satisfies (20) and (21)} \}.$$

Obviously, any interior contract \mathcal{W} incentive feasible according to (2) - i.e. $\mathcal{W} \in \Omega$ - is such that $\mathcal{W} \in \Omega_{FOC}$. For the purposes of this paper, we should look at conditions under which an optimal solution \mathcal{W}^* to the following *relaxed* planner's problem

$$\begin{aligned} \sup_{\mathcal{W} \in \Omega_{FOC}} \quad & \Pi_0^T(\mathcal{W}; h^0) \\ \text{s.t.} \quad & \mathbf{U}_0^T(\mathcal{W}; h^0) \geq U_0, \end{aligned} \quad (22)$$

³⁴One way to guarantee interiority, is to assume appropriate Inada conditions (as we did in the simulations). In general, since there are no ("ad hoc") borrowing constraints, the set of first order conditions is

$$u_e(\mathbf{c}_t(h^t), \mathbf{e}_t(h^t)) + \beta \sum_i p'_i(\mathbf{e}_t(h^t)) \mathbf{U}_{t+1}^T(\mathcal{W}; (h^t, y^i)) \geq 0$$

with equality if $\mathbf{e}_t(h^t) > 0$, and

$$u'_c(\mathbf{c}_t(h^t), \mathbf{e}_t(h^t)) \leq \beta(1+r) \sum_i p_i(\mathbf{e}_t(h^t)) u'_c(\mathbf{c}_{t+1}(h^t, y^i), \mathbf{e}_{t+1}(h^t, y^i)),$$

with equality if $\mathbf{c}_{t+1}(h^t, y^i) > 0$. When $T < \infty$, we have that $\mathbf{e}_{T-1}(h^{T-1}) = 0$.

is such that $\mathcal{W}^* \in \Omega$. In that case, the solution to problem (22) satisfies incentive compatibility, hence we have actually derived the second best allocation.

A direct application of this argument is used in Section 7.5. After computing the optimal contract according to (22), we will take advantage of the recursive formulation and construct a parsimonious numerical procedure to verify *ex-post* whether the solution to the relaxed problem is actually incentive compatible. The procedure allows the agent to re-maximize his lifetime utility by choosing effort, consumption and bond holdings, taking the optimal (relaxed) transfer scheme as given. Then, we check whether the optimal decisions of this re-maximization problem coincide with the ones implied by the relaxed optimal contract \mathcal{W}^* .

6.2 The Value Function

In this section we will focus on the infinite horizon case, and in order to simplify notation, we do not report the superindex $T = \infty$. The “true” value function of our problem is thus defined as follows

$$\begin{aligned} V_{foc}^*(y_0, U_0) &= \sup_{\mathcal{W} \in \Omega_{FOC}} \Pi_0(\mathcal{W}; h^0) \\ \text{s.t. } \mathbf{U}_0(\mathcal{W}; h^0) &\geq U_0, \end{aligned}$$

where $y_0 = h^0$ is the initial history.

According to the standard recursive contracting literature,³⁵ the use of the agents’s expected discounted lifetime utility as state variable has been proven to be a very valuable tool. However, one can easily realize that when consumers can secretly borrow and lend, agents’ lifetime utilities are no longer sufficient to describe constrained efficient allocations. We will show how recursivity can be recovered with the introduction of an additional endogenous state: the marginal utility of consumption $u_c(c, e)$. The adoption of $u_c(c, e)$ to keep track of agents’ asset decisions becomes natural if one recalls that (i) agents choose bond holdings by equating the marginal rate of transformation ($1 + r$) to the expected marginal rate of substitution between consumption levels at two adjacent dates: $\frac{u_c(c_t, e_t)}{\beta \mathbf{E}[u_c(c_{t+1}, e_{t+1})]}$, and that (ii) in our framework both r and β are known and fixed (in particular, and in contrast with Ramsey taxation problems, the planner cannot impose taxes on capital income).

³⁵Leading examples are Spear and Srivastava (1987), Green (1987), Fudenberg et al. (1990) and Phelan and Townsend (1991).

Our state space will hence become two-dimensional, since we will use both the agent's lifetime utility U , and his marginal utility of consumption $x \equiv u'_c(c, e)$. However, before applying the recursive techniques of Stokey et al. (1989) (SLP) to our problem, we have to solve another difficulty. The feasibility correspondence turns out to be empty for some combinations of the states. To overcome this complication, we could basically follow two alternative procedures. The first possibility is the use of infinite penalizations. This option basically sets the value of the planner to minus infinity for each combination of states (U, x) such that there does not exist an incentive feasible contract that can implement it.³⁶ We choose to follow a second approach, which typically delivers a continuous value function. This second procedure is divided into two main steps. It first derives the state space (or domain restriction) M^* . The set M^* contains all the combinations of (U, x) for which there exists an incentive feasible contract that delivers lifetime utility level U , and a marginal utility of consumption equal to x to the agent in period zero. In the second step of the procedure, the problem is rephrased within the dynamic programming framework, and solved using usual recursive techniques.

The set of incentive feasible states is formally defined as follows³⁷

$$M^* \equiv \{(U, x) \in \mathfrak{R}^2; \exists \mathcal{W} \in \Omega_{FOC}; u'_c(\mathbf{c}_0(h^0), \mathbf{e}_0(h^0)) = x, \text{ and } \mathbf{U}_0(\mathcal{W}; h^0) = U\}.$$

We discuss the properties of this set, and describe how M^* can be computed below. However, it is easy to see that M^* is non empty,³⁸ closed,³⁹ although possibly unbounded. Given M^* we have the following

Proposition 5 *Given an initial history $h^0 = y_0$ and an ex-ante expect discounted utility level U_0 guaranteed to the agent, the value function V_{foc}^* solves*

$$V_{foc}^*(y_0, U_0) = \sup_{x, U \geq U_0; (x, U) \in M^*} V(y_0, U, x)$$

³⁶See Rockafellar (1975), Rustichini (1998), and Messner and Pavoni (2002).

³⁷Notice that, in our repeated framework with *i.i.d.* shocks the exogenous state y does not affect the state space M^* .

³⁸For example, the full insurance contract with $\mathbf{e}_t(h^t) \equiv 0$ for any h^t is always feasible. In this case, for any given $x \in \mathfrak{R}_{++}$, the corresponding utility level is obtained by $U = \frac{u(g(x), 0)}{1-\beta}$, where g is the consumption component of the inverse of the marginal utility of c when $e = 0$. Obviously, self-insurance is incentive feasible as well, and can be implemented with $\tau_t(h^t) \equiv 0$.

³⁹Since the constraint set is formed by either equalities or weak inequalities.

where $V : Y \times M^* \rightarrow \Re$ is a solution to the following functional equation

$$V(y, U, x) = \sup_{\substack{(x^i, U^i) \in M^* \\ e \in E, c \geq 0}} y - c + \frac{1}{1+r} \sum p_i(e) V(y^i, U^i, x^i) \quad (23)$$

s.t.

$$U = u(c, e) + \beta \sum_i p_i(e) U^i \quad (24)$$

$$-u'_e(c, e) = \beta \sum_i p'_i(e) U^i \quad (25)$$

$$u'_c(c, e) = \beta(1+r) \sum_i p_i(e) x^i \quad (26)$$

$$x = u'_c(c, e). \quad (27)$$

Conversely, if a bounded function V defined on $Y \times M^*$ satisfies the functional equation (23)-(27) and M^* is bounded then $V_{foc}^*(y_0, U_0) = \max_{x, U \geq U_0; (x, U) \in M^*} V(y_0, U, x)$.

The function V_{foc}^* is the constrained Pareto frontier, whereas the value function $V(y_0, U, x)$ can be interpreted as the “conditional” utility possibility frontier.

Let us now turn to the domain restriction M^* . According to Abreu et al. (1990), the set M^* can be derived by starting from a sufficiently large set $\overline{M} \supset M^*$, and computing the largest fixed point of the following operator

$$\mathbf{B}(M) \equiv \{(U, x) \in \Re^2; \exists (U^i, x^i) \in M, e \in E \text{ and } c \geq 0; (24)-(27) \text{ are satisfied}\}. \quad (28)$$

It turns out that \mathbf{B} is monotone⁴⁰ and maps closed sets into closed sets. Moreover, since the sequence $M_n = \mathbf{B}^n(\overline{M})$ is monotone, it must converge to the set $M_\infty = \lim_{n \rightarrow \infty} \mathbf{B}^n \overline{M} = \bigcap_{n=1}^{\infty} M_n$, which is closed as intersection of closed sets. It can be shown that if \overline{M} is chosen sufficiently large, we have $M_\infty = M^*$ since the sequence converges to the largest fixed point of the operator \mathbf{B} .

In the finite horizon version of our model, the domain restriction sets can be similarly computed following a backward procedure.⁴¹ Some authors provided numerical techniques to compute sets analogous to M^* (Chang, 1998; Judd et al., 2000; and Cronshaw, 1997). For our computations, we will adopt a discretization procedure which recalls the one used by Chang (1998).

⁴⁰That is, for any two sets $M, M' \subseteq \Re^2$ if $M \subseteq M'$, then $\mathbf{B}(M) \subseteq \mathbf{B}(M')$.

⁴¹One can compute the domain restriction set M_T for the T -horizon problem by applying the same map

Proposition 5 asserts that, under some conditions, the optimal contract can be fully characterized by analyzing the properties of the value function associated to the maximization problem (23)-(27). Unfortunately, it is by no means simple to identify the shape of V . Nevertheless, the recursive formulation can be still profitably used to study the main characteristics of the constrained efficient allocation. To this extent, consider the Lagrangian associated with the *extended static* problem (23)-(27)

$$\begin{aligned} V(y, U, x) = & \sup_{c \geq 0, e \in E, (x^i, U^i) \in M^*} \inf_{\lambda, \mu, \phi, \xi} y - c + \frac{1}{1+r} \sum_i p_i(e) V(y^i, U^i, x^i) + \\ & + \lambda \left[u(c, e) + \beta \sum_i p_i(e) U^i - U \right] + \mu \left[u'_e(c, e) + \beta \sum_i p'_i(e) U^i \right] \\ & + \xi \left[u'_c(c, e) - \beta(1+r) \sum_i p_i(e) x^i \right] + \phi [x - u'_c(c, e)]. \end{aligned}$$

Assuming $\beta(1+r) = 1$, interiority with respect to M^* , and validity of the envelope theorem, we obtain⁴²

$$V'_U(y, U, x) = -\lambda \quad (29)$$

$$V'_x(y, U, x) = \phi. \quad (30)$$

We might assume that V decreases with U and increases with in x , hence $\lambda, \phi \geq 0$. The multiplier λ should be interpreted as the agent's "temporary" Pareto weight. Now consider the first order conditions with respect to U^i :

$$\lambda^i = -V'_U(y^i, U^i, x^i) = \lambda + \mu \frac{p'_i(e)}{p_i(e)}. \quad (31)$$

Let suppose that $\mu > 0$. In this case, condition (31) describes the optimal incentive provision characteristic typical of efficient allocations under moral hazard. Because of its informational content on the supplied level of effort, the sign and magnitude of the likelihood ratio

B we defined above as follows

$$M_T = \mathbf{B}(M_{T-1}) = \mathbf{B}^{T-1}(M_1),$$

starting from the set of implementable states in a one-period problem:

$$M_1 = \{(x, U) : u_c(c, 0) = x; U = u(c, 0) \ c \geq 0\}.$$

⁴²It is easy to see that in our repeated model neither V_U nor V_x depend on y .

$\frac{p'_i(e)}{p_i(e)}$ (and only this) induce proportional variations on the Pareto weight λ^i assigned to the agent. Now, take expectations on (31), and use (29) to get

$$V'_U(y, U, x) = \mathbf{E} [V'_U(y^i, U^i, x^i) | e]. \quad (32)$$

Condition (32) should also be familiar. It can be interpreted as the planner's Euler equation, where U is the relevant asset, and (24) is the relevant budget constraint. According to (32), the planner allocates agent's lifetime utilities *intertemporally* in order to maximize his net surplus. While doing so, he must respect the "cross-sectional" restriction implied by the effort incentive constraint (25). This restriction makes U an imperfect asset similar to a risk-free bond. Indeed, (25) roughly requires that the different U^i 's move altogether. In a first best world, the planner would be able to allocate U^i optimally also across different states and therefore U^i would become constant in i .

The necessary conditions for the optimal choice of x^i are

$$V'_x(y^i, U^i, x^i) = (1 + r)\xi \text{ for any } i = 1, 2, \dots, N. \quad (33)$$

According to (33), the planner allocates marginal utilities x^i so that to equalize (marginal) costs *across* the different states, subject to the intertemporal constraint (26) which require an average level of marginal utility to be delivered to the agent. A positive ξ forces the planner to deliver in each state a marginal utility of consumption which is "distorted" with respect to the promised level U . In a pure moral hazard model with observable assets, ξ would always equal zero and the marginal utility of consumption would be perfectly aligned with U .

Finally, we briefly analyze the necessary condition for c

$$-1 + \lambda u'_c(c, e) + \mu u''_{ec}(c, e) + (\xi - \phi)u''_{cc}(c, e) = 0. \quad (34)$$

Assume for simplicity that $\xi - \phi \geq 0$ and linear separability ($u''_{ec}(c, e) = 0$). Since $u''_{cc}(c, e) < 0$, condition (34) suggests that in order to discourage agent's savings the planner is forced to reduce today's consumption below the level that would be implied by the Pareto weight λ . Notice moreover, that only the difference between distortion needed today (ξ) and the one already inherited from the past (ϕ) matters. For example, suppose that $\xi = \phi > 0$. In this case, the planner does not need any further "correction," on c , as all the required distortions are fully summarized by λ , and we have $-V'_U(y, U, x) = \frac{1}{x} = \frac{1}{u'_c(c)}$.

7 Quantitative Analysis

In this section we present quantitative results regarding the normative and positive characteristics of the second best allocation. Our main purpose is to explore further the properties of the efficient allocation of consumption and quantify the welfare gains. In Section 7.3 we study the properties of the state space (or domain restriction set) M^* . In Section 7.4 we use our results to analyze what are the consequences of the introduction of a benevolent social planner to our bond economy. In particular, we explore how the characteristics of the consumption path may vary between the bond economy, the second best and a model of pure moral hazard (when assets are observable). We also provide quantitative evidence on the efficiency implications. In Section 7.5 we present the ex-post verification procedure of the first-order approach based upon our numerical solution.

7.1 Numerical Methodology

To reduce computational difficulties we assume only two possible realizations of output ($N = 2$). Hereafter we will refer to the “high” output realization ($i = 2$) with the superindex h and to the low output realizations ($i = 1$) with superindex l , and we define $p(e) \equiv \Pr(y = y^h \mid e)$, and $1 - p(e) \equiv \Pr(y = y^l \mid e)$. Throughout this section we also assume the following representation for agents’ preferences: $u(c, e) = u(c) - v(e)$. Note that in this case concavity implies a one-to-one relationship between consumption c and marginal utility $x = u'(c)$. As a consequence, we can use consumption itself as a state variable, together with the lifetime utility U and the exogenous state y .

Recall that our aim is to solve the following constrained optimization problem:

$$V(y, U, c) = \sup_{\substack{(c^h, U^h), (c^l, U^l) \in M^* \\ e \in E}} y - c + \frac{1}{1+r} [p(e)V(y^h, U^h, c^h) + (1-p(e))V(y^l, U^l, c^l)]$$

s.t.

$$U = u(c) - v(e) + \beta [p(e)U^h + (1-p(e))U^l] \tag{35}$$

$$v'(e) = \beta p'(e) (U^h - U^l) \tag{36}$$

$$u'(c) = \beta(1+r) [p(e)u'(c^h) + (1-p(e))u'(c^l)]. \tag{37}$$

Our first step is to compute of the domain restriction set M^* . Recall that M^* is the set of (U, c) couples for which there exists $(U^l, c^l) \in M^*$ and $(U^h, c^h) \in M^*$ such that (35)-

(37) are all satisfied. In order to construct M^* , we used a modification of the algorithm proposed by Chang (1998).⁴³ We defined a two dimensional grid $M_0 = \{U_k\}_{k=1}^n \times \{c_j\}_{j=1}^m$ where c_j ranged between $\underline{c} = \underline{\alpha}y^l$ and $\bar{c} = \bar{\alpha}y^h$ and U_k ranged between $\underline{U} = \frac{u(\underline{c}) - v(\max(E))}{1-\beta}$ and $\bar{U} = \frac{u(\bar{c}) - v(0)}{1-\beta}$.⁴⁴ We then applied the set valued operator \mathbf{B} defined in (28) on M_0 iteratively until we obtained a fixed point.

Since the probability shifting specification makes the numerical solutions methods involving first order conditions more difficult to implement, in order to compute the optimal program we used a value function iterations method. We approximated the value function with Chebyshev polynomials. For a more detailed explanation of our approach, the interested reader should go to Appendix C. However, notice that from (35) and (36) we can express U^l and U^h as a (possibly implicit) function of e and our states. Similarly, (37) gives c^l and c^h as a function of e , $\Delta u' = u'(c^h) - u'(c^l)$ and c . Therefore, we can substitute U^l , U^h , c^l and c^h into the principal's objective function and implement numerical maximization with respect to e and $\Delta u'$ for any given U, c and y .

Finally notice, that the problem of the initial period differs from the subsequent ones, since in this case the planner is not constrained to keep the promise on consumption (or equivalently on marginal utility). However, according to Proposition 5, the knowledge of the value function $V(y, U, c)$ is enough to determine the value of the planner at period 0. The initial period's problem then can be solved by *choosing the initial values* for c and U performing the following numerical maximization

$$\max_{\substack{c, U \geq U_0 \\ (c, U) \in M^*}} V(y, U, c). \quad (38)$$

If we call $U^*(U_0), x^*(U_0)$ the result of such a maximization program, we have $V_{foc}^*(y, U_0) = V(y, U^*(U_0), x^*(U_0))$.

7.2 Functional Forms and Parameters

Since the purpose of our numerical analysis is to compare welfare gains of different economic environments, and because our model is not linked directly to any particular “empirical”

⁴³The details of the procedure are explained in Appendix C.

⁴⁴It turned out that the actual value of the highest possible effort level - $\max(E)$ - had very little influence on the shape of the feasibility set. The only possibly harmful grid restriction is the upper limit on attainable consumption \bar{c} . This problem will be discussed below.

problem, our calibration process is not linked to real world data either. Our parametrization is closest to an unemployment insurance problem where any particular agent faces a very low income when he is unemployed and a relatively high income when he is employed, therefore his effort decision can be interpreted as search effort when unemployed and “work effort” when he is employed.⁴⁵

We chose a standard CRRA utility for consumption: $u(c) \equiv \frac{c^{1-\sigma}}{1-\sigma}$; and quadratic disutility of effort: $v(e) \equiv \eta e^2$. The probability of the high output realization is given by $p(e) \equiv 1 - \exp(-\rho e)$. It is easy to see that this specification of $p(e)$ satisfies both full support, CDFC and MLRC; moreover (36) implies that positive effort is implemented if and only if $U^h > U^l$.

The table below contains the parameter values we used for our calculations. We used two values of risk aversion for our welfare comparisons and we chose η in a way that in the first best economy we obtained the same optimal allocation with the two risk-aversion parameters. This choice was motivated by the fact that in Section 7.4 we use the compensating differential in terms of consumption for comparing welfare across different models. Therefore, in order to have a meaningful comparison across different risk aversion cases, we need comparable levels of consumptions as well.

| | β | σ | ρ | y^l | y^h | η |
|--------------------|---------|----------|--------|-------|-------|--------|
| Low risk aversion | 0.99 | 0.5 | 0.1 | 0.1 | 100 | 0.5 |
| High risk aversion | 0.99 | 2 | 0.1 | 0.1 | 100 | 0.0067 |

7.3 The Domain Restriction

For the practical implementation of the iterative procedure that determines M^* we used as benchmark $\underline{\alpha} = 0.1^{-4}$, $\bar{\alpha} = 1$ and r such that $\beta(1+r) = 1$.⁴⁶

Figure 1A here

⁴⁵If asset holding were observable, this could be seen as an *i.i.d.* shocks version of the UI model with moral hazard also associated with job retention proposed by Zhao (2001), who in turn builds on Hopenhayn and Nicolini (1997).

⁴⁶Since we chose $\beta = 0.99$, we can interpret a time period in our model as one quarter. This implies an annual interest rate equal to 4.1 percent.

Figure 1A shows our resulting domain restriction set and a set of realized (U, c) couples for 1000 time periods.⁴⁷ Notice that the realizations are in the interior of the set and far from the upper limit of consumption. In fact, we checked robustness by increasing \bar{c} and we found that, although the feasible set might expand upwards significantly, the realized equilibrium path does not change in any noticeable way. The set is a relatively nicely shaped and connected set. In particular, and in contrast with the results of some other authors for different models (see Chang, 1998), it does not present “holes”.⁴⁸

There are two types of promised-utility/promised-consumption combinations (U, c) which are not feasible. First, notice that at the bottom right of the figure, bundles containing high life-time utility and low consumption are excluded. However, this is a pure numerical artifact due to the upper bound on consumption.⁴⁹ We show in Appendix C that if we modify our algorithm so that the upper limit on consumption is somehow relaxed, these bundles are not excluded from the feasibility set. More importantly, at the top left of the grid, one can notice that high promised consumptions are not compatible with low levels of life-time utility. This is so since if today’s consumption level is high and at the same time lifetime utility must be kept low, future consumption must decrease, as a consequence agents would have incentive to save, and this cannot be incentive compatible (i.e. violates the Euler’s equation). Figure 1B shows that M^* is shrinking when the interest rate increases. This is natural and confirms our previous intuition about the shape of M^* . When the interest rate increases incentive costs related to unobservable asset accumulation increase as well and therefore, the planner is able to implement only a smaller set of allocations.

Figure 1B here

7.4 Comparing Different Allocations

In this section we will show some qualitative and quantitative differences between our model, the bond economy, the first-best, and the pure moral hazard model (with observable

⁴⁷All the results presented in the quantitative section refer to the low risk aversion case unless stated otherwise.

⁴⁸However, we also find that the procedure applied to very coarse grids does “artificially” eliminate some points in the interior of the set.

⁴⁹In this case, in order to satisfy the promise-keeping constraint and the Euler equation simultaneously, large upward jumps are needed in consumption, but they are not available because $0 \leq c \leq \bar{c}$.

assets). In our set-up, the bond economy allocation with $T = \infty$ solves

$$\begin{aligned}
U^{Bond}(z) &= \max_{e \in E, c \geq 0, b \geq -B} u(c) - v(e) + \beta [p(e)U^{Bond}(z^h) + (1 - p(e))U^{Bond}(z^l)] \\
&\text{s.t.} \\
c + b &\leq z \\
z^i &= y^i + (1 + r)b; \quad i = h, l \text{ and}
\end{aligned}$$

The method of value function iterations using Chebyshev polynomials was applied again with z as a state variable. We have initialized our benchmark bond economy model with $b_{-1} = 0$ and we imposed the “natural debt limit” $B = \frac{y^l}{r}$. This is the debt limit which guarantees that the agent is able to pay back his debt even under the “worst” circumstances (see Aiyagari, 1994).⁵⁰

Following the standard recursive contracts literature, we also solved the following infinite-horizon pure moral hazard model, using the lifetime utility U as the only endogenous state variable

$$\begin{aligned}
W(U, y) &= \max_{c \geq 0, e \in E, U^h, U^l \in D^*} y - c + \beta [p(e)W(y^h, U^h) + (1 - p(e))W(y^l, U^l)] \\
&\text{s.t.} \quad : \\
U &= u(c) - v(e) + \beta [p(e)U^h + (1 - p(e))U^l] \tag{39}
\end{aligned}$$

$$v'(e) = \beta p'(e) (U^h - U^l). \tag{40}$$

We used value function iterations. The domain restriction set D^* on U have been computed starting from $D_0 = \left[\frac{u(\underline{c}) - v(\max(E))}{1 - \beta}, \frac{u(\bar{c}) - v(0)}{1 - \beta} \right]$, and applying numerical techniques similar to the ones used to determine M^* .⁵¹

Finally, the first best allocation is straightforward, and it involves a deterministic path for c_t and e_t .

Policy Functions We show the policy functions for the second best, the first best and for the bond economy on Figure 2 and 3, assuming $\beta(1 + r) = 1$. In each case, we used

⁵⁰For comparison purposes we imposed the same upper limit on consumption as in the model of moral hazard with hidden savings, this restriction implies an upper bound on asset holdings.

⁵¹The set valued operator in this case is

$$\mathbf{B}^{MH}(D) = \{U : \exists c, e \geq 0, U^h, U^l \in D; \text{ both (39) and (40) hold}\}.$$

For the low risk aversion case, we got $D^* = \left[0, \frac{u(\bar{c})}{1 - \beta} \right]$.

promised utility (U) as the independent variable and assumed that the agent has y^l as his initial output. Moreover, to eliminate the dependence on promised consumption c , we consider period zero policy functions, where c is a choice variable for all models. Finally, note that the agent's lifetime utility in the bond economy is strictly increasing in the level of wealth z , therefore we can express policies unambiguously in terms of $U(= U^{Bond}(z)$ for some z) in the self-insurance case as well.

Figure 2A and 2B here

As expected, the profile of consumption is much steeper in the case of the bond economy than in the case of the first and second best. Overall, the main reason is the usual precautionary saving motive. The incompleteness of the assets markets in self insurance induces agents to save a high proportion of their income when their asset level is low, and increase the consumption rate as their wealth accumulates. In a second best allocation, precautionary saving motives are less important, as agents face a larger set of assets. In the region where lifetime utility is low, there is an additional effect that tends to increase the steepness of the self insurance consumption plan with respect to the second best. Recall that, in the bond economy, low U imply low asset holdings. When lifetime utility is low self insurance consumption is depressed also because the agent is close to the borrowing limit. In the second best, the planner's reallocation of resources effectively relaxes the liquidity constraint and permits low wealth agents to increase consumption. Finally, notice that consumption in the first best is always above the second best consumption policy. This is intuitive since in a second best allocation the Euler equation (together with the Jensen inequality) implies an increasing average consumption path, while the first best allocation offers constant consumption. This different intertemporal behavior reduces the second best level of consumption in period zero, and increases it in future periods with respect to the first best.

The profile of effort is practically the mirror image of the pattern of consumption. This is not surprising for two main reasons. First, notice that "leisure" is normal good in our examples, therefore as wealth (or welfare) increases effort decreases ("leisure" increases). In a second best allocation, this is the main effect. In a bond economy, there is a second reason for this pattern. Recall that in self insurance a large c means high wealth level. In this case, as wealth increases, the agent is less worried about output fluctuations since his effective *absolute* risk aversion is reduced. Therefore the dispersion in lifetime utilities

decreases and, as a consequence, the agent's effort incentives are reduced further as well. Finally, notice that at any U the first best effort level is larger than the one implemented in the second best, even though consumption is higher in the former case. This property reflects the non-observability of effort in a second best allocation: additional incentive costs are associated with the implementation of a given effort level.

Figure 3A and 3B here

Figure 3 shows next period's promised consumption (3A) and promised utility (3B), as a function of utility today. The first best case in 3B is just the 45 degree line, as the planner offers constant lifetime utility. We see a clear pattern in terms of risk here: when U is low, the agent faces more risk in the bond economy (in terms of the differences in next period's utility or consumption). However as U increases, that is as the agent accumulates assets, this risk diminishes in terms of life-time utility and eventually it becomes smaller than the one faced by the agent in the second best. That is, in a second best economy the planner smooths the agents' consumption across states (insurance efficiency) for lower levels of wealth. However, this insurance property is weakened as the agent becomes more wealthy because he appreciates leisure more and more. Total output efficiency requires here that the planner offers sufficient amount of risk to offset this tendency. In summary, at low levels of utility the differences between self insurance and the second best are mainly driven by insurance efficiency. While for high U the dominating factor is total output efficiency. This characteristic will become more apparent in the following time series analysis.

Some Time Series Aspects of the Optimal Contract We can see the properties of the second-best allocation from time series perspective on Figure 4. There, we present the average of 1000 independent simulations, each 500 periods long, for the same three models (first-best, pure moral hazard and second-best). For all three cases, U_0 was chosen in a way that the planner makes zero profits given the same initial output distribution.

Figure 4 here

This figure shows the consequences of the introduction of different informational problems on the optimal allocation. The first best obviously represents the case where there are no incentive problems and, in this case, the planner implements a constant consumption/effort path. When the planner faces moral hazard only with respect to effort, the first

best allocation cannot be implemented and the result is a decreasing average consumption path (and consequently decreasing life-time utility), and increasing effort. This result confirms the martingale property of the inverse of the marginal utility displayed in equation (13).⁵² This property leads to the well-known gradual immiseration, or forever decreasing lifetime utility, which is apparent in Figure 4.⁵³ On the other hand, recall that this allocation cannot be incentive compatible when the agent has access to unobservable assets because this pattern of consumption implies that the agents would be willing to save and decrease their current consumption. The planner is hence forced to adjust the allocation in order to be compatible with the agents' saving incentives (the Euler equation). Notice, that in our example this adjustment changes the time series characteristics of the optimal allocation dramatically. In particular, the second-best allocation (with hidden assets) inherits the increasing pattern of consumption and lifetime utility, key properties of the bond economy.

In Figure 5 we compare the average time pattern of the second best with the bond economy.⁵⁴ As it was mentioned above, qualitatively, these patterns are somehow similar. However, the second best allocation allows for much higher degree of consumption smoothing than self insurance does. That is, Figure 5 shows that the second best allocation provides not only cross-state insurance (see Figure 2A) but intertemporal one as well. Recall that asset holdings are the only insurance device for the agent in the bond economy and in order to obtain higher level of insurance they need to accumulate a significant amount of assets. Initially, when they have low levels of assets, they need to use their effort to “buy” some extra insurance: with higher effort they can obtain higher expected income available for wealth accumulation. This explains the much higher initial average effort levels in the bond economy. In a second best allocation the planner can cross-insure agents, and there is no need for this steep wealth accumulation. The consequence is a much flatter consumption path and more stable effort decisions.

Figure 5 here

⁵²In our benchmark case (with $\sigma = 1/2$ and $\beta = \frac{1}{1+r}$) (13) becomes $\sqrt{c_t} = \sum_i p_i(e) \sqrt{c_{t+1}^i}$, which implies $c_t > \sum_i p_i(e) c_{t+1}^i$ by the usual Jensen inequality.

⁵³See Thomas and Worrall, 1990.

⁵⁴Recall that we assumed $b_{-1} = 0$ for every agent in the bond economy.

On the other hand, as agents become more wealthy in the bond economy they exert smaller and smaller effort, which results in decreasing level of average output. Therefore, they rely more and more on their interest income for financing their consumption. This explains why the asset accumulation path becomes flatter.⁵⁵ This is a force specific to our model since it is coming through the channel that income distribution is dependent on wealth through effort (leisure) choice.⁵⁶ This mechanism leads to a situation where agents in the bond economy are in average “wealthy and lazy”.

Observe, that the forces operating in the second best economy place the consumption patterns of the second best allocation *in-between* self insurance and pure moral hazard. Consumption in the pure moral hazard is frontloaded. On the other hand it is clearly backloaded in the case of the bond economy. It is also backloaded in the case of the second best, but to much less extent.⁵⁷

The fourth panel of Figure 4 and 5 completes the picture. In the bond economy, agents accumulate assets very heavily in early periods due to precautionary saving motives. Therefore they achieve high level of lifetime utilities in average much faster than agents in the second best. We can determine the implied asset holdings in the planner problems as well.⁵⁸ Note that asset holdings are constant in the first best. This is natural, but contrasts with the second best allocation. In this case the planner accumulates assets in order to support an (in average) increasing consumption path. However, since the planner can cross-insure agents much less asset accumulation is required than in the bond economy. In the case of pure moral hazard the average asset holdings decrease as consumption is frontloaded. This picture suggests an interesting cross-country relationship between informational restrictions on domestic asset accumulation and the net asset position against the rest of the world. In particular, economies where the credit market is largely monitored tend to accumulate foreign debt as they attract foreign capital.

⁵⁵We believe that the contribution to this pattern of the upper bound on assets is minimal since we set it around 10,000, and the asset accumulation path starts to flatten out around the level of 850.

⁵⁶This effect is for example not present in the standard models of self insurance with heterogeneous agents (e.g. Huggett, 1993; Aiyagari, 1994; or Krusell Smith, 1998) since in these economies agents face exogenous shocks and supply a fixed amount of labour.

⁵⁷Notice that, when $\sigma > 1$, average consumption would increase with time in the pure moral hazard allocation as well; although it would present a much less steep path than the one followed by the average consumption in the second best allocation.

⁵⁸See Appendix D for details.

Quantify Welfare Gains We can measure welfare gains in several ways. Table 2 shows the results of our calculations for welfare gains at the first best, at the model of pure moral hazard and at the model of moral hazard with hidden savings for different interest rates and for different risk aversion parameters. The first panel contains the efficiency gains for our benchmark model with low risk aversion and $\beta(1+r) = 1$.

Table 2 here

We assumed everywhere that $y_0 = y^l$. In each panel, the first two rows show gains in terms of the level of initial utility (U_0). These are the utility levels which would guarantee zero total surplus (profits) to the planner. In other words, they would be the agents' expected life-time utility when incentive-compatible insurance contracts are available to him. The third row contains the initial saving levels b_{-1}^S which would induce the life-time utility obtained by the particular planner problem in self insurance. For example, if U_0 is the second-best utility level as described above, then b_{-1}^{SBest} is such that $U^{Bond}((1+r)b_{-1}^{SBest} + y_0) = U_0$, where the value function $U^{Bond}(\cdot)$ has been defined above. In the benchmark case of Table 2, if an agent in self insurance receives a lump-sum transfer of 62.3 in period -1 (or, equivalently, $62.3(1+r)$ in period 0) then he would obtain the same life time utility (645.4) which he can obtain in second best with zero initial wealth. The total surplus of the planner is given by $W(y_0, U_0)$ and by $V(y_0, U^*(U_0), x^*(U_0))$ for the pure moral hazard and for our problem respectively (the first best case is straightforward). These figures are in the fourth row of each panel.⁵⁹

Finally, we also calculated compensating differentials in terms of consumption for the three problems and reported them in the last row of each panel. In order to obtain these results, we fixed the policy function of effort to the one of the bond economy. This normalization reinforces the interpretation of this compensating variation, since this means that we express all welfare gains in terms of consumption changes. The first panel shows that welfare gains in our model can be expressed by a 5.7 percentage change in consumption in all states and time periods compared to self insurance.

In general, we can say that, in this benchmark case, welfare gains are big compared to self insurance, since they are actually very close to the gains of the first best. On the

⁵⁹Note, that the self insurance allocation is available for the planner in all scenarios, therefore in the worst case he has zero surplus.

other hand, since the welfare gains of the second best are close to the ones of pure moral hazard, we can claim that the welfare losses due to the additional informational constraint are relatively small when risk aversion is low, which is intuitive.

We saw that even when $\beta = \frac{1}{1+r}$ consumption is not stationary either in the bond economy or in the second-best. However, because of the precautionary savings motive agents in the bond economy experience a very steep asset accumulation path initially. It is important to see how much of the total welfare gains are due to the longer run behavior and how much are due to this phase of intensive asset accumulation. For this reason we computed the welfare gains for the same economies also for a higher level of initial utility/asset holdings. We expect smaller welfare gains in this case because agents in the bond economy can self insure themselves much better with the corresponding higher initial asset holdings. This conjecture is confirmed, however the second best still improve significantly on the bond economy allocation (3.5 percent in terms of consumption growth).

We calculated welfare gains for different interest rates as well. Somewhat surprisingly, welfare gains increased both when we increased and decreased the interest rate. It is not surprising in the case of a smaller interest rate because agents are worse off in the bond economy (they can obtain less self insurance) and the state space for the planner increased, as the asset accumulation incentive constraint is relaxed. In contrast, according to Figure 1B, the state space became smaller when we increased the interest rate, because it is harder to control the agent's incentive to save. Moreover, when r is high self insurance itself performs better compared to the benchmark case. So the increase in welfare gains in this latter case is somehow surprising. However, there is another force which is operating here. In contrast to the bond economy, in a second best allocation the agent has access to a complete set of contingent commodities. An increase in r reduces the (discounted) price of such history contingent commodities and the agent can make additional profits with respect to self insurance by intertemporally trading assets related to these goods. This latter force locally dominated the other ones, and we see increasing welfare gains in this direction as well.

Finally, we checked how these measures of welfare gains change if we increase the risk aversion of the agents. If they are more risk averse, limited insurance opportunities hurt them more, therefore full-insurance (the first best) will imply significantly higher welfare gains (33.2 percent versus 6.6 percent in consumption terms when $\sigma = 1/2$). On the other

hand, the sharp decline in welfare in both the second best and the pure moral hazard allocation compared to the first best suggests that incentive problems are more severe in this case. In addition, notice that welfare gains of the second best are largely reduced also in comparison to the pure moral hazard allocation. The non-observability of asset holdings in the second best implies a further significant reduction in welfare because the incentive constrains on savings are more effective now. A large σ forces the planner to implement a consumption path which is steeper than the low σ case. This, in turn, hurts more the agent as he dislikes more intertemporal variations.

Decomposition of welfare gains We have already pointed out that the main sources of welfare gains can be decomposed between insurance and total output efficiency. For this reason, it is important to separately quantify these two components. To get some intuition about our measure of production efficiency, consider the first best condition (8),⁶⁰ for a given level U_0 of lifetime utility guaranteed to the agent in period zero. In a first best allocation, once full insurance is provided to him ($c^*(U_0)$ is constant), the effort choice $e^*(U_0)$ is chosen so as to maximize total output efficiency (expected output minus costs):

$$\frac{1}{1+r} \sum_i p_i(e) y^i - \frac{v(e)}{u'(c^*(U_0))}.$$

We are interested in measuring the deviations of our different allocations from this first best benchmark. The expected discounted production differential between model $S = SBest$, $PureMHazard$, $Bond$, and the first best can be computed as follows⁶¹

$$\Delta_{\text{prod}}^S(U_0) = \frac{\sum_i p_i(e^*(U_0)) y^i}{r} - \frac{1}{1+r} \mathbf{E}_0^S \left[\sum_{t=0}^{\infty} \left(\frac{1}{1+r} \right)^t y_t^S \mid U_0 \right],$$

where, for the bond economy we simply chose b_{-1} so that $U^{Bond}((1+r)b_{-1} + y_0) = U_0$. We then compute the effort costs differentials with respect to first best

$$\Delta_{\text{cost}}^S(U_0) = \frac{\frac{v(e^*(U_0))}{u'(c^*(U_0))}}{1-\beta} - \mathbf{E}_0^S \left[\sum_{t=0}^{\infty} \beta^t \frac{v(e_t^S)}{u'(c_t^S)} \mid U_0 \right].$$

⁶⁰The separable utility version of equation (8) is

$$\frac{v'(e^*(U_0))}{u'(c^*(U_0))} = \frac{1}{1+r} \sum_i p'_i(e^*(U_0)) y^i,$$

where $e^*(U_0)$ and $c^*(U_0)$ are such that $\frac{u(c^*(U_0)) - v(e^*(U_0))}{1-\beta} = U_0$.

⁶¹The notation “ $\mid U_0$ ” recalls the fact that we used the same U_0 in all models.

Then, $\Delta_{\text{prod}}^S(U_0) - \Delta_{\text{cost}}^S(U_0)$ measures the welfare loss of model S due to the different choice of effort, hence it measures total output loss of efficiency with respect to the first best.⁶² Notice that we express welfare gains in present consumption goods, so that they are also directly comparable to the planners's surplus. Our calculations show that after the phase of intensive asset accumulation ($U_0 = 720$), roughly half of the welfare gains are due to total output efficiency gains in the case of the first best and second best. This proportion is much smaller for the low assets case ($U_0 = 622.6$). This confirms our previous discussions. For low asset levels, the main welfare improvements on self insurance are obtained through insurance efficiency.

Finally, we should mention that total output efficiency plays a more important role in a pure moral hazard allocation than it does in a second best allocation, in explaining the efficiency gains with respect to self insurance. This is so because the introduction of hidden assets increases the effort implementation costs, since a larger consumption dispersion (if e is high) also implies a steeper consumption path (because of Jensen inequality applied to the Euler equation). The result is that in a second best allocation the agent is more insured than in the case of pure moral hazard, but this effect of course reduces effort incentives, and total output efficiency.

7.5 Ex-post Verification Procedure of the First-Order Approach

The numerical solution enables us to develop a verification procedure of the first-order condition approach. Recall, that the validity of this approach is a prerequisite of our recursive reformulation of the constrained efficient problem. The basic idea of the procedure is coming from the fact that for any particular realized output sequence we can compute the implied transfer series (assuming $b_t = 0 \forall t$ for example) for our recursive second best. Then, we can verify the first-order approach by allowing the agent to remaximize, assuming that he receives the transfer series defined by the second best problem. The first-order approach is valid if agents choose the same consumption and effort (and bond

⁶²For example, in order to determine the total output gains between the second best and the bond economy we just have to compute

$$\Delta_{\text{prod}}^{\text{Bond}}(U_0) - \Delta_{\text{cost}}^{\text{Bond}}(U_0) - [\Delta_{\text{prod}}^{\text{SBest}}(U_0) - \Delta_{\text{cost}}^{\text{SBest}}(U_0)].$$

holdings) as the one implied by the initial solution.

Given a transfer scheme $\tau = \{\tau_t(h^t)\}$, the agents' incentive constraint is described by the fact that the agent chooses contingent plans σ of effort, consumption and bond holdings solving the following (re)maximization problem

$$U_0^R(\tau) = \sup_{\sigma} \mathbf{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t, e_t) \mid \alpha \right] \quad (41)$$

s.t.

$$c_t + b_t = y_t + \tau_t(h^t) + (1+r)b_{t-1} \text{ for any } h^t; \text{ and } b_{-1} = 0$$

where α is the effort plan implied by σ .

Definition 2 *If $U^*(U_0)$ is the expected discounted lifetime utility level obtained by the agent in the second best allocation according to (38), and if τ is the associated second best transfer scheme, we say that the first order condition approach is justified if $U_0^R(\tau) = U^*(U_0)$.*

This verification procedure seems to require a formidable task, since τ is an infinite dimensional ‘‘parameter.’’ The key simplification of our procedure comes from the observation that the recursive approach implies that the history dependence of the endogenous variables can be summarized by the states U_t, x_t .⁶³ In turn, within the dynamic programming framework, the states evolve according to the second best *time invariant* policies

$$U_{t+1}^i = g(U_t, x_t, y_{t+1}^i) \quad (42)$$

$$x_{t+1}^i = h(U_t, x_t, y_{t+1}^i), \quad i = 1, 2, \dots, N. \quad (43)$$

As a consequence, the re-maximization problem (41) can be written in recursive form as follows

$$V^R(U, x, b) = \max_{c^R, b^R, e^R} u(c^R, e^R) + \beta \sum_i p_i(e^R) V^R(U^i, x^i, b^R) \quad (44)$$

s.t.

⁶³Since the shocks are i.i.d., the relevant policies do not depend on y_t either. Using the budget constraint (and fixing bond holdings equal to zero) the planner's transfers can be recovered as follows

$$\tau_t = f(U_t, x_t, y_t) = c(U_t, x_t) - y_t.$$

$$\begin{aligned}
c^R + b^R &\leq c(U, x) + (1 + r)b \\
U^i &= g(U, x, y^i) \\
x^i &= h(U, x, y^i).
\end{aligned}$$

Therefore, given that the transfer scheme τ is exogenous, the problem to be solved is very similar to the self insurance one, since the *only endogenous state variable* is the bond level b . According to Definition 2, the first order approach is verified if we have $V^R(U^*(U_0), x^*(U_0), 0) = U^*(U_0)$, where $U^*(U_0), x^*(U_0)$ are the initial values for the states, we derived using the maximization problem (38) above and the initial level of assets is $b_{-1} = 0$.

We implemented this procedure by defining a grid M on U and x and discretize the optimal policy rules (42) and (43). Then we defined a grid $G = \{b_s\}_{s=1}^q$ on asset holdings and solved (44) as a discrete dynamic programming problem. This approach has one particular advantage. We know that if the first order approach is valid then the agent will have no incentive to hold any assets since we calculated the transfer scheme assuming zero asset holdings. Therefore by fixing the grid for b to the singleton $G = \{0\}$ we could test whether the agent has incentives to deviate when he can choose only his effort level, as consumption is determined by the budget constraint.

Recall that in the separable case we used c as a state. Therefore, using agent's budget constraint the procedure computes the limit point (as $t \rightarrow \infty$) of the sequence of functions

$$V_{t+1}^R(U, c, b) = \max_{b^R \in G, e^R \in E} u(c + (1 + r)b - b^R) - v(e^R) + \beta \sum_i p_i(e^R) V_t^R(g(U, c, y^i), h(U, c, y^i), b^R),$$

with $(U, c, b) \in M \times G$, and starting function

$$V_1^R(U, c, b) = \max_{b^R \in G, e^R \in E} u(c + (1 + r)b - b^R) - v(e^R) + \beta \sum_i p_i(e^R) g(U, c, y^i).$$

Note that $U^*(U_0)$ and the laws of motions for U and x are approximated. Moreover the discretization adds an additional error to their approximation. Therefore we cannot expect $U_0^R(\tau) = U^*(U_0)$ hold exactly. But these are the very errors which identify our verification test. For our final tests we used the following grid

$$G = [-0.5, -0.25, 0, 0.25, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 19, 25, 35].$$

We claim that our verification results validate the first-order approach, because the discrepancy between $U_0^R(\tau)$ and $U^*(U_0)$ is low compared to the welfare gains. In particular,

for the low risk aversion case, this discrepancy is below 0.01 percent for all levels of initial utility where we evaluated the welfare gains (see the second line of the first panel in Table 2 for the actual levels of utilities in percentage terms). The discrepancy is substantially larger for the high risk aversion case: 0.3 and 0.2 percent for the low initial utility and high initial utility cases respectively. However, these values are still significantly below the magnitude of the welfare gains (6.4 percent and 1.7 percent respectively).⁶⁴

The errors are not sensitive to the range of the grid. As we expected, they are more sensitive to how fine the grid is. Around the value $b = 0$, we chose values on G consistently with the grid of the states (U, c) (in particular with the grid we chose for c). Otherwise, the agent could simply gain on the numerical imprecision of the scheme or, when G is too coarse, could be excessively constrained by the discrete available levels of savings.

8 Conclusions

This paper develops a tractable recursive framework which is suitable for studying efficient allocations in the dynamic moral hazard problem with non-observable asset accumulation.

Virtually all existing models on optimal dynamic contracting assume perfect observability on agents' asset holdings. Relaxing this assumption is potentially very valuable. However, previous literature suggests that this exercise could be uninteresting. ACK show that in a pure endowment economy with adverse selection the efficient allocation coincides with the allocation of a pure bond economy (self insurance), that is the planner cannot improve on welfare. We find that in our model this is not the case. In a (hidden action) dynamic moral hazard framework, a second best allocation improves on self insurance, even in the case where the agent can secretly borrow and lend. We believe that this result is very important since it leaves the possibility of government interventions open in examples such as unemployment insurance or other welfare programs, where the presence of informational asymmetries between the government and the agents prevent the implementation of first best policies. Moreover, we find that welfare gains are substantial and that they do not vary monotonically with the interest rate.

The presence of welfare gains also suggests that the efficient allocation may possess

⁶⁴We also computed the discrepancy between $U_0^R(\tau)$ and $U^*(U_0)$ for the whole range of initial utilities the agent can obtain in the bond economy. We never found discrepancies larger than 0.02 percent for $\sigma = 1/2$ and 0.32 percent for $\sigma = 2$.

characteristics which are new with respect to existing models. In this paper, we discuss how the second best allocation can be empirically distinguished both from self insurance, from the pure moral hazard model with observable savings and from the first best allocation. Our testable implications involve data on individual consumption and income alone.

The study of the efficient allocation has been enormously simplified by the fact that we were able to solve the model within the recursive contracts framework, where together with the promised lifetime utility we used the agent's marginal utility of consumption as additional endogenous state variable. This reformulation of the problem has been made feasible by the use of a generalized first-order condition approach. In turn, the recursive formulation allowed us to implement a parsimonious numerical ex-post verification procedure of the first-order approach.

With an appropriate calibration exercise the model can be used to study welfare implications of optimal unemployment insurance schemes in presence of hidden asset holdings. Our results can be compared to the ones previously obtained in the literature, where the complication of the model forced the authors to consider suboptimal transfer schemes (e.g. Hansen and İmrohoroglu, 1992; and Abdulkadirođlu et al. 2002).

The model also proposes a new mapping from income (or wage) inequality into consumption inequality, which is suitable both for a more sophisticated positive analysis, and for intergenerational (or between group) welfare comparisons. In future developments, with a better calibrated model, we can study the properties of the joint distribution of wealth, consumption, labor (effort), and output in a general equilibrium model.

References

- [1] Abdulkadiroğlu, A., B. Kuruşçu, and A. Şahin, (2002) “Unemployment Insurance and the Role of Self-Insurance,” *Review of Economic Dynamics*, 5(3), July.
- [2] Abreu, D., D. Pearce and E. Stacchetti (1990), “Towards a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica*, 58, 5, September: 1041-1064.
- [3] Aiyagari, S. R. (1994), ”Uninsured Idiosyncratic Risk and Aggregate Savings,” *Quarterly Journal of Economics*. August: 659-684.
- [4] Allen, F. (1985), “Repeated Principal-Agent Relationship with Lending and Borrowing,” *Economic Letters*, 17: 27-31.
- [5] Atkeson, A. (1991), ”International Lending with Moral Hazard and Risk of Repudiation,” *Econometrica* 59(4), July: 1069-1089.
- [6] Attanasio, O. and S.J. Davis (1996), “Relative Wage Movements and the Distribution of Consumption,” *Journal of Political Economy*, 104:1227-1262.
- [7] Bizer, D. S. and P. M. DeMarzo (1999), “Optimal Incentive Contracts When Agents can Save, Borrow and Default” *Journal of Financial Intermediation*, 8: 241-269.
- [8] Chang, R. (1998) “Credible Monetary Policy in an Infinite Horizon Model: Recursive Approaches,” *Journal of Economic Theory*, 81: 431-461.
- [9] Chiappori, P.A., I. Macho, P. Rey and B. Salanie’ (1994), “Repeated Moral Hazard: The Role of Memory, Commitment, and Access to Credit Markets,” *European Economic Review*, 1527-1553.
- [10] Cochrane, J. (1991), ”A Simple Test of Consumption Insurance,” *Journal of Political Economy*, 99: 957-976.
- [11] Cole H.L. and N. R. Kocherlakota (2001), “Dynamic Games with Hidden Actions and Hidden States,” *Journal of Economic Theory*, 98, 114-126.
- [12] Cole, H.L. and N. R. Kocherlakota (2001), “Efficient Allocations with Hidden Income and Hidden Storage,” *Review of Economic Studies*, 68(3): 523-542.

- [13] Cronshaw, M.B. (1997), "Algorithms for Finding Repeated Game Equilibria," *Computational Economics*, 10: 139-168.
- [14] Doepke, M. and R. M. Townsend (2002), "Dynamic Mechanism Design with Hidden Income and Hidden Actions," *mimeo*, UCLA and University of Chicago.
- [15] Fernandes, A. and C. Phelan (2000) "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, 91: 223-247.
- [16] Fudenberg, D., O. Holmström and P. Milgrom (1990), "Short-Term Contracts and Long-Term Agency Relationships," *Journal of Economic Theory*, 51(1): 1-31.
- [17] Green, E.J. (1987), "Lending and Smoothing of Uninsurable Income," in *Contractual Arrangements for Intertemporal Trade*. (E.C. Prescott and N. Wallace, Eds.), Minnesota, Minnesota Press.
- [18] Grossman, S.J. and O.D. Hart (1983), "An Analysis of the Principal-Agent Problem," *Econometrica*, 51: 7-45.
- [19] Hansen, G.D. and A. İmrohoroglu (1992) "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard," *Journal of Political Economy*, 100, February: 118-142.
- [20] Hayashi, F., J. Altonji and L. Kotlikoff (1996), "Risk-sharing Between and within Families," *Econometrica*, 64: 261-294.
- [21] Hopenhayn, H. and J.P. Nicolini (1997), "Optimal Unemployment Insurance," *Journal of Political Economy*, 105(2): 412-438.
- [22] Huggett, M. (1993), "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17: 953-969.
- [23] Jewitt, I. (1988), "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, 56(5): 1177-1190.
- [24] Judd, K.L. (1998), *Numerical Methods in Economics*, MIT Press, Cambridge, MA.
- [25] Judd, K.L., S. Yeltekin and J. Conklin (2000), "Computing Supergame Equilibria," *mimeo*, Hoover Institution, Stanford CA.

- [26] Kydland, F. E. and E. C. Prescott (1980), “Dynamic Optimal Taxation, Rational Expectations and Optimal Control.” *Journal of Economic Dynamics and Control*, 2(1): 79-91
- [27] Kocherlakota, N. (2002), “Simplifying Optimal Unemployment Insurance: The Impact of Hidden Savings,” *mimeo*, Stanford University.
- [28] Krusell, P. and A.A. Smith (1998), “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, 106(5): 867-896.
- [29] Ligon, E. (1998), “Risk Sharing and Information in Village Economies,” *Review of Economic Studies*, 65: 847-864.
- [30] Mace, B.J. (1991), “Full Insurance in the Presence of Aggregate Uncertainty,” *Journal of Political Economy*, 99: 928-956.
- [31] Marcet, A. and R. Marimon (1999), “Recursive Contracts,” *mimeo* Universitat Pompeu Fabra.
- [32] Messner, M. and N. Pavoni (2002), “On Recursive Contracts,” *mimeo*, UPF and UCL.
- [33] Pavoni, N. (1999), “*Recursive Methods and Dynamic Contracts Under Complete Information: Theory and Applications*,” Ph.D. dissertation, in Mathematics for Economic Decisions, January, University of Trieste, Italy.
- [34] Pavoni, N. (2002a), “On Optimal Unemployment Compensation,” *mimeo*, University College London.
- [35] Pavoni, N. (2002b), “Optimal Unemployment Insurance, with Human Capital Depreciation, and Duration Dependence,” *mimeo*, University College London.
- [36] Phelan, C. and E. Stacchetti (2001), “Sequential Equilibria in a Ramsey Tax Model,” *Econometrica*, 69(6), November: 1491-1518.
- [37] Phelan, C. and R.M. Townsend (1991), “Computing Multi-Period, Information-Constrained Optima,” *Review of Economic Studies*, 58: 853-881.
- [38] Rogerson, W. (1985a), “Repeated Moral Hazard,” *Econometrica*, 53: 69-76.

- [39] Rogerson, W. (1985b), “The First-Order Approach to Principal-Agent Problems,” *Econometrica*, 53: 1357-1368.
- [40] Rockafellar, R.T. (1975), *Convex Analysis*, Princeton University Press, Princeton, N.J.
- [41] Rustichini, A. (1998), “Dynamic Programming Solution of Incentive-Constrained Problems,” *Journal of Economic Theory*, 78, February: 329-354.
- [42] Spear, S.E. and S. Srivastava (1987), “On Repeated Moral Hazard with Discounting,” *Review of Economic Studies*, 53: 599-617.
- [43] Stokey N.L., and R.E. Lucas, with E. C. Prescott (1989), *Recursive Methods in Economic Dynamics*, Cambridge (MA), Harvard University Press.
- [44] Thomas, J. and T. Worrall (1990), “Income Fluctuations and Asymmetric Information: An Example of Repeated Principal-Agent Problem,” *Journal of Economic Theory*, 51: 367-390.
- [45] Townsend, R.M. (1982), “Optimal Multiperiod Contracts and The Gains from Enduring Relationships Under Private Information,” *Journal of Political Economy*, 90(6): 1166-1186.
- [46] Townsend, R.M. (1994), “Risk and Insurance in Village India,” *Econometrica*, 62: 535-592.
- [47] Werning, I. (2002), “Optimal Unemployment Insurance with Unobservable Savings,” *mimeo*, University of Chicago and UTDT.
- [48] Zhao, R. (2001), “The Optimal Unemployment Insurance Contract: Why a Replacement Ratio?” *mimeo*, University of Illinois at Urbana-Campaign.

9 Appendices

9.1 Appendix A: Proofs

Proof of Proposition 1 Part (i) is immediate since first-best implies full insurance. However, in a moral hazard framework full insurance is associated with the cost minimizing effort each period (i.e. $e_t = 0$ for any t); this contradicts our assumptions. (ii) Notice first that the self insurance allocation is (by definition) incentive feasible. Moreover, a necessary condition of optimality for the pure moral hazard model is condition (13) which is incompatible with the Euler equation (6). Now a simple revealed preference argument on the planner's allocation choices implies that pure moral hazard improves on self insurance (iia). Notice that since the agent can secretly borrow and lend, condition (6) is required for feasibility in a second best allocation. As a consequence, the incompatibility between condition (13) and the Euler equation also shows that the optimal contract under monitorable assets (pure moral hazard) is not incentive feasible, hence the optimal second best contract (with non-monitorable assets) must be welfare reducing with respect to pure moral hazard (iib). (For this last point see also Bizer and DeMarzo, 1999.) **Q.E.D.**

Proof of Proposition 2 The bond economy allocation can be seen as a planner transfer scheme where period t transfers are h^{t-1} measurable. Consider now the last two periods of the program. To replicate self insurance, the continuation plan $\tau \setminus h^{T-2} = \{\tau_{T-2}(h^{T-2}), \tau_{T-1}(h^{T-2}, y_{T-1}^i)\}_{i=1}^N$ must be such that $\tau_{T-1}(h^{T-2}, y_{T-1}^i) = \tau_{T-1}^* \forall i$. We now show that a transfer scheme with τ_{T-1} that does not change with y_{T-1}^i cannot be optimal. Notice that this will imply our result since the planner cannot loose with respect to self insurance in previous periods. At this point, our line of proof is in part an adaptation of Proposition 5 in Grossman and Hart (1983). To simplify notation denote $\tau_{T-2}(h^{T-2}) = \tau_{T-2}^*$. Now assume that u is unbounded below and consider the following modification of the scheme:⁶⁵ leave τ_{T-2}^* unchanged, set $\hat{\tau}_{T-1}^i(\varepsilon) = \tau_{T-1}^* + \varepsilon$ for $i = 1, 2, \dots, N - 1$, and $\hat{\tau}_{T-1}^N(\varepsilon) = \tau_{T-1}^* - \mu^\varepsilon \varepsilon$; with $\varepsilon > 0$. Where, for any ε , we choose μ^ε

⁶⁵When u is unbounded above, one can modify the original self insurance scheme as follows: $\hat{\tau}_{T-2} = \tau_{T-2}^*$, $\hat{\tau}_{T-1}^i(\varepsilon) = \tau_{T-1}^* + \mu^\varepsilon \varepsilon$ for $i = 1, 2, \dots, N - 1$, and $\hat{\tau}_{T-1}^N(\varepsilon) = \tau_{T-1}^* - \varepsilon$.

so that the agent is indifferent, that is⁶⁶

$$\begin{aligned}
& \max_{e \in E, b \geq -B(\varepsilon)} u(y_{T-2} + \tau_{T-2}^* - b, e) + \beta \sum_i p_i(e) u(y_{T-1}^i + \hat{\tau}_{T-1}^i(\varepsilon) + (1+r)b, 0) \\
= & \max_{e \in E, b \geq -B(\varepsilon)} u(y_{T-2} + \tau_{T-2}^* - b, e) + \beta \sum_i^{N-1} p_i(e) u(y_{T-1}^i + \tau_{T-1}^* + \varepsilon + (1+r)b, 0) \\
& + \beta p_N(e) u(y_{T-1}^N + \tau_{T-1}^* - \mu^\varepsilon \varepsilon + (1+r)b, 0) \\
= & u(y_{T-2} + \tau_{T-2}^*, e_{T-2}^*) + \beta \sum_i p_i(e_{T-2}^*) u(y_{T-1}^i + \tau_{T-1}^*, 0) \quad (\text{recall that } b_{T-2}^* = 0).
\end{aligned}$$

Where for each ε , $B(\varepsilon) = \frac{1}{1+r} \min \{y_{T-1}^1 + \tau_{T-1}^* + \varepsilon, y_{T-1}^N - \varepsilon \mu^\varepsilon\}$. We want to show that the planner could gain with this scheme for ε small enough. Call $\hat{e}_{T-2}^\varepsilon$ the effort taken by the agent under the new scheme. It is easy to see that the difference in the returns for the planner between the two schemes is

$$\delta(\varepsilon) = p_N(\hat{e}_{T-2}^\varepsilon) \mu^\varepsilon \varepsilon - (1 - p_N(\hat{e}_{T-2}^\varepsilon)) \varepsilon,$$

we want to show that for ε small enough $\delta(\varepsilon) > 0$. To this extent we can take the limit as $\varepsilon \rightarrow 0$ of

$$\frac{\delta(\varepsilon)}{\varepsilon} = p_N(\hat{e}_{T-2}^\varepsilon) \mu^0 - (1 - p_N(\hat{e}_{T-2}^\varepsilon)).$$

and show that $\delta'(0) = p_N(\hat{e}_{T-2}^0) \mu^0 - (1 - p_N(\hat{e}_{T-2}^0)) > 0$, where \hat{e}_{T-2}^0 , \hat{b}_{T-2}^0 and μ^0 are the limit effort, asset and μ choices when ε is driven to zero. The Theorem of the Maximum implies that the optimal correspondence of choices $\hat{e}_{T-2}^\varepsilon$ and $\hat{b}_{T-2}^\varepsilon$ is upper hemicontinuous if considered as a function of ε . The consequence is that \hat{e}_{T-2}^0 and \hat{b}_{T-2}^0 are also optimal under the original scheme, hence for any $\varepsilon > 0$ we have

$$\begin{aligned}
& u(y_{T-2} + \tau_{T-2}^* - \hat{b}_{T-2}^0, \hat{e}_{T-2}^0) + \beta \sum_i p_i(\hat{e}_{T-2}^0) u(y_{T-1}^i + \hat{\tau}_{T-1}^i(\varepsilon) + (1+r)\hat{b}_{T-2}^0, 0) \\
\leq & u(y_{T-2} + \tau_{T-2}^* - \hat{b}_{T-2}^\varepsilon, \hat{e}_{T-2}^\varepsilon) + \beta \sum_i p_i(\hat{e}_{T-2}^\varepsilon) u(y_{T-1}^i + \hat{\tau}_{T-1}^i(\varepsilon) + (1+r)\hat{b}_{T-2}^\varepsilon, 0) \\
= & u(y_{T-2} + \tau_{T-2}^* - \hat{b}_{T-2}^0, \hat{e}_{T-2}^0) + \beta \sum_i p_i(\hat{e}_{T-2}^0) u(y_{T-1}^i + \tau_{T-1}^* + (1+r)\hat{b}_{T-2}^0, 0),
\end{aligned}$$

⁶⁶Such a μ^ε exists by the Maximum Theorem. Notice indeed that both u and \mathbf{p} are continuous and E is compact. Finally, notice that we can, without loss of generality, require $b \in B(\varepsilon, \mu)$, where $B(\varepsilon, \mu) \equiv [-\frac{1}{1+r} \min \{y_{T-1}^1 + \tau_{T-1}^* + \varepsilon, y_{T-1}^N + \tau_{T-1}^* - \varepsilon \mu\}, y_{T-2} + \tau_{T-2}^*]$ is a non-empty, compact and continuous correspondence (notice that $\varepsilon \mu \leq y_{T-1}^N + \tau_{T-1}^*$).

or

$$\sum_i p_i(\hat{e}_{T-2}^0) \left[u \left(y_{T-1}^i + \hat{\tau}_{T-1}^i(\varepsilon) + (1+r)\hat{b}_{T-2}^0, 0 \right) - u \left(y_{T-1}^i + \tau_{T-1}^* + (1+r)\hat{b}_{T-2}^0, 0 \right) \right] \leq 0. \quad (45)$$

Since u is concave (45) implies that

$$0 > \sum_i^{N-1} p_i(\hat{e}_{T-2}^0) \varepsilon u'_c \left(y_{T-1}^i + \tau_{T-1}^* + (1+r)\hat{b}_{T-2}^0, 0 \right) - p_N(\hat{e}_{T-2}^0) \mu^\varepsilon \varepsilon u'_c \left(y_{T-1}^N + \tau_{T-1}^* + (1+r)\hat{b}_{T-2}^0 - \mu^\varepsilon \varepsilon, 0 \right).$$

Dividing by ε and taking the limit $\varepsilon \rightarrow 0$ we get

$$0 \geq \sum_i^{N-1} p_i(\hat{e}_{T-2}^0) u'_c \left(y_{T-1}^i + \tau_{T-1}^* + (1+r)\hat{b}_{T-2}^0, 0 \right) - p_N(\hat{e}_{T-2}^0) \mu^0 u'_c \left(y_{T-1}^N + \tau_{T-1}^* + (1+r)\hat{b}_{T-2}^0, 0 \right). \quad (46)$$

It is now easy to realize that by strict concavity (46) implies $\delta'(0) > 0$. **Q.E.D.**

Proof of Proposition 3 Proposition 3 and 4 are proved by assuming non-negativity of the multipliers. Our strategy partially recalls the one used by Rogerson (1985b). We first consider a (double) relaxed problem where proper inequality versions of the first order conditions are considered. We then show that if the first order conditions approach is justified, the optimal scheme in fact solves the constraints with equality. So consider the problem

$$\max_{\substack{e_0 \in E, \\ c_0, c_1^i \geq 0}} y_0 - c_0 + \frac{1}{1+r} \sum_i p_i(e_0) [y_1^i - c_1^i] \quad (47)$$

subject to

$$\mu : v'(e_0) \leq \beta \sum_i p_i'(e_0) u(c_1^i), \quad (48)$$

$$\xi : u'(c_0) \geq \beta(1+r) \sum_i p_i(e_0) u'(c_1^i); \quad (49)$$

and the participation condition (we normalize $v(0) = 0$)

$$\lambda : u(c_0) - v(e_0) + \beta \sum_i p_i(e_0) u(c_1^i) \geq U_0. \quad (50)$$

Lemma 1 *Assume that the problem can be solved using the first order approach. If the contract is such that $c_0 > 0$ and $c_1^i > 0 \forall i$, the original problem (16)-(19) can be replaced by the double relaxed problem (47)-(50).*

The proof will be given below. Let us now first calculate the first order conditions:

$$\frac{1}{u'(c_0)} = \lambda + \xi \frac{u''(c_0)}{u'(c_0)} \quad (51)$$

$$\frac{1}{u'(c_1^i)} \frac{1}{(1+r)\beta} = \lambda + \mu \frac{p'_i(e)}{p_i(e)} - (1+r)\xi \frac{u''(c_1^i)}{u'(c_1^i)} \quad (52)$$

The proof of **Proposition 3** is now immediate from (52) since $\mu, \xi \geq 0$. Recalling that concavity and prudence implies that both $\frac{1}{u'}$ and $\frac{u''}{u'}$ increase with c_1^i . **Q.E.D.**

Proof of Proposition 4 Taking the difference between (52) in two adjacent states, we get

$$\left[\frac{1}{u'(c_1^{i+1})} - \frac{1}{u'(c_1^i)} - \xi(1+r)(a(c_1^{i+1}) - a(c_1^i)) \right] = \mu \left[\frac{p'_{i+1}(e)}{p_{i+1}(e)} - \frac{p'_i(e)}{p_i(e)} \right].$$

By assumption the right hand side increases with i . Since the functions in the left hand side are increasing and concave both in c_1^i and c_1^{i+1} and - by Proposition 1 - both c_1^i and c_1^{i+1} increase with i , we must have that $c_1^{i+1} - c_1^i > 0$ increases with i . **Q.E.D.**

Proof of Lemma 1

Lemma 2 *In an optimal contract it must be that (49) is satisfied with equality.*

Proof. If $\xi > 0$ we are done. If $\xi = 0$, an optimal contract necessarily generates equation (13) which from Jensen inequality is either incompatible with (49) (as long as there is some consumption dispersion) so $\xi > 0$, or it is compatible because the agent is fully insured and (49) is satisfied with equality. **Q.E.D.**

Lemma 3 *In an optimal contract it must be that $v'(e_0) = \beta \sum_i p'_i(e_0) u(c_1^i)$.*

Proof. If $\mu > 0$ we are done. So consider the case where $\mu = 0$. In this case, the planner fully insures the agent. But then (49) must be satisfied with equality since $v'(e_0) \geq 0 \forall e_0 \in E$. **Q.E.D.**

Proof of Proposition 5 Notice that to each contract \mathcal{W} , we can associate the states as follows

$$u'_c(\mathbf{c}_t(h^t), \mathbf{e}_t(h^t)) = \mathbf{x}_t(h^t) \quad (53)$$

$$\mathbf{U}_t(\mathcal{W}; (h^t)) = \mathbf{U}_t(h^t) \quad (54)$$

where $h^0 = y_0$ is given. From the definition of $\mathbf{U}_t(\mathcal{W}; (h^t))$, (53) and (54) satisfy (24) and (27) by construction. Moreover, if the resulting plan of states is generated by an incentive feasible contract $\mathcal{W} \in \Omega_{FOC}$, then the sequence of states must also satisfy (25), (26) and the domain restriction. According to the Bellman principle, V_{foc}^* can be decomposed as follows

$$V_{foc}^*(y_0, U_0) = \sup_{x, U \geq U_0, (x, U) \in M^*} \left\{ \begin{array}{l} \sup_{\mathcal{W} \in \Omega_{FOC}} \Pi_0(\mathcal{W}; h^0) \\ \text{s.t. } \mathbf{U}_0(\mathcal{W}; h^0) = U \\ u'_c(\mathbf{c}_0(h^0), \mathbf{e}_0(h^0)) = x. \end{array} \right\}$$

Now, if for any $(x, U) \in M^*$ and $y \in Y$ we define

$$\begin{aligned} W_{foc}^*(y, x, U) &\equiv \sup_{\mathcal{W} \in \Omega_{FOC}} \Pi_0(\mathcal{W}; h^0) \\ \text{s.t.} \quad &: \mathbf{U}_0(\mathcal{W}; h^0) = U; u'_c(\mathbf{c}_0(h^0), \mathbf{e}_0(h^0)) = x; h^0 = y \end{aligned}$$

we can apply Theorem 4.2 of Stokey et al. (1989) (SLP) to show that W_{foc}^* solves the functional equation (23)-(27). To show the converse, first notice that the contract generated recursively must be (relaxed) incentive compatible (in particular, notice that the relaxed incentive compatibility consider only one-period deviations). In addition, since M^* is bounded the optimal policies always generate bounded values of lifetime utilities, hence (24) always generates well defined agent's lifetime utilities. As a consequence, when V is bounded we can use Theorem 4.3 of SLP to show our result. In this case, we can also apply Theorem 4.6 of SLP and prove that V is continuous. **Q.E.D.**

9.1.1 Other Propositions (Related to Section 3)

Lemma 4 *When b_t is contractible and monitorable, the optimal contract implies a consumption/effort plan, which is identical to the one implemented by the standard dynamic moral hazard model, where the agent has no access to the credit market. Moreover, there is a continuum of optimal contracts with the same consumption and effort allocation but with different transfer and asset holding plans.*

Proof. We will first show that when b_t is contractible, the incentive constraint (2) involving the budget constraint (1) can be replaced by (11) and (12).

It is of course key to notice that since b_t is contractible and observable the agent will never deviate from the planner's prescription $\mathbf{b} = \{\mathbf{b}_t(h^t)\}_{t=0}^{T-1}$ in σ . As a consequence, the planner can always replicate (11) from (1) by setting $\mathbf{b}_t(h^t) \equiv 0$ for $\forall h^t$. Moreover, note that from (11) the planner has perfect control on the agent's consumption. Therefore the agent can only deviate from the planner's prescriptions in the effort decisions. Consequently, (2) can be replaced by (12).

To finish the proof we just need to show that the planner cannot gain by choosing a plan for bonds different from $\mathbf{b}_t(h^t) \equiv 0$ for $\forall h^t$. First of all notice that the planner is indifferent among all the possible bond/transfers plans, as long as they imply the same consumption and effort plans. Since we have already pointed out that the observability of \mathbf{b} implies that through τ the planner has complete control on \mathbf{c} , it remains to show that the planner cannot lose any control on effort decisions by choosing the zero bounds plan. To this extent, it suffice to realize that the agent cares only about consumption, in particular he does not care about the particular timing of payments in τ and/or \mathbf{b} as long as \mathbf{c} is the same. Hence the control on \mathbf{c} is enough to implement any effort plan. **Q.E.D.**

Proposition 6 (Rogerson 1985a) *Assume $u(c, e) = u(c) - v(e)$ and that the optimal contract is interior in consumption (i.e. $\mathbf{c}_t(h^t) > 0 \forall t, h^t$). The optimal allocation with moral hazard and observable assets satisfies*

$$\frac{1}{1+r} \sum_i p_i(e_t) \frac{1}{u'(c_{t+1}^i)} = \beta \frac{1}{u'(c_t)} \quad \forall t, h^t.$$

Proof. Consider the following feasible deviation from the optimal contract on period t and $t+1$ payments (hence on consumption) after history h^t . Consider the possibility of choosing deviation $d_t(\delta)$ from the optimal c_t , so that

$$u(c_t + d_t(\delta)) = u(c_t) - \beta\delta,$$

or $d_t(\delta) = u^{-1}(u(c_t) - \beta\delta) - c_t$. Moreover, take the vector of deviations $\{d_{t+1}^i(\delta)\}_{i=1}^N$ such that for any i

$$u(c_{t+1}^i + d_{t+1}^i(\delta)) = u(c_{t+1}^i) + \delta,$$

which gives $d_{t+1}^i(\delta) = u^{-1}(u(c_{t+1}^i) + \delta) - c_{t+1}^i$. One can easily check that this deviation from the optimal plan will satisfy incentive compatibility with exactly the same effort plans as before. In addition, this deviation does not violate the initial participation constraint either. Optimality requires that planner cannot gain by such deviations. Therefore, $\delta = 0$ should be the planner's optimal solution to the following maximization problem:⁶⁷

$$\max_{\delta} -(\tau_t + d_t(\delta)) - \frac{1}{1+r} \sum_i p_i(e_t) (\tau_{t+1}^i + d_{t+1}^i(\delta)).$$

The first-order condition evaluated at $\delta = 0$ gives the desired result. **Q.E.D.**

9.2 Appendix B: Adverse Selection versus Moral Hazard

To rephrase the adverse selection framework of ACK (hidden information) into a hidden action moral hazard notation, one should first consider models with a timing of events slightly different than the one we use in this paper, and more consistent with the repeated games framework. It is straightforward to see that all the results we presented in the main text apply to this version of the model as well. Consider the following timing of the model: in each period t effort e_t is taken, which affects the output realization y_t according to $p_i(e_t)$. Payments and consumption are decided *at the end* of each period, conditional on the realized y_t . For example, in the two period version of the model, each agent's preferences become

$$\sum_{i=1}^N p_i(e_0) \left[u(\tau_0(y_0^i) + y_0^i - \mathbf{b}_0(y_0^i), e_0) + \beta \sum_{j=1}^N p_j(\mathbf{e}_1(y_0^i)) u(\tau_1(y_0^i, y_1^j) + y_1^j + (1+r)\mathbf{b}_0(y_0^i), \mathbf{e}_1(y_0^i)) \right].$$

Now consider a T period model of adverse selection. Models of this type are T repetitions of games where in each period, each type receives private information determining his type $\theta_t \in \Theta$ with Θ finite. The distribution $\pi(\theta)$ over Θ is common knowledge. After learning his type, the agent chooses his move $y_t^i \in Y$ which is publicly observable. An action \mathbf{e}_t is a complete mapping from Θ into Y , and the agent's *realized* payoff $g(\theta_t, \tau_t(h^{t-1}, y_t^i))$ depends both on the signal and on the public outcome, since usually the transfer scheme $\tau_t : Y^t \rightarrow \mathfrak{R}$ depends on the latter.

The return function used by ACK to study the trade-off between consumption insurance and incentives is $g(\theta_t + \tau_t) = g(\theta_t + y_t + \tau_t - \mathbf{e}_t(\theta_t))$, with g being a concave function. The

⁶⁷Payoffs at dates different from t and $t+1$ are not affected by δ , as a consequence they can be ignored.

agent's *ex-ante* payoff function over effort, assets, planner transfers and public outcomes becomes

$$u(y_t + \tau_t + (1+r)b_{t-1} - b_t, \mathbf{e}_t) \equiv \frac{\sum_{\theta \in \mathbf{e}_t^{-1}(y_t)} \pi(\theta) g(\theta + y_t + \tau_t + (1+r)b_{t-1} - b_t - \mathbf{e}_t(\theta))}{\sum_{\theta \in \mathbf{e}_t^{-1}(y_t)} \pi(\theta)}.$$

Formally $u(y_t + \tau_t + (1+r)b_{t-1} - b_t, \mathbf{e}_t)$ is not defined for y_t if $\mathbf{e}_t^{-1}(y_t)$ is empty. However, since these public outcomes occur with zero probability, we can define it arbitrarily for these events also without affecting the analysis.

The conditional distribution of public outcomes given effort profile \mathbf{e} , $p_i(\mathbf{e}) = \Pr\{y = y^i \mid \mathbf{e}\}$ is defined as follows

$$\begin{aligned} p_i(\mathbf{e}) &\equiv \sum_{\theta \in \mathbf{e}^{-1}(y^i)} \pi(\theta) \quad \text{if } \mathbf{e}^{-1}(y^i) \text{ is non-empty.} \\ p_i(\mathbf{e}) &\equiv 0 \quad \text{otherwise.} \end{aligned}$$

Allen (1985) studies the continuous output version of this model. He specifies a production function of the form $y = \theta + e$, with an exogenous density $\pi(\theta)$ defined on the support Θ ; and a utility function where consumption and effort are perfect substitutes: $u(c, e) = u(c - e)$. He then assumes that the effort choice e is taken *after* having observed the exogenous shock θ , which is of course equivalent to assuming that the agent can take *ex-ante* the contingent map $\mathbf{e} : \Theta \rightarrow E$ as before.

9.3 Appendix C: Computation Strategy

The Domain Restriction Let us rewrite the definitions of the feasible set for our separable case with c as a state variable. The domain restriction set M^* is the set of (U, c) couples such that there exists $(U^l, c^l) \in M^*$, $(U^h, c^h) \in M^*$ and $e \in E$ such that (55)-(57) are all satisfied:

$$U = u(c) - v(e) + \beta(U^l + p(e)(U^h - U^l)) \quad (55)$$

$$v'(e) = \beta p'(e)(U^h - U^l) \quad (56)$$

$$u'(c) = \beta(1+r)(u'(c^l) + p(e)(u'(c^h) - u'(c^l))). \quad (57)$$

We defined a two dimensional grid $M_0 = \{U_k\}_{k=1}^n \times \{c_j\}_{j=1}^m$ where c_j ranged between \underline{c} and \bar{c} and U_k ranged between $(u(\underline{c}) - v(\bar{e})) / (1 - \beta)$ and \bar{U} , where $\bar{e} \equiv \max(E)$ and

$\bar{U} = (u(\bar{c}) - v(0)) / (1 - \beta)$. Then, we applied the set valued operator \mathbf{B} defined below on M_0 iteratively until we obtained a fixed point:

$$\begin{aligned} M_t &= \mathbf{B}(M_{t-1}) \equiv & (58) \\ &\equiv \{ (U, c) \in M_{t-1}; \exists (U^i, c^i) \in M_{t-1}, e \in E; (55)-(57) \text{ are satisfied} \}. \end{aligned}$$

Note, that conditions (55)-(57) impose 3 constraints on our 5 endogenous variables $[U^l, c^l, U^h, c^h, e]$. Since M_{t-1} is a discrete set, for any $(U, c) \in M_{t-1}$ it is not always possible to have $(U^l, c^l) \in M_{t-1}$ and $(U^h, c^h) \in M_{t-1}$ satisfying (55)-(57) exactly for any $e \in E$. Note, that this is solely due to the fact that we are working with a discrete initial set, this has nothing to do with the actual feasibility of these bundles. For this reason, we had to modify the algorithm defined by (58). We used the following modified operator:

$$\tilde{\mathbf{B}}(M_{t-1}) \equiv \left\{ \begin{array}{l} (U, c) \in M_{t-1}; \exists (U^i, c^i) \quad i = l, h \text{ and } e \in E \\ \text{such that (55)-(57) are satisfied and} \\ \forall i \exists (U_k^i, c_j^i) \in M_{t-1} \text{ s.t. } \|(U^i, c^i) - (U_k^i, c_j^i)\| < \varepsilon_{j,k} \end{array} \right\} \quad (59)$$

$\tilde{\mathbf{B}}(M_{t-1})$ excludes (U, c) bundles for which there is no such solution of (55)-(57) that has a close enough point in the discrete set M_{t-1} . Generally, we chose $\varepsilon_{j,k}$ being one half of the distance between grid points around (U_k^i, c_j^i) . Practically, this way we guaranteed that we are not excluding bundles just because of discretization.

We had to modify $\tilde{\mathbf{B}}(\cdot)$ for some practical reasons around the boundaries. First, note that $\underline{c} = 0$ would have raised numerical problems, since in this case the marginal utility of consumption (included in the Euler equation) is equal to minus infinity. For this reason we chose the lowest consumption in the grid being $\underline{c} = 0.0001y^l$. However, we chose $\varepsilon_{j,1}$ in a way that we did not exclude points where the only feasible continuation included $c^l \in (0, \underline{c}]$.

The fixed point of this modified operator is given by $\tilde{M}^* = \tilde{\mathbf{B}}(\tilde{M}^*)$. \tilde{M}^* is the discrete approximation of the feasibility set we used throughout the paper. The feasibility sets presented in the main text are the results of this procedure.

However, we wanted to study a bit more what are the consequences of imposing the upper bound on consumption. As we discussed in the main text the upper bound on c and U was chosen because of computational reasons. Therefore we modified the procedure given by $\tilde{\mathbf{B}}(\cdot)$ in a way that feasible (high output) continuations could exceed the upper bounds. On the other hand, we had to impose some consistency requirement for these bundles.

We expected (and it was confirmed by our results) that the higher is promised utility the lower is the implied effort. For this reason we imposed that a continuation (U^h, c^h) where $U^h > \bar{U}$ or $c^h > \bar{c}$ can be feasible only if $U^h = u(c^h)/(1 - \beta)$ and the other conditions of (59) are satisfied.⁶⁸ These points are always feasible because they represent a full insurance plan from that period on with consumption equal to c_h and effort equal to zero. It is also a good approximated behavior around the upper bound of U and c , because around (\bar{U}, \bar{c}) the optimal effort is actually very close to zero. We show the obtained feasibility sets of this modified procedure on Figure A1.

Figure A1 here

Value Function Iteration We could have used directly the discrete feasibility set \widetilde{M}^* for solving for the value function by discrete value function iterations. This approach was not really feasible because generically we cannot satisfy (55)-(57) with (U^h, c^h) and (U^l, c^l) bundles on the grid. Therefore we chose a different strategy. We calculated the feasible set with a fine grid as explained in the previous subsection. Then we implied value function iteration with a projection method: we approximated the value function with Chebyshev polynomials. Traditionally, Chebyshev polynomials were applied to approximate functions of one variable. Judd (1998) shows how to extend this methodology for functions of two variables. Since $V(y^h, U, c) = V(y^l, U, c) + (y^h - y^l)$ we had to approximate a function of two variables only. Typically, this methodology is applied on the set $\mathcal{M} = [\underline{U}, \bar{U}] \times [\underline{c}, \bar{c}]$, however the methodology can be modified straightforwardly for our feasibility set which does not have a squared shape.

However, we still had to face a technical difficulty. The value function becomes increasingly steep around the boundaries of the feasibility set. Even with a high-order polynomial it is quite difficult to capture this feature. In order to have this feature we approximated different functional forms of the value function over different regions of the feasibility set. In order to achieve consistency between the different regions, the projections took into account the neighboring functional form at the borderline. Figure A2 shows how the actual feasibility set was partitioned (in case of $\beta(1 + r) = 1$).

⁶⁸Note, that by construction $(\bar{U}, \bar{c}) \in M^*$ since $U^l = U^h = \bar{U}$, $c^l = c^h = \bar{c}$ and $e = 0$ satisfies all the conditions. This represents a full insurance plan with consumption equal to \bar{c} . The Euler equation, $c \leq \bar{c}$ and $c^h > \bar{c}$ will imply that c^l has to be smaller than \bar{c} , therefore (U^l, c^l) has to be in the interior of M_t .

Technical and detailed description of this procedure is available from the authors by request.

Figure A2 here

9.4 Appendix D: Computing asset holding in planner problems

In this section we describe how we computed the agents asset level in the different planner problems. There is an obvious way in which this procedure also describes a first step toward a full decentralization of our allocation.

We have already pointed out that there is a continuum of combinations of asset holdings and transfer payments implementing the same (constrained) efficient allocation of consumption and effort. However, notice that the way we defined our allocation the planner's value function $V(y_t, U_t, x_t)$ evaluated in period t (at the state implied by history h^t) is the expected discounted profits $\Pi_t(\mathcal{W}^*; h^t)$ corresponding to the optimal contract \mathcal{W}^* such that at any node h^t , $\mathbf{b}_t^*(h^t) = 0$, and $\boldsymbol{\tau}_t^*(h^t) = \mathbf{c}_t^*(h^t) - y_t$. That is

$$-V(y_t, U_t, x_t) = -\Pi_t(\mathcal{W}^*; h^t) = \mathbf{E}_t \sum_{n=0}^{\infty} \left[\left(\frac{1}{1+r} \right)^t \tau_{t+n}^* \right]$$

is the (aggregate) net present value of transfers agents will receive from period t onwards. According to this interpretation, we can determine a new asset plan $\{\hat{\mathbf{b}}_t(h^t)\}_{t=0}^{\infty}$ as follows

$$\hat{\mathbf{b}}_t(h^t) \equiv -\frac{1}{1+r} \mathbf{E}_t \Pi_{t+1}(\mathcal{W}^*; h^{t+1}) = \mathbf{E}_t \sum_{n=1}^{\infty} \left[\left(\frac{1}{1+r} \right)^t (\mathbf{c}_{t+n}^*(h^{t+n}) - y_{t+n}) \right]. \quad (60)$$

Considered in (cross section) average terms, the number \hat{b}_t represents the net position of the economy with respect to the rest of the world, the economywide level of savings in period t . Note, that this definition is consistent with our initial values. That is $0 = \hat{b}_{-1} \equiv -\frac{1}{1+r} \mathbf{E}_{-1} \Pi_0(\mathcal{W}^*; y_0)$, where we interpret $\mathbf{E}_{-1} \Pi_0(\mathcal{W}^*; y_0)$ as the (cross-sectional) average profit at time 0 (what we fixed to zero by picking U_0 properly).

This definition has a clear intuitive meaning from the agent's budget constraint as well. The definition of profits and the agent's budget constraint together with the above definition of implied asset holdings will result in the following transfer scheme

$$\begin{aligned} \hat{\boldsymbol{\tau}}_t(h^t) &= \boldsymbol{\tau}_t^*(h^t) + \mathbf{E}_{t-1} \Pi_t(\mathcal{W}^*; h^t) - \frac{1}{1+r} \mathbf{E}_t \Pi_{t+1}(\mathcal{W}^*; h^{t+1}) \\ &= \mathbf{c}_t^*(h^t) - y_t + (1+r) \hat{\mathbf{b}}_{t-1}(h^{t-1}) - \hat{\mathbf{b}}_t(h^t). \end{aligned}$$

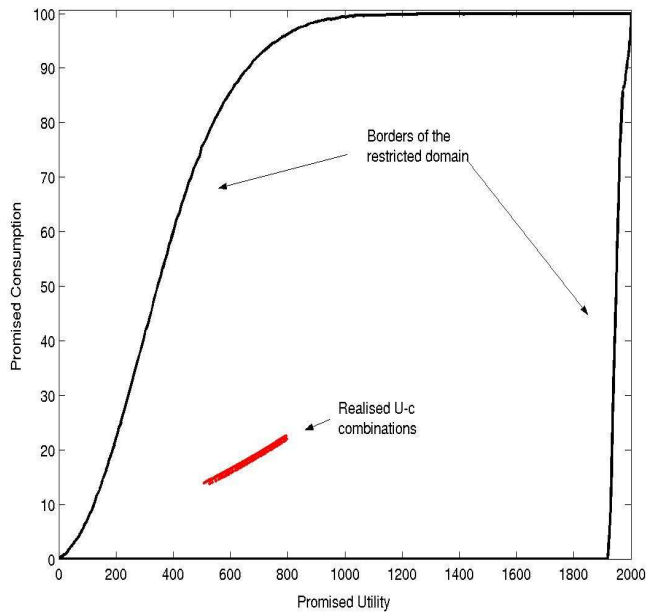


Figure 1A: The domain restriction and the realized U-c combinations

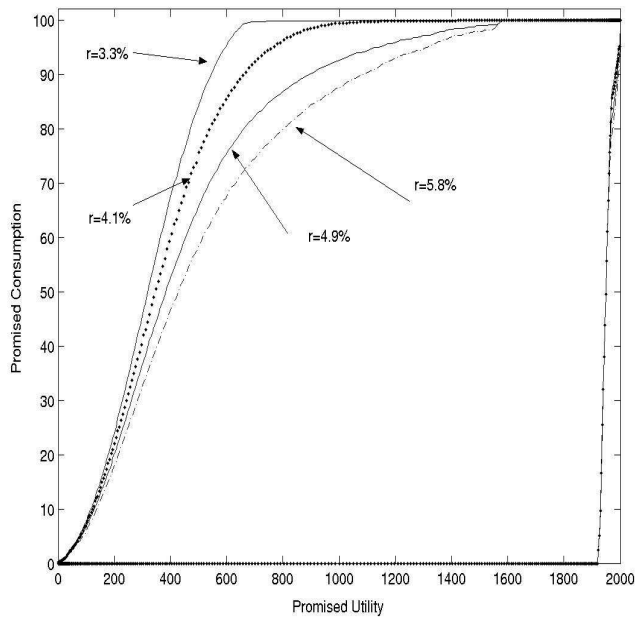


Figure 1B: The domain restriction for different interest rates

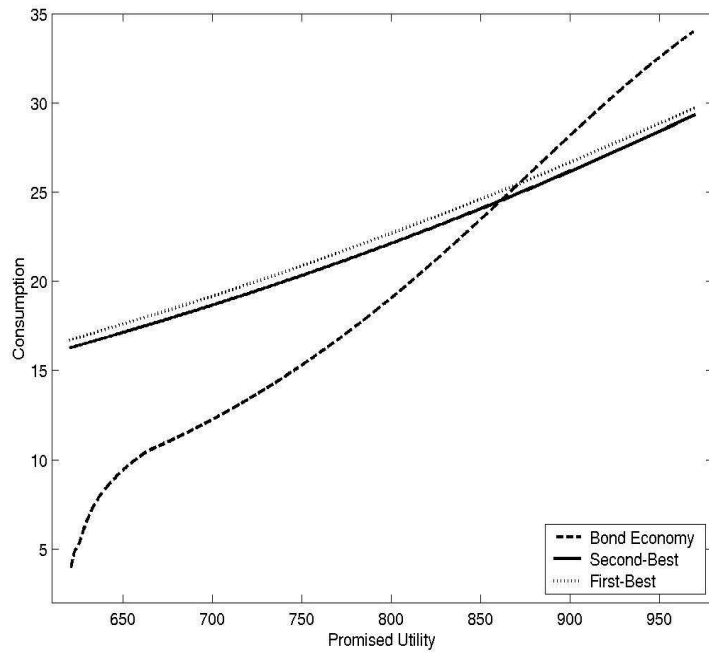


Figure 2A: Current period consumption as a function of lifetime utility

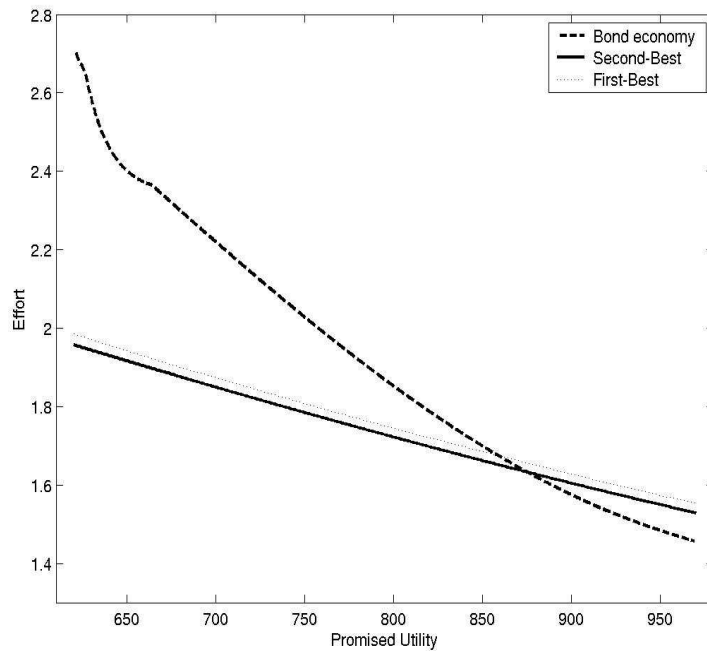


Figure 2B: Effort level as a function of lifetime utility

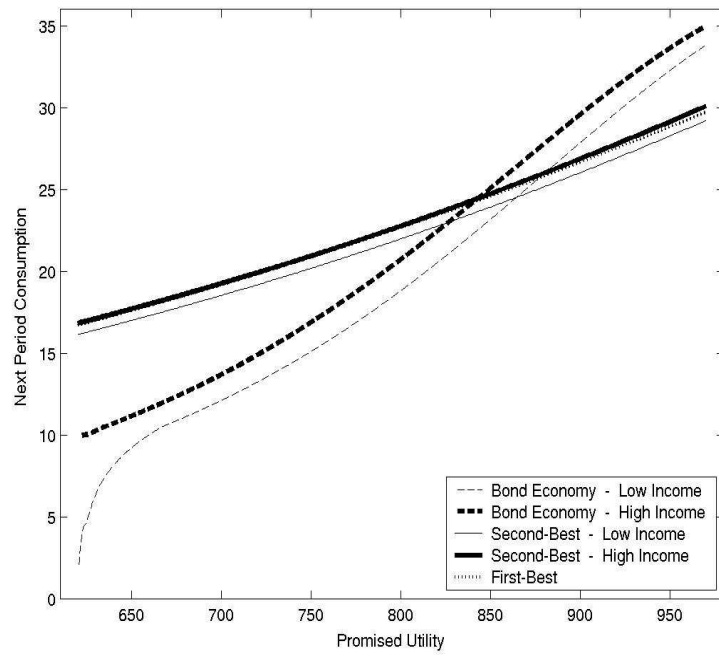


Figure 3A: Next period consumption as a function of lifetime utility

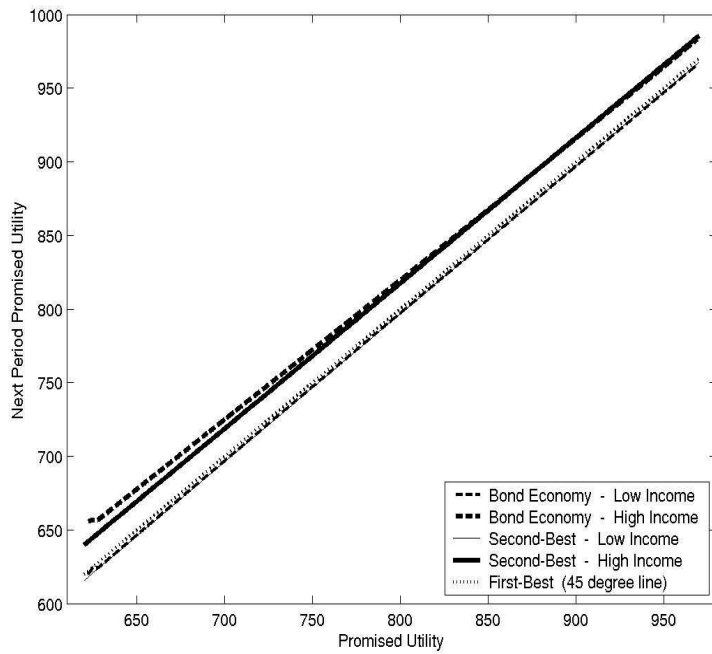


Figure 3B: Next period utility as a function of lifetime utility

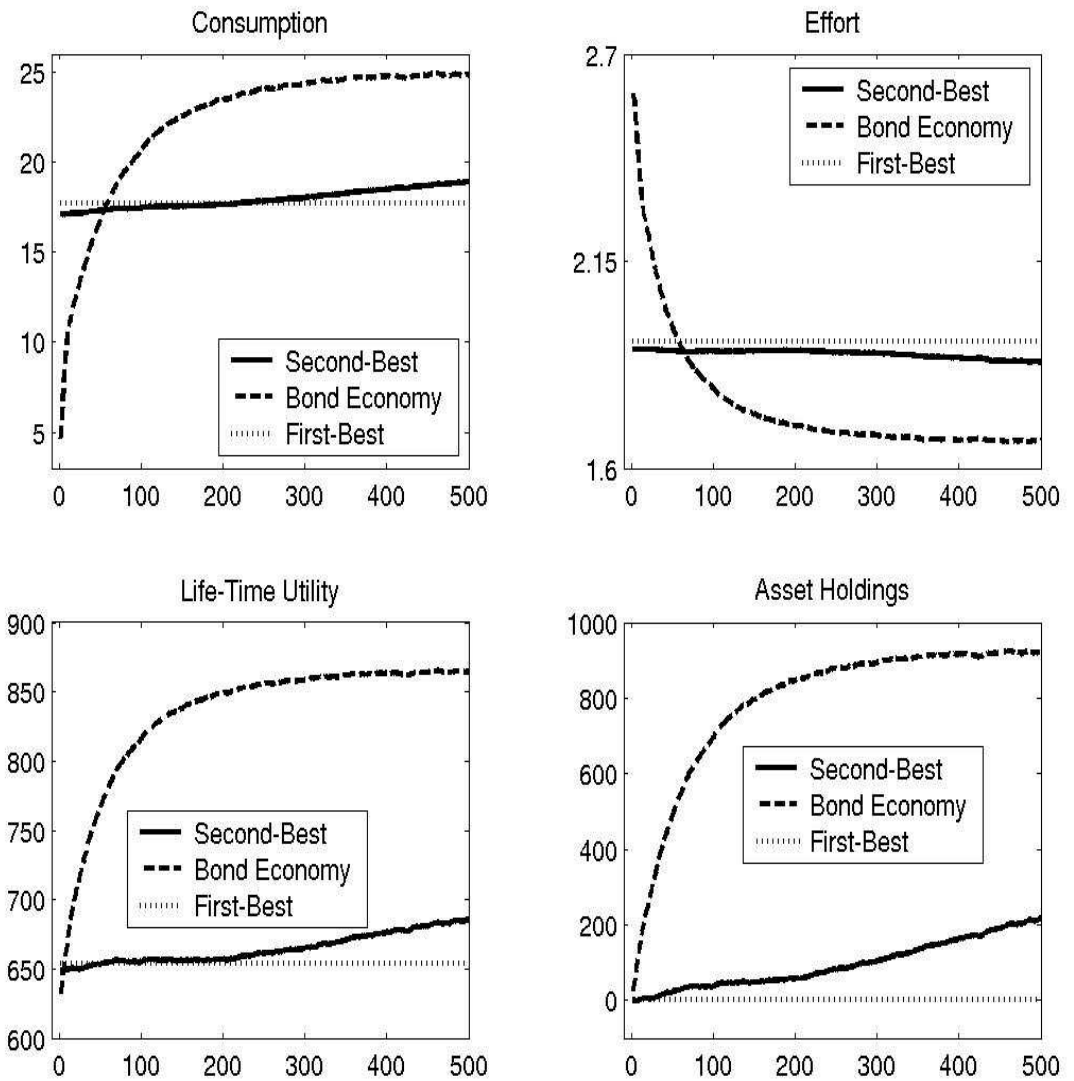


Figure 4: Average path for the first 500 periods ($b_{-1}=0$ for the bond economy and U_0 such that the planner has zero expected surplus with the same given initial distribution of output)

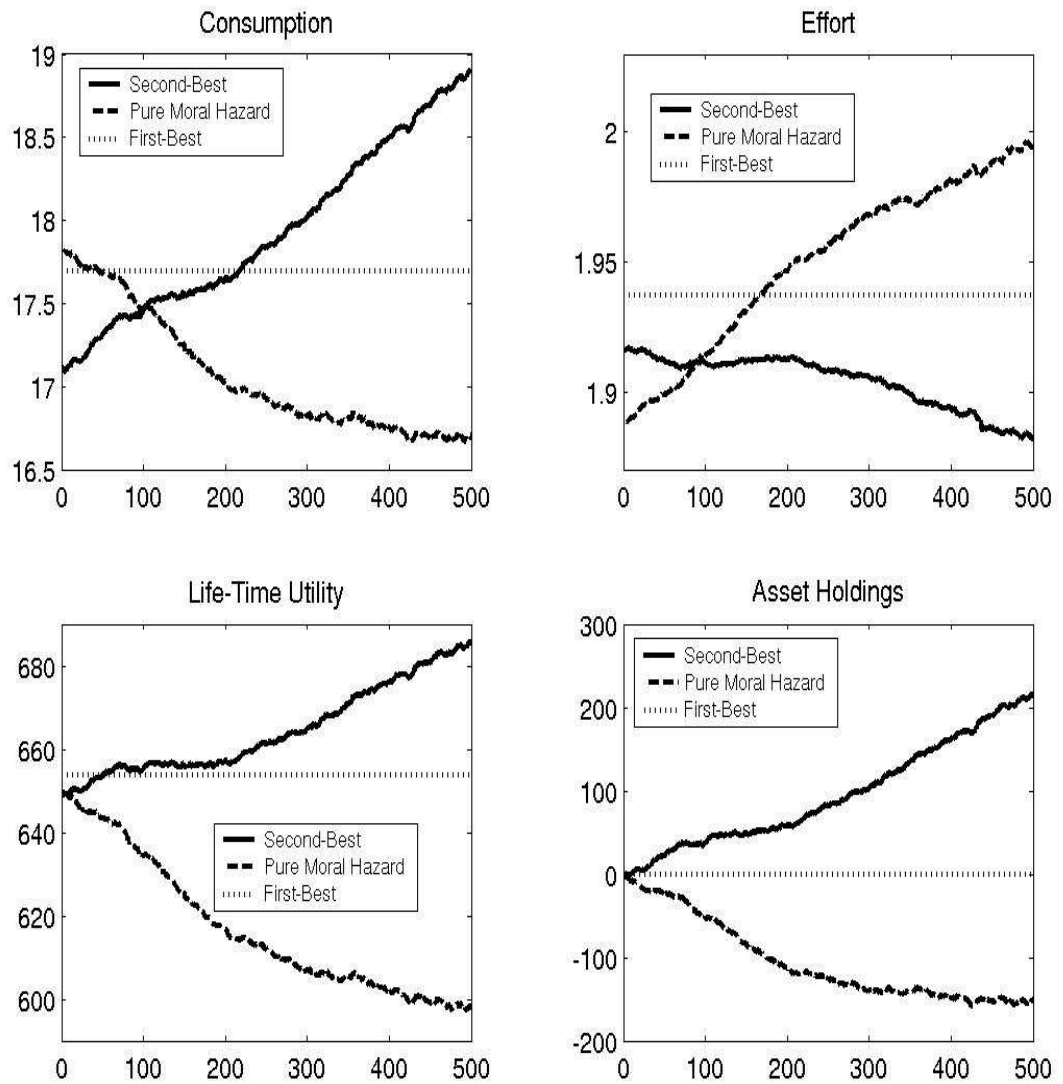


Figure 5: Average path for the first 500 periods (U_0 such that the planner has zero expected surplus with the same given initial distribution of output)

Table 2: Efficiency gains

| | Bond Economy | First-Best | Pure Moral Hazard | Second-best |
|---|-----------------|------------|----------------------|-------------|
| <i>Low risk aversion ($\sigma=0.5$)</i> | | | | |
| <i>$\beta(1+r)=1$ ($r=4.1\%$)</i> | | | | |
| Agent's life-time utility | 622.6 | 649.4 | 646.6 | 645.4 |
| Agent's life-time utility (%) | 100 | 104.3 | 103.9 | 103.7 |
| Initial asset holdings | 0 | 74.4 | 65.9 | 62.3 |
| Planner's surplus | 0 | 111.2 | 100.0 | 93.6 |
| Consumption increase (%) | 0 | 6.6 | 6.0 | 5.7 |
| <i>$\beta(1+r)<1$ ($r=3.3\%$)</i> | | | | |
| Agent's life-time utility (%) | 100 | 110.5 | 109.2 | 107.8 |
| Initial asset holdings | 0 | 203.5 | 178.3 | 151.2 |
| Planner's surplus | 0 | 306.1 | 282.1 | 261.1 |
| Consumption increase (%) | 0 | 16.1 | 14.1 | 11.8 |
| <i>$\beta(1+r)>1$ ($r=4.9\%$)</i> | | | | |
| Agent's life-time utility (%) | 100 | 107.3 | 106.9 | 105.3 |
| Initial asset holdings | 0 | 136.6 | 128.6 | 102.1 |
| Planner's surplus | 0 | 154.9 | 149.4 | 120.4 |
| Consumption increase (%) | 0 | 12.2 | 11.5 | 8.8 |
| <i>Welfare gains with higher initial utility $\beta(1+r)=1$ ($r=4.1\%$)</i> | | | | |
| Agent's life-time utility (%) | 100 | 102.1 | 102.0 | 102.0 |
| Initial asset holdings | 369.1 | 444.1 | 441.2 | 438.7 |
| Planner's surplus | -354.6 | -304.5 | -316.1 | -319.5 |
| Consumption increase (%) | 0 | 3.7 | 3.6 | 3.5 |
| <i>High risk aversion ($\sigma=2$)</i> | | | | |
| <i>$\beta(1+r)=1$ ($r=4.1\%$)</i> | | | | |
| Agent's life-time utility (%) | 100 | 116.1 | 110.1 | 106.4 |
| Initial asset holdings | 0 | 138.8 | 42.6 | 13.4 |
| Planner's surplus | 0 | 443.0 | 324.3 | 183.2 |
| Consumption increase (%) | 0 | 33.2 | 18.5 | 10.9 |
| <i>Welfare gains with higher initial utility $\beta(1+r)=1$ ($r=4.1\%$)</i> | | | | |
| Agent's life-time utility (%) | 100 | 103.1 | 102.4 | 101.7 |
| Initial asset holdings | 163.8 | 226.3 | 211.0 | 197.2 |
| Planner's surplus | -162.1 | -80.9 | -111.8 | -137.2 |
| Consumption increase (%) | 0 | 4.7 | 3.6 | 2.5 |

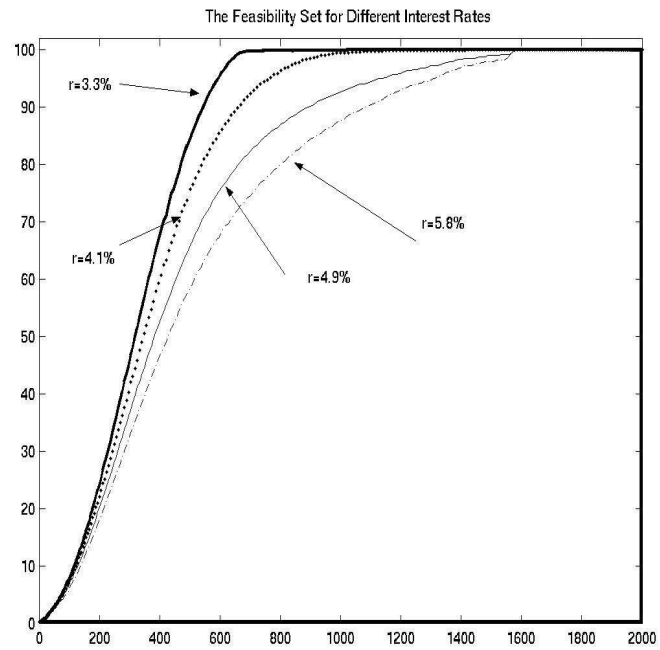


Figure A1: The domain restriction for different interest rates with a less restrictive upper bound on consumption

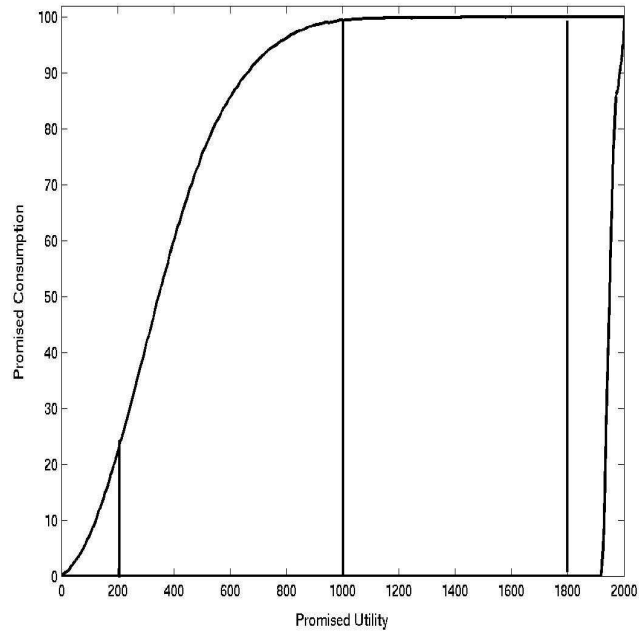


Figure A2: The partition of the restricted domain for the implementation of value function iterations