Dynamic Mechanism Design with Hidden Income and Hidden Actions *

Matthias Doepke
UCLA

Robert M. Townsend
University of Chicago

April 2002

Abstract

We develop general recursive methods to solve for optimal contracts in dynamic principal-agent models with hidden income and hidden actions. We further show that a curse of dimensionality which arises from the interaction of hidden endowments and hidden actions can be overcome by specifying behavior and utility promises off the equilibrium path. In an application to banking, we formulate the problem of providing optimal incentive-compatible credit insurance to banks who invest in risky loans, and find the optimal contract. We show that public reserves with a low return and a partial credit-guarantee scheme can be optimal if banks face uncertain and unobserved returns on their unobserved investments.

*We are grateful to Harold Cole, Hugo Hopenhayn, Ken Judd, Marek Kapicka, Roger Myerson, Ned Prescott, Phil Reny, Tom Sargent, Steve Tadelis, Ivan Werning, Ruilin Zhou, and seminar participants at Stanford, Chicago, MIT, Boston University, IIES, Richmond Fed, the SED meeting in San José, and the SIEPR conference on "Credit Market Frictions in the Macroeconomy" for helpful comments. Doepke: Department of Economics, UCLA, Box 951444, Los Angeles, CA 90095. E-mail address: doepke@econ.ucla.edu. Townsend: Department of Economics, The University of Chicago, 1126 East 59th Street, Chicago, IL 60637. E-mail address: rtownsen@midway.uchicago.edu.

1 Introduction

We consider a class of dynamic principal-agent problems in which a risk-averse agent receives unobserved income shocks and can take unobserved actions which influence future income realizations. The principal wants to provide optimal incentive-compatible insurance against the income shocks. We formulate a general planning problem allowing for history dependence and unrestricted communication and prove that this problem can be reduced to a recursive version with direct mechanisms and vectors of utility promises as the state variable. In this recursive formulation, however, a "curse of dimensionality" leads to an excessively large number of incentive constraints, which makes it impossible to compute optimal allocations numerically. We solve this problem by providing alternative equivalent formulations of the planning problem in which the planner can specify behavior off the equilibrium path. This leads to a dramatic reduction in the number of constraints that need to be imposed when computing the optimal contract. We provide numerical examples which apply our methods to hidden-investment and banking environments.

The design of optimal incentive-compatible mechanisms in environments with privately observed income shocks has been analyzed by a number of previous authors, including Townsend (1979), Green (1987), Thomas and Worrall (1990), Wang (1995), and Phelan (1995). In such environments, the principal extracts information from the agents about their endowments and uses this information to provide optimal incentive-compatible insurance. The existing literature has concentrated on environments in which the agent is not able to take unobserved actions that influence the probability distribution over future endowments. The reason for this limitation is mainly technical; with hidden actions, the agent and the planner do not have common knowledge over probabilities of future states, which renders standard methods of computing optimal mechanisms inapplicable.

Theoretical considerations suggest that the presence of hidden actions can have large effects on the constrained-optimal mechanism. To induce truthful reporting of endowments, the planner needs control over the agent's intertemporal rate of substitution. When hidden actions are introduced, the planner has less control over rates of substitution, so that providing insurance becomes more difficult. Indeed, there are some special cases where the presence of hidden actions causes insurance to break down completely. Allen (1985) shows that if an agent with hidden income shocks has unobserved access to a perfect credit market, the principal cannot offer any insurance beyond the borrowing-lending solution. The reason is that any agent will choose the reporting scheme that yields the highest net discounted transfer regardless of what the actual endowment is. Cole and Kocherlakota (2001b) consider a related environment in which the agent can only save, but not borrow. They find that as long as the return on storage or saving is sufficiently high, the optimal allocation is equivalent to access to a perfect credit market. As in Allen's result, no insurance is possible beyond self-insurance using borrowing and lending.[1]

The methods developed in this paper can be used to compute optimal incentive-constrained mechanisms in more general environments with hidden actions in which insurance does not break down completely. This happens when self-insurance is an imperfect substitute for financial saving, because the return is either low or uncertain. In such environments, optimal information-constrained insurance improves over the borrowing-and-lending solution despite the severe information problems. As in Townsend (1982), the agent can be given incentives to correctly announce the underlying state, since he is aware of his own intertemporal rate of substitution. Unobserved investment with low or random returns does not undercut this intuition.

The existing literature on dynamic mechanism design has been restricted to environments in which the curse of dimensionality that arises in our model is avoided from the outset. The paper most closely related to ours is Fernandes and Phelan (2000). Fernandes and Phelan develop recursive methods to deal with dynamic incentive problems with Markov income links between the periods, and as we do, they use a vector of utility promises as the state variable. The main difference is that Fernandes and Phelan do not consider environments in which both actions and states are unobservable to the planner, which is the main source of complexity in our setup. Werning (2001) develops a recursive first-order approach to compute a dynamic moral hazard problem with storage. Werning's method applies to a more restricted setting in which output is observed, the return on storage does not exceed the credit-market return, and storage is not subject to random shocks. We regard his methods as complementary to ours, since the computational cost of the first-order approach is lower as long as it is justified. Haegler (2001) employs a method similar to Werning's in an environment with hidden storage which can be monitored at a cost. Our analysis differs from all these papers in that we allow for randomization and unrestricted communication, and show from first principles that our recursive formulations are equivalent to the general formulation of the planning problem.

[1] The issue of credit-market access is also analyzed in Fudenberg, Holmstrom, and Milgrom (1990), who show that if principal and agent can access the credit market on equal terms, the optimal dynamic incentive contracts can be implemented as a sequence of short-term contracts.

We put our methods to work in two different applications. First, we consider an environment where an infinitely-lived agent receives unobserved income shocks, and can invest part of the endowment in a storage technology at a specified expected return. We compute the utility frontier for a range of returns, and find that the utility of the agent decreases in the return on storage as long as the return is lower than the credit-market interest rate. Thus an increased ability to self-insure through hidden actions makes insurance through the principal more difficult, leaving the agent worse off. However, insurance is not eliminated entirely. Indeed, the further utility loss from a restriction to autarky is almost as large as the loss from introducing private information into the full information environment. Here our results are similar to Ligon, Thomas, and Worrall (2000), who derive a parallel finding in an environment with limited commitment. The result is also related to the literature on the optimality of public insurance, or ironically, how public insurance can be non-optimal. For example, in Attanasio and Rios-Rull (2000) the provision of public insurance for aggregate shocks can undercut self-sustaining private insurance against idiosyncratic shocks. In Perri and Krueger (1999), similar effects arise from progressive taxation. The point is that private and public insurance must be considered jointly.

Second, we use our methods to analyze the problem of providing optimal incentive-compatible insurance in an illustrative environment in which banks face uncertain returns on their loans or investments, and the level of investment and the actual returns on investment are unobservable to anyone but the bank itself. We find that the optimal incentive-compatible insurance scheme has the feature that the planner or "central bank" builds up from collected premia a stock of reserves which have a low return. The reserves can be maintained over intermediate periods if indemnities for credit losses are balanced by premia, and can be run down in subsequent periods if payouts are sufficiently high. Our model provides a new rationale for bank reserves. Reserves do not have the traditional function of providing liquidity, but arise as a feature of the optimal credit insurance contract in an environment characterized by moral hazard. The feature which distinguishes our approach from the literature is that we do not suppose that the public aspect of insurance is imposed exogenously and suboptimally onto the privately optimizing agents. Rather, we solve a mechanism design problem which derives the optimal public policy, that is, the optimal credit insurance scheme. In effect there is little distinction between the public and the private sector. We might equally well expect the reserve-fund scheme we describe to be adopted by the private sector or to be imposed by an optimizing government. We only know that some kind of collectivity is required.

2 Outline of Methods

We consider a dynamic mechanism design problem with hidden income and hidden actions. The agent realizes an unobserved endowment at the beginning of the period, then the planner makes a transfer to the agent, and at the end of period the agent takes an action which influences the probability distribution over future income realizations. We formulate a general planning problem of providing optimal incentive-constrained insurance to the agent. Our ultimate aim is to derive a formulation of the planning problem which can be solved numerically, using linear-programming techniques as in Phelan and Townsend (1991). In order to make the problem tractable, we employ dynamic programming techniques to convert the general planning problem to problems which are recursive and of relatively low dimension. Dynamic programming is applied at two different levels. First, building on Spear and Srivastava (1987), we use utility promises as a state variable to gain a recursive formulation of the planning problem. In addition, we apply similar techniques within the period as a method for reducing the dimensionality of the resulting programming problem.

A key feature of our approach, central to what we do, is that we start from a general setup which allows randomization, full history dependence, and unrestricted communication. We formulate a general planning problem in the unrestricted setup (Section 3.2), and show from first principles that our recursive formulations are equivalent to the original formulation. Rather than imposing truth-telling and obedience from the outset, we prove a version of the revelation principle for our environment (Proposition 2).[2] Truth-telling and obedience are thus derived as endogenous constraints capturing all the information problems inherent in our setup. This is done not only to ensure that the revelation principle applies in our dynamic setting (as is often just assumed), but also because we actually apply two different versions of the revelation principle, one with single and one with double reporting. We thus derive two different direct mechanisms from the same

[2] Initially, derivations of the revelation principle as in Harris and Townsend (1977), Harris and Townsend (1981), Myerson (1979), and Myerson (1982) were for the most part static formulations. Here we are more explicit about the deviation possibilities in dynamic games. Unlike Myerson (1986), we do not focus on zero-probability events, but instead concentrate on within-period maximization operators, either using dynamic programming to work backwards from the end of the period, or off-path utility bounds to summarize all possible deviation behavior.

general planning problem. Even though both mechanisms deliver exactly the same equilibrium allocation, and it that sense double reporting is not necessary, deriving separate mechanisms is still useful, since it can lead to enormous computational gains. Which of the mechanisms is used determines to a large extent whether the resulting program is computable.

After proving the revelation principle, the next step is to work towards a recursive formulation of the planning problem. Given that in our model both endowments and actions are hidden, so that in fact the planner does not observe anything, standard methods need to be extended to be applicable to our environment. The main problem is that with hidden endowments and actions a scalar utility promise is not a sufficient state variable, since agent and planner do not have common knowledge over probabilities of current and future states. However, our underlying model environment is Markov in the sense that the unobserved action only affects the probability of tomorrow's income. Once that income is realized, it completely describes the current state of affairs for the agent, apart from any history dependence generated by the mechanism itself. We thus show that the planning problem can still be formulated recursively by using a vector of endowment-specific utility promises as the state variable (see Proposition 4 and the ensuing discussion).

It is a crucial if well-understood result that the equilibrium of a mechanism generates utility realizations. That is, along the equilibrium path a utility vector is implicitly being assigned, a scalar number for each possible endowment (though a function of the realized history). If the planner were to take that vector of utility promises as given and reoptimize so as to potentially increase surplus, the planner could do no better and no worse than under the original mechanism. Thus, equivalently, we can assign utility promises explicitly, and allow the planner to reoptimize at the beginning of each date. Using a vector of utility promises as the state variable introduces an additional complication, since the set of feasible utility vectors is not known in advance. To this end, we show that the set of feasible utility vectors can be computed recursively as well, by applying a variant the methods developed in Abreu, Pearce, and Stacchetti (1990) (Propositions 7 and 8 in Appendix A.2).[3]

Starting from our recursive formulation, we discretize the state space to formulate a version of the planning problem which can be solved using linear programming and value-function iteration (Program 1 in Section 4.3 below). However, we now face the problem

[3]See also Cole and Kocherlakota (2001a) in an application to dynamic games with hidden states and actions.

that in this "standard" recursive formulation a "curse of dimensionality" arises, in the sense that the number of constraints that need to be imposed when computing the optimal mechanism becomes very large. The problem is caused by the truth-telling constraints, which describe that reporting the true endowment at the beginning of the period is optimal for the agent. In these constraints, the utility resulting from truthful reporting has to be compared to all possible deviations. When both endowments and actions are unobserved, the number of such deviations is large. A deviation consists of lying about the endowment, combined with an "action plan" which specifies which action the agent will take, given any transfer and recommendation he may receive from the planner before taking the action. The number of such action plans is equal to the number of possible actions taken to the power of the product of the number of actions and the number of transfers (recall that there are finite grids for all choice variables to allow linear programming). The number of constraints therefore grows exponentially in the number of transfers and the number of actions. Thus even for moderate sizes for these grids, the number of constraints becomes too large to be handled on any computer.

To deal with this problem, we show that the number of constraints can be reduced dramatically by allowing the planner to specify outcomes off the equilibrium path. The intuition is that we use a maximization operator which makes it unnecessary to check all possible deviations. This can be done either by having the agent report the endowment twice, or by imposing off-path utility bounds. The advantage of specifying behavior off the equilibrium path is that optimal behavior will be at least partly defined even if the agent misreports, so that not all possible deviations need to be checked. In the version with double reporting (Program 2 in Section 5.1 below), the agent reports his endowment a second time after receiving a transfer from the planner. Incentive constraints ensure that the second report is correct, regardless of the first report. In this double-reporting version, at the first report there is just one truth-telling constraint for each alternative endowment, since the agent already knows that even conditional on lying it will be optimal to tell the truth at the second report and follow the recommendations after that. There are additional truth-telling constraints for the second report, but these are simplified since the report is made after a specific transfer has been realized. In particular, the number of possible transfers no longer enters the number of deviation action plans. The result is that with double reporting the number of constraints is approximately linear in the number of possible transfers, and grows exponentially only in the number of possible actions. In environments with a fine grid for the transfer, the number of constraints in Program

2 is dramatically lower than in Program 1. Program 2 is derived from the general planning problem with the same methods as Program 1, and the two programs are therefore equivalent.

Notice that the advantages of using double reporting are similar to the advantages of using a recursive formulation in the first place. One of the key advantages of a recursive formulation with utility promises as a state variable is that only one-shot deviations need to be considered. The agent knows that his future utility promise will be delivered, and therefore does not need to consider deviations that extend over multiple periods. Double reporting applies a similar intuition to the incentive constraints within a period. The agent knows that it will be in his interest to return to the equilibrium path in the second part of the period (from the second report on), which simplifies the incentive constraints in the first part of the period.

An alternative method for reducing the number of truth-telling constraints is to let the planner choose off-path utility bounds directly (see Program 3 in Section 5.2 below). This technique derives from Prescott (1997), who uses the same approach in a static moral-hazard framework. The planner now specifies upper bounds to the utility an agent can get by lying and receiving a specific transfer and recommendation afterwards. The truth-telling constraints can then be formulated in a particularly simple way by summing over the utility bounds. Additional constraints ensure that the utility bounds hold, i.e., the actual utility of deviating must not exceed the utility bound regardless what action the agent takes. The number of such constraints is equal to the product of the number of transfers and the square of the number of actions. The total number of constraints in Program 3 is approximately linear in the number transfers, and quadratic in the number of actions. Once again, putting structure off the equilibrium path leads to a reduction in the number of constraints. In Proposition 6, we show that Program 1 and Program 3 are equivalent. With Program 3, the planning problem can be solved with fine grids for all choice variables. In Appendix A.3, we describe how the dimensionality of this Program can be reduced even further by introducing subperiods as in Program 2. However, instead of reporting the endowment a second time, the period is subdivided completely, and utility promises in the middle of the period are assigned as an additional state variable. The reduction in the dimensionality of the programs comes at the expense of an increase in the number of programs that need to be computed.

The remainder of the paper is organized as follows. In Section 3 we introduce the economic environment that underlies our mechanism design problem, and formulate a general planning problem with unrestricted communication and full history dependence. In Section 4, we invoke and reprove the revelation principle to reformulate the planning problem, using direct message spaces and enforcing truth-telling and obedience. We then provide a recursive formulation of this problem with a vector of utility promises as the state variable (Program 1). In Section 5, we develop alternative formulations which allow the planner to specify behavior and utility promises off the equilibrium path, either by allowing the agent to report the endowment twice (Program 2), or by imposing off-path utility bounds (Program 3). These methods lead to a dramatic reduction in the number of incentive constraints that need to be imposed when computing the optimal contract. Section 6 applies our methods to a hidden-storage and a banking environment, and Section 7 concludes. All proofs are contained in the mathematical appendix.

3 The Model

In the following sections we develop a number of recursive formulations for a general mechanism design problem. For maximum generality, when deriving the different recursive formulations we concentrate on the case of infinitely many periods with unobserved endowments and actions in every period. With little change in notation, the formulations can be adapted to models with finitely many periods and/or partially observable endowments and actions.

3.1 Physical Setup

The physical setup is identical for all programs that we consider. At the beginning of each period the agent receives an income or endowment e from a finite set E. The income cannot be observed by the planner. Then the planner gives a transfer τ from a finite set T to the agent. A positive transfer can be interpreted as an indemnity and negative transfer as a premium. At the end of the period, the agent takes an action a from a finite set A. Again, the action is unobservable for the planner. In most examples below we will concentrate on positive a and interpret that as storage or investment, but without any changes in the setup we could also allow for a to be negative, which can be interpreted as borrowing. The interpretation is the usual small-economy one, with unrestricted access to outside credit markets. The agent consumes the amount $e + \tau - a$ and enjoys period utility $u(e + \tau - a)$. Our methods do not require any specific assumptions on the utility function $u(\cdot)$, apart from it being real-valued.

The action a influences the probability distribution over the income or endowment in the next period. Probability $p(e|a)$ denotes the probability of endowment e if the agent took action a in the previous period. The word "endowment" is thus a misnomer as income next period is endogenous, a function of investment or unobserved credit-market activity. It is only in the initial period that the probability $p(e)$ of endowment e does not depend on any prior actions. For tractability, and consistent with the classic formulation of a moral-hazard problem, we assume that all states occur with positive probability, regardless of the action:

Assumption 1 *The probability distribution over the endowment e satisfies $p(e|a) > 0$ for all $e \in E$, all $a \in A$.*

Otherwise, we place no restrictions on the investment technology. Note that since yesterday's action affects probabilities over endowments only in the current period[4], the income realization completely describes the current environment as far as the agent is concerned. Likewise, with time-separable utility, previous actions and endowments do not enter current utility as a state variables. Apart from physical transactions, there is also communication taking place between the agent and the planner. We do not place any prior restrictions on this communication, in order not to limit the attainable outcomes. At a minimum, the agent has to be able send a signal about his beginning-of-period endowment, and the planner has to be able to send a recommendation for the investment or unobserved action.

In what follows Q is the discount factor of the planner, and β is the discount factor of the agent. The planner is risk-neutral and minimizes the expected discounted transfer, while the agent maximizes expected discounted utility. The discount factor Q is given by $Q = \frac{1}{1+r}$, where r is taken to be the outside credit-market interest rate for borrowers and lenders in this small open economy. We assume that both discount factors are less than one so that utility is finite and our problem is well defined.

Assumption 2 *The discount factors Q and β of the planner and the agent satisfy $0 < Q < 1$ and $0 < \beta < 1$.*

[4]This assumption could be weakened as in Fernandes and Phelan (2000). If the action had effects over multiple periods, additional state variables need to be introduced to recover a recursive formulation.

When there are only finitely many periods, we only require that both discount factors be bigger than zero, because utility will still be well defined.

While we formulate the model in terms of a single agent, another powerful interpretation is that there is a continuum of agents with mass equal to unity. In that case, the probability of an event represents the fractions in the population experiencing that event. Here the planner is merely a programming device to compute an optimal allocation: when the discounted surplus of the continuum of agents is zero, then we have attained a Pareto optimum.

3.2 The Planning Problem

We now want to formulate the Pareto problem of the planner maximizing surplus subject to providing reservation utility to the agent. Since the planner does not have any information on endowments and actions of the agent, we need to take a stand on what kind of communication is possible between planner and agent. In order not to impose any constraints from the outset, we start with a general communication game with arbitrary message spaces and full history-dependence. At the beginning of each period the agent realizes an endowment e. Then the agent sends a message or report m_1 to the planner, where m_1 is in a finite set M_1. Given the message, the planner assigns a transfer $\tau \in T$, possibly at random. Then the agent sends a second message m_2, where m_2 is in some finite set M_2. The planner responds by sending a message or recommendation $m_3 \in M_3$ to the agent, and M_3 is finite as well. Finally, the agent takes an action $a \in A$. In the direct mechanisms that we will introduce later, m_1 and m_2 will be reports on the endowment e, while m_3 will be a recommendation for the action a.[5]

We will use h_t to denote the realized endowment and all choices by planner and agent in period t :

$$h_t \equiv \{e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, a_t\}.$$

We denote the space of all possible h_t by H_t. The history of up to time t will be denoted by h^t:

$$h^t \equiv \{h_{-1}, h_0, h_1, \ldots, h_t\}.$$

[5]It is customary in the literature to start with a direct mechanism from the outset, assuming that the revelation principle holds. We start with the more general setup, since we are going to derive two different direct mechanisms and need to show that they are equivalent to each other and to the more general setup. Specifically, Program 2 relies on the presence of the second report m_2.

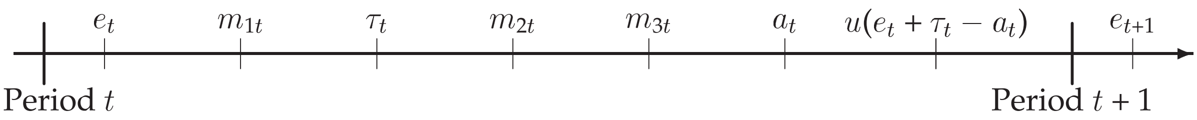

Figure 1: The Sequence of Events and Messages in Period t

Here $t = 0$ is the initial date. The set of all possible histories up to time t is denoted by H^t and is thus given by:

$$H^t \equiv H_{-1} \times H_0 \times H_1 \times \ldots \times H_t.$$

At any time t, the agent knows the entire history up to time $t-1$. On the other hand, the planner sees neither the true endowment nor the true action. We will use s_t and s^t to denote the part of the history known to the planner. We therefore have

$$s_t \equiv \{m_{1t}, \tau_t, m_{2t}, m_{3t}\},$$

where the planner's history of the game up to time t will be denoted by s^t, and the set S^t of all histories up to time t is defined analogously to the set H^t above. Since the planner sees a subset of what the agent sees, the history of the planner is uniquely determined by the history of the agent. We will therefore write the history of the planner as a function $s^t(h^t)$ of the history h^t of the agent. There is no information present at the beginning of time, and consequently we define $h_{-1} \equiv s_{-1} \equiv \emptyset$.

The choices by the planner are described by a pair of outcome functions $\pi(\tau_t|m_{1t}, s^{t-1})$ and $\pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1})$ which map the history up to the last period as known by the planner and events (messages and transfers) that already occurred in the current period into a probability distribution over transfer τ_t and a report m_{3t}. The choices of the agent are described by a strategy. A strategy consists of a function $\sigma(m_{1t}|e_t, h^{t-1})$ which maps the history up to the last period as known by the agent and the endowment into a probability distribution over the first report m_{1t}, a function $\sigma(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1})$ which determines a probability distribution over the second report m_{2t}, and a function $\sigma(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1})$ which determines the action.

We use $p(h^t|\pi, \sigma)$ to denote the probability of history h^t under a given outcome function and strategy. The probabilities over histories are defined recursively, given history h^{t-1}

and action $a_{t-1}(h^{t-1})$, by:

$$p(h^t|\pi,\sigma) = p(h^{t-1}|\pi,\sigma)\ p(e_t|a_{t-1}(h^{t-1}))\ \sigma(m_{1t}|e_t,h^{t-1})\ \pi(\tau_t|m_{1t},s^{t-1}(h^{t-1}))$$
$$\sigma(m_{2t}|e_t,m_{1t},\tau_t,h^{t-1})\ \pi(m_{3t}|m_{1t},\tau_t,m_{2t},s^{t-1}(h^{t-1}))\ \sigma(a_t|e_t,m_{1t},\tau_t,m_{2t},m_{3t},h^{t-1}).$$

Also, $p(h^t|\pi,\sigma,h^k)$ is the conditional probability of history h^t given that history h^k with $k \leq t$ occurred, and conditional probabilities are defined analogously.

For a given outcome function π and strategy σ, the expected utility of the agent is given by:

$$U(\pi,\sigma) \equiv \sum_{t=0}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi,\sigma) u(e_t + \tau_t - a_t)\right]. \tag{1}$$

The expression above represents the utility of the agent as of time zero. We will also require that the agent use a maximizing strategy at all other nodes, even if they occur with probability zero. The utility of the agent given that history h^k has already been realized is given by:

$$U(\pi,\sigma|h^k) \equiv \sum_{t=k+1}^{\infty} \beta^{t-1-k} \left[\sum_{H^t} p(h^t|\pi,\sigma,h^k) u(e_t + \tau_t - a_t)\right]. \tag{2}$$

We now define an *optimal strategy* σ for a given outcome function π as a strategy that maximizes the utility of the agent at all nodes. The requirement that the strategy be utility maximizing can be described by a set of inequality constraints. Specifically, for a given outcome function π, for any alternative strategy $\hat{\sigma}$, and any history h^k, an optimal strategy σ has to satisfy:

$$\forall \hat{\sigma}, h^k: \quad U(\pi,\hat{\sigma}|h^k) \leq U(\pi,\sigma|h^k). \tag{3}$$

Inequality (3) thus imposes or describes optimization from any history h^k on.

In addition, we also require that the strategy be optimal at any node that starts after an arbitrary first report in a period is made, i.e., even if in any period $k+1$ the first report was generated by a strategy $\hat{\sigma}$, it is optimal to revert to σ from the second report in period $k+1$ on. For any alternative strategy $\hat{\sigma}$ and any history h^k, an optimal strategy σ therefore also has to satisfy:

$$\forall \hat{\sigma}, h^k: \quad U(\pi, \hat{\sigma}|h^k) \\ \leq \sum_{h_{k+1}} p(e_{k+1}|h^k)\ \hat{\sigma}(m_{1k+1}|e_{k+1}, h^k)\ p(\tau_{k+1}, m_{2k+1}, m_{3k+1}, a_{k+1}|e_{k+1}, m_{1k+1}, \pi, \sigma, h^k) \\ \left[u(e_{k+1} + \tau_{k+1} - a_{k+1}) + \beta U(\pi, \sigma|h^{k+1})\right]. \quad (4)$$

Notice that on the right-hand side the first report m_{1k+1} is generated by strategy $\hat{\sigma}$, the remaining messages, transfers, and actions are generated under π and σ, as captured by the second term $p(\cdot|\cdot)$, and the future is generated by π and σ as well. This condition is not restrictive, since by (3) even without this condition the agent chooses the second report optimally conditional on any first report that occurs with positive probability. The only additional effect of condition (4) is to impose or describe that the agent chooses the second report and action optimally even conditional on first reports that occur with zero probability under σ, but could be generated under a counterfactual strategy. Describing optimal behavior off the equilibrium path will help us later in deriving recursive formulations of the planning problem that can be computed efficiently.

We are now able to provide a formal definition of an optimal strategy:

Definition 1 *Given an outcome function π, an optimal strategy σ is a strategy such that inequalities* (3) *and* (4) *are satisfied for all, k, all $h^k \in H^k$, and all alternative strategies $\hat{\sigma}$.*

Of course, for $h^k = h^{-1}$ this condition includes the maximization of expected utility (1) at time zero.

We imagine the planner as choosing an outcome function and a corresponding optimal strategy subject to the requirement that the agent realize at least reservation utility, W_0:

$$U(\pi, \sigma) \geq W_0. \quad (5)$$

Definition 2 *An equilibrium $\{\pi, \sigma\}$ is an outcome function π together with a corresponding optimal strategy σ such that* (5) *holds, i.e., the agent realizes at least his reservation utility. A feasible allocation is a probability distribution over endowments, transfers and actions that is generated by an equilibrium.*

The set of equilibria is characterized by the promise-keeping constraint (5), by the optimality conditions (3) and (4), and of course a number of adding-up constraints that ensure that both outcome function and strategy consist of probability measures. For brevity these latter constraints are not written explicitly.

The objective function of the planner is:

$$V(\pi, \sigma) \equiv \sum_{t=0}^{\infty} Q^t \left[\sum_{H^t} p(h^t|\pi, \sigma)(-\tau_t) \right] \tag{6}$$

When there is a continuum of agents, there is no aggregate uncertainty, and (6) is the actual surplus of the planner, or equivalently, the surplus of the community as a whole. In the single-agent interpretation, there is uncertainty about the realization of transfers, and (6) is the expected surplus. In either case, the planner's problem is to choose an equilibrium that maximizes (6). By construction, this will be Pareto optimal. The Pareto frontier can be traced out by varying reservation utility W_0.

Definition 3 *An optimal equilibrium is an equilibrium that solves the planner's problem.*

Proposition 1 *There are reservation utilities $W_0 \in R$ such that an optimal equilibrium exists.*

4 Deriving a Recursive Formulation

4.1 The Revelation Principle

Our ultimate aim is to find a computable, recursive formulation of the planning problem. We begin by showing that without loss of generality we can restrict attention to a direct mechanism where there is just one message space each for the agent and the planner. The message space of the agent will be equal to the space of endowments E, and the agent will be induced to tell the truth. The message space for the planner will be equal to the space of actions A, and it will be in the interest of the agent to follow the recommended action. Since we fix the message spaces and require that truth-telling and obedience be optimal for the agent, instead of allowing any optimal strategy as before, it has to be the case that the set of feasible allocations in this setup is no larger than in the general setup with arbitrary message spaces. The purpose of this section is to show that the set of feasible allocations is in fact identical. Therefore there is no loss of generality in restricting attention to truth-telling and obedience from the outset.[6]

More formally, we consider the planning problem described above under the restriction that $M_1 = E$ and $M_3 = A$. M_2 is set to be a singleton, so that the agent does not have an actual choice over the second report. For simplicity, we will suppress m_2 in the notation below. We can then express the contemporary part of the history of the planner as:

$$s_t \equiv \{e_t, \tau_t, a_t\},$$

with history s^t up to time t defined as above. Notice that since we are considering the history of the planner, e_t is the *reported*, not necessary actual, endowment, and a_t is the *recommended* action, not necessarily the one actually taken.

As before, the planner chooses an outcome function consisting of probability distributions over transfers and reports. For notational convenience, we express the outcome function as the joint probability distribution over combinations of transfer and recommendation. This is equivalent to choosing marginal probabilities as above. The planner therefore chooses probabilities $\pi(\tau_t, a_t|e_t, s^{t-1})$ that determine the transfer τ_t and the recommended action a_t as a function of the reported endowment e_t and the history up to the last period s^{t-1}.

We now impose constraints on the outcome function that ensure that the outcome function together with a specific strategy of the agent, namely truth-telling and obedience, are an equilibrium. First, the outcome function has to define probability measures. We require that $\pi(\tau_t, a_t|e_t, s^{t-1}) \geq 0$ for all transfers, actions, endowments and histories, and that:

$$\forall e_t, s^{t-1}: \quad \sum_{T,A} \pi(\tau_t, a_t|e_t, s^{t-1}) = 1. \tag{7}$$

Given an outcome function, we define probabilities $p(s^t|\pi)$ over histories in the obvious way, where the notation for σ is suppressed on the premise that the agent is honest and obedient. Given these probabilities, as in (5), the outcome function has to deliver reservation utility W_0 to the agent, provided that the agent reports truthfully and takes the

[6]We still prefer to start from the general setup (as opposed to just assuming that the revelation principle holds) since we are deriving two different direct mechanisms from the same general setup, and we need to show that they are equivalent.

recommended actions:

$$\sum_{t=0}^{\infty} \beta^t \left[\sum_{S^t} p(s^t|\pi) u(e_t + \tau_t - a_t) \right] \geq W_0. \tag{8}$$

Finally, it has to be optimal for the agent to tell the truth and follow the recommended action, so that (3) holds for the outcome function and the maximizing strategy σ of the agent, which is to be truthful and obedient. In particular, the the utility of honesty and obedience must weakly dominate the utility derived from any possible deviation strategy mapping any realized history, which may be generated by possible earlier lies and disobedient actions, into possible lies and alternative actions today, with plans for possible deviations in the future. We write a possible deviation strategy δ, which is allowed to be fully history-dependent, as a set of functions $\delta_e(h^{t-1}, e_t)$ that determine the reported endowment as a function of the actual history h^{t-1} and the true endowment e_t, and functions $\delta_a(h^{t-1}, e_t, \tau_t, a_t)$ that determine the actual action as a function of the history h^{t-1}, endowment e_t, transfer τ_t, and recommended action a_t. Since the actual action may be different from the recommendation, this deviation also changes the probability distribution over histories and states. The agent takes this change into account, and the changed probabilities are denoted as $p(h^t|\pi, \delta)$, with the inclusion of other conditioning elements where appropriate. In particular, we require that the actions of the agent be optimal from any history s^k on, and it will also be useful to write down separate constraints for each possible endowment e_{k+1} in period $k+1$. Then for every possible deviation (δ_e, δ_a), any history s^k, and any e_{k+1}, the outcome function has to satisfy:

$$\begin{aligned} \forall \delta, s^k, e_{k+1}: \quad & \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] \\ & \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \end{aligned} \tag{9}$$

Here $p(h^t|\pi, \delta, s^k, e_{k+1})$ on the left-hand side is the probability of actual history h^t implied by outcome function π and deviation δ, conditional on the planner's history s^k and realized endowment e_{k+1}, and the $p(s^t|\pi, s^k, e_{k+1})$ on the right-hand side is the probability under truth-telling and obedience as above, but now conditioned on s^k and e_{k+1}. Condition (9) imposes or describes honesty and obedience on the equilibrium path, similar to (3).

It might seem at first sight that (9) is less restrictive than (3), because only a subset of possible deviations is considered. Specifically, deviations are nonrandom, and a constraint is imposed only at every s^t node, instead of every node h^k of the agent's history. However, none of these limitations are restrictive. Allowing for randomized deviations would lead to constraints which are linear combinations of the constraints already imposed. Imposing (9) is therefore sufficient to ensure that the agent cannot gain from randomized deviations. Also, notice that the conditioning history s^k enters (9) only by affecting probabilities over future states s^t. These probabilities are identical for all h^k that coincide in the s^k part once e_{k+1} is realized, since the agent's private information on past endowments and actions affects the present only through the probabilities over different endowments. Imposing a separate constraint for each h^k therefore would not put additional restrictions on π.

Definition 4 *An outcome function is an equilibrium outcome function under truth-telling and obedience if it satisfies the constraints* (7), (8) *and* (9) *above. A feasible allocation in the truth-telling mechanism is a probability distribution over endowments, transfers and actions that is implied by an equilibrium outcome function.*

Feasible allocations under truth-telling and obedience are a subset of the feasible allocations in the general setup, since (8) implies that (5) holds, (9) implies that (3) holds, and (4) does not constrain allocations and could be satisfied by specifying off-path behavior appropriately. In fact, we can show that the set of feasible allocations in the general and the restricted setup are in fact identical.

Proposition 2 (Revelation Principle) *For any message spaces M_1, M_2, and M_3, any allocation that is feasible in the general mechanism is also feasible in the truth-telling-and-obedience mechanism.*

The proof (outlined in the appendix) takes the usual approach of mapping an equilibrium of the general setup into an equilibrium outcome function in the restricted setup. Specifically, given an equilibrium $(\pi^\star, \sigma^\star)$ in the general setup, the corresponding outcome function in the restricted setup is gained by prescribing the outcomes on the equilibrium path, while integrating out all the message spaces:

$$\pi(\tau_t, a_t|e_t, s^{t-1}) \equiv \sum_{H^{t-1}(s^{t-1}),M_1,M_2,M_3} p(h^{t-1}|s^{t-1})\, \sigma^\star(m_{1t}|e_t, h^{t-1})\, \pi^\star(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\sigma^\star(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1})\, \pi^\star(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\, \sigma^\star(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}).$$

The proof then proceeds by showing that the outcome function π on the left-hand side satisfies all the required constraints. The essence of the matter is that lying or deviating under the new outcome function would be equivalent to using the optimizing strategy function under the original outcome function, but evaluated at a counterfactual realization. For example, an agent who has endowment e but reports $\hat{e}$ will face the same probability distribution over transfers and recommendations as an agent who under the original outcome function behaved "as if" the endowment were $\hat{e}$. The agent can never gain this way, since $\sigma^\star$ is an optimal strategy, and it is therefore preferable to receive the transfers and recommendations intended for endowment e instead of $\hat{e}$.

We are therefore justified in continuing with the restricted setup which imposes truth-telling and obedience. The objective function of the planner is now:

$$V(\pi) \equiv \sum_{t=0}^{\infty} Q^t \left[\sum_{S^t} p(s^t|\pi)(-\tau_t) \right], \tag{10}$$

and the original planning problem can be expressed as maximizing (10) subject to (7), (8), and (9) above.

4.2 Utility Vectors as State Variables

We now have a representation of the planning problem that requires truth-telling and obedience, and yet does not constitute any loss of generality. However, we still allow fully history-dependent outcome functions π. The next step is to reduce the planning problem to a recursive version with a vector of promised utilities as the state variable.

We wish to work towards a problem in which the planner has to deliver a vector of promised utilities at the beginning of period k, with elements depending on the endowment e_k. It will be useful to consider an auxiliary problem in which the planner has to deliver a vector of reservation utilities $\mathbf{w}_0$, depending on the endowment in the initial period. The original planning problem can then be cast, as we shall see below, as choosing the vector of initial utility assignments $\mathbf{w}_0$ which yields the highest expected surplus for the planner, given the initial exogenous probability distribution over states $e \in E$ at time $t = 0$.

In the auxiliary planning problem, we impose the same probability constraints (7) and incentive constraints (9) as before. However, instead of a single promise-keeping constraint (8) there is now a separate promise-keeping constraint for each possible initial endowment. For all e_0, we require:

$$\forall e_0 :$$
$$\sum_{T,A} \pi(\tau_0, a_0 | \mathbf{w}_0, e_0) \left[u(e_0 + \tau_0 - a_0) + \sum_{t=1}^{\infty} \beta^t \left[\sum_{S^t} p(s^t | \pi, s_0) u(e_t + \tau_t - a_t) \right] \right] = w_0(e_0). \quad (11)$$

Here the vector $\mathbf{w}_0$ of endowment-specific utility promises $w_0(e_0)$ is taken as given. Notice that we write the outcome function π as a function of the vector of initial utility promises $\mathbf{w}_0$. In period 0, there is no prior history, but in a subsequent period t the outcome function also depends on the history up to period $t-1$, so that the outcome function would be written as $\pi(\tau_t, a_t | \mathbf{w}_0, e_t, s^{t-1})$.

In principle, specifying a separate utility promise for each endowment is more restrictive than a requiring that a scalar utility promise be delivered in expected value across endowments. However, the original planning problem can be recovered by introducing an initial stage at which the initial utility vector is chosen by the planner. Since the vector of promised utilities $\mathbf{w}_0$ will serve as our state variable, it will be important to show that the set of all feasible utility vectors has nice properties.

Definition 5 *The set* $\mathbf{W}$ *is given by all vectors* $\mathbf{w}_0 \in R^{\#E}$ *that satisfy constraints* (7), (9), *and* (11) *for some outcome function* $\pi(\tau_t, a_t | e_t, s^{t-1})$.

Proposition 3 *The set* $\mathbf{W}$ *is nonempty and compact.*

Now we consider the problem of a planner who has promised utility vector $\mathbf{w}_0 \in \mathbf{W}$ and has received report e_0 from the agent. In the auxiliary planning problem, the maximized surplus of the planner is given by:

$$V(\mathbf{w}_0, e_0) = \max_{\pi} \sum_{T,A} \pi(\tau_0, a_0 | \mathbf{w}_0, e_0) \left[-\tau_0 + \sum_{t=1}^{\infty} Q^t \left[\sum_{S^t} p(s^t | \pi, s_0)(-\tau_t) \right] \right], \quad (12)$$

where the maximization over current and future π is subject to constraints (7), (9), and (11) above, for a given $\mathbf{w}_0 \in \mathbf{W}$ and $e_0 \in E$.

We want to show that this problem has a recursive structure. To do this, we need to define on-path future utilities that result from a given choice of π. For all s^{k-1}, e^k, let:

$$w(e_k, s^{k-1}, \pi) =$$
$$\sum_{T,A} \pi(\tau_k, a_k | \mathbf{w}_0, e_k, s^{k-1}) \left[u(e_k + \tau_k - a_k) + \sum_{t=k+1}^{\infty} \beta^{t-k} \left[\sum_{S^t} p(s^t | \pi, s^k) u(e_t + \tau_t - a_t) \right] \right], \quad (13)$$

and let $\mathbf{w}(s^{k-1}, \pi)$ be the vector of these utilities over all e_k. We can now show a version of the principle of optimality for our environment:

Proposition 4 *For all $\mathbf{w}_0 \in \mathbf{W}$ and $e_0 \in E$, and for any s^{k-1} and e_k, there is an optimal contract $\pi^\star$ such that the remaining contract from s^{k-1} and e_k is an optimal contract for the auxiliary planning problem with $e_0 = e_k$ and $\mathbf{w}_0 = \mathbf{w}(s^{k-1}, \pi^\star)$.*

Thus the planner is able to reoptimize the contract at any future node. For Proposition 4 to go through, it is essential that we chose a vector of utility promises as the state variable, as opposed to the usual scalar utility promise which is realized in expected value across states. If the planner reoptimized given a scalar utility promise at a given date, the distribution of expected utilities across states might be different than in the original contract. Such a reallocation of utilities would change the incentives for lying and disobedience in the preceding period, so that incentive-compatibility of the complete contract would no longer be guaranteed. This problem is avoided by specifying a separate utility promise for each possible endowment. Likewise, in implementing the utility promises it does not matter whether the agent lied or was disobedient in the past, since the agent has to report the realized endowment anyway, and once the endowment is realized past actions have no further effects.[7]

Given Proposition 4, we know that the maximized surplus of the planner can be written as:

$$V(\mathbf{w}_0, e_0) = \sum_{A,T} \pi^\star(\tau_0, a_0 | \mathbf{w}_0, e_0) \left[-\tau_0 + Q \sum_E p(e_1 | s^0) V(\mathbf{w}(s^0, \pi^\star), e_1) \right]. \quad (14)$$

[7]The state space would have to be extended further if the action affected outcomes for more than one period into the future.

In light of (14), we can cast the auxiliary planning problem as choosing transfers and actions in the initial period, and choosing continuation utilities from the set $\mathbf{W}$, conditional on history $s^0 = \{e_0, \tau_0, a_0\}$.

We are now close to the recursive formulation of the planning problem that we are looking for. We will drop the time subscripts from here on, and write the choices of the planner as a function of the current state, namely the vector of promised utilities $\mathbf{w}$ that has to be delivered in the current period, and the reported endowment e. The choices are functions $\pi(\tau, a|\mathbf{w}, e)$ and $\mathbf{w}'(\mathbf{w}, e, \tau, a)$, where $\mathbf{w}'$ is the vector of utilities promised from tomorrow on, which is restricted to lie in $\mathbf{W}$. Assuming that the value function V is known (it needs to be computed in practice), the auxiliary planning problem can be solved by solving a static optimization problem for all vectors in $\mathbf{W}$. An optimal contract for the non-recursive auxiliary planning problem can be found by assembling the appropriate solutions of the static problem.

We still need to determine which constraints need to be placed on the static choices $\pi(\tau, a|\mathbf{w}, e)$ and $\mathbf{w}'(\mathbf{w}, e, \tau, a)$ in order to guarantee that the implied contract satisfies probability measure constraints (7), maximization (9), and promise keeping (11) above. In order to reproduce (7), we need to impose:

$$\sum_{T,A} \pi(\tau, a|\mathbf{w}, e) = 1. \tag{15}$$

The promise-keeping constraint (11) will be satisfied if we impose:

$$\sum_{T,A} \pi(\tau, a|\mathbf{w}, e) \left[u(e + \tau - a) + \beta \sum_{E} p(e'|a) w'(\mathbf{w}, e, \tau, a)(e') \right] = w(e) \tag{16}$$

where along the equilibrium path, honesty and obedience prevails in reports e and actions a. The incentive constraints are imposed in two parts. We first require that the agent cannot gain by following another action strategy $\delta_a(\tau, a)$, assuming that the reported endowment e was correct. Note that e enters the utility function as the actual value and as the conditioning element in π as the reported value.

$$\sum_{T,A} \pi(\tau, a|\mathbf{w}, e) \left[u(e + \tau - \delta_a(\tau, a)) + \beta \sum_{E} p(e'|\delta_a(\tau, a)) w'(\mathbf{w}, e, \tau, a)(e') \right] \leq w(e). \tag{17}$$

A similar constraint on disobedience is also required if the initial report was e, but the true state was $\hat{e}$, i.e., false reporting. Note that $\hat{e}$ enters the utility function as the actual value but e is the conditioning element in π on the left-hand side, and $w(\hat{e})$ is the on-path realized utility under honesty and obedience at $\hat{e}$.

$$\sum_{T,A} \pi(\tau, a|\mathbf{w}, e) \left[u(\hat{e} + \tau - \delta_a(\tau, a)) + \beta \sum_{E} p(e'|\delta_a(\tau, a)) w'(\mathbf{w}, e, \tau, a)(e') \right] \leq w(\hat{e}). \tag{18}$$

Conditions (17) and (18) impose a sequence of period-by-period incentive constraints on the implied full contract. The constraints rule out that the agent can gain from disobedience or misreporting in any period, given that he goes back to truth-telling and obedience from the next period on. Equations (17) and (18) therefore imply that (9) holds for one-shot deviations. We still have to show that (17) and (18) are sufficient to prevent deviations in multiple periods, but the argument follows as in Phelan and Townsend (1991). That is, for a finite number of deviations, we can show that the original constraints are satisfied by backward induction. The agent clearly does not gain in the last period when he deviates, since this is just a one-time deviation and by (17) and (18) is not optimal. Going back one period, the agent has merely worsened his future expected utility by lying or being disobedient in the last period. Since one-shot deviations do not improve utility, the agent cannot make up for this. Going on this way, we can show by induction that any finite number of deviations does not improve utility. Lastly, consider an infinite number of deviations. Let us assume that there is a deviation that gives a gain of ϵ. Since $\beta < 1$, there is a period T such that at most $\epsilon/2$ utils can be gained from period T on. This implies that at least $\epsilon/2$ utils have to be gained until period T. But this contradicts our result that there cannot be any gain from deviations with a finite horizon.

Thus we are justified to pose the auxiliary planning problem as solving:

$$V(\mathbf{w}, e) = \max_{\pi \geq 0, \mathbf{w}'} \sum_{A,T} \pi(\tau, a|\mathbf{w}, e) \left[-\tau + Q \sum_{E} p(e'|a) V(\mathbf{w}'(\mathbf{w}, e, \tau, a), e') \right] \tag{19}$$

by choice of π and $\mathbf{w}'$, subject to constraints (15) to (18) above. Program 1 below is a version of this problem with a discrete grid for promised utilities as an approximation. We have assumed that the function $V(\mathbf{w}, e)$ is known. In practice, $V(\mathbf{w}, e)$ can be computed with standard dynamic programming techniques. Specifically, the right-hand side of (19) defines an operator T that maps functions $V(\mathbf{w}, e)$ into $TV(\mathbf{w}, e)$. It is easy to show, as in Phelan and Townsend (1991), that T maps bounded continuous functions into bounded continuous functions, and that T is a contraction. It then follows that T has a unique fixed point, and the fixed point can be computed by iterating on the operator T.

The preceding discussion was based on the assumption that the set $\mathbf{W}$ of feasible utility vectors is known in advance. In practice, $\mathbf{W}$ is not known and needs to be computed alongside the value function $V(\mathbf{w}, e)$. $\mathbf{W}$ can be computed with the dynamic-programming methods described in detail in Abreu, Pearce, and Stacchetti (1990). An outline of the method is given in Appendix A.2.

Finally, the entire discussion is easily specialized to the case of a finite horizon T. V_T would be the value function for period T, V_{T-1} for period $T-1$, $\mathbf{W_{T-1}}$ the set of feasible promised utilities at time $T-1$, and so on.

4.3 The Discretized Version

For numerical implementation of the recursive formulation of the planning problem, we require finite grids for all choice variables in order to employ linear programming techniques. $\#E$ is the number of grid points for the endowment, $\#T$ is the number of possible transfers, and $\#A$ is the number of actions. The vector of promised utilities is also assumed to be in a finite set $\mathbf{W}$, and the number of possible choices is $\#\mathbf{W}$. To stay in the linear programming framework, we let the planner choose a probability distribution over vectors of utility promises, instead of choosing a specific utility vector.[8] That is, τ, a, and $\mathbf{w}'$ are chosen jointly under π. Notice that while the finite grids for endowment, transfer, and action are features of the physical setup of the model, the finite grid for utility promises is merely a numerical approximation of the continuous set in our theoretical formulation.

With finite grids, the optimization problem of a planner who has promised vector $\mathbf{w}$ and *has received report* e is:

Program 1:

$$V(\mathbf{w}, e) = \max_{\pi \geq 0} \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[-\tau + Q \sum\nolimits_E p(e'|a) V(\mathbf{w}', e')\right] \tag{20}$$

[8]This imposes no loss of generality, since choosing probabilities over utility promises is equivalent to choosing the corresponding expected utility vector directly.

subject to the constraints (21) to (24) below. The first constraint is that the $\pi(\cdot)$ sum to one to form a probability measure, as in (15):

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) = 1. \tag{21}$$

Second, the contract has to deliver the utility that was promised for state e, as in (16):

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(e + \tau - a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right] = w(e). \tag{22}$$

Third, the agent needs incentives to be obedient. Corresponding to (17), for each transfer τ and recommended action a, the agent has to prefer to take action a over any other action $\hat{a} \neq a$:

$$\begin{aligned} \forall \tau, a, \hat{a} \neq a: \qquad & \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(e + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a})w'(e')\right] \\ & \leq \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(e + \tau - a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right]. \end{aligned} \tag{23}$$

Finally, the agent needs incentives to tell the truth, so that no agent with endowment $\hat{e} \neq e$ would find this branch attractive. Under the promised utility vector $\mathbf{w}$, agents at $\hat{e}$ should get $w(\hat{e})$. Thus, an agent who actually has endowment $\hat{e}$ but says e nevertheless must not get more utility than was promised for state $\hat{e}$. This has to be the case regardless whether the agent follows the recommendations for the action or not. Thus, for all states $\hat{e} \neq e$ and all functions $\delta : T \times A \to A$ mapping transfer τ and recommended action a into an action $\delta(\tau, a)$ actually taken, we require as in (18):

$$\begin{aligned} & \forall \hat{e} \neq e, \delta: \\ & \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e} + \tau - \delta(\tau, a)) + \beta \sum\nolimits_E p(e'|\delta(\tau, a))w'(e')\right] \leq w(\hat{e}). \end{aligned} \tag{24}$$

Note that similar constraints are written for the $\hat{e}$ problem, so that agents with $\hat{e}$ receive $w(\hat{e})$ from a constraint like (22). For a given vector of utility promises, there are $\#E$ Program 1's to solve.

Program 1 allows us to numerically solve the auxiliary planning problem for a given vector of utility promises, by using linear programming and iteration on the value function. To recover the original planning problem with a scalar utility promise W_0, we let the planner offer a lottery $\pi(\mathbf{w}|W_0)$ over utility vectors $\mathbf{w}$ before the first period starts and before e is known. The problem of the planner at this initial stage is:

$$V(W_0) = \max_{\pi \geq 0} \sum_{\mathbf{W}} \pi(\mathbf{w}|W_0) \ \left[\sum\nolimits_E p(e) V(e, \mathbf{w})\right] \tag{25}$$

subject to a probability and a promise-keeping constraint:

$$\sum_{\mathbf{W}} \pi(\mathbf{w}|W_0) = 1. \tag{26}$$

$$\sum_{\mathbf{W}} \pi(\mathbf{w}|W_0) \left[\sum\nolimits_E p(e)\ w(e)\right] \geq W_0. \tag{27}$$

The same methods can be used for computing models with finitely many periods. With finitely many periods, the value functions carry time subscripts. The last period T would be computed first, by solving Program 1 with all terms involving utility promises omitted. The computed value function $V_T(\mathbf{w}, e)$ for period T is then an input in the computation of the value function for period $T-1$. Moving backward in time, the value function for the initial period is computed last.

An important practical limitation of the approach outlined this far is that the number of truth-telling constraints in Program 1 is very large, which makes computation practically infeasible even for problem with relatively small grids. For each state $\hat{e}$ there is a constraint for each function $\delta : T \times A \rightarrow A$, and there are $(\#A)^{(\#T \times \#A)}$ such functions. Unless the grids for τ and a are rather sparse, memory problems make the computation of this program infeasible. The total number of variables in this formulation, the number of objects under $\pi(\cdot)$, is $\#T \times \#A \times \#\mathbf{W}$. There is one probability constraint (21) and one promise-keeping constraint (22). The number of obedience constraints (23) is $\#T \times \#A \times (\#A - 1)$, and the number of truth-telling constraints (24) is $(\#E - 1) \times (\#A)^{(\#T \times \#A)}$. Thus, the number of constraints grows exponentially with the product of the grid sizes for actions and transfers.

As an example, consider a program with two states e, ten transfers τ, two actions a, and ten utility vectors $\mathbf{w}'$. With only ten points, the grids for transfers and utility promises are rather sparse. Still, for this example and a given vector of utility promises $\mathbf{w}$ and realized state e Program 1 is a linear program with 200 variables and 1,048,598 constraints. If we increase the number of possible actions a to ten, the number of truth-telling constraints alone is 10^{100}. Clearly, such programs will not be computable now or any time in the future. It is especially harmful that the grid size for the transfer causes computational problems, as it does here because of the dimensionality of $\delta(\tau, a)$. One can imagine economic environments in which there are only a small number of options for actions available, but it is much harder to come up with a reason why the planner should be restricted to a small set of transfers. In the next section we present alternative formulations, equivalent to the one developed here, that can be used for computing solutions in environments with many transfers and many actions.

5 Computationally Efficient Formulations

In this section we develop a series of alternative recursive formulations of the planning problem which are equivalent to Program 1, but require a much smaller number of constraints. The key method for reducing the number of constraints in the program is to allow the planner to specify behavior and utility promises off the equilibrium path. This can be achieved by multiple reporting, or by incorporating off-path utility bounds as choice variables for the planner. Further efficiency gains are possible if the period is divided into two subperiods with separate planning problems. We will address each method in turn.

5.1 A Version with Double Reporting

The basic idea of this section is to let the agent report the endowment a second time after the transfer is received, but before a recommendation for the action is received. On the equilibrium path, the agent will make the correct report twice and follow the recommended action, and the optimal allocation will be the same as in the first formulation. At first sight, it might therefore appear that the second report is not necessary, since in equilibrium it will always coincide with the first report. The advantage of double reporting is that it allows the planner to specify behavior off the equilibrium path, because outcomes are determined even if the two reports differ. We will see that this possibility leads to a significant reduction in the number of incentive constraints.

To see how the version with double reporting works, it is instructive to retrace some of the steps which led from the general setup to Program 1. We started by applying the revelation principle to argue that without loss of generality we can restrict attention to direct mechanisms with truth-telling and obedience. The proof of the revelation principle proceeds by showing that any equilibrium outcome in the general setup can be mapped into an equivalent outcome function in the restricted setup. The outcome function π in the restricted setup is derived from an equilibrium outcome function and strategy $(\pi^\star, \sigma^\star)$ in the general setup by integrating out the message spaces M_1, M_2, and M_3, and prescribing the outcomes that occur on the equilibrium path (see equation (45) in the Appendix):

$$\pi(\tau_t, a_t | e_t, s^{t-1}) \equiv \sum_{H^{t-1}(s^{t-1}), M_1, M_2, M_3} p(h^{t-1}|s^{t-1})\, \sigma^\star(m_{1t}|e_t, h^{t-1})\, \pi^\star(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\sigma^\star(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1})\, \pi^\star(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\, \sigma^\star(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}). \quad (28)$$

While outcomes on the equilibrium path are unchanged, the original equilibrium $(\pi^\star, \sigma^\star)$ contains information on outcomes off the equilibrium path which are lost in the switch to the truth-telling-and-obedience outcome function π. Specifically, the original equilibrium strategy prescribes optimal behavior following any initial report m_{1t}, even after those reports which never actually occur in equilibrium. This would include reports that are generated under $\sigma^\star$, but at a counterfactual e_t.

In the restricted setup with double reporting the two message spaces of the agent are given by the space of endowments, $M_1 = M_2 = E$, and the message space of the planner is the space of possible actions, $M_3 = A$. The planner chooses an outcome function $\pi(\tau_t, a_t|e_{1t}, e_{2t}, s^{t-1})$ which determines the transfer and the action as a function of the two reported endowments by the agent and the history up to period $t-1$. Notice that since this function is also specified for the case that the two reports differ, the planner in effect specifies behavior off the equilibrium path. Exactly what happens if the two reports differ is tightly linked to the prescriptions for off-path behavior of the equilibrium $(\pi^\star, \sigma^\star)$ in the general setup. Specifically, $\pi(\tau_t, a_t|e_{1t}, e_{2t}, s^{t-1})$ describes the outcome if the actual endowment is e_{2t}, but the first report of the agent was governed by e_{1t}, as if the agent acted as if the true endowment were e_{1t} when making the first report. Formally, an equilibrium $(\pi^\star, \sigma^\star)$ of the original game is transformed into a new outcome function π under double reporting in the following way:

$$\pi(\tau_t, a_t|e_{1t}, e_{2t}, s^{t-1}) \equiv \sum_{H^{t-1}(s^{t-1}) M_1, M_2, M_3} p(h^{t-1}|s^{t-1})\, \sigma^\star(m_{1t}|e_{1t}, h^{t-1})\, \pi^\star(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\sigma^\star(m_{2t}|e_{2t}, m_{1t}, \tau_t, h^{t-1})\, \pi^\star(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\, \sigma^\star(a_t|e_{2t}, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}). \quad (29)$$

Notice that the first report m_{1t} is governed by e_{1t}, while the second report m_{2t} and the action a_t are governed by e_{2t}. If e_{1t} and e_{2t} coincide, the prescriptions are the usual on-path equilibrium outcomes, and (28) and (29) are the same. If e_{1t} and e_{2t} are different, the outcomes are as if the agent mistakenly assumed that the endowment was e_{1t} when making the first report, a non-maximizing strategy, but then realized that the endowment was actually e_{2t} when making the second report and choosing the action. In the general setup, such off-path behavior is always well defined, since the agent needs to specify a strategy for all possible initial reports, even those that occur with zero probability.

The advantage of using this information in a version with truth-telling and obedience is that the resulting outcome function satisfies a version of constraint (4), which states that the actions of the agent from the second report on have to be optimal *regardless* of what the first report was. Translated into the version with truth-telling and obedience, the constraint requires on the right-hand side of (30) below that after any history h^k, even if the first report in period $k+1$ were wrong and generated by $\delta_{e1}(h^k, e_{k+1})$, it would be optimal for the agent to tell the truth e_{k+1} at the second report in $k+1$, follow the recommended action a_{k+1}, and be honest and obedient in the future, instead of following some deviation δ in the present and future as on the left-hand side:

$$\begin{aligned}\forall \delta, s^k, e_{k+1}: \quad & \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] \\ & \leq \beta^{k+1} \left[\sum_{T,A} \pi\left(\tau_{k+1}, a_{k+1} | \delta_{e1}(h^k, e_{k+1}), e_{k+1}, s^k\right) u(e_{k+1} + \tau_{k+1} - a_{k+1}) \right] \\ & \qquad + \sum_{t=k+2}^{\infty} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \end{aligned} \tag{30}$$

This constraint is imposed in addition to a version of the usual, beginning-of-the-period incentive constraint (9). Imposing (30) simplifies the incentive constraints for being truthful at the first report, since the agent already knows that he will be truthful and obedient from the second report on, regardless of what he says at the first report. This leads to a dramatic reduction in the number of incentive constraints.

The derivation of the version with double reporting follows the same outline taken above for Program 1. Since the steps and proofs are virtually identical, we omit them here. The point is that both programs are derived from first principles. The apparently more general double-reporting version is however easily reduced to the single-reporting version along the equilibrium path. Thus both are equivalent.[9]

We go directly to the discretized recursive version of the planning problem with double reporting. We will concentrate on the differences between single and double reporting, and the efficiency gains resulting from specifying behavior off the equilibrium path. As in Program 1, the agent comes into the period with a vector of promised utilities $\mathbf{w}$. At the beginning of the period, the agent observes the state e and makes a first report to the planner. Then the planner delivers the transfer τ, and afterwards the agent reports the endowment e again. Incentive-compatibility constraints, based on (30), ensure that this second report will be correct, even if the first report is false. Because now the planner receives a report after the transfer, the number of possible transfers does not affect the number of truth-telling constraints, as it did in (24). This is the main source of the efficiency gain of double reporting.

Program 2:
The optimization problem of a planner who promised utility vector $\mathbf{w}$ and *has already received first report* e is:

$$V(\mathbf{w}, e) = \max_{\pi \geq 0} \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, e) \ \left[-\tau + Q \sum\nolimits_{E} p(e'|a)V(\mathbf{w}', e')\right] \tag{31}$$

subject to constraints (32)-(37) below. Notice that the contract $\pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, e)$ is conditioned on two reports e, unlike in (20). The first constraint, much like (21), is that the $\pi(\cdot)$ form a probability measure for any second report $\hat{e}$. Note again that in the program endowment e is a fixed state or parameter.

$$\forall \hat{e}: \quad \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) = 1. \tag{32}$$

Since the second report is made *after* the transfer, we have to enforce that the transfer does not depend on the second report. For all $\hat{e} \neq \tilde{e}$ and all τ, we require:

$$\forall \hat{e} \neq \tilde{e}, \tau: \quad \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) = \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \tilde{e}). \tag{33}$$

Given that the agent told the truth twice, the contract has to deliver the promised utility

[9] The full derivations are carried out in Doepke and Townsend (2002).

$w(e)$ for state e from vector $\mathbf{w}$. That is, much like (22) above:

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, e) \ \left[u(e + \tau - a) + \beta \sum\nolimits_E p(e'|a) w'(e')\right] = w(e). \tag{34}$$

Next, the agent needs to be obedient. Given that the second report is true, it has to be optimal for the agent to follow the recommended action a. For each true state $\hat{e}$, transfer τ, recommended action a, and alternative action $\hat{a} \neq a$ we require much like (23):

$$\forall \hat{e}, \tau, a, \hat{a} \neq a : \quad \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \ \left[u(\hat{e} + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e')\right]$$
$$\leq \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \ \left[u(\hat{e} + \tau - a) + \beta \sum\nolimits_E p(e'|a) w'(e')\right]. \tag{35}$$

We also have to ensure that the agent prefers to tell the truth at the second report, no matter what he reported the first time around. *Since the transfer is already known at the time the second report is made, the number of deviations from the recommended actions that we have to consider does not depend on the number of possible transfers*. For each actual $\hat{e}$, transfer τ, second report $\hat{\hat{e}} \neq \hat{e}$, and action strategy $\delta : A \to A$, we require:

$$\forall \hat{e}, \tau, \hat{\hat{e}} \neq \hat{e}, \delta : \quad \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{\hat{e}}) \left[u(\hat{e} + \tau - \delta(a)) + \beta \sum\nolimits_E p(e'|\delta(a)) w'(e')\right]$$
$$\leq \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \left[u(\hat{e} + \tau - a) + \beta \sum\nolimits_E p(e'|a) w'(e')\right]. \tag{36}$$

This last constraint is derived from constraint (30) above, and is specific to the version with double reporting. Finally, we also have to ensure that the first report e be correct. That is, an agent who is truly at state $\hat{e}$ and should get $w(\hat{e})$, but made a counterfactual first report e, cannot get more utility than was promised for state $\hat{e}$. For all $\hat{e} \neq e$ we require:

$$\forall \hat{e} \neq e : \quad \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \left[u(\hat{e} + \tau - a) + \beta \sum\nolimits_E p(e'|a) w'(e')\right] \leq w(\hat{e}). \tag{37}$$

Notice that these latter truth-telling constraints do not involve deviations in the action a. At the time of the first report the agent knows that the second report will be correct and that he will take the recommended action, because constraints (35) and (36) hold. In Program 1 the agent had to consider a much more complicated set of deviations when deciding on the the first report, resulting in a much higher number of constraints. Despite the differences in the sets of constraints, Program 1 and Program 2 are equivalent.

Proposition 5 *Program 1 and Program 2 are equivalent.*

The proof for this proposition proceeds by showing that Program 2 is equivalent to the general planning problem, using the same arguments as in the derivation of Program 1 above.[10] Since both Program 1 and Program 2 are equivalent to the general planning problem, they are also equivalent to each other.

The number of variables in this formulation is $\#E \times \#T \times \#A \times \#\mathbf{W}$. Thus, the number of variables increased relative to Program 1, since π now also depends on the second report $\hat{e}$. The number of constraints, however, is much lower than in Program 1. There are $\#E$ probability constraints (32), an increase, $(\#E-1) \times \#T$ independence constraints (33), entirely new, and there is one promise-keeping constraint (34), as before. The total number of obedience constraints (35) is $\#E \times \#T \times \#A \times (\#A-1)$, an increase. The number of truth-telling constraints for the second report (36) is $\#E \times \#T \times (\#E-1) \times (\#A)^{(\#A)}$, and the number of truth-telling constraints for the first report (37) is $\#E-1$. This is where we obtain a huge reduction in the number of constraints.

Going back to our example, consider a program with two states e, ten transfers τ, two actions a, and ten utility vectors $\mathbf{w}'$. For this example Program 2 has 400 variables and 134 constraints. Compared to Program 1, the number of variables increases by 200, but the number of constraints decreases by more than one million. This makes it possible to solve Program 2 on a standard personal computer. Program 2 does less well if the number of actions is large. If we increase the number of actions a to ten, Program 2 has 2000 variables and more than 10^{11} truth-telling constraints. This is many orders of magnitude smaller than Program 1, but still too big to be handled by standard computer hardware. The key advantage of Program 2 relative to Program 1 is that the number of constraints does not increase exponentially with the number of possible transfers τ. As long as the number of possible actions a is small, this formulation allows computation with fine grids for the other variables. However, the number of constraints still increases exponentially with the number of actions. In the next section we will present yet another formulation of our original program which solves this problem, once again by specifying outcomes off the equilibrium path.

[10]The proof is carried out in detail in Doepke and Townsend (2002).

5.2 A Version With Off-Path Utility Bounds

We saw already in the last section that specifying behavior off the equilibrium path can lead to a reduction in the number of incentive-compatibility constraints. We will now exploit this idea in a way similar to Prescott (1997) in order to reduce the number of truth-telling constraints. The choice variables in the new formulation include utility bounds $v(\cdot)$ that specify the maximum utility (that is, current utility plus expected future utility) an agent can get when lying about the endowment and receiving a certain recommendation. Specifically, for a given reported endowment e, $v(\hat{e}, e, \tau, a)$ is an upper bound for the utility of an agent who actually has endowment $\hat{e} \neq e$, reported endowment e nevertheless, and received transfer τ and recommendation a. This utility bound is already weighted by the probability of receiving transfer τ and recommendation a. Thus, in order to compute the total expected utility that can be achieved by reporting e when the true state is $\hat{e}$, we simply have to sum the $v(\hat{e}, e, \tau, a)$ over all possible transfers τ and recommendations a. The truth-telling constraint is then that the utility of saying e when being at state $\hat{e}$ is no larger than the utility promise $w(\hat{e})$ for $\hat{e}$.

Program 3:
The optimization problem of the planner in this formulation given report e and promised utility vector $\mathbf{w}$ is:

$$V(\mathbf{w}, e) = \max_{\pi \geq 0, v} \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[-\tau + Q \sum\nolimits_E p(e'|a) V(\mathbf{w}', e')\right] \tag{38}$$

subject to the constraints (39)-(43) below. Apart from the addition of the utility bounds $v(\cdot)$ the objective function (38) is identical to (20). The first constraint is the probability measure constraint, identical with (21):

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) = 1. \tag{39}$$

The second constraint is the promise-keeping constraint, identical with (22):

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(e + \tau - a) + \beta \sum\nolimits_E p(e'|a) w'(e')\right] = w(e). \tag{40}$$

We have to ensure that the agent is obedient and follows the recommendations of the planner, given that the report is true. For each transfer τ, recommended action a, and alternative action $\hat{a} \neq a$, we require as in (23):

$$\forall \tau, a, \hat{a} \neq a: \quad \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(e+\tau-\hat{a}) + \beta \sum\nolimits_{E} p(e'|\hat{a}) w'(e')\right]$$
$$\leq \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(e+\tau-a) + \beta \sum\nolimits_{E} p(e'|a) w'(e')\right]. \quad (41)$$

Next, the utility bounds have to be observed. An agent who reported state e, is in fact at state $\hat{e}$, received transfer τ, and got the recommendation a, cannot receive more utility than $v(\hat{e}, e, \tau, a)$, where again $v(\hat{e}, e, \tau, a)$ incorporates the probabilities of transfer τ and recommendation a. For each state $\hat{e} \neq e$, transfer τ, recommendation a, and all possible actions $\hat{a}$ we require:

$$\forall \hat{e} \neq e, \tau, a, \hat{a}: \quad \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e}+\tau-\hat{a}) + \beta \sum\nolimits_{E} p(e'|\hat{a}) w'(e')\right] \leq v(\hat{e}, e, \tau, a). \quad (42)$$

Finally, the truth-telling constraints are that the utility of an agent who is at state $\hat{e}$ but reports e cannot be larger than the utility promise for $\hat{e}$. For each $\hat{e} \neq e$ we require:

$$\forall \hat{e} \neq e: \quad \sum_{T,A} v(\hat{e}, e, \tau, a) \leq w(\hat{e}). \quad (43)$$

The number of variables in this problem is $\#T \times \#A \times \#\mathbf{W}$ under $\pi(\cdot)$ plus $(\#E-1) \times \#T \times \#A$, where the latter terms reflect the utility bounds $v(\cdot)$ that are now choice variables. There is one probability constraint (39) and one promise-keeping constraint (40). The number of obedience constraints (41) is $\#T \times \#A \times (\#A-1)$. There are $(\#E-1) \times \#T \times (\#A)^2$ constraints (42) to implement the utility bounds, and $(\#E-1)$ truth-telling constraints (43). Notice that the number of constraints does not increase exponentially in any of the grid sizes. The number of constraints is approximately quadratic in $\#A$ and approximately linear in all other grid sizes. This makes it possible to compute models with a large number of actions. In Program 3, our example with two states e, ten transfers τ, two actions a, and ten utility vectors $\mathbf{w}'$ is a linear program with 220 variables and 63 constraints, even smaller than Program 2. For the first time, the program is still computable if we increase the number of actions a to ten. In that case, Program 3 has 1100 variables and 1903 constraints, which is still sufficiently small to be solved on a personal computer.

We now want to show that Program 3 is equivalent to Program 1. In both programs, the planner chooses lotteries over transfer, action, and promised utilities. Even though in

Program 3 the planner also chooses utility bounds, in both programs the planner's utility depends only on the lotteries, and not on the bounds. The objective functions are identical. In order to demonstrate that the two programs are equivalent, it is therefore sufficient to show that the set of feasible lotteries is identical. We therefore have to compare the set of constraints in the two programs.

Proposition 6 *Program 1 and Program 3 are equivalent.*

The proof (contained in the Appendix) consists of showing that constraints (39)-(43) in Program 3 place the same restrictions on the outcome function $\pi(\cdot)$ as the constraints (21)-(24) of Program 1. The probability constraints (21) and (39), the promise-keeping constraints (22) and (40), and the obedience constraints (23) and (41) are in fact identical. Therefore one only needs to show that for any $\pi(\cdot)$ that satisfies the incentive constraints (24) in Program 1, one can find utility bounds such that the same outcome function satisfies (42) and (43) in Program 3 (and vice versa). Since the objective function is identical, it then follows that the programs are equivalent.

In Program 3, the number of constraints gets large if both the grids for transfer τ and action a are made very fine. In practice, this may lead to memory problems when computing. Further reductions in the number of constraints are possible in a formulation in which *the transfer and the recommendation are assigned at two different stages.* In other words, we are subdividing the period into two parts as in Program 2. However, instead of having a second report of the endowment, we subdivide the period completely and assign interim utility promises as an additional state variable. This procedure yields even smaller programs. However, the reduction comes at the expense of an increase in the number of programs that needs to be computed. The programs for both stages are described in Appendix A.3.

6 Computed Examples

6.1 Storage with Fixed Return

In this section, we present numerical examples in which the hidden action is designed to resemble storage at a fixed return. In all example environments the agent has logarithmic utility. There are two possible endowments e (2.1 or 3.1), three storage levels a (from 0 to 0.8), and thirty transfers τ (from -1.2 to 1) in every period. Planner and agent have the same discount factor $\beta = Q = .952$. The discount factor of the planner corresponds to a risk-free interest rate of 5%. The utility grids have fifty elements. In the initial period, each endowment occurs with probability .5. In all other periods, the probability distribution over endowments depends on the investment level in the previous period. If the investment level is zero, the high endowment occurs with probability .5. The probability of the high endowment increases linearly with the investment level. The expected return on storage R determines how fast the probability of the high endowment increases. For example, for $R = 1$ an increase in storage of 0.1 results in an increase in the expected endowment of 0.1. Since the difference between the two endowments equals one ($3.1 - 2.1$), with $R = 1$ increasing the storage level a by 0.1 results in an increase in the probability of the high endowment of 3.1 by 0.1, and an equal decrease in the probability of the low endowment of 2.1. We used Program 3 for all computations.

We computed optimal policy functions for a variety of returns R *in an infinite-period setting* as a function of the (scalar) initial utility promise W_0. The policy functions are computed by iterating on the Bellman operator until convergence, and then solving the initial planning problem (25). As described in Appendix A.2, we start by choosing a large initial guess for the value set $\mathbf{W}$ (a set consisting of vectors of utility promises), and then iterate backwards in time, using the initial guess $\mathbf{W}_0$ as the set of possible utility promises in the first iteration. The initial value set is based on the bounds for consumption implied by the grids for endowments, transfers, and actions. To reduce the number of feasible utility promises and facilitate computation, we also impose some additional restrictions on $\mathbf{W}_0$. For example, it follows from the incentive constraints that an agent can never receive lower utility by realizing the high endowment instead of the low endowment. Therefore, we disregard all initial utility vectors which promise higher utility when the low endowment is realized.

In all computed solutions to the initial planning problem (25), the surplus of the planner $V(W_0)$ is a strictly decreasing function of the utility promise W_0, implying that the promise-keeping constraint (27) is always binding. This also means that in future periods it is not in the interest of both the agent and the principal to tear up the contract and start over in such a way as to make both better off. However, either party may be worse off in expected utility than they were at the initial date.

Figure 2 shows the policy functions for $R = 1$ as a function of the initial utility promise

W_0. It turns out that the choices of the planner are similar for all values of R that are lower than the associated risk-free interest rate. Therefore we use Figure 2 at $R = 1$ to represent general features of the optimal incentive-compatible contract for $R < 1.05$. The choices of the planner are shown as a function of the initial utility promise to the agent W_0. In all graphs a vertical line marks the initial utility level at which the expected discounted surplus of the planner is zero. As before, if there is a continuum of agents, the planner can be interpreted as a mere programming device to determine the optimal allocation for the community, and the zero-surplus utility level corresponds to the constrained Pareto-optimal allocation for a continuum of agents.

As long as $R < 1.05$ (the market rate), the planner never recommends a positive investment or storage level; only the outside credit market is used. The planner uses transfers and utility promises to provide incentive-compatible insurance every period. The transfer for the low endowment is higher than for the high endowment, and at zero expected surplus high-income agents today pay premia, while low-income agents receive indemnities. The difference in consumption between low- and high-endowment agents is only about .2 over most of the range of W_0, even though the endowment differs by 1. In order to ensure incentive compatibility (i.e., induce the high-endowment agent to report truthfully), the planner promises a schedule of higher future utility from tomorrow on to agents with the high endowment. Note that separate utilities are promised for each of the two possible incomes tomorrow. Today, naturally, both consumption and utility promises increase with initial promised utility. The dynamics are implicit in the utility promise transitions. For example, agents with high income today are pushed towards higher expected utility tomorrow.

Figure 3 shows the dynamics of consumption for a number of agents who start out with the same utility promise W_0 (corresponding to zero expected surplus for the planner) and realize the same initial endowment.[11] Consequently, in the first period all agents have the same consumption. In the second period the consumers split in two groups, with the ones who receive the high income enjoying higher consumption as well. The distribution of consumption disperses over time as each consumer is hit by a different series of shocks. However, since there are just two income levels and the transfers are bounded, consumption always moves within a finite interval.

Indeed, we can use the optimal policy functions to compute the long-run distribution of

[11] We plot outcomes for eight different agents, thus the graph does not represent all possible consumption paths. With a higher number of agents, the paths start to fill the space and become hard to distinguish.

consumption in an economy with a continuum of agents.[12] Figure 4 shows that the long-run distribution of consumption is twin-peaked. Consumption disperses over time, and after a series of good or bad shocks either corner for consumption can be reached. The corners are not absorbing states, however. For example, agents who consume minimum consumption are thrown back into the distribution once a positive shock hits. Our interpretation for the twin-peaks distribution is that the lower peak is accounted for by agents who were deflected away from the lower bound for consumption.

Figure 5 shows how the expected utility of the agent varies with the return on storage, subject to the requirement that the discounted surplus of the planner be zero.[13] The utility of the agent is shown for values of R between 0 and 1.10. The top line is the utility the agent gets under full information, that is, when both endowment and storage are observable by the planner. In this case the planner provides full insurance, and consumption and utility promises are independent of the endowment. It turns out that the utility of the agent does not vary with the return on storage as long as $R \leq 1.05$. This is not surprising, since 1.05 is the risk-free interest rate. As long as the return on storage does not exceed the credit-market return, raising the return on storage does not extend the unconstrained Pareto frontier. Storage is never used, and the utility of the agent does not depend on the return on storage. When R exceeds 1.05, the Pareto frontier expands and the utility of the agent increases dramatically. In effect, when $R > 1.05$ an arbitrage opportunity arises which is only limited by the finite grids.

The lower line shows the utility of the agent under autarky. When the return on storage is sufficiently high the agent uses storage to self-insure against the income shocks. Since under autarky storage is the only insurance for the agent, utility increases with the return on storage.

The middle line shows the utility of the agent with hidden endowments and hidden action. Once the return on storage exceeds some critical value, the utility of the agent decreases with the return on storage, instead of increasing as it does under autarky. As long as the planner never recommends positive private storage levels, a higher return on

[12] Atkeson and Lucas, Jr. (1992) found that in a class of dynamic incentive problems the efficient allocation has the property that the distribution of consumption diverges over time, i.e., a diminishing fraction of the population receiving an increasing fraction of resources. Phelan (1998) examines the long-run implications of incentive problems more generally. He finds that the limiting distribution for consumption and utility crucially depends on features of the utility function, namely whether marginal utility is bounded away from zero as consumption approaches infinity. In our setup, the set of possible consumption and utility values is bounded, since finite grids are used.

[13] To speed up computations, we computed a three-period model for this graph.

storage has no positive effects. On the contrary, with a high return on storage it becomes harder to satisfy the obedience constraints which require that the agent does not prefer a positive storage level when the planner recommends zero storage. Therefore raising the return on storage shrinks the set of feasible allocations, and the utility of the agent falls. Once R exceeds the credit-market return by a sufficient margin, however, private storage is used in the constrained solution, and consequently further increases in R raise utility. Unlike under full information, $R > 1.05$ does not imply an arbitrage opportunity, since insurance is constrained by incentives.

The optimal allocations in this application are computed subject to two types of incentive constraints: obedience constraints force the agent to take the right action, and truth-telling constraints ensure that the agent reports the actual endowment. What happens if only one type of constraint is introduced? If we remove the truth-telling constraints (i.e., we assume that the actual endowment can be observed by the planner) and assume that the return on storage does not exceed the credit-market return, the situation is simple: The planner can provide full insurance against income shocks. In this case the obedience constraints are not binding, since the optimal storage level is zero, which is also what the agent prefers if income shocks are fully insured. Conversely, if only the obedience constraints are removed, as if storage could be observed, full insurance cannot be achieved. Since the planner will require zero private storage, the outcome is the same as in the situation where the storage technology has zero gross return, as in Figure 5 for $R = 0$. Without the obedience constraints, more insurance is possible, but it still falls well short of full insurance.

We have not computed a version of Phelan and Townsend (1991) or Atkeson and Lucas, Jr. (1992) with unobserved storage as these models are not equivalent to the one analyzed here. We conjecture however that the welfare consequences of Figure 5 would carry over in the sense that there would be a welfare loss but still nontrivial insurance relative to autarky.

6.2 Risky Banking and the Optimality of Public Reserves

In this section, we consider an environment in which the return on the hidden action, interpreted as investment, varies with the level of investment. The agents, or investors, will be interpreted as banks. For simplicity, we assume that banks (the agents) enter the first period with an observed initial endowment or net worth (a special version of the general setup), and that the world lasts for three periods.[14] The initial endowment can be withdrawn from the bank as profit and consumed by the owner, or it can be invested in varying quantities of loans. Investment and consumption are unobserved. The return on investment, or the income in the second period, is random, and can be observed only by the bank itself, so the returns to investment are unobservable to outsiders. In the second period the bank can invest once again, and in the third period the final wealth is consumed.

Bankers are assumed to be risk-averse, so that other things being equal, they would like to smooth out income fluctuations and insure against the uncertain returns of their investments. If the bankers are on their own, their only mechanism for smoothing income between the two periods is to invest in risky loans. They have no possibility of insuring against the uncertain returns on their investments. But in addition to the banks, there is also a planner or insurance fund who has access to a risk-free storage technology with zero net yield. We solve the problem of providing optimal incentive-compatible insurance to the banks, subject to the requirement that the planner receives a zero discounted surplus, thus the fund breaks even. The planner could be identified with the central bank. An alternative assumption would be that there is a continuum of banks who face idiosyncratic, but not aggregate, risk. In that case, the outcome of the planning problem can be interpreted as a mutual insurance scheme run by the banks themselves. The results do not depend on specific institutional assumptions; that is, the optimal contract can be implemented in different ways.

The known endowment in the first period is 1.135. In the second period, the possible income levels are 0.9 and 1.5, and in the third period income can be 0.7 or 1.3. The difference is 0.6 in each instance. The incomes are chosen such that with investment at the efficient level expected income is constant across the periods. In addition to zero investment, three levels of private investment (.1, .2, and .3) are possible. The expected incremental marginal returns for investing at level .1, .2, and .3 are given by 1.31, 1.01, and .9. Reflecting these returns, Table 1 shows the probabilities over the two income levels, depending on investment.

The first column corresponds to zero investment, the second column to an investment of

[14]Even with the methods described in this paper, computing this model is not a trivial task. The main difficulty in the banking example is that public and private investment are close substitutes, so that inaccuracies which arise from using grids for transfers and utility promises can lead to jumps in the policy functions for private and public investment. When the two forms of investment are close substitutes, fine grids for transfers and utility promises need to be chosen, and computation time is high.

	Investment Level			
	0	0.1	0.2	0.3
p(Low Income)	.9950	.7767	.6083	.4583
p(High Income)	.0050	.2233	.3917	.5417

Table 1: Probabilities Over Income Given Investment

.1, and so on. With zero investment, the probability of getting the low income level is .995, and the probability for the high income is .005. Here, in this example, the income variance increases with investment, i.e., higher expected returns are naturally associated with more risk. The grids for the transfers in period one and two have 30 elements, and there are 70 grid points for the transfer in the third period. The grid for the promised utility vector which is assigned in period one has 3600 elements, and the grid for the utility assignment in the second period has 3600 elements as well.[15] Storage by the planner has zero net return (i.e., the discount factor of the planner equals one, $Q = 1$). The discount factor β of the bankers equals one as well. This assumption is made for simplicity. The qualitative results do not change if the discount factor of the bankers is less than one. Notice that the return on private investments exceeds the unit return on public storage up to an investment level of .2, but is lower after that. If the aim were to maximize the expected endowment in the second period as in a neoclassical world, we therefore would expect the banks to invest .2.

Figure 6 shows the outcome under full information (income and investment level are observed in each period). The outcome is computed using the same grids on transfers and utility promises that were used for the constrained solution. The policy functions are shown as a function of the initial utility promise W_0 to the agent, and the vertical line in all graphs marks the initial utility promise at which the expected surplus of the planner is exactly zero. Since full insurance is possible, investment is at the first-best level of .2 throughout (see the graph in the upper right-hand corner). Through income-contingent transfers, the planner provides full insurance, i.e., in periods 2 and 3 the difference between the transfer for low- and high-income bankers equals their difference in income in that period. At zero expected surplus for the planner, the average transfer turns out to be zero in every period. In other words, there are no public reserves created at the initial date.

Figure 7 shows the results under private information (income and investment are unob-

[15]However, the program dismisses utility vectors which turn out to infeasible, so the dimension of the linear programs which are actually computed is smaller than the one implied by the grid sizes.

served). Here the usual tradeoff between insurance and incentives arises: If the insurance fund provides full insurance, banks will prefer not to invest, and will report low income in every period. Both the level of investment and the return on investment are unobserved by the planner, and the planner has to provide incentives for obedience (investing properly) and truthful reporting. Under the optimal contract with private information, investment by the bank falls to .1 throughout, below the first-best level of .2. Note that private investment is still positive and has a high marginal rate of return well above the outside interest rate. The transfers in the second and third period and the promised utilities for the second period are contingent on income in the second period. The utility promises for the third period are contingent on income in periods 2 and 3 ('LH' means low income in 2 and high income in 3, and so on). At zero expected surplus, the transfer is negative in the first period, which implies that the community is creating a safe investment fund. In the second period, the average transfer (i.e., averaging over successful and unsuccessful banks) is close to zero, meaning that the fund is maintained. In the third period the average transfer is positive, so the community is returning the safe investment from the fund to the bankers. In essence, the planner now uses public storage to smooth the income of the banks. Note that the planner can provide a substantial amount of insurance against the income shock in the second period. The difference between the transfer for bankers with high and low income (solid and dotted line) is about .3, or half of the income difference. But even though a substantial amount of insurance is provided, it is optimal to replace high-yield private investment with low-yield, low-risk public reserves. The income shock in the third period cannot be insured with shock-contingent transfers, since the absence of another period makes it impossible for the planner to trade off transfers against future promised utilities.

The results can be reversed if we assume, counterintuitively, that investment lowers risk, i.e., the variance of output decreases with the level of investment. This can be achieved by modifying the probabilities over endowments in Table 1 appropriately. In such an environment, in the private-information solution the planner recommends to invest more than the first-best level. Returns on investment are driven down due to over-investment, but the agents are willing to accept lower returns to reduce risk. In order to finance the additional investment, the planner borrows from the outside to finance transfers to the agents in the early periods. The transfers are reversed in the later periods, when the planner has to pay off the outside lenders. Thus, the relationship between investment and risk is central for explaining our results.

In summary, we find that public reserves with a low return can be optimal if banks face a moral-hazard problem. The key underlying assumption is that increased lending or investment increases risk, as seems natural. Using public reserves lowers overall risk and is therefore optimal in the information-constrained solution, even though public storage would be inefficient in the first-best solution.

7 Conclusions

In this paper we show how a general dynamic mechanism design problem with hidden income and hidden actions can be formulated in a way that is suitable for computation. We start from a general planning problem which allows arbitrary message spaces, lotteries, and history dependence. The planning problem is reduced to a recursive version which imposes truth-telling and obedience, and uses vectors of utility promises as state variables. We also develop methods to reduce the number of constraints that need to be imposed when computing the optimal contract. The main theme of these methods is to allow the planner to specify utility and behavior off the equilibrium path. In Program 2 this occurs because the agent reports his endowment more than once, so that the planner has to specify what happens if there are conflicting reports. In Program 3 the planner chooses bounds that limit the utility the agent can get on certain branches off the equilibrium path. By applying these methods, we can solve problems on a standard PC that otherwise would be impossible to compute on any computer.

In an application to a model with hidden storage, we use our methods to examine how the optimal insurance contract varies with the return of the private storage technology. Specifically, we observe that in certain cases the utility of the agent can actually decrease as the return of the private technology rises. This occurs when the return on public storage is sufficiently high so that in the constrained optimum only public storage is used. If now the return to the private technology rises, it becomes more difficult for the planner to provide incentives for truth-telling and obedience, and consequently less insurance is possible. Thus the effect of a rising return of private investment on the agent's utility in the constrained optimum is exactly the opposite of the effect that prevails under autarky, where a higher return raises utility.

In another illustrative application, we solve for the optimal credit-insurance scheme in a model where banks face uncertain and privately observed returns on their investments and are subject to moral hazard. We find that in the constrained optimum low-yielding public reserves and high-yielding private investments can coexist, even though public reserves would not be used in the first-best solution. To this end, our model provides a new rationale for bank reserves that relies on moral hazard only, without reference to traditional justifications for reserves such as the need for liquidity.

While in this paper we have presented our recursive formulations with double reporting and utility bounds as tools to speed up computations, we also believe that the results are suggestive for the design of actual games and mechanisms. In dealing with the government or private businesses, we often find ourselves reporting the same information over and over again. While at first sight this may appear as unnecessary and inefficient, our results suggest that multiple reporting may be a useful feature of real-life mechanisms, because implementation is simplified. We plan to explore this perspective on the characteristics of actual mechanisms in future research.

A Mathematical Appendix

A.1 Proofs for all Propositions

Proposition 1 *There are reservation utilities $W_0 \in R$ such that an optimal equilibrium exists.*

Proof of Proposition 1 We need to show that for some W_0 the set of equilibria is nonempty and compact, and the objective function is continuous. To see that the set of equilibria is nonempty, notice that the planner can assign a zero transfer in all periods, and always send the same message. If the strategy of the agent is to choose the actions that are optimal under autarky, clearly all constraints are trivially satisfied for the corresponding initial utility W_0. The set of all contracts that satisfy the probability-measure constraints is compact in the product topology, since π and σ are probability measures on finite support. Since only equalities and weak inequalities are involved, it can be shown that the constraints (3), (4), and (5) define a closed subset of this set. Since closed subsets of compact sets are compact, the set of all feasible contracts is compact. We still need to show that the objective function of the planner is continuous. Notice that the product topology corresponds to pointwise convergence, i.e., we need to show that for a sequence of contracts that converges pointwise, the surplus of planner converges. This is easy to show since we assume that the discount factor of the planner is smaller than one, and that the set of transfers is bounded. Let π_n, σ_n, be a sequence of contracts that converges pointwise to π, σ, and choose $\epsilon > 0$. We have to show that there is an N such that $|V(\pi_n, \sigma_n) - V(\pi, \sigma)| < \epsilon$. Since the transfer τ is bounded and $Q < 1$, there is an T such that the discounted surplus of the planner from time T on is smaller than $\epsilon/2$. Thus we only have to make the difference for the first T periods smaller than $\epsilon/2$, which is the usual Euclidian finite-dimensional case. □

Proposition 2 (Revelation Principle) *For any message spaces M_1, M_2, and M_3, any allocation that is feasible in the general mechanism is also feasible in the truth-telling mechanism.*

Proof of Proposition 2 (Outline) Corresponding to any feasible allocation in the general setup there is a feasible contract that implements this allocation. Fix a feasible allocation and the corresponding contract $\{\pi, \sigma\}$. We will now define an outcome function for the truth-telling mechanism that implements the same allocation. To complete the proof, we then have to show that this outcome function satisfies constraints (7) to (9).

We will define the outcome function such that the allocation is the one implemented by (π, σ) along the equilibrium path. To do that, let $H^t(s^t)$ be the set of histories h^t in the general game such that the sequence of endowments, transfers, and actions in h^t coincides with the sequence of reported endowments, transfers, and recommended actions in history s^t in the restricted game. Likewise, define $p(h^t|s^t)$ as the probability of history h^t conditional on s^t:

$$p(h^t|s^t) \equiv \frac{p(h^t)}{\sum_{H^t(s^t)} p(h^t)} \tag{44}$$

If s^t has zero probability (that is, if the sequence s^t of endowments, transfers, and actions occurs with probability zero in the allocation implemented by $\{\pi, \sigma\}$), the definition of $p(h^t|s^t)$ is irrelevant, and is therefore left unspecified. Now we define an outcome function for the truth-telling mechanism by:

$$\pi(\tau_t, a_t|e_t, s^{t-1}) \equiv \sum_{H^{t-1}(s^{t-1}), M_1, M_2, M_3} p(h^{t-1}|s^{t-1})\, \sigma^\star(m_{1t}|e_t, h^{t-1})\, \pi^\star(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$

$$\sigma^\star(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1})\, \pi^\star(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\, \sigma^\star(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}). \tag{45}$$

Basically, the outcome function is gained by integrating out the message spaces M_1, M_2, and M_3 and prescribing the outcomes that occur on the equilibrium path.

We now have to verify that with this choice of an outcome function conditions (7) to (9) above are satisfied. In showing this, we can make use of the fact that $\{\pi, \sigma\}$ are probability measures and satisfy (3), (4), and (5). The proof proceeds by substituting (45) into (3), (4), and (5), and showing that the resulting equations imply (7) to (9). This is carried out in detail in Doepke and Townsend (2002). □

Proposition 3 *The set* $\mathbf{W}$ *is nonempty and compact.*

Proof of Proposition 3 To see that $\mathbf{W}$ is nonempty, notice that the planner can always assign a zero transfer in every period, and recommend the optimal action that the agent would have chosen without the planner. For the $w_0(e)$ that equals the expected utility of the agent under autarky under state e, all constraints are satisfied. To see that $\mathbf{W}$ is bounded, notice that there are finite grids for the endowment, the transfer, and the action. This implies that in every period consumption and therefore utility is bounded from above and from below. Since the discount factor β is smaller than one, total expected utility is also bounded. Since each $w_0(e)$ has to satisfy a promise-keeping constraint with equality, the set $\mathbf{W}$ must be bounded. Finally, we can show that $\mathbf{W}$ is closed by a contradiction argument. Assume that $\mathbf{W}$ is not closed. Then there exists a converging sequence $\mathbf{w}_n$ such that each element of the sequence is in $\mathbf{W}$, but its limit $\mathbf{w}$ is not. Corresponding to each $\mathbf{w}_n$ there is a contract $\pi(\tau_t, a_t|e_t, s^{t-1})_n$ satisfying constraints (7), (8), and (9). Since the contracts are within a compact subset of R^∞ with respect to the product topology, there is a convergent subsequence with limit $\pi(\tau_t, a_t|e_t, s^{t-1})$. It then follows that $\mathbf{w}$ must satisfy (7), (8), and (9) when $\pi(\tau_t, a_t|e_t, s^{t-1})$ is the chosen contract. □

Proposition 4 *For all $\mathbf{w}_0 \in \mathbf{W}$ and $e_0 \in E$, and for any s^{k-1} and e_k, there is an optimal contract $\pi^\star$ such that the remaining contract from s^{k-1} and e_k is an optimal contract for the auxiliary planning problem with $e_0 = e_k$ and $\mathbf{w}_0 = \mathbf{w}(s^{k-1}, \pi^\star)$.*

Proof of Proposition 4 We will first construct $\pi^\star$ from a contract which is optimal from time zero and another contract which is optimal starting at s^{k-1} and e_k. We will then show by a contradiction argument that $\pi^\star$ is an optimal contract from time zero as well.

We have shown earlier that an optimal contract exists. Let π be an optimal contract from time zero, and π_k an optimal contract for $e_0 = e_k$ and $\mathbf{w}_0 = \mathbf{w}(s^{k-1}, \pi)$, with the elements of vector $\mathbf{w}(s^{k-1}, \pi)$ defined in (13). Now construct a new contract $\pi^\star$ that is equal to π_k from (e_k, s^{k-1}) on, and equals π until time k and on all future branches other than e_k, s^{k-1}. First, notice that by the way $\pi^\star$ is constructed, we have $\mathbf{w}(s^{k-1}, \pi) = \mathbf{w}(s^{k-1}, \pi^\star)$, and since $\pi^\star$ equals π_k from (e_k, s^{k-1}), $\pi^\star$ fulfills the reoptimization requirement of the proposition. We now claim that $\pi^\star$ is also an optimal contract. To show this, we have to demonstrate that $\pi^\star$ satisfies constraints (7), (8), and (9), and that it maximizes the surplus of the planner subject to these constraints. To start, notice that the constraints that are imposed if we compute an optimal contract taking $e_0 = e_k$ and $\mathbf{w}_0 = \mathbf{w}(s^{k-1}, \pi)$ as the starting point also constrain the choices of the planner in the original program from (e_k, s^{k-1}) on. By reoptimizing at (e_k, s^{k-1}) in period k as if the game were restarted, the planner clearly cannot lower his surplus, since no additional constraints are imposed. Therefore the total surplus from contract $\pi^\star$ cannot be lower than the surplus from π. Since π is assumed to be an optimal contract, if $\pi^\star$ satisfies (7), (8), and (9), it must be optimal as well. Thus we only have to show that (7), (8), and (9) are satisfied, or in other words, that reoptimizing at e_k, s^{k-1} does not violate any constraints of the original problem.

The probability constraints (7) are satisfied by contract $\pi^\star$, since the reoptimized contract is subject to the same probability constraints as the original contract. The promise-keeping constraint (8) is satisfied since the new contract delivers the same on-path utilities by construction. We still have to show that the incentive constraints (9) are satisfied. We will do this by contradiction. Suppose that (9) is not satisfied by contract $\pi^\star$. Then there is a deviation δ such that for some s^l, e_{l+1}:

$$\sum_{t=l+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] > \sum_{t=l+1}^{\infty} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right]. \quad (46)$$

Consider first the case $l + 1 \geq k$. On any branch that starts at or after time k, contract $\pi^\star$ is entirely identical to either π or π_k. But then (46) implies that either π or π_k violates incentive-compatibility (9), a contradiction. Consider now the case $l + 1 < k$. Here the contradiction is not immediate, since the remaining contract is a mixture of π and π_k. Using $w(e_k, s^{k-1}, \delta)$ to denote the continuation utility of the agent from time k on under the deviation strategy, we can rewrite (46) as:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, \delta, s^l, e_{l+1}) w(e_k, s^{k-1}, \delta)$$
$$> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, s^l, e_{l+1}) w(e_k, s^{k-1}). \quad (47)$$

Notice that for s^{k-1} that are reached with positive probability under the deviation we have:

$$w(e_k, s^{k-1}, \delta) \leq w(e_k, \delta(s^{k-1})), \quad (48)$$

where $\delta(s^k)$ is the history as seen by the planner (reported endowments, delivered transfers, and recommended actions) under the deviation strategy. Otherwise, either π or π_k would violate incentive constraints. To see why, assume that e_k, s^{k-1} is a branch after which $\pi^\star$ is identical to π. If we had $w(e_k, s^{k-1}, \delta) > w(e_k, \delta(s^{k-1}))$, an agent under contract π who reached history $\delta(s^{k-1})$ could gain by following the deviation strategy δ afterwards. This cannot be the case since π is assumed to be an optimal contract, and therefore deviations are never profitable. Using (48) in (47) gives:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t|\pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi^\star, \delta, s^l, e_{l+1}) w(e_k, \delta(s^{k-1}))$$
$$> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t|\pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi^\star, s^l, e_{l+1}) w(e_k, s^{k-1}). \tag{49}$$

The outcome function $\pi^\star$ enters (49) only up to time $k-1$. Since up to time $k-1$ the outcome function $\pi^\star$ is identical to π, and since by construction of $\pi^\star$ continuation utilities at time k are the same under $\pi^\star$ and π, we can rewrite (49) as:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi, \delta, s^l, e_{l+1}) w(e_k, \delta(s^{k-1}))$$
$$> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi, s^l, e_{l+1}) w(e_k, s^{k-1}). \tag{50}$$

But now the left-hand side of (50) is the utility that the agent gets under plan π from following the deviation strategy until time k, and following the recommendations of the planner afterwards. Thus (50) contradicts the incentive compatibility of π. We obtain a contradiction, $\pi^\star$ actually satisfies (9). This shows that plan $\pi^\star$ is within the constraints of the original problem. Since $\pi^\star$ yields at least as much surplus as π and π is an optimal contract, $\pi^\star$ must be optimal as well. □

Proposition 5 *Program 1 and Program 2 are equivalent.*

Proof of Proposition 5 (This proof is carried out in Doepke and Townsend (2002). We derive Program 2 from the general planning problem, taking the same steps as in the derivation of Program 1. Since both Program 1 and Program 2 are equivalent to the general planning problem, they are also equivalent to each other.)

Proposition 6 *Program 1 and Program 3 are equivalent.*

Proof of Proposition 6 We want to show that constraints (39)-(43) in Program 3 place the same restrictions on the outcome function $\pi(\cdot)$ as the constraints (21)-(24) of Program 1. The probability constraints (21) and (39), the promise-keeping constraints (22) and (40), and the obedience constraints (23) and (41) are identical. This leaves us with the truth-telling constraints. Let us first assume we have found a lottery $\pi(\tau, a, \mathbf{w}'|\mathbf{w}, e)$ that satisfies the truth telling constraint (24) of Program 1 for all $\hat{e}$ and $\delta : T \times S \to A$. We have to show that there exist utility bounds $v(\hat{e}, e, \tau, a)$ such that the same lottery satisfies (42) and (43) in Program 3. For each $\hat{e}$, τ, and a, define $v(\hat{e}, e, \tau, a)$ as the maximum of the left hand side of (42) over all $\hat{a}$:

$$v(\hat{e}, e, \tau, a) \equiv \max_{\hat{a}} \left\{ \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e') \right] \right\}. \tag{51}$$

Then clearly (42) is satisfied, since the left-hand side of (42) runs over $\hat{a}$. Now for each τ and a, define $\hat{\delta}(\cdot)$ by setting $\hat{\delta}(\tau, a)$ equal to the $\hat{a}$ that maximizes the left-hand side of (42):

$$\hat{\delta}(\tau, a) \equiv \text{argmax}_{\hat{a}} \left\{ \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e') \right] \right\}. \tag{52}$$

Since $\pi(\tau, a, \mathbf{w}'|\mathbf{w}, e)$ satisfies (24) for any function $\delta(\cdot)$ by assumption, we have for our particular $\hat{\delta}(\cdot)$:

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e} + \tau - \hat{\delta}(\tau, a)) + \beta \sum\nolimits_E p(e'|\hat{\delta}(\tau, a)) w'(e') \right] \leq w(\hat{e}). \tag{53}$$

By the way we chose $\hat{\delta}(\cdot)$ and the $v(\cdot)$, we have from (42):

$$\sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{\delta}(\tau, a)) + \beta \sum\nolimits_E p(e'|\hat{\delta}(\tau, a)) w'(e') \right] = v(\hat{e}, e, \tau, a). \tag{54}$$

Substituting the left-hand side into (53), we get:

$$\sum_{T,A} v(\hat{e}, e, \tau, a) \leq w(\hat{e}). \tag{55}$$

which is (43).

Conversely, suppose we have found a lottery $\pi(\tau, a, \mathbf{w}'|\mathbf{w}, e)$ that satisfies (42) and (43) in Program 3 for some choice of $v(\hat{e}, e, \tau, a)$. By (42), we have then for any $\hat{e}$ and $\hat{a}$ and hence any $\delta : T \times S \to A$:

$$\sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e} + \tau - \delta(\tau, a)) + \beta \sum\nolimits_E p(e'|\delta(\tau, a))w'(e')\right] \leq v(\hat{e}, e, \tau, a). \tag{56}$$

Substituting the left-hand side of (56) into the assumed (43) for the $v(\hat{e}, e, \tau, a)$, we maintain the inequality:

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e} + \tau - \delta(\tau, a)) + \beta \sum\nolimits_E p(e'|\delta(\tau, a))w'(e')\right] \leq w(\hat{e}). \tag{57}$$

But this is (24) in Program 1. Therefore the sets of constraints are equivalent, which proves that Program 1 and Program 3 are equivalent. □

A.2 Computing the Value Set

Our analysis was based on the assumption that the set $\mathbf{W}$ of feasible utility vectors is known in advance. In practice, $\mathbf{W}$ is not known and needs to be computed alongside the value function $V(\mathbf{w}, e)$. $\mathbf{W}$ can be computed with the dynamic-programming methods described in detail in Abreu, Pierce, and Stachetti (1990), henceforth APS. An outline of the method follows.

We start by defining an operator B that maps nonempty compact subsets of $\mathbf{R}^{\#E}$ into nonempty compact subsets of $\mathbf{R}^{\#E}$. Let $\mathbf{W}'$ be a nonempty compact subset of $\mathbf{R}^{\#E}$. Then $B(\mathbf{W}')$ is defined as follows:

Definition 6 *A utility vector $\mathbf{w} \in B(\mathbf{W}')$ if there exist probabilities $\pi(\tau, a|\mathbf{w}, e)$ and future utilities $w'(\mathbf{w}, e, \tau, a) \in \mathbf{W}'$ such that* (15) *to* (18) *hold.*

The key point is that utility promises are chosen from the set $\mathbf{W}'$ instead of the true value set $\mathbf{W}$. Intuitively, $B(\mathbf{W}')$ consists of all utility vectors $\mathbf{w}$ that are feasible today (observing all incentive constraints), given that utility vectors from tomorrow on are drawn from the set $\mathbf{W}'$. The fact that B maps compact set into compact sets follows from the fact that all constraints are linear and involve only weak inequalities. Clearly, the true set of feasible utility vectors $\mathbf{W}$ satisfies $\mathbf{W} = B(\mathbf{W})$, thus $\mathbf{W}$ is a fixed point of B. The computational approach described in APS consists of using B to define a shrinking sequence of sets that converges to $\mathbf{W}$.

To do this, we need to start with a set $\mathbf{W}_0$ that is known to be larger than $\mathbf{W}$ a priori. In our case, this is easy to do, since consumption is bounded and therefore lifetime utility is bounded above and below. We can choose $\mathbf{W}_0$ as an interval in $\mathbf{R}^{\#E}$ from a lower bound that is lower than the utility from receiving the lowest consumption forever to a number that exceeds utility from consuming the highest consumption forever. We can now define a sequence of sets $\mathbf{W}_n$ by defining $\mathbf{W}_{n+1}$ as $\mathbf{W}_{n+1} = B(\mathbf{W}_n)$. We have the following results:

Proposition 7

- *The sequence $\mathbf{W}_n$ is shrinking, i.e., for any n, $\mathbf{W}_{n+1}$ is a subset of $\mathbf{W}_n$.*
- *For all n, $\mathbf{W}$ is a subset of $\mathbf{W}_n$.*
- *The sequence $\mathbf{W}_n$ converges to a limit $\bar{\mathbf{W}}$, and $\mathbf{W}$ is a subset of $\bar{\mathbf{W}}$.*

Proof of Proposition 7 To see that $\mathbf{W}_n$ is shrinking, we only need to show that $\mathbf{W}_1$ is a subset of $\mathbf{W}_0$. Since $\mathbf{W}_0$ is an interval, it suffices to show that the upper bound of $\mathbf{W}_1$ is lower than the upper bound of $\mathbf{W}_0$, and that the lower bound of $\mathbf{W}_1$ is higher than the lower bound of $\mathbf{W}_0$. The upper bound of $\mathbf{W}_1$ is reached by assigning maximum consumption in the first period and the maximum utility vector in $\mathbf{W}_0$ from the second period on. But the maximum utility vector $\mathbf{W}_0$ by construction corresponds to consuming more than maximum consumption every period, and since utility is discounted, the highest utility vector in $\mathbf{W}_1$ therefore is smaller than the highest utility vector in $\mathbf{W}_0$.

To see that $\mathbf{W}$ is a subset of all $\mathbf{W}_n$, notice that by the definition of B, if $\mathbf{C}$ is a subset of $\mathbf{D}$, $B(\mathbf{C})$ is a subset of $B(\mathbf{D})$. Since $\mathbf{W}$ is a subset of $\mathbf{W}_0$ and $\mathbf{W} = B(\mathbf{W})$, we have that $\mathbf{W}$ is a subset of $\mathbf{W}_1 = B(\mathbf{W}_0)$, and correspondingly for all the other elements.

Finally, $\mathbf{W}_n$ has to converge to a nonempty limit since it is a decreasing sequence of compact sets, and the nonempty set $\mathbf{W}$ is a subset of all elements of the sequence. □

Up to this point, we know that $\mathbf{W}_n$ converges to $\bar{\mathbf{W}}$ and that $\mathbf{W}$ is a subset of $\bar{\mathbf{W}}$. What we want to show is that $\bar{\mathbf{W}}$ and $\mathbf{W}$ are actually identical. What we still need to show, therefore, is that $\bar{\mathbf{W}}$ is also a subset of $\mathbf{W}$.

Proposition 8 *The limit set $\bar{\mathbf{W}}$ is a subset of the true value set $\mathbf{W}$.*

Proof of Proposition 8 The outline of the proof is as follows. To show that an element $\mathbf{w}$ of $\bar{\mathbf{W}}$ is in $\mathbf{W}$, we have to find $\pi(\tau_t, a_t|e_t, s^{t-1})$ that satisfy constraints (7), (9), and (11) for $\mathbf{w}$. These $\pi(\tau_t, a_t|e_t, s^{t-1})$ can be constructed period by period from the π that are implicit in the definition of the operator B. Notice that in each period continuation utilities are drawn from the same set $\bar{\mathbf{W}}$, since $\bar{\mathbf{W}}$ as the limit of the sequence $\mathbf{W}_n$ satisfies $\bar{\mathbf{W}} = B(\bar{\mathbf{W}})$. By definition of B, the resulting $\pi(\tau_t, a_t|e_t, s^{t-1})$ satisfy the period-by-period constraints (15) to (18). We therefore need to show that satisfying the period-by-period constraints (with a given set of continuation utilities) is equivalent to satisfying the original constraints (7), (9), and (11), which we have done above in Section 4.2. □

A.3 A Program With Two Subperiods

The period is divided into two parts. In the first subperiod the agent reports the endowment, and the planner assigns the transfer, an interim utility when the agent is telling the truth, as well as a vector of utility bounds in case the agent was lying. In the second subperiod the planner assigns an action and a vector of promised utilities for the next period. The solution to the problem in the second subperiod is computed for each combination of endowment e, transfer τ, interim utility $w_m(e)$ (m for "middle" or interim) along the truth-telling path, and vector of utility bounds for lying, $\bar{\mathbf{w}}_m(\hat{e}, e)$. Here $\bar{w}_m(\hat{e}, e)$ is an upper bound on the utility *an agent can get who has endowment $\hat{e}$, but reported e nevertheless.* $\bar{\mathbf{w}}_m(\hat{e}, e)$ is the vector of $\bar{w}_m(\hat{e}, e)$ with components running over endowments $\hat{e} \neq e$. The choice variables in the second subperiod are lotteries over the action a and the vector of promised utilities $\mathbf{w}'$ for the next period. We use $V_m[e, \tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e)]$ to denote the utility of the planner if the true state is e, the transfer is τ, and $w_m(e)$ and $\bar{\mathbf{w}}_m(\hat{e}, e)$ are the assigned interim utility and utility bounds for lying. The function $V_m(\cdot)$ is determined in the second subperiod (Program 4b below).

In the first subperiod the planner assigns transfers, interim utilities, and utility bounds. We use $\mathcal{W}(e, \tau)$ to denote the set of feasible utility assignments for a given state e and transfer τ. The agent comes into the first subperiod with a vector of promised utilities $\mathbf{w}$, to be effected depending on the realized state e. The planner assigns a transfer τ, on-path interim utilities $w_m(e)$, and off-path utility bounds $\bar{\mathbf{w}}_m(\hat{e}, e)$ subject to promise-keeping and truth-telling constraints.

Program 4a:
The maximization problem of the planner given reported endowment e and promised utility vector $\mathbf{w}$ is:

$$V(\mathbf{w}, e) = \max_{\pi \geq 0} \sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e) | \mathbf{w}, e) \; V_m[e, \tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e)] \tag{58}$$

subject to the constraints (59)-(61) below:

$$\sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e) | \mathbf{w}, e) = 1, \tag{59}$$

$$\sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e) | \mathbf{w}, e) \; w_m(e) = w(e), \tag{60}$$

$$\forall \hat{e} \neq e: \quad \sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e) | \mathbf{w}, e) \; \bar{w}_m(\hat{e}, e) \leq w(\hat{e}). \tag{61}$$

The number of variables in this program is $\sum_T \#\mathcal{W}(e, \tau)$, and there are $1+(\#E)$ constraints: one probability constraint (59), one promise-keeping constraint (60), and (#E-1) truth-telling constraints (61).

We now turn to the second subperiod. As in the last section, apart from the lotteries over actions and utilities, the program in the second subperiod also assigns utility bounds. $v(\hat{e}, e, \tau, a)$ is an upper bound on the utility an agent can get who has true endowment $\hat{e}$, reported endowment e, and receives transfer τ and recommendation a. These utility bounds are weighted by the probability of receiving recommendation a.

Program 4b:
The following program, given endowment e, transfer τ, and interim utilities and utility bounds, determines $V_m[\cdot]$:

$$V_m[e, \tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e)] = \max_{\pi \geq 0, v} \sum_{A, \mathbf{W}'} \pi(a, \mathbf{w}') \; [-\tau + Q \sum\nolimits_E p(e'|a) V(\mathbf{w}', e')] \tag{62}$$

subject to constraints (63)-(67) below:

$$\sum_{A, \mathbf{W}'} \pi(a, \mathbf{w}') = 1, \tag{63}$$

$$\sum_{A,\mathbf{W}'} \pi(a,\mathbf{w}') \left[u(e+\tau-a)+\beta \sum\nolimits_E p(e'|a)w'(e')\right] = w_m(e), \tag{64}$$

$$\begin{aligned} \forall a, \hat{a} \neq a: \quad & \sum_{\mathbf{W}'} \pi(a,\mathbf{w}') \left[u(e+\tau-\hat{a})+\beta \sum\nolimits_E p(e'|\hat{a})w'(e')\right] \\ & \leq \sum_{\mathbf{W}'} \pi(a,\mathbf{w}') \left[u(e+\tau-a)+\beta \sum\nolimits_E p(e'|a)w'(e')\right], \end{aligned} \tag{65}$$

$$\forall \hat{e} \neq e, a, \hat{a}: \quad \sum_{\mathbf{W}'} \pi(a,\mathbf{w}') \left[u(\hat{e}+\tau-\hat{a})+\beta \sum\nolimits_E p(e'|\hat{a})w'(e')\right] \leq v(\hat{e},e,\tau,a), \tag{66}$$

$$\forall \hat{e} \neq e: \quad \sum_{A} v(\hat{e},e,\tau,a) \leq \bar{w}_m(\hat{e},e). \tag{67}$$

The number of variables in this program is $\#A \times \#\mathbf{W}$ under $\pi(\cdot)$ plus $(\#E-1)\times\#A$ under $v(\cdot)$, where again e and τ are fixed at this stage of the problem. In our example with two endowments e, ten transfers τ, ten actions a, and ten utility vectors $\mathbf{w}'$, Program 4b has 110 variables and 193 constraints, while Program 4a has 3 constraints. The number of variables in Program 4a and the number of Programs 4b that need to be computed depends on the grid for interim utilities. For example, if the grid for interim utilities for each τ and e has ten values, Program 4a has 100 variables, and 1000 Programs 4b need to be computed. In practice, it is generally faster to use Program 3 as long as it is feasible to do so. Large programs can only be computed using Program 4a and 4b, since the individual programs are smaller and require less memory than Program 3.

We still have to define the set $\mathcal{W}(e,\tau)$ of feasible interim utility assignments in Program 4a. It is the set of all assignments for which Program 4b has a solution. With $\mathbf{W}_m$ being the utility grid that is used for $w_m(e)$ and $\bar{\mathbf{w}}_m(\hat{e},e)$, define:

$$\mathcal{W}(e,\tau) = \left\{(w_m(e),\bar{\mathbf{w}}_m(\hat{e},e)) \in (\mathbf{W}_m)^{(\#E)} \mid V_m[e,\tau,w_m(e),\bar{\mathbf{w}}_m(\hat{e},e)] \text{ is defined}\right\} \tag{68}$$

In other words, we vary utility assignments as parameters or states in Program 4b and rule out the ones for which there is no feasible solution.

The equivalency of Programs 3 and Program 4a/4b can be shown by directly comparing the sets of constraints. This is carried out in detail in Doepke and Townsend (2002).

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica* 58 (5): 1041–1063.

Allen, Franklin. 1985. "Repeated Principal-Agent Relationships with Lending and Borrowing." *Economics Letters* 17:27–31.

Atkeson, Andrew and Robert E. Lucas, Jr. 1992. "On Efficient Distribution with Private Information." *Review of Economic Studies* 59:427–453.

Attanasio, Orazio and Jose-Victor Rios-Rull. 2000. "Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?" *European Economic Review* 44 (7): 1225–58.

Cole, Harold L. and Narayana Kocherlakota. 2001a. "Dynamic Games with Hidden Actions and Hidden States." *Journal of Economic Theory* 98 (1): 114–26.

———. 2001b. "Efficient Allocations with Hidden Income and Hidden Storage." *Review of Economic Studies* 68 (3): 523–42.

Doepke, Matthias and Robert M. Townsend. 2002. "Dynamic Mechanism Design with Hidden Income and Hidden Actions: Technical Appendix." UCLA Department of Economics Working Paper No. 819.

Fernandes, Ana and Christopher Phelan. 2000. "A Recursive Formulation for Repeated Agency with History Dependence." *Journal of Economic Theory* 91 (2): 223–47.

Fudenberg, Drew, Bengt Holmstrom, and Paul Milgrom. 1990. "Short-Term Contracts and Long-Term Agency Relationships." *Journal of Economic Theory* 51:1–31.

Green, Edward J. 1987. "Lending and the Smoothing of Uninsurable Income." In *Contractual Arrangements for Intertemporal Trade*, edited by E. Prescott and N. Wallace, 3–25. Minneapolis: University of Minnesota Press.

Haegler, Urs. 2001. "Recursive Mechanisms and Costly Storage Audits to Insure Privately Informed Agents." Unpublished Manuscript, LSE.

Harris, Milton and Robert M. Townsend. 1977. "Allocation Mechanisms for Asymetrically Informed Agents." Manuscript presented at the NBER-CME Conference Honoring Herb Simon.

———. 1981. "Resource Allocation under Asymmetric Information." *Econometrica* 49 (1): 33–64.

Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall. 2000. "Mutual Insurance, Individual Savings, and Limited Commitment." *Review of Economic Dynamics* 3 (2): 216–246.

Myerson, Roger B. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47 (1): 61–73.

——— 1982. "Optimal Coordination Mechanisms in Generalized Principal Agent Problems." *Journal of Mathematical Economics* 10 (1): 67–81.

——— 1986. “Multistage Games with Communication.” *Econometrica* 54 (2): 323–58.

Perri, Fabrizio and Dirk Krueger. 1999. “Risk Sharing: Private Insurance Markets or Redistributive Taxes?” Federal Reserve Bank of Minneapolis Staff Report 262.

Phelan, Christopher. 1995. “Repeated Moral Hazard and One-Sided Commitment.” *Journal of Economic Theory* 66 (2): 488–506.

———. 1998. “On the Long Run Implications of Repeated Moral Hazard.” *Journal of Economic Theory* 79:174–91.

Phelan, Christopher and Robert M. Townsend. 1991. “Computing Multi-Period, Information-Constrained Optima.” *Review of Economic Studies* 58:853–881.

Prescott, Edward S. 1997. “Computing Private Information Problems with Dynamic Programming Methods.” Federal Reserve Bank of Richmond.

Spear, S. and S. Srivastava. 1987. “On Repeated Moral Hazard with Discounting.” *Review of Economic Studies* 54:599–618.

Thomas, Jonathan and Tim Worrall. 1990. “Income Fluctuations and Asymmetric Information: An Example of a Repeated Principal-Agent Problem.” *Journal of Economic Theory* 51 (2): 367–390.

Townsend, Robert M. 1979. “Optimal Contracts and Competitive Markets with Costly State Verification.” *Journal of Economic Theory* 21 (2): 265–93.

——— 1982. “Optimal Multiperiod Contracts and the Gain from Enduring Relationships under Private Information.” *Journal of Political Economy*, no. 6:1166–86.

Wang, Cheng. 1995. “Dynamic Insurance with Private Information and Balanced Budgets.” *Review of Economic Studies* 62 (4): 577–95.

Werning, Iván. 2001. “Repeated Moral Hazard with Unmonitored Wealth: A Recursive First-Order Approach.” Unpublished Manuscript, University of Chicago.

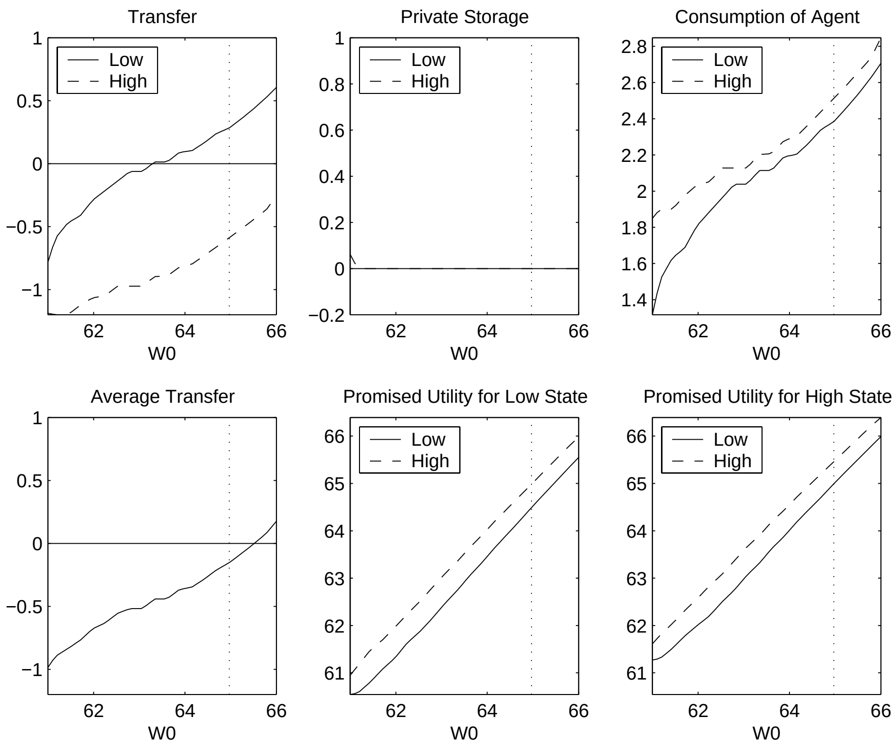

Figure 2: Policy Functions, R=1

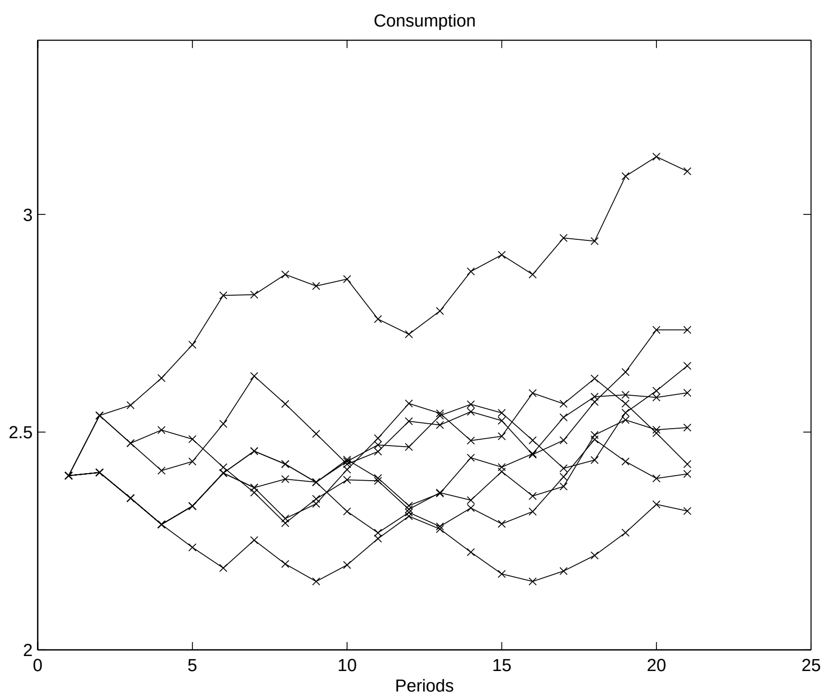

Figure 3: Time-Series for Consumption

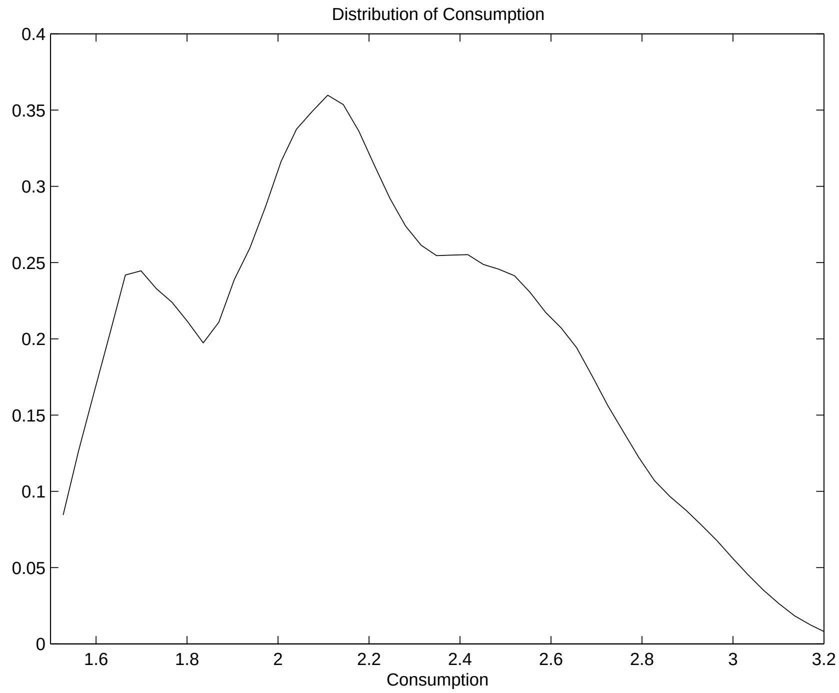

Figure 4: The Long-Run Distribution of Consumption

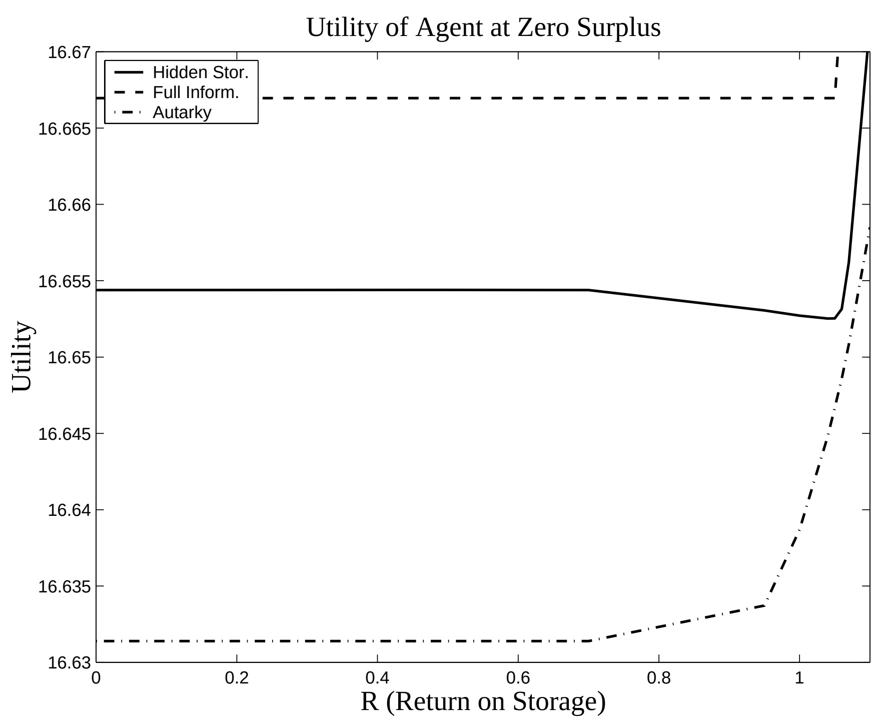

Figure 5: Utility of the Agent with Full Information, Private Information, and Autarky

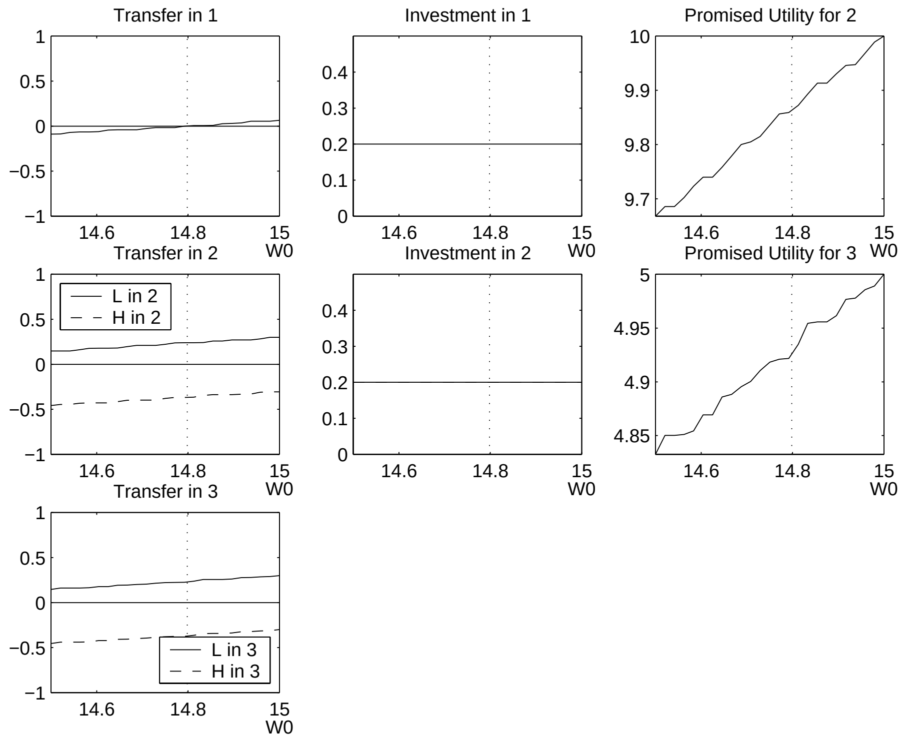

Figure 6: Policy Functions, Unconstrained Solution

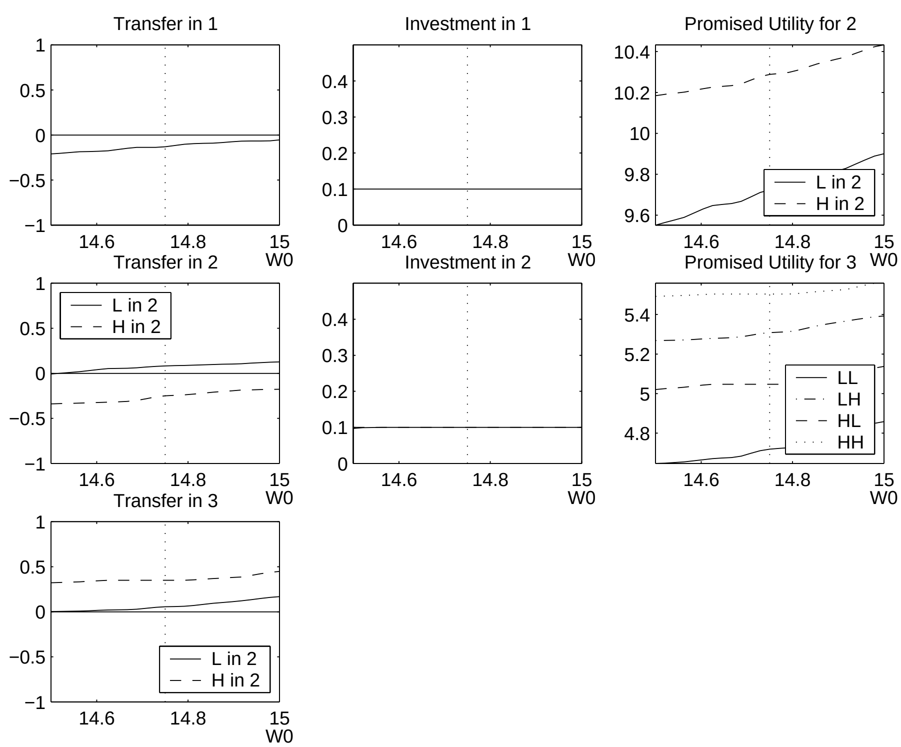

Figure 7: Policy Functions, Constrained Solution