Dynamic Mechanism Design with Hidden Income and Hidden Actions: Technical Appendix*

Matthias Doepke
UCLA

Robert M. Townsend
University of Chicago

April 2002

Abstract

In our paper "Dynamic Mechanism Design with Hidden Income and Hidden Actions," we develop general recursive methods to solve for optimal contracts in dynamic principal-agent models with hidden income and hidden actions. This appendix provides the detailed derivations of all recursive formulations presented in the paper, as well as proofs for all propositions.

*Matthias Doepke: Department of Economics, UCLA, Box 951444, Los Angeles, CA 90095. E-mail address: doepke@econ.ucla.edu. Robert Townsend: Department of Economics, The University of Chicago, 1126 East 59th Street, Chicago, IL 60637. E-mail address: rtownsen@midway.uchicago.edu.

1 Introduction

In "Dynamic Mechanism Design with Hidden Income and Hidden Actions," we consider a class of dynamic principal-agent problems in which the 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 show how this problem can be reduced to a recursive version with direct mechanisms and vectors of utility promises as the state variable. In this Technical Appendix we provide the full derivations of all recursive formulations provided in the paper, as well as proofs for all propositions. For ease of exposition, this document is self-contained and repeats the general setup and all programs contained in the original paper.

In Section 2 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 3, we invoke 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. In Sections 4, 5, and 6, we develop alternative formulations which allow the planner to specify behavior off the equilibrium path. All proofs are contained in the mathematical appendix.

2 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.

2.1 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. 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 $0 < \beta < 1$.*

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.

2.2 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.[1]

[1] 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.

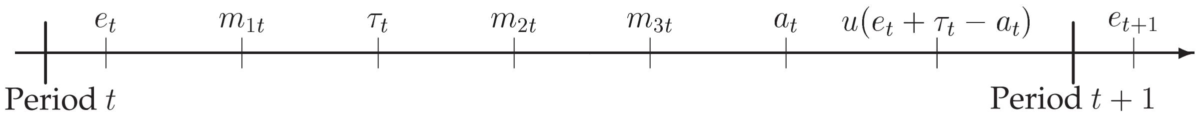

Figure 1: The Sequence of Events and Messages in Period t

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:

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

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:

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

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:

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

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.*

3 Deriving a Recursive Formulation

3.1 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.[2]

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

[2]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.

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 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. We write a possible deviation strategy δ 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 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.

3.2 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:

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

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:

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

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:

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

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.[3]

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

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

[3]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 more subtle. 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 $\hat{e}$.

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

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:

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

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 Section 7.

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.

3.3 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.[4] 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:

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

[4]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 $\hat{a} \neq a$:

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

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.

4 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 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.

4.1 The Revelation Principle Once Again

We now will follow the same steps as above to derive another recursive formulation of the original planning problem. The first step is to apply the revelation principle to a restricted setup in which the agent reports the endowment twice. We therefore consider the original problem under the restriction that the two message spaces of the agent are given by the space of endowments, $M_1 = M_2 = E$, and that the message space of the planner is the space of possible actions, $M_3 = A$. The contemporary part s_t of the history of the planner is now given by:

$$s_t \equiv \{e_{1t}, e_{2t}, \tau_t, a_t\},$$

where e_{1t} and e_{2t} are the two reports on the endowment, τ_t is the transfer, and a_t is the recommended action. 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.

The planner chooses an outcome function subject to a number of constraints which ensure that the outcome function and truth-telling and obedience on the part of the agent are an equilibrium. To ensure that the outcome function forms a probability measure, we require that $\pi(\tau_t, a_t | e_t, s^{t-1}) \geq 0$ for all transfers, actions, endowments and histories, and that:

$$\forall e_{1t}, e_{2t}, s^{t-1}: \quad \sum_{T,A} \pi(\tau_t, a_t | e_{1t}, e_{2t}, s^{t-1}) = 1. \tag{28}$$

Next, the transfer τ cannot depend on the second report e_{2t}, since the second report is made only after the agent receives the transfer. This implies the following condition:[5]

$$\forall \tau_t, e_{1t}, e_{2t}, \bar{e}_{2t}, s^{t-1}: \quad \sum_A \pi(\tau_t, a_t | e_{1t}, e_{2t}, s^{t-1}) = \sum_A \pi(\tau_t, a_t | e_{1t}, \bar{e}_{2t}, s^{t-1}). \tag{29}$$

Given an outcome function, probabilities $p(s^t|\pi)$ of histories are defined in the obvious way. Given these probabilities, as in (8) the outcome function has to deliver reservation utility W_0 to the agent, provided that the agent reports truthfully twice and takes the recommended actions:

$$\sum_{t=0}^{\infty} \beta^t \left[\sum_{S^t} p(s^t|\pi) u(e_t + \tau_t - a_t) \right] \geq W_0. \tag{30}$$

Also, it has to be optimal for the agent to tell the truth each time and follow the recommended action. We will write a possible deviation strategy δ as a function $\delta_{e1}(h^{t-1}, e_t)$ that determines the first reported endowment as a function of the history and the true endowment, a function $\delta_{e2}(h^{t-1}, e_t, \tau_t)$ that determines the second reported endowment as a function of the history h^{t-1}, the true endowment e_t, and the transfer τ_t, and a function $\delta_a(h^{t-1}, e_t, \tau_t, a_t)$ that determines the action as a function of history h^{t-1}, endowment e_t, transfer τ_t, and recommended action a_t. Since the action may be different, 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 $p(h^t|\pi, \delta)$.

We require that truth-telling and obedience be optimal for the agent, starting in any period and from any history. In addition, we also require that the prescribed actions be optimal for the agent at the node in the middle of a period, i.e., after the transfer has been assigned, regardless of what the first report was. Thus even if the agent misreports the endowment at the beginning of the period, it has to be optimal to report truthfully at the second report. As we will see later on, this requirement leads to major simplifications in the computation of the optimal outcome function.

The first requirement, akin to (9) is that the prescribed reports and actions be optimal, given any history h^k that has been already realized, inducing honesty and obedience in

[5] Alternatively, we could have formulated the outcome functions in terms of marginal probabilities for τ_t and a_t, in which case (29) would be implied.

the future:

$$\forall \delta, s^k, e_{k+1}: \quad \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right]$$
$$\leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \quad (31)$$

In fact, we place an even more strict restriction on the outcome function by reproducing (4). Given history h^k, we require that even if the first report in period $k+1$ was wrong, it is optimal for the agent to tell the truth at the second report in $k+1$ and follow the recommended action, and to be honest and obedient in the future. This leads to the following condition:

$$\forall \delta, s^k, e_{k+1}: \quad \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right]$$
$$\leq \beta^{k+1} \left[\sum_{T,A} \pi \left(\tau_{k+1}, a_{k+1} | \delta_{e1}(h^k, e_{k+1}), e_{k+1}, s^k \right) u(e_{k+1} + \tau_{k+1} - a_{k+1}) \right]$$
$$+ \sum_{t=k+2}^{\infty} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \quad (32)$$

We can now define feasible allocations under double reporting, analogous to Definition 4:

Definition 6 *An outcome function is an equilibrium outcome function under truth telling and obedience with double reporting if it satisfies the constraints* (28) *to* (32) *above. A feasible allocation is a probability distribution over endowments, transfers and actions that is implied by an equilibrium outcome function.*

Feasible allocations under truth-telling and obedience with double reporting are a subset of the feasible allocations in the general setup. The next proposition states that the set of constraints (28) to (32) does not put more restrictions on feasible allocation as the original constraints (3), (4), and (5), so that the sets of feasible allocations are in fact identical.

Lemma 1 *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 with double reporting.*

We are therefore justified in framing the maximization problem of the planner as maximizing:

$$V(\pi) \equiv \sum_{t=0}^{\infty} Q^t \left[\sum_{S^t} p(s^t|\pi)(-\tau_t) \right], \tag{33}$$

subject to (28) to (32) above.

4.2 Utility Vectors as State Variables

As before with Program 1, the next step is to reduce the current game to a recursive version with a vector of promised utilities as the state variable. The result will be Program 2.

Again, we consider an auxiliary planning problem in which the planner has to deliver a vector of reservation utilities $\mathbf{w}_0$, depending on the endowment in the first period. The choice object of the planner is given by $\pi(\tau_t, a_t|e_{1t}, e_{2t}, s^{t-1})$. For brevity of notation, we suppress the dependence of the outcome function on the initial utility vector $\mathbf{w}_0$. The planner picks the contract conditional on the initial report e_{10} (the first report in period 0) by the agent. The probability and consistency constraints (28) and (29) are imposed as before. Instead of one promise-keeping constraint (30) which is realized in expected value across states, there now is a promise-keeping constraint for each possible initial endowment. For all e_0, we require:

$$\begin{aligned}\forall e_0 : \quad & \sum_{T,A} \pi(\tau_0, a_0|e_0, e_0) \Big[u(e_0 + \tau_0 - a_0) \\ & + \sum_{t=1}^{\infty} \beta^t \sum_{S^{t-1},E,T,A} p(s^{t-1}|s_0) p(e_t|s^{t-1}) \pi(\tau_t, a_t|e_t, e_t, s^{t-1}) u(e_t + \tau_t - a_t) \Big] = w_0(e_0). \end{aligned} \tag{34}$$

The incentive constraints (31) and (32) are imposed as well. Since the vector of promised utilities $\mathbf{w}$ will serve as our state variable, it will be important to show that the set of all feasible utility vectors has nice properties.

Definition 7 *The set* $\mathbf{W}$ *is given by all vectors* $\mathbf{w} \in R^{\#E}$ *that satisfy constraints* (28), (29), (31), (32), *and* (34) *for some outcome function* $\pi(\tau_t, a_t|e_t, e_t, s^{t-1})$.

Lemma 2 *The set* $\mathbf{W}$ *is nonempty and compact.*

(The proof is the same as for Proposition 3)

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

$$V(\mathbf{w}_0, e_0) = \max_{\pi} \sum_{T,A} \pi(\tau_0, a_0|e_0, e_0) \left[-\tau_0 + \sum_{t=1}^{\infty} Q^t \Big[\sum_{S^t} p(s^t|\pi, s_0)(-\tau_t) \Big] \right], \tag{35}$$

where the maximization is subject to constraints (28), (29), (31), (32), and (34) above, for a given $e_0 \in E$ and $\mathbf{w} \in \mathbf{W}$. We want to show that this problem has a recursive structure.

To do this, we need to define on-path future utilities. For all s^{k-1}, e^k, let:

$$w(e_k, s^{k-1}) =$$
$$\sum_{T,A} \pi(\tau_k, a_k|e_k, e_k, s^{k-1}) \left[u(e_k + \tau_k - a_k) + \sum_{t=k+1}^{\infty} \beta^{t-k} \Big[\sum_{S^t} p(s^t|\pi, s_k) u(e_t + \tau_t - a_t) \Big] \right], \tag{36}$$

and let $\mathbf{w}(s^{k-1}, \pi)$ be the vector of these utilities over all e_k. We can now show the following result:

Lemma 3 *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 $\mathbf{w}_0 = \mathbf{w}(s^{k-1}, \pi^\star)$.*

Given this result, we know that the maximized surplus of the planner can be written as:

$$V(\mathbf{w}_0, e_0) = \sum_{A,T} \pi^\star(\tau_0, a_0|e_0, e_0) \left[-\tau_0 + Q \sum_{E} p(e_1|s^0) V(\mathbf{w}(s^0, \pi^\star), e_1) \right]. \tag{37}$$

In light of (37), we can cast the planner's problem as choosing transfers and actions in the present period, and choosing continuation utilities from the set $\mathbf{W}$ from tomorrow on. We will write the choices of the planner as a function of the vector of promised utilities $\mathbf{w}$ that has to be delivered in the current period, and the current state e. The choices of the planner are therefore functions $\pi(\tau, a|\mathbf{w}, e_1, e_2)$ and $\mathbf{w}'(\mathbf{w}, e_1, e_2, \tau, a)$, where $\mathbf{w}'$ is the vector of utilities promised from tomorrow on, and is restricted to lie in $\mathbf{W}$. We still need to determine which constraints need to be placed on these choices in order to guarantee that the implied contract satisfies (28), (29), (31), (32), and (34). In order to reproduce (28), we need to impose:

$$\forall e_2 : \quad \sum_{T,A} \pi(\tau, a|\mathbf{w}, e, e_2) = 1. \tag{38}$$

The consistency constraint (29) is satisfied if we impose:

$$\forall \tau, e_2, \bar{e}_2 : \quad \sum_{A} \pi(\tau, a|\mathbf{w}, e, e_2) = \sum_{A} \pi(\tau, a|\mathbf{w}, e, \bar{e}_2). \tag{39}$$

The promise-keeping constraint (34) will be satisfied if we impose:

$$\sum_{T,A} \pi(\tau, a|\mathbf{w}, e, e) \left[u(e + \tau - a) + \beta \sum_{E} p(e'|a) w'(\mathbf{w}, e_1, e_2, \tau, a)(e') \right] = w(e). \tag{40}$$

We now get to the incentive constraints. As above in the derivation of Program 1, we only write down incentive constraints for one-shot deviations, which can be shown to be equivalent to the full incentive constraints (31) and (32). It will also be useful to write down separate constraints for deviations in actions and reports. First of all, we require that it is optimal to take the required action, assuming that the truth was reported at the second time. This constraint implements (32) for one-shot deviations in terms of taking the right action:

$$\begin{aligned} \forall \hat{e}, \tau, a, \hat{a} : \qquad & u(\hat{e} + \tau - \hat{a}) + \beta \sum_{E} p(e'|\hat{a}) w'(\mathbf{w}, e, \hat{e}, \tau, a)(e') \\ & \leq u(\hat{e} + \tau - a) + \beta \sum_{E} p(e'|a) w'(\mathbf{w}, e, \hat{e}, \tau, a)(e'). \end{aligned} \tag{41}$$

In order to reproduce (32) in terms of reporting, and therefore ensure that the second report be correct and the recommended action is taken even if the first report were wrong, we need to impose:

$$\begin{aligned} & \forall \hat{e}, \tau, \hat{\hat{e}} \neq \hat{e}, \delta : \\ & \quad \sum_{A,\mathbf{W}'} \pi(\tau, a, |\mathbf{w}, e, \hat{\hat{e}}) \left[u(\hat{e} + \tau - \delta(a)) + \beta \sum_{E} p(e'|\delta(a)) w'(\mathbf{w}, e, \hat{\hat{e}}, \tau, a)(e') \right] \\ & \qquad \leq \sum_{A,\mathbf{W}'} \pi(\tau, a|\mathbf{w}, e, \hat{e}) \left[u(\hat{e} + \tau - a) + \beta \sum_{E} p(e'|a) w'(\mathbf{w}, e, \hat{e}, \tau, a)(e') \right]. \end{aligned} \tag{42}$$

Finally, we still need to ensure that the truth be reported the first time, thereby implementing (31):

$$\forall \hat{e} \neq e: \quad \sum_{T,A} \pi(\tau, a, |\mathbf{w}, e, \hat{e}) \left[u(\hat{e} + \tau - a) + \beta \sum_{E} p(e'|a) w'(\mathbf{w}, e, \hat{e}, \tau, a)(e') \right] \leq w(\hat{e}). \quad (43)$$

In this last constraint, we assume that the agent will tell the truth and follow the recommendations afterwards, instead of checking over all possible reporting and action strategies. We are justified in doing so because constraint (42), which is imposed simultaneously, guarantees that truthful reporting at the second report and obedience will be the optimal strategy even if the first report were wrong. Thus allowing for other strategies merely lowers the utility of the agent on the left-hand side, and does not introduce additional constraints. It is here were the gains from double reporting arise.

The constraints above rule out that the agent can gain from misreporting or disobedience in any period, given that he goes back to truth-telling and obedience from the next period on. The constraints therefore imply that (31) and (32) hold for one-shot deviations. The argument used in the derivation of Program 1, however, applies here as well, so that (41) to (43) imply (31) and (32) even for a finite or infinite number of deviations.

Thus we are justified to pose the problem of the planner as solving:

$$V(\mathbf{w}, e) = \max_{\pi \geq 0} \sum_{A,T} \pi(\tau, a|\mathbf{w}, e, e) \left[-\tau + \beta \sum_{E} p(e'|a) V(\mathbf{w}'(\mathbf{w}, e, e, \tau, a), e') \right] \quad (44)$$

by choice of π and $\mathbf{w}'$, subject to constraints (38) to (43) above. Program 2 is a version of this problem with a discrete grid for promised utilities. Of course, we now assume that the function $V(\mathbf{w}, e)$ is already known. $V(\mathbf{w}, e)$ can be computed with standard dynamic programming techniques. Specifically, the right-hand side of (44) defines an operator T that maps functions $V(\mathbf{w}, e)$ into $TV(\mathbf{w}, e)$. It is easy to show 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.

4.3 The Discretized, Recursive Version

The planning problem with double reporting can now further be reduced to a recursive version with a vector of utility promises as a state variable by following the same steps as above in the derivation of Program 1. Since the steps and proofs are virtually identical (they are contained in the working paper version of the article), we now go directly to the discretized recursive version of the planning problem with double reporting.

As before, 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 report to the planner. Then the planner delivers the transfer τ, and afterwards the agent reports the endowment e again. The incentive-compatibility constraints will ensure that this second report be correct, even if the first report was 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).

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

$$V(\mathbf{w}, e) = \max_{\pi \geq 0} \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, e) \left[-\tau + Q \sum\nolimits_E p(e'|a) V(\mathbf{w}', e')\right] \tag{45}$$

subject to constraints (46)-(51) 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 $\hat{e}$:

$$\forall \hat{e}: \quad \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) = 1. \tag{46}$$

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 $\hat{e} \neq e$ and all τ, we require:

$$\forall \hat{e} \neq e, \tau: \quad \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, e) = \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}). \tag{47}$$

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

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

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, e) \left[u(e+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right] = w(e). \tag{48}$$

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. For each true state $\hat{e}$, transfer τ, recommended action a, and alternative action $\hat{a} \neq a$ we require much like (23):

$$\begin{aligned} \forall \hat{e}, \tau, a, \hat{a} \neq a : \quad & \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \left[u(\hat{e}+\tau-\hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a})w'(e')\right] \\ & \leq \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \left[u(\hat{e}+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right]. \end{aligned} \tag{49}$$

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:

$$\begin{aligned} \forall \hat{e}, \tau, \hat{\hat{e}} \neq \hat{e}, \delta : \quad & \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{\hat{e}}) \left[u(\hat{e}+\tau-\delta(a)) + \beta \sum\nolimits_E p(e'|\delta(a))w'(e')\right] \\ & \leq \sum_{A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \left[u(\hat{e}+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right]. \end{aligned} \tag{50}$$

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:

$$\forall \hat{e} \neq e : \quad \sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e, \hat{e}) \left[u(\hat{e}+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right] \leq w(\hat{e}). \tag{51}$$

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 (49) and (50) hold.

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

The number of variables in this formulation is $\#E \times \#T \times \#A \times \#\mathbf{W}$. Thus, the number of variables increased relative to the first version, since π now also depends on the second report $\hat{e}$. Similarly, there are $\#E$ probability constraints (46), $(\#E-1)\times\#T$ independence constraints (47), and there is one promise-keeping constraint (48). The total number of obedience constraints (49) is $\#E\times\#T\times\#A\times(\#A-1)$. The number of truth-telling constraints for the second report (50) is $\#E\times\#T\times(\#E-1)\times(\#A)^{(\#A)}$, and the number of truth-telling constraints for the first report (51) is $\#E-1$. 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 address methods to deal with this problem.

5 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 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 this 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:

$$V(\mathbf{w},e)=\max_{\pi\geq 0,v}\sum_{T,A,\mathbf{W}'}\pi(\tau,a,\mathbf{w}'|\mathbf{w},e)\ \left[-\tau+Q\sum\nolimits_{E}p(e'|a)V(\mathbf{w}',e')\right] \tag{52}$$

subject to the constraints (53)-(57) below. Apart from the addition of the utility bounds $v(\cdot)$ the objective function (52) is identical to (20). The first constraint is the probability measure constraint, identical with (21):

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) = 1. \tag{53}$$

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

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(e+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right] = w(e). \tag{54}$$

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):

$$\begin{aligned}\forall \tau, a, \hat{a} \neq a: \quad & \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(e+\tau-\hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a})w'(e')\right] \\ & \leq \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(e+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right]. \end{aligned} \tag{55}$$

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 that $v(\hat{e}, e, \tau, a)$, 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:

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

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:

$$\forall \hat{e} \neq e: \quad \sum_{T,A} v(\hat{e}, e, \tau, a) \leq w(\hat{e}). \tag{57}$$

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 (53) and one promise-keeping constraint (54). The number of obedience constraints (55) is $\#T \times \#A \times (\#A - 1)$. There are $(\#E - 1) \times \#T \times (\#A)^2$ constraints (56) to implement the utility bounds, and $(\#E - 1)$ truth-telling constraints (57). 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.

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.*

6 Subdividing the Problem

In the last formulation, the number of constraints gets large if both the grids for transfer τ and action a are made very fine. In practice, this leads to memory problems when computing. In this section we develop a formulation in which *the transfer and the recommendation are assigned at two different stages.* At each stage, the number of variables and constraints is relatively small. This enables us to compute problems with fine grids for actions and transfers.

As in Section 4 on double reporting, 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.* In order to make the exposition more transparent, we write $w_m(e)$ and the vector $\bar{\mathbf{w}}_m(\hat{e}, e)$ as a function of e to indicate the state for which the interim utility is assigned. $\bar{\mathbf{w}}_m(\hat{e}, e)$ 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), but as is typical of dynamic programs, we now take that function $V_m(\cdot)$ as given.

We will analyze the first subperiod. In the first subperiod the planner assigns transfers, interim utilities, and utility bounds. When choosing the utility assignments, the planner is restricted to assignments that can actually be implemented in the second subperiod. We use $\mathcal{W}(e, \tau)$ to denote the set of feasible utility assignments for a given state e and transfer τ. For now we will take this set as given, and define it precisely below when we turn to the second subperiod.

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. The first constraint is the probability constraint:

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

Promises have to be kept:

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

Finally, truth telling has to be observed, so that this e branch is not made too tempting. An agent who is at state $\hat{e}$ should not be able to gain by claiming to be at state e. For all $\hat{e} \neq e$, we require:

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

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 before in Section 3.3, 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 $V_m[\cdot]$:

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

subject to constraints (63)-(67) below. The first constraint is the usual probability constraint:

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

The promise-keeping constraint requires that the interim utility $w_m(e)$ that was promised is actually delivered:

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

The obedience constraints ensure that given that the reported endowment is correct, the agent carries out the recommended action. For all recommended a and alternative actions

$\hat{a} \neq a$, we require:

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

The next set of constraints implements, or respects, the utility bounds. The utility of having reported e, being at state $\hat{e}$, having received transfer τ and recommendation a cannot be larger than $v(\hat{e}, e, \tau, a)$. For all $\hat{e} \neq e$, all recommendations a, and all actions $\hat{a}$ we require:

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

where as usual e and τ are givens for Program 4b. Finally, the utility bounds for misreporting the endowment have to be observed. For each endowment $\hat{e} \neq e$ we require:

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

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. There is one probability constraint (63) and one promise-keeping constraint (64), and there are $\#A \times (\#A - 1)$ obedience constraints (65). The number of constraints (66) to implement the utility bounds is $(\#E - 1) \times (\#A)^2$, and the number of truth-telling constraints (67) is $(\#E - 1)$. Notice that the number of constraints and variables does not depend on the grid for the transfer. More possible transfers will *increase the number of programs to be computed* at this second stage (Program 4b), but *the programs will not change in size*.

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.

We now will show that Program 3 and Program 4 are equivalent in the sense that the same allocations are feasible in both programs. By allocation we mean a probability distribution over transfer, storage, and promised future utilities. In order to prove this claim, we will start with an allocation that satisfies the constraints in Program 3 and then show that this allocation is also feasible in Program 4. Afterwards, we will start with an allocation that is feasible in Program 4, and then show that it is also feasible in Program 3. Since in both Programs the planner has the same objective of maximizing surplus, this result also implies that the maximizing allocation will be the same in both programs, as will be the utility of the planner.

Lemma 4 *The allocations that can be implemented in Program 3 and 4 are identical.*

We have now shown that Programs 1 and 2 are equivalent, that Programs 1 and 3 are equivalent, and that Programs 3 and 4 are equivalent. We therefore have:

Corollary 1 *Programs 1, 2, 3, and 4 are all equivalent.*

7 Computing the Value Set

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, 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 8 *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 $\bar{\mathbf{W}}$.*

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 $\mathbf{W}$.*

A 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. $\square$

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.*

Proof of Proposition 2 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 $\pi^\star$ 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:

$$p(h^t|s^t) \equiv \frac{p(h^t)}{\sum_{H^t(s^t)} p(h^t)} \tag{69}$$

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:

$$\pi^\star(\tau_t, a_t|e_t, s^{t-1}) \equiv \sum_{H^{t-1}(s^{t-1}),M_1,M_2,M_3} p(h^{t-1}|s^{t-1})\, \sigma(m_{1t}|e_t, h^{t-1})\, \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\sigma(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1})\, \pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\, \sigma(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}). \tag{70}$$

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).

We start with the probability-measure constraint (7). Fix s^{t-1} and e_t. Definition (69) implies:

$$\sum_{H^{t-1}} p(h^{t-1}|s^{t-1}) = 1. \tag{71}$$

Since the elements of $\{\pi, \sigma\}$ form probability measures, they sum up to one. We can therefore insert the sum over $\{\pi, \sigma\}$ into (71) to get:

$$\sum_{H^{t-1}} p(h^{t-1}|s^{t-1}) \left[\sum_{M_1,T,M_2,M_3,A} \sigma(m_{1t}|e_t, h^{t-1})\, \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1})) \right.$$
$$\left. \sigma(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1})\, \pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\, \sigma(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}) \right] = 1.$$

By a change in the order of summation, this is:

$$\sum_{T,A}\Bigg[\sum_{H^{t-1}(s^{t-1}),M_1,M_2,M_3} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_t,h^{t-1})\ \pi(\tau_t|m_{1t},s^{t-1}(h^{t-1}))$$
$$\sigma(m_{2t}|e_t,m_{1t},\tau_t,h^{t-1})\ \pi(m_{3t}|m_{1t},\tau_t,m_{2t},s^{t-1}(h^{t-1}))\ \sigma(a_t|e_t,m_{1t},\tau_t,m_{2t},m_{3t},h^{t-1})\Bigg]=1.$$

The inner sum is the right-hand side of (70). Replacing the sum by the left-hand side of (70) we get:

$$\sum_{T,A}\pi^\star(\tau_t,a_t|e_t,s^{t-1})=1,$$

which is (7).

The promise-keeping constraint (8) is next. Writing out (5) gives:

$$\sum_{t=0}^{\infty}\beta^t\left[\sum_{H^t}p(h^t|\pi,\sigma)u(e_t+\tau_t-a_t)\right]\geq W_0.$$

Because of (69) and (70), the probability of any history is the same under $\pi^\star$ as under (π,σ). Thus the constraint can be written as:

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

Rearranging the sum gives:

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

which can also be written as:

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

Because of (69), we have $\sum_{H^t(s^t)}p(h^t|s^t)=1$. We therefore get:

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

which is (8).

The last step is to show that the incentive constraint (9) is satisfied. Fix a node s^k and e_{k+1} and a deviation δ_e, δ_a. Given this deviation, we define an equivalent deviation strategy $\hat{\sigma}$ in the original setup, namely:

$$\hat{\sigma}(m_{1t}|e_t, h^{t-1}) \equiv \sigma(m_{1t}|\delta_e(h^{t-1}, e_t), h^{t-1}),$$

$$\hat{\sigma}(m_{2t}|e_t, m_{1t}, \tau_t, h^{t-1}) \equiv \sigma(m_{2t}|\delta_e(h^{t-1}, e_t), m_{1t}, \tau_t, h^{t-1}),$$

and:

$$\hat{\sigma}(a_t|e_t, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}) \equiv \sum_{\{\tilde{a} \in A | a_t = \delta_a(h^{t-1}, e_t, \tau_t, \tilde{a}_t)\}} \sigma(\tilde{a}_t|\delta_e(h^{t-1}, e_t), m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}).$$

On all branches other than s^k and e_{k+1}, we define $\hat{\sigma}$ as equal to σ, i.e., the agent does not actually deviate on other branches. Writing out (3) for this deviation and a given $h^k \in H^k(s^k)$ gives:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \hat{\sigma}, h^k) u(e_t + \tau_t - a_t) \right] \\ \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \sigma, h^k) u(e_t + \tau_t - a_t) \right]. \quad (72)$$

Since (72) holds for all $h^k \in H^k(s^k)$, we can multiply by $p(h^k|s^k)$ and sum over $H^k(s^k)$ on both sides, resulting in a version in which probabilities are conditioned on s^k instead of h^k:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \hat{\sigma}, s^k) u(e_t + \tau_t - a_t) \right] \\ \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \sigma, s^k) u(e_t + \tau_t - a_t) \right]. \quad (73)$$

Also, on branches other than e_{k+1} by the construction of $\hat{\sigma}$ the left- and right-hand-side terms are identical. Thus the inequality remains intact if we also condition on e_{k+1}:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \hat{\sigma}, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right] \\ \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \sigma, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \quad (74)$$

Because of (69) and (70), the probability of any history on the right-hand side of the equation is the same under (π, σ) as under $\pi^\star$ in the restricted setup. On the left-hand side, by the construction of $\hat{\sigma}$ the probabilities are the same as under the deviation $\delta = (\delta_e, \delta_a)$ in the restricted setup. We can therefore rewrite (74) as:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi^\star, \delta, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] \\ \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi^\star, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \quad (75)$$

Rewriting the sum on the right-hand side gives:

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

and because of (69) we get:

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

which is (9).

In summary, the contract defined in (70) satisfies constraints (7) to (9), which completes the proof. □

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. $\square$

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}:

$$\sum_{t=l+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] > \sum_{t=l+1}^{\infty} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right]. \tag{76}$$

Consider first the case $l + 1 \geq k$. On any branch that starts at or after time k, contract $\pi^\star$ is by construction entirely identical to either π or π_k. But then (76) 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 (76) as:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, \delta, s^l, e_{l+1}) w(e_k, s^{k-1}, \delta)$$
$$> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, s^l, e_{l+1}) w(e_k, s^{k-1}). \tag{77}$$

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

$$w(e_k, s^{k-1}, \delta) \leq w(e_k, \delta(s^{k-1})), \tag{78}$$

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 (78) in (77) gives:

$$\begin{aligned}
&\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t|\pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t))\right] + \\
&\qquad\qquad \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi^\star, \delta, s^l, e_{l+1}) w(e_k, \delta(s^{k-1})) \\
&> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t|\pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t)\right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi^\star, s^l, e_{l+1}) w(e_k, s^{k-1}).
\end{aligned} \tag{79}$$

The outcome function $\pi^\star$ enters (79) 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 (79) as:

$$\begin{aligned}
&\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t))\right] + \\
&\qquad\qquad \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi, \delta, s^l, e_{l+1}) w(e_k, \delta(s^{k-1})) \\
&> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^l, e_{l+1}) u(e_t + \tau_t - a_t)\right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi, s^l, e_{l+1}) w(e_k, s^{k-1}).
\end{aligned} \tag{80}$$

But now the left-hand side of (80) 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 (80) 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. □

Lemma 1 *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 with double reporting.*

Proof of Lemma 1 As above, we will start with a feasible allocation in the general setup, and then find the outcome function in the restricted setup which implements the same allocation. Fix this contract (π, σ) corresponding to the feasible allocation in the general setup. We will now define an outcome function $\pi^\star$ for the truth-telling-and-obedience mechanism that implements the same allocation. To complete the proof, we then will have to show that this outcome function satisfies constraints (28) to (32).

We will define the outcome function such that the allocation is the one implemented by (π, σ) along the equilibrium path. Let $H^t(s^t)$ be the set of the histories h^t in the general setup such that the sequence of endowments, transfers, and actions in h^t coincides with the sequence of *second* reported endowments, transfers, and recommended actions in history s^t of the restricted game. For $h^t \in H^t(s^t)$, $p(h^t|s^t)$ is the probability of agent's history h^t conditional on on the planner's history s^t, as in (69) above:

$$p(h^t|s^t) \equiv \frac{p(h^t)}{\sum_{H^t(s^t)} p(h^t)} \tag{81}$$

For the case that both reports coincide, the probabilities reproduce the original equilibrium path. For differing reports, we will define the outcome function (and therefore probabilities) such that the second report corresponds to the true endowment, while the agent was "mistaken" at the first report. This is also the reason that the second report needs to coincide to the endowment in the history h^t in the definition of $H^t(s^t)$ above. The probabilities $p(h^t|s^t)$ are then defined recursively as:

$$\begin{aligned} p(h^t|s^t) \equiv\ & p(h^{t-1}(h^t)|s^{t-1}(h^t))p(e_{2t}|h^{t-1}) \\ & \sum_{M_1,M_2,M_3} \sigma(m_{1t}|e_{1t}, h^{t-1})\ \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1})) \\ & \sigma(m_{2t}|e_{2t}, m_{1t}, \tau_t, h^{t-1})\ \pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\ \sigma(a_t|e_{2t}, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}). \end{aligned} \tag{82}$$

Notice that the actual endowment e_{2t} governs the second report, while the first report is governed by a potentially different endowment e_{1t}. Analogously, we define an outcome function for the truth-telling-and-obedience mechanism by:

$$\pi^\star(\tau_t, a_t | e_{1t}, e_{2t}, s^{t-1}) \equiv \sum_{H^{t-1}(s^{t-1})M_1, M_2, M_3} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t}, h^{t-1})\ \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\sigma(m_{2t}|e_{2t}, m_{1t}, \tau_t, h^{t-1})\ \pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\ \sigma(a_t|e_{2t}, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1}). \quad (83)$$

We now have to verify that with this choice of an outcome function conditions (28) to (32) above are satisfied.

(83) is identical to (70) as long as the same endowment is reported twice. Since the promise-keeping constraint (30) only involves on-path predictions, the proof for showing that (30) is identical to the one for the promise-keeping constraint (8) in the proof of Proposition 2. The proof for the probability constraint (28) is identical as well. We therefore proceed to the consistency constraint (29). Fix τ_t, e_{1t}, e_{2t}, $\bar{e}_{2t}$, and s^{t-1}. We start with the following identity (left- and right-hand sides are identical):

$$\sum_{H^{t-1}(s^{t-1}), M_1} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t}, h^{t-1})\ \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$= \sum_{H^{t-1}(s^{t-1})M_1} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t}, h^{t-1})\ \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1})).$$

Since the elements of (π, σ) form probability measures, they sum to one. Therefore the equation remains intact if we insert sums over (π, σ) on both sides:

$$\sum_{H^{t-1}(s^{t-1}), M_1} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t}, h^{t-1})\ \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\left[\sum_{M_2, M_3, A} \sigma(m_{2t}|e_{2t}, m_{1t}, \tau_t, h^{t-1})\pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\sigma(a_t|e_{2t}, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1})\right]$$
$$= \sum_{H^{t-1}(s^{t-1}), M_1} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t}, h^{t-1})\ \pi(\tau_t|m_{1t}, s^{t-1}(h^{t-1}))$$
$$\left[\sum_{M_2, M_3, A} \sigma(m_{2t}|\bar{e}_{2t}, m_{1t}, \tau_t, h^{t-1})\pi(m_{3t}|m_{1t}, \tau_t, m_{2t}, s^{t-1}(h^{t-1}))\sigma(a_t|\bar{e}_{2t}, m_{1t}, \tau_t, m_{2t}, m_{3t}, h^{t-1})\right].$$

Changing the order of summation gives:

$$\sum_A \Big[\sum_{H^{t-1}(s^{t-1}),M_1,M_2,M_3} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t},h^{t-1})\ \pi(\tau_t|m_{1t},s^{t-1}(h^{t-1}))$$
$$\sigma(m_{2t}|e_{2t},m_{1t},\tau_t,h^{t-1})\ \pi(m_{3t}|m_{1t},\tau_t,m_{2t},s^{t-1}(h^{t-1}))\ \sigma(a_t|e_{2t},m_{1t},\tau_t,m_{2t},m_{3t},h^{t-1})\Big]$$
$$= \sum_A \Big[\sum_{H^{t-1}(s^{t-1}),M_1,M_2,M_3} p(h^{t-1}|s^{t-1})\ \sigma(m_{1t}|e_{1t},h^{t-1})\ \pi(\tau_t|m_{1t},s^{t-1}(h^{t-1}))$$
$$\sigma(m_{2t}|\bar{e}_{2t},m_{1t},\tau_t,h^{t-1})\ \pi(m_{3t}|m_{1t},\tau_t,m_{2t},s^{t-1}(h^{t-1}))\ \sigma(a_t|\bar{e}_{2t},m_{1t},\tau_t,m_{2t},m_{3t},h^{t-1})\Big].$$

The inner sum is the right-hand side of (83). Replacing it by the left-hand side gives:

$$\sum_A \pi^\star(\tau_t,a_t|e_{1t},e_{2t},s^{t-1}) = \sum_A \pi^\star(\tau_t,a_t|e_{1t},\bar{e}_{2t},s^{t-1}),$$

which is (29).

We now move to the incentive constraints. We will follow the same approach as in the proof of Proposition 2. Fix a node s^k and e_{k+1} and a deviation $\delta_{e1},\delta_{e2},\delta_a$. Given this deviation, we define an equivalent deviation strategy $\hat{\sigma}$ in the original setup, namely:

$$\hat{\sigma}(m_{1t}|e_t,h^{t-1}) \equiv \sigma(m_{1t}|\delta_{e1}(h^{t-1},e_t),h^{t-1}),$$

$$\hat{\sigma}(m_{2t}|e_t,m_{1t},\tau_t,h^{t-1}) \equiv \sigma(m_{2t}|\delta_{e2}(h^{t-1},e_t,\tau_t),m_{1t},\tau_t,h^{t-1}),$$

and:

$$\hat{\sigma}(a_t|e_t,m_{1t},\tau_t,m_{2t},m_{3t},h^{t-1}) \equiv \sum_{\{\tilde{a}\in A|a_t=\delta_a(h^{t-1},e_t,\tau_t,\tilde{a}_t)\}} \sigma(\tilde{a}_t|\delta_{e2}(h^{t-1},e_t,\tau_t),m_{1t},\tau_t,m_{2t},m_{3t},h^{t-1}).$$

On all branches other than s^k and e_{k+1}, we define $\hat{\sigma}$ as equal to σ, i.e., the agent does not actually deviate on other branches. We now start by deriving (31), and then turn to (32).

Writing out (3) for deviation $\hat{\sigma}$ and a given $h^k \in H^k(s^k)$ gives:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \hat{\sigma}, h^k) u(e_t + \tau_t - a_t) \right] \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \sigma, h^k) u(e_t + \tau_t - a_t) \right]. \quad (84)$$

By the same argument used in Proposition 2, inequality (84) implies:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \hat{\sigma}, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right] \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \sigma, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \quad (85)$$

Because of (81) and (83), the probability of any history on the right-hand side of the equation is the same under (π, σ) as under $\pi^\star$ in the restricted setup. On the left-hand side, by the construction of $\hat{\sigma}$ the probabilities are the same as under the deviation $\delta = (\delta_{e1}, \delta_{e2}, \delta_a)$ in the restricted setup. We can therefore rewrite (85) as:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi^\star, \delta, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] \leq \sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi^\star, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right]. \quad (86)$$

Rewriting the sum on the right-hand side as above we get:

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

which is (31). The argument for the second incentive constraint (32) follows the same outline. Starting from (4), we get by the same steps as above:

$$\sum_{t=k+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \hat{\sigma}, s^k, e_{k+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right]$$
$$\leq \beta^{k+1} \Bigg[\sum_{H^k, M_1, T, M_2, M_3, A} p(h^k|s^k) \hat{\sigma}(m_{1k+1}|e_{k+1}, h^k)\ \pi(\tau_{k+1}|m_{1k+1}, s^k(h^k))$$
$$\sigma(m_{2k+1}|e_{k+1}, m_{1k+1}, \tau_{k+1}, h^k)\ \pi(m_{3k+1}|m_{1k+1}, \tau_{k+1}, m_{2k+1}, s^k(h^k))$$
$$\sigma(a_{k+1}|e_{k+1}, m_{1k+1}, \tau_{k+1}, m_{2k+1}, m_{3k+1}, h^k) u(e_{k+1} + \tau_{k+1} - a_{k+1}) \Bigg]$$
$$+ \sum_{t=k+2}^{\infty} \beta^t \left[\sum_{H^t} p(h^t|\pi, \sigma, s^k, e_{k+1}) u(e_t + \tau_t - a_t) \right].$$

Using (83) on the right-hand side and otherwise the same steps, this can be transformed into:

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

which is (32). □

Lemma 2 *The set* $\mathbf{W}$ *is nonempty and compact.*

Proof of Lemma 2 (See proof of Proposition 3)

Lemma 3 *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 Lemma 3 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 at 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 $\mathbf{w}_0 = \mathbf{w}(s^{k-1}, \pi)$. 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 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 an optimal contract at time zero. To show this, we have to demonstrate that $\pi^\star$ satisfies constraints (28), (29), (31), (32), and (34), 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}) 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 (28), (29), (31), (32), and (34), it must be optimal as well. Thus we only have to show that (28), (29), (31), (32), and (34) 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 and consistency constraints (28) and (29) 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 (34) is satisfied since the new contract delivers the same on-path utilities by construction. We still have to show that the incentive constraints (31) and (32) are satisfied. We will do this by contradiction.

Suppose that (31) is not satisfied by contract $\pi^\star$. Then there is a deviation $\delta_{e1}(e_t, s^{t-1})$, $\delta_{e2}(e_t, s^{t-1})$, $\delta_a(s^t)$ such that:

$$\sum_{t=l+1}^{\infty} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] > \sum_{t=l+1}^{\infty} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right]. \quad (87)$$

Consider first the case $l + 1 \geq k$. On any branch that starts at or after time k, contract $\pi^\star$ is by construction entirely identical to either π or π_k. But then (87) implies that either π or π_k violates incentive-compatibility (31), 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 (87) as:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \\ \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, \delta, s^l, e_{l+1}) w(e_k, s^{k-1}, \delta) \\ > \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, s^l, e_{l+1}) w(e_k, s^{k-1}). \tag{88}$$

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

$$w(e_k, s^{k-1}, \delta) \leq w(e_k, \delta(s^{k-1})), \tag{89}$$

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 (89) in (88) gives:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t | \pi^\star, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \\ \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, \delta, s^l, e_{l+1}) w(e_k, \delta(s^{k-1})) \\ > \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t | \pi^\star, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1} | \pi^\star, s^l, e_{l+1}) w(e_k, s^{k-1}). \tag{90}$$

The outcome function $\pi^\star$ enters (90) 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 (90) as:

$$\sum_{t=l+1}^{k-1} \beta^t \left[\sum_{H^t} p(h^t|\pi, \delta, s^l, e_{l+1}) u(e_t + \tau_t - \delta_a(h^{t-1}, e_t, \tau_t, a_t)) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi, \delta, s^l, e_{l+1}) w(e_k, \delta(s^{k-1}))$$
$$> \sum_{t=l+1}^{k-1} \beta^t \left[\sum_{S^t} p(s^t|\pi, s^l, e_{l+1}) u(e_t + \tau_t - a_t) \right] + \beta^k \sum_{E, s^{k-1}} p(e_k, s^{k-1}|\pi, s^l, e_{l+1}) w(e_k, s^{k-1}). \tag{91}$$

But now the left-hand side of (91) 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 (91) contradicts the incentive compatibility of π. We obtain a contradiction, $\pi^\star$ actually satisfies (31). The proof for constraints (32) follows the same lines. 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 Propositions 2 and 4 established that Program 1 is equivalent to the general planning problem, and Lemmas 1 and 3 do the same for Program 2. The programs are therefore equivalent to each other.

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

Proof of Proposition 6 We want to show that constraints (53)-(57) 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 (53), the promise-keeping constraints (22) and (54), and the obedience constraints (23) and (55) 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 (56) and (57) 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 (56) over all $\hat{a}$:

$$v(\hat{e}, e, \tau, a) \equiv \max_{\hat{a}} \left\{ \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e') \right] \right\}. \tag{92}$$

Then clearly (56) is satisfied, since the left-hand side of (56) 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 (56):

$$\hat{\delta}(\tau, a) \equiv \text{argmax}_{\hat{a}} \left\{ \sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e') \right] \right\}. \tag{93}$$

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

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e} + \tau - \hat{\delta}(\tau, a)) + \beta \sum\nolimits_E p(e'|\hat{\delta}(\tau, a)) w'(e') \right] \leq w(\hat{e}). \tag{94}$$

By the way we chose $\hat{\delta}(\cdot)$ and the $v(\cdot)$, we have from (56):

$$\sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{\delta}(\tau, a)) + \beta \sum\nolimits_E p(e'|\hat{\delta}(\tau, a)) w'(e') \right] = v(\hat{e}, e, \tau, a). \tag{95}$$

Substituting the left-hand side into (94), we get:

$$\sum_{T,A} v(\hat{e}, e, \tau, a) \leq w(\hat{e}). \tag{96}$$

which is (57).

Conversely, suppose we have found a lottery $\pi(\tau, a, \mathbf{w}'|\mathbf{w}, e)$ that satisfies (56) and (57) in Program 3 for some choice of $v(\hat{e}, e, \tau, a)$. By (56), we have then for any $\hat{e}$ and $\hat{a}$ and hence any $\delta : T \times S \to A$:

$$\sum_{\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \left[u(\hat{e} + \tau - \delta(\tau, a)) + \beta \sum\nolimits_E p(e'|\delta(\tau, a)) w'(e') \right] \leq v(\hat{e}, e, \tau, a). \tag{97}$$

Substituting the left-hand side of (97) into the assumed (57) for the $v(\hat{e}, e, \tau, a)$, we maintain the inequality:

$$\sum_{T,A,\mathbf{W}'} \pi(\tau,a,\mathbf{w}'|\mathbf{w},e)\left[u(\hat{e}+\tau-\delta(\tau,a))+\beta\sum\nolimits_{E} p(e'|\delta(\tau,a))w'(e')\right]\leq w(\hat{e}). \tag{98}$$

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

Lemma 4 *The allocations that can be implemented in Program 3 and 4 are identical.*

Proof of Lemma 4 Assume that for a given vector of promised utilities $\mathbf{w}$ and a given true state e we have found an allocation $\pi^{\star}(\tau,a,\mathbf{w}'|\mathbf{w},e)$ that satisfies the constraints (53)-(57) of Program 3 for some specification of utility bounds $v^{\star}(\hat{e},e,\tau,a)$. Here and below the * denotes a solution to Program 3. We now need to find contracts and utility bounds for Programs 4a and 4b that implement the same allocation. The contracts that we choose have the property that the planner does not randomize over intermediate utilities in Program 4a. Given a transfer τ, one specific vector of intermediate utilities is assigned in Program 4a. Given this transfer and intermediate utility vector, in Program 4b the correct distribution over storage and future utilities is implemented. We will first describe which lotteries over storage and promised utilities to choose in stage 4b, and then move backwards to Program 4a.

At stage 4b, the planner recommends an action and assigns a vector of promised utilities, conditional on the transfer and the vector of interim utilities that were assigned at 4a. Since we will choose contracts at stage 4a that do not randomize over interim utilities, for each transfer τ we have to specify only one contract at stage 4b. Fix the transfer τ. We want to find a contract $\pi(a,\mathbf{w}')$ that implements the same distribution over a and $\mathbf{w}'$, conditional on τ, as $\pi^{\star}(\tau,a,\mathbf{w}'|\mathbf{w},e)$. In order to do this, define:

$$\pi(a,\mathbf{w}')\equiv\frac{\pi^{\star}(\tau,a,\mathbf{w}'|\mathbf{w},e)}{\sum_{A\times\mathbf{W}'}\pi^{\star}(\tau,a,\mathbf{w}'|\mathbf{w},e)}. \tag{99}$$

We also have to define the utility bounds. Our choices are:

$$v(\hat{e},e,\tau,a)\equiv\frac{v^{\star}(\hat{e},e,\tau,a)}{\sum_{A\times\mathbf{W}'}\pi^{\star}(\tau,a,\mathbf{w}'|\mathbf{w},e)}. \tag{100}$$

These choices also correspond to a specific vector of interim utilities. We choose these utilities to ensure that all constraints in Program 4b are satisfied. We write these utilities as a function of τ so that we can identify which interim utilities correspond to which transfers in Program 4a. Specifically, we let:

$$w_m(e)(\tau) \equiv \sum_{A,\mathbf{W}'} \pi(a,\mathbf{w}') \left[u(e+\tau-a) + \beta \sum\nolimits_E p(e'|a)w'(e')\right], \tag{101}$$

where the $\pi(a,\mathbf{w}')$ are defined in (99). For each $\hat{e}$, we define

$$\bar{w}_m(\hat{e},e)(\tau) \equiv \sum_A v(\hat{e},e,\tau,a), \tag{102}$$

where the $v(\hat{e},e,\tau,a))$ are defined in (100). We now have to show that all constraints Program 4b are satisfied. We can use the fact that $\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)$ and the $v^\star(\hat{e},e,\tau,a)$ satisfy all constraints in Program 3.

We will start with the probability constraint (63). Substituting (99) into the left-hand side of (63), we get:

$$\frac{\sum_{A\times\mathbf{W}'}\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)}{\sum_{A\times\mathbf{W}'}\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)} = 1, \tag{103}$$

which on the left-hand side clearly equals unity. Therefore (63) is satisfied. The promise-keeping constraint (64) is identical to (101) and therefore also satisfied. We now move to the obedience constraint (65). Substituting our definition (99) into what one would like to show, namely (65), we need to establish:

$$\begin{aligned}&\sum_{\mathbf{W}'} \frac{\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)}{\sum_{A\times\mathbf{W}'}\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)} \left[u(e+\tau-a) + \beta\sum\nolimits_E p(e'|a)w'(e')\right] \\ &\quad\geq \sum_{\mathbf{W}'} \frac{\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)}{\sum_{A\times\mathbf{W}'}\pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e)} \left[u(e+\tau-\hat{a}) + \beta\sum\nolimits_E p(e'|\hat{a})w'(e')\right],\end{aligned}$$

or, equivalently, by pulling out the common constant in the denominator:

$$\begin{aligned}&\sum_{\mathbf{W}'} \pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e) \left[u(e+\tau-a) + \beta\sum\nolimits_E p(e'|a)w'(e')\right] \\ &\quad\geq \sum_{\mathbf{W}'} \pi^\star(\tau,a,\mathbf{w}'|\mathbf{w},e) \left[u(e+\tau-\hat{a}) + \beta\sum\nolimits_E p(e'|\hat{a})w'(e')\right].\end{aligned} \tag{104}$$

This is (55) in Program 3 and therefore by assumption satisfied. Next, substituting our definitions (99) and (100) into what we would like to show, namely (66), and multiplying by $\sum_{A\times\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)$ on both sides we need to establish:

$$\sum_{\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e) \ \left[u(\hat{e} + \tau - \hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e')\right] \leq v^\star(\hat{e}, e, \tau, a), \tag{105}$$

But this is (56) and therefore by assumption true. Finally, the constraints (67) are satisfied because we defined the utility bounds in (102) accordingly. Therefore in Program 4b all constraints are satisfied.

We now move to stage 4a. As mentioned earlier, for a given τ, the contract we use does not randomize over interim utilities. Since we have already determined which interim utilities correspond to which transfer τ, the only thing left to do is to assign probability mass to different transfers and their associated interim utilities. We choose the contract such that the probability of each transfer equals the probability of that transfer in contract $\pi^\star(\cdot)$. For each τ, we have:

$$\pi(\tau, w_m(e)(\tau), \bar{\mathbf{w}}_m(\hat{e}, e)(\tau)) \equiv \sum_{A,\mathbf{W}} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e). \tag{106}$$

We set the probability of all other combinations of transfer and interim utilities to zero. Clearly, by choosing this contract we implement the same allocation over transfer, storage, and promised utilities as $\pi^\star(\cdot)$ does. We still have to check whether the constraints (59)-(61) of Program 4a are satisfied. The first constraint (59) is the probability constraint. We have:

$$\begin{aligned} \sum_T \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e)) &= \sum_T \pi(\tau, w_m(e)(\tau), \bar{\mathbf{w}}_m(\hat{e}, e)(\tau)) \qquad (107) \\ &= \sum_{T,A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e). \end{aligned}$$

since the right-hand side of (107) is just a sum over the left-hand side of (106). Since the $\pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)$ satisfy (53) we have, as required:

$$= 1. \tag{108}$$

For the promise-keeping constraint (60) we start on its left-hand side and use (99) and

(101) to get:

$$\begin{aligned}
&\sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e))\ w_m(e) \\
&\qquad = \sum_{T} \pi(\tau, w_m(e)(\tau), \bar{\mathbf{w}}_m(\hat{e}, e)(\tau))\ w_m(e)(\tau) \\
&\qquad = \sum_{T,A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ w_m(e)(\tau) \\
&\qquad = \sum_{T,A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ \left[u(e + \tau - a) + \beta \sum\nolimits_{E} p(e'|a) w'(e')\right] \\
&\qquad = w(e),
\end{aligned} \tag{109}$$

where the last equality follows because the $\pi^\star(\cdot)$ satisfy (54). Finally, for the truth-telling constraint (61), using (106) and starting on its left-hand side, given any $\hat{e} \neq e$ we get:

$$\begin{aligned}
&\sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi(\tau, w_m(e), \bar{\mathbf{w}}_m(\hat{e}, e))\ \bar{w}_m(\hat{e}, e) \\
&\qquad = \sum_{T} \pi(\tau, w_m(e)(\tau), \bar{\mathbf{w}}_m(\hat{e}, e)(\tau))\ \bar{w}_m(\hat{e}, e)(\tau) \\
&\qquad = \sum_{T,A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ \bar{w}_m(\hat{e}, e)(\tau) \\
&\qquad = \sum_{T,A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ \left[\sum_{A} v(\hat{e}, e, \tau, a)\right]
\end{aligned}$$

by (102),

$$= \sum_{T,A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ \left[\frac{\sum_{A} v^\star(\hat{e}, e, \tau, a)}{\sum_{A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)}\right]$$

because of (100). The $\pi(\cdot)$ will cancel, that is:

$$\begin{aligned}
&= \sum_{T} \left[\sum_{A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ \left[\frac{\sum_{A} v^\star(\hat{e}, e, \tau, a)}{\sum_{A,\mathbf{W}'} \pi^\star(\tau, a, \mathbf{w}'|\mathbf{w}, e)}\right]\right] \\
&= \sum_{T,A} v^\star(\hat{e}, e, \tau, a) \\
&\leq w(\hat{e}).
\end{aligned} \tag{110}$$

The last inequality follows because the $v^\star(\cdot)$ by assumption satisfy (57). Thus all constraints are satisfied. Therefore an allocation that is feasible in Program 3 can always be implemented in Program 4.

We will now show conversely that an allocation that is feasible in Program 4 is also feasible in Program 3. For a given vector of promised utilities $\mathbf{w}$ and a given true endowment, assume that we have found contracts and utility bounds $\pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)$, $\pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m)$, and $v^\star(\hat{e}, e, \tau, a)(\tau, w_m, \bar{\mathbf{w}}_m)$ that satisfy all constraints in Programs 4a and 4b. We use * to denote a solution to Program 4a and 4b and write the contracts and utility bounds for Program 4a as a function of transfer and interim utilities to indicate to which transfer and interim utility vector the contracts applies to. In this section, we drop the arguments from the interim utilities in order to fit the equations on the page. We now have to find contracts and utility bounds for Program 3 that implement the same allocation and that satisfy constraints (53) to (57). For each τ, a, $\mathbf{w}'$, define as the obvious guess:

$$\pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) \equiv \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\ \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m). \tag{111}$$

Basically, we integrate out the interim utilities. We guess the following utility bounds:

$$v(\hat{e}, e, \tau, a) \equiv \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\ v^\star(\hat{e}, e, \tau, a). \tag{112}$$

Clearly, the chosen contract implements the same allocation. We have to show that the contract and the utility bounds satisfy the constraints (53)-(57) of Program 3.

We start with the probability constraint (53). Using definition (111), we get:

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e) = \sum_{T,A,\mathbf{W}} \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\ \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m).$$

Rewriting the order of summation:

$$= \sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m) \left[\sum_{A,\mathbf{W}'} \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m) \right].$$

Using the probability constraint (63) in Program 4b gives:

$$=\sum_{T}\sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m) = 1, \tag{113}$$

where, because of the probability constraint (59) in Program 4a, the left-hand side of (113) equals unity, and thus (53) is satisfied. We now will show that the promise-keeping constraint (54) holds. Starting on its left-hand side and using definition (111) we get:

$$\sum_{T,A,\mathbf{W}'} \pi(\tau, a, \mathbf{w}'|\mathbf{w}, e)\ \left[u(e+\tau-a) + \beta \textstyle\sum_{E} p(e'|a)w'(e')\right]$$
$$= \sum_{T,A,\mathbf{W}'}\sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m)\ \pi^{\star}(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m)\ \left[u(e+\tau-a) + \beta \textstyle\sum_{E} p(e'|a)w'(e')\right].$$

Rearranging the order of summation gives:

$$= \sum_{T}\sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m)$$
$$\left[\sum_{A,\mathbf{W}'} \pi^{\star}(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m)\ \left[u(e+\tau-a) + \beta \textstyle\sum_{E} p(e'|a)w'(e')\right]\right].$$

Using the promise-keeping constraint (64) from Program 4b gives:

$$= \sum_{T}\sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m)\ w_m = w(e), \tag{114}$$

where for the last step we used the promise-keeping constraint (60) from Program 4a.

Thus promise keeping is satisfied as well. Substituting definition (111) into both sides of what we hope will be the obedience constraint (55), we get:

$$\sum_{\mathbf{W}'}\sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m)\ \pi^{\star}(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m)\ \left[u(e+\tau-a) + \beta \textstyle\sum_{E} p(e'|a)w'(e')\right]$$
$$\geq \sum_{\mathbf{W}'}\sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m)\ \pi^{\star}(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m)\ \left[u(e+\tau-\hat{a}) + \beta \textstyle\sum_{E} p(e'|\hat{a})w'(e')\right]. \tag{115}$$

Rearranging the order of summation yields:

$$\sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m) \left[\sum_{\mathbf{W}'} \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m) \left[u(e+\tau-a) + \beta \sum\nolimits_E p(e'|a) w'(e')\right]\right]$$
$$\geq \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m) \left[\sum_{\mathbf{W}'} \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m) \left[u(e+\tau-\hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e')\right]\right]. \tag{116}$$

This inequality holds because of (65) term by term before summing over $\mathcal{W}(e, \tau)$. Next, substituting our definitions (111) and (112) into what we hope will be constraint (56), we want to establish that:

$$\sum_{\mathbf{W}'} \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\, \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m) \left[u(\hat{e}+\tau-\hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e')\right]$$
$$\leq \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\, v^\star(\hat{e}, e, \tau, a). \tag{117}$$

Rearranging gives:

$$\sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m) \left[\sum_{\mathbf{W}'} \pi^\star(a, \mathbf{w}')(w_m, \bar{\mathbf{w}}_m) \left[u(\hat{e}+\tau-\hat{a}) + \beta \sum\nolimits_E p(e'|\hat{a}) w'(e')\right]\right]$$
$$\leq \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\, v^\star(\hat{e}, e, \tau, a), \tag{118}$$

which holds because by (66) it is assumed to hold pointwise in Program 4b. Finally, starting from the left-hand side of the hoped for (57), we use definition (112) to get:

$$\sum_{T,A} v(\hat{e}, e, \tau, a) = \sum_{T,A} \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m)\, v^\star(\hat{e}, e, \tau, a).$$

Changing the order of summation gives:

$$= \sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi^\star(\tau, w_m, \bar{\mathbf{w}}_m) \left[\sum_{A} v^\star(\hat{e}, e, \tau, a)\right],$$

because of the assumed (67) in Program 4b this is:

$$
\begin{aligned}
&\leq \sum_{T} \sum_{\mathcal{W}(e,\tau)} \pi^{\star}(\tau, w_m, \bar{\mathbf{w}}_m)\ \bar{w}_m(\hat{e}, e) \\
&\leq w(\hat{e}),
\end{aligned}
\tag{119}
$$

where the last inequality follows from (61). Thus all constraints are satisfied, which completes the proof. Since the feasible allocations are identical in both Programs, and the objective function is the same, this also implies that Program 3 and 4 have the same solution. □

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. □

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

Proof of Proposition 8 To show that an element $\mathbf{w}$ of $\bar{\mathbf{W}}$ is in $\mathbf{W}$, we have to find outcome probabilities $\pi(\tau_t, a_t|e_t, s^{t-1})$ 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 3.2. $\square$

References

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

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