# Sequential multidimensional screening

Sina Litterscheid and Dezsö Szalay*
University of Bonn

July 8, 2013
Preliminary

**Abstract**

We study a sequential screening problem where the agent produces multiple items and has a multidimensional type that he learns over time. With multiple payoff relevant parameters and action choices, the optimal contract does not necessarily induce truthtelling off equilibrium path. Instead, sequentially optimal lying strategies need to be discouraged. The resulting optimal mechanism displays nonstandard features such as upward distortions. We apply our results to the optimal timing of productive decisions; if postponing all production decisions is feasible, then it is always preferred unless the principal's utility of consuming the goods is independent of the amount of other goods consumed.

## 1 Introduction

Contracting is often plagued by problems of asymmetric information. E.g., sellers have typically better information about costs of production. A vast literature has studied such problems. The majority of contributions is built around two central assumptions. First, the seller's private information can be captured by a one-dimensional parameter. Second,

*Send communications to szalay@uni-bonn.de or to sina.litterscheid@uni-bonn.de or by ordinary mail to Dezsö Szalay, Institute for Microeconomics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany

the seller is endowed from the outset with his private information. The literature on multidimensional mechanism design relaxes the first assumption. The literature on sequential and dynamic mechanism design relaxes the second assumption. We study a problem that combines both approaches.

Specifically, we have the following contracting problem in mind. A buyer wishes to consume a bundle of two goods and contracts with a seller to trade these goods. The seller has some private information about his costs of production already at the outset; however, as time goes by, he learns more about his costs of production. More precisely, the bundle of goods consists of $x$ units of a first good and $y$ units of a second good. At the outset, the seller knows the marginal costs of raising quanitity $x$, but he does not know yet what the marginal cost of raising $y$ will be. However, knowing the marginal cost of $x$ may provide some information as to the likely marginal costs of $y$. Building on the Revelation Principle for dynamic games (Myerson (1986)), the optimal way to write contracts in the present context can be analyzed by means of a game where the seller is asked to announce each piece of information as it arrives. So, at the outset, the seller is asked to announce his marginal costs of producing the first good in quantity $x$; in the second period, the seller is asked to announce his marginal costs of producing the second good in quantity $y$.

This model is a convex combination of the multidimensional model analyzed by Armstrong and Rochet (1999) and the sequential screening models studied by Baron and Besanko (1984) and Courty and Li (2000). As in Armstrong and Rochet (1999) we take the marginal cost parameters from binary distributions, so that there are 2x2 cost realizations in the model. In contrast to their model, the seller has only part of his information at the outset, so that the information that arrives in each period is one-dimensional. As in Courty and Li (2000), the early information may carry some information about future parameter realizations. More, precisely, Courty and Li (2000) study the case where consumers learn early on a distribution of their taste parameter and later on refine this information to a particular parameter value. We depart from this assumption in that we allow the early information to have direct payoff consequences in addition to affecting the distribution of preference parameters.

Among the questions that our model allows us to address are the following:

How does the presence of an additional payoff relevant parameter affect the solution procedure? What is the set of implementable outcomes for our model? What behaviour does an optimal contract induce in our model; that is, in what sense is truthtelling at all nodes of the game still an equilibrium outcome of the game? A key insight of our modeling exercise is that truthtelling is not always optimal off the equilibrium path - obviously, truthtelling is always optimal on equilibrium path. What new features of contracts can be explained if we allow for this richer setup, where complicated, dynamic lying strategies present a binding constraint on contract design?

Off-equilibrium-path-lies have been analyzed in only a small literature, so it is in order to explain their relevance in detail. In our model, there are two rounds of reporting. In the first round of reporting, the seller faces a choice between being truthful or not. The buyer designs contracts in a way that make sure that the seller has incentives to announce his true marginal cost of raising $x$. Given he does so, then contracts make sure that he announces the second piece of information - his marginal cost of raising $y$ - again truthtully in the second round of reporting. To put this differently, on equilibrium path, reporting is truthful. However, what if the seller chose to lie in the first round of reporting? What will the seller chose to report in the second round of reporting? Truthful reporting is nothing natural at all to expect in that case. The reason is that truthfulness usually is a shortcut for saying that choices reflect the preference parameter of the agent with private information. The key insight is that truthfulness in the second round of reporting only reflects the preference parameter of the agent in case he was truthful already in the first period. If in contrast the agent lied in the first round, then his preference parameters are different from the one of the agent for whom contracts are designed to match his preferences.

At first glance, one is tempted to argue that off path lies are irrelevant to contract design, precisely because they are not supposed to be observed if all players stick to their equilibrium strategies. However, this is not quite right. As in every screening model, the seller receives rents arising from his private information, in particular the private information that prevails already at the outset of the contracting game. The level of rents that the seller is able to capture depends on the mechanism and in particular on how attractive a lie in the first round of communication is. When the seller contemplates a possible deviation in the first period, he needs to make a plan about an optimal report also in the second round of communication. In other words, he needs to formulate a sequentially rational lying strategy.

Our contribution is a systematic analysis of optimal mechanisms in this largely unstudied framework. In particular, we are able to delineate precisely under what circumstances dynamic lying strategies influence optimal contracts and under what circumstances the best strategy available to the seller satisfies a one-period deviation principle. To fix ideas we look at the case of positively correlated types. In this case, the optimal mechanism can be found imposing truthtelling constraints on equilibrium path exclusively only when the goods are complementary to the buyer, but not too stongly so. For the case of relatively weak substitutability, the mechanism still satisfies truthtelling constraints. However, among the binding constraints is now also a truthtelling constraint off equilibrium path. Finally, for the case of strong complements or strong substitutes, the optimal mechanism induces lying off equilibrium path.

The intuition for these results is very simple. Truthtelling off equilibrium path is neither a necessary implication of incentive compatibility nor is it any desirable feature per se. Rather, the buyer trades off the benefits and costs of inducing different kinds of behavior off path. It may simply be impossible to implement truthtelling after a first round lie or it may be cheaper to implement a sequentially rational lie.

There are two main lessons from this excercise. Firstly, we provide conditions such that the standard contracts (that are derived from truthtelling constraints on equilibrium path only) are fully optimal. This finding is reassuring in that it says that not everything is automatically different as soon as we depart from the standard context of the sequential screening literature. Secondly, our exercise allows us to learn something about qualitative properties of contracting solutions that we would not be able to spot in the standard framework. In particular, while the standard approach delivers (in our context) the familiar downward distortions in economic activity, the case of binding off path constraints as well as the case where lying is optimal off path usually display some upward as well as some downward distortions. Thus, this paper can be seen at the same time as a robustness analysis of the benchmark case where early information is not ultimately payoff relevant later on, as well as providing a taxonomy of other phenomena that the standard framework does not allow us to explain. We believe these results are an important complement to other approaches that study the more tractable cases.

Going beyond that, we apply our methodology to address optimal timing in the sequential screening problem. The literature usually looks at the case where the timing of allocation choices is given. Using our methods, we can endogenize the optimal timing of decisions, if that is flexible. In particular, we show that there is an option value to waiting. If allocation variables are chosen based on early information only, then the designer loses the flexibility that comes with designing the whole production scheme jointly. We show that this option value is strictly positive except for the case where the goods are independent. In that case, the timing of production becomes irrelevant.

The related literature is as follows. The simplest way to see our problem is as a combination of Armstrong and Rochet (1999), who study a tractable model of multidimensional screening[1], and Courty and Li (2000), who analyze a model of sequential screening, where a consumer first learns the distribution of his taste parameter and later on learns the precise realization of the taste parameter. In contrast to Armstrong and Rochet (1999), the agent in our context learns his information over time, as in Courty and Li (2000). However, unlike in the latter paper, the information that the agent receives early on is not only about the distribution of his preference parameter but also directly payoff relevant. While the formulation in Courty and Li (2000) definitely comes in more handy, we point out that the model can rationalize a substantially wider variety of allocations once we allow for direct payoff effects of early information.

Sequentially optimal lies are also analyzed in Eső and Szentes (2007a,b, 2013). In their framework, an agent who misreported early information will also misreport information that arrives later on. More specifically, the agent undoes his earlier lie so that he receives the same allocation as if he had been truthful at each instance. This is different in our context, and hence optimal allocations reflect different trade-offs. Krähmer and Strausz (2008) provide an analysis of the case where it is impossible to undo an earlier lie in the Courty and Li (2000) model, because the support of late information depends on the realization of early information. As in our model, the optimal mechanims induces at times lies off the equilibrium path. However, the model and questions they address are quite different from ours. For more recent analyses of sequential screening models, see also Boleslavsky and Said (2012), Krähmer and Strausz (2012,2013) and Li and Shi (2013).

Obviously, our results are closely related to Armstrong and Rochet (1999) who provide

[1]For a survey of multidimensional screening, see Rochet and Stole (2003).

a taxonomy of binding constraints in the static two-by-two-dimensional model. In contrast to their analysis, we stick to the case of positively correlated types. On the other hand, we allow for the case of complements and substitutes, while Armstrong and Rochet (1999) look at independent goods. We are not aware of any literature that attacks the standard multidimensional model (that is twodimensional) model sequentially.

A question related to our timing application is addressed in Krähmer and Strausz (2012), where it is shown that ex post participation constraints eliminate the value of sequential screening in that there is bunching with respect to early information. In that sense, the principal could simply wait for definite information to arrive and not screen until then. Note that this is different in our context where early information is directly payoff relevant; not screening early would expose the principal to a static multidimensional screening problem later on; hence, this is suboptimal in our model.

Closely related to sequential screening is the literature on dynamic mechanism design. Baron and Besanko (1984) and Battaglini (2005) provide the first general analysis of optimal contracts in this dynamic framework. Battaglini (2005) studies monopolistic selling in context where consumer's tastes follow a Markov process. He shows that allocations satisfy a generalized no distortion at the top property. Moreover, a central building block of the analysis is that the most tempting deviations are single period deviations. Pavan et al. (2012) provide a general model of dynamic mechanism design. In each period, new information arrives and the designer chooses a set of allocation variables as a function of current information and past reports. Again, a central block of their analysis is to establish a version of the one-stage-deviation-principle. More specifically, they show that in their context, the most tempting deviation strategy consists in a single period lie and reverting to truthful reporting after that. We complement this approach by looking at a case where the one stage deviation principle does not always apply. In particular, it applies for the case of weak complements in our model, but not otherwise. As a result, we are able to rationalize allocations that would be impossible to explain when the one-stage deviation principle does apply.

Complementary to this paper is contemporanous work by Battaglini and Lamba (2013) who argue that there are important interactions between the regularity conditions imposed on the screening problem and the length of the time horizon. In particular, in the dynamic screening problem separation may not be feasible even though it would be feasible in the static counterpart of the model. Unlike in our model it is the within period incentive constraints that become binding beyond the local ones; in our model, within period incentive compatibility poses no problem, but the across periods incentive constraints become binding beyond the local ones. Similar to the present approach, their analysis allows them to explain allocations that could not be rationalized using local constraints only.

We view our results as an important complement to the general approaches based on local incentive constraints: our results demonstrate the robustness and the limitations of models that lack the richness we allow for. We hope our results sharpen our awareness of what we are usually assuming away. However, since our analysis cannot be extended to richer type spaces, there is clearly no hope to develop a general approach based on our methodology. So, our approach is definitely a complement, not a substitute for more tractable approaches.

We clearly do not do justice to many papers that we do not mention here. The reader is deferred to Pavan et al. (2012) for an extensive survey of the literature on dynamic mechanism design.

The paper is organized as follows. In section two, we present the model and state the buyer's problem. Sections three presents the buyer's problem and presents some preliminary observations. Section 4 breaks up the problem into two steps, where in the first step we search for the minimal transfers and optimal off path behaviour that implements given allocations. Section 5 presents and solves a reduced problem where some of the incentive constraints are dropped and provides conditions under which the reduced problem solves the buyer's overall problem. Section 6 discusses the structure of optimal allocations in regular cases where the strength of complementarity/substitutability of goods in the buyer's utility function is limited. Section 7 gives an example that is outside this regular structure. In Section 8, we discuss an application of our approach to the optimal timing of productive decisions in the sequential context. The final section concludes. The final section concludes. Lengthy proofs are gathered in the appendix.

## 2 The Model

A buyer contracts with a supplier to obtain two goods in quantities $x$ and $y$. The buyer's utility is

$$V(x, y) - T,$$

where $T$ is a transfer made to the seller. The seller's payoff is

$$T - \theta x - \eta y,$$

where $\theta$ and $\eta$ are cost shifters.

Contracting is a sequential process. At date 1, the seller knows $\theta$, whereas the buyer only knows that $\theta \in \{\underline{\theta}, \overline{\theta}\}$, where $\overline{\theta} > \underline{\theta} > 0$, and that $\Pr(\theta = \underline{\theta}) = \alpha$. $\eta$ is not known initially, not even to the seller. However, both the seller and the buyer know that $\eta \in \{\underline{\eta}, \overline{\eta}\}$. The seller's date 1 information also carries some information about $\eta$ in the sense that the conditional distribution of $\eta$ given $\theta$ depends on $\theta$; let $\lambda(\theta) \equiv \Pr\{\eta = \underline{\eta} | \theta\}$. At date 2, $\eta$ becomes known to the seller but not to the buyer. Finally, the goods are produced and traded in exchange for transfer $T$.

We place no assumptions on $V(x, y)$ for the time being except that $V(x, y)$ is jointly concave in $x$ and $y$ and that the first unit of consumption is extremely valuable to the buyer, that is $\lim_{x \to 0} V_x(x, y) = \infty$ for all $y$ and $\lim_{y \to 0} V_y(x, y) = \infty$ for all $x$. Further assumptions will be discussed as we go along.

## 3 The Buyer's Problem

Invoking the appropriate revelation principle (Myerson (1986)), it is without loss of generality to analyze optimal contracting in terms of direct, incentive compatible mechanisms. Our contracting game is dynamic. In the first period of contracting, incentive compatibility is equivalent to truthfulness about the agent's first period information, $\theta$. In the second period, the revelation principle implies that without loss of generality, we can restrict messages to the incremental information that arrives in period two; that is, in period two, the sender chooses a message $\hat{\eta} \in \{\underline{\eta}, \overline{\eta}\}$. Moreover, messages are truthful on equilibrium path; that is, after a truthful report $\hat{\theta} = \theta$ in the first period, the revelation principle implies truthfulness of the second period report $\hat{\eta} = \eta$. The revelation principle does not have any implications as to reporting off equilibrium path, except for the fact that the agent chooses the optimal report to send as part of his strategy. So, to assess the value of a first period deviation, we need to consider the possibility that the optimal thing to do in the second period after a first period lie is to lie again. Since the second period optimal behavior of the agent depends on the first period report, the first period true type, and the second period true type, we need to distinguish between the incremental information that arrives in period two and the agent's private information. That is, in the second period, the agent privately knows which node, identified by the triple $\left(\theta, \hat{\theta}, \eta\right)$, in the game tree has been reached. We let $\hat{\eta}^*\left(\theta, \hat{\theta}, \eta\right)$ denote the optimal report at node $\left(\theta, \hat{\theta}, \eta\right)$. It is easy to show that the optimal mechanism is nonstochastic. This is because the principal is risk averse (with respect to lotteries over $x$ and/or $y$) while the agent only cares for the expected values of such lotteries. Even though the equilibrium concept is a bit different, the proof essentially follows from Myerson (1986).

We can now state the buyer's problem:

$$\max_{x(\cdot,\cdot),y(\cdot,\cdot),T(\cdot,\cdot),\hat{\eta}^*\left(\theta,\hat{\theta},\cdot\right)} \mathbb{E}_{\theta}\mathbb{E}_{\eta|\theta}\left[V\left(x\left(\theta,\eta\right),y\left(\theta,\eta\right)\right) - T\left(\theta,\eta\right)\right] \tag{1}$$

s.t.

$$T\left(\overline{\theta},\underline{\eta}\right) - \overline{\theta}x\left(\overline{\theta},\underline{\eta}\right) - \underline{\eta}y\left(\overline{\theta},\underline{\eta}\right) \geq T\left(\overline{\theta},\overline{\eta}\right) - \overline{\theta}x\left(\overline{\theta},\overline{\eta}\right) - \underline{\eta}y\left(\overline{\theta},\overline{\eta}\right), \tag{2}$$

$$T\left(\overline{\theta},\overline{\eta}\right) - \overline{\theta}x\left(\overline{\theta},\overline{\eta}\right) - \overline{\eta}y\left(\overline{\theta},\overline{\eta}\right) \geq T\left(\overline{\theta},\underline{\eta}\right) - \overline{\theta}x\left(\overline{\theta},\underline{\eta}\right) - \overline{\eta}y\left(\overline{\theta},\underline{\eta}\right) \tag{3}$$

$$T\left(\underline{\theta},\underline{\eta}\right) - \underline{\theta}x\left(\underline{\theta},\underline{\eta}\right) - \underline{\eta}y\left(\underline{\theta},\underline{\eta}\right) \geq T\left(\underline{\theta},\overline{\eta}\right) - \underline{\theta}x\left(\underline{\theta},\overline{\eta}\right) - \underline{\eta}y\left(\underline{\theta},\overline{\eta}\right) \tag{4}$$

$$T\left(\underline{\theta},\overline{\eta}\right) - \underline{\theta}x\left(\underline{\theta},\overline{\eta}\right) - \overline{\eta}y\left(\underline{\theta},\overline{\eta}\right) \geq T\left(\underline{\theta},\underline{\eta}\right) - \underline{\theta}x\left(\underline{\theta},\underline{\eta}\right) - \overline{\eta}y\left(\underline{\theta},\underline{\eta}\right) \tag{5}$$

$$\begin{aligned} & \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\underline{\theta},\eta\right) - \underline{\theta}x\left(\underline{\theta},\eta\right) - \eta y\left(\underline{\theta},\eta\right)\right] \\ \geq\ & \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right) - \underline{\theta}x\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right) - \eta y\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right)\right], \end{aligned} \tag{6}$$

$$\begin{aligned} & \mathbb{E}_{\eta|\overline{\theta}}\left[T\left(\overline{\theta},\eta\right) - \overline{\theta}x\left(\overline{\theta},\eta\right) - \eta y\left(\overline{\theta},\eta\right)\right] \\ \geq\ & \mathbb{E}_{\eta|\overline{\theta}}\left[T\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right) - \overline{\theta}x\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right) - \eta y\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right)\right], \end{aligned} \tag{7}$$

$$\mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\underline{\theta},\eta\right) - \underline{\theta}x\left(\underline{\theta},\eta\right) - \eta y\left(\underline{\theta},\eta\right)\right] \geq 0, \tag{8}$$

$$\mathbb{E}_{\eta|\overline{\theta}}\left[T\left(\overline{\theta},\eta\right)-\overline{\theta}x\left(\overline{\theta},\eta\right)-\eta y\left(\overline{\theta},\eta\right)\right]\geq 0, \tag{9}$$

and for $\theta \neq \hat{\theta}$

$$\hat{\eta}^*\left(\theta,\hat{\theta},\eta\right)\in \arg\max_{\hat{\eta}} T\left(\hat{\theta},\hat{\eta}\right)-\theta x\left(\hat{\theta},\hat{\eta}\right)-\eta y\left(\hat{\theta},\hat{\eta}\right) \tag{10}$$

for all $\theta, \hat{\theta} \in \left\{\underline{\theta}, \overline{\theta}\right\}$ and $\eta \in \left\{\underline{\eta}, \overline{\eta}\right\}$.

Constraints (2) through (3) are the second period on equilibrium path constraints: after a truthful report in period one, the seller must find it optimal to be truthful about $\eta$ as well. (6) and (7) are the first period incentive constraints. As of date one, the seller anticipates that he chooses the second period report optimally, as captured by (10). (8) and (9) are the participation constraints.

The problem is relatively rich and requires careful analysis. To avoid useless case distinctions, we place enough structure on the problem to ensure that the low cost producer in the first period is better to the buyer than the high cost producer. This obviously depends on the correlation between costs. We impose the following assumption:

**Assumption 1:** costs are weakly positively correlated, that is $\lambda\left(\underline{\theta}\right) \geq \lambda\left(\overline{\theta}\right)$.

The assumption amounts to a regularity condition commonly used in the sequential screening literature. The alternative implies a lot of hassle due to further case distinctions one has to go through; however, our solution procedure can be readily adapted to this case.

Assumption 1 implies the following Lemma, which is of course the reason to impose it in the first place:

**Lemma 1** *If* $\lambda\left(\underline{\theta}\right) \geq \lambda\left(\overline{\theta}\right)$, *then* (8) *is automatically satisfied if* (9) *is.*

The argument is essentially the same as in a static two-type model. We can use the first period incentive constraint (6) to show that an allocation that satisfies (9) automatically also satisfies (8).

Clearly, at least one participation constraint must be binding; otherwise all payments could be lowered and the buyer's payoff could be increased. From Lemma 1 we can deduce that constraint (9) is binding at the optimum. Likewise, at least one of the ex ante incentive constraints must be binding. Otherwise we could again reduce some payments in a way that keeps incentive compatibility satisfied and increases the buyer's expected payoff. It is easy to see that the critical constraint is $(6)$. Which other constraints bind is a relatively complex matter. The reason is that the implications of optimal off-path reporting are quite intricate. We begin with a discussion of the implications of the on-equilibrium path incentive constraints.

**Lemma 2** *$\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\underline{\eta}$ for $x\left(\overline{\theta},\underline{\eta}\right)\geq x\left(\overline{\theta},\overline{\eta}\right)$ and $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\overline{\eta}$ for $x\left(\underline{\theta},\underline{\eta}\right)\geq x\left(\underline{\theta},\overline{\eta}\right)$. Likewise, $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)=\overline{\eta}$ for $x\left(\overline{\theta},\underline{\eta}\right)\leq x\left(\overline{\theta},\overline{\eta}\right)$ and $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\underline{\eta}$ for $x\left(\underline{\theta},\underline{\eta}\right)\leq x\left(\underline{\theta},\overline{\eta}\right)$.*

The on path constraints have some, however limited, implications for the optimal reports off path. In particular, it is *never* the case that *all* off path types find it optimal to lie in the second period. Depending on the monotonicity properties of the $x-$allocation, there is always some types that will automatically - that is, by implications of the on-path constraints - find it optimal to report their second period incremental information truthfully. The difficulty at this stage is of course that the monotonicity of the $x-$allocation with respect to $\eta$ is not known and endogenous.

As should be obvious from Lemma 2, the optimal mechanism does not necessarily induce truthtelling about $\eta$ off equilibrium path. Technically, this is due to the fact that both $\theta$ and $\eta$ directly enter the agent's payoff in period two. Therefore, the optimal report off equilibrium path - where optimal carries both the meaning of incentive compatible from the perspective of the agent and cost minimizing from the perspective of the principal - becomes a design variable in addition to quantities and payments.

We solve our problem as follows. We aim for a reduced problem, where constraints (9) and (6) hold as equalities, while (7) is slack. Moreover, we solve the reduced problem in a two step procedure, where we determine at step one the cheapest way to implement a given allocation and then determine the optimal allocation in step two. In the problem solved in the first step we simultaneously optimize over payments and off path reports.

# 4 Implementing given allocations at lowest cost

Payments to types $\left(\overline{\theta}, \eta\right)$ and optimal off-path reporting by types $\left(\underline{\theta}, \overline{\theta}, \eta\right)$ for $\eta \in \left\{\underline{\eta}, \overline{\eta}\right\}$ solve the following problem:

$$\Delta \equiv \min_{\left\{T\left(\overline{\theta},\eta\right),\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right\}_{\eta\in\left\{\underline{\eta},\overline{\eta}\right\}}} \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right) - \underline{\theta}x\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right) - \eta y\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right)\right)\right] \tag{11}$$

$$s.t.$$

$$\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\eta\right) \in \arg\max_{\hat{\eta}} T\left(\overline{\theta},\hat{\eta}\right) - \underline{\theta}x\left(\overline{\theta},\hat{\eta}\right) - \eta y\left(\overline{\theta},\hat{\eta}\right) \text{ for } \eta \in \left\{\underline{\eta},\overline{\eta}\right\}$$

$$\mathbb{E}_{\eta|\overline{\theta}}\left[T\left(\overline{\theta},\eta\right) - \overline{\theta}x\left(\overline{\theta},\eta\right) - \eta y\left(\overline{\theta},\eta\right)\right] = 0,$$

$$(2)\,, \text{ and } (3)\,.$$

The buyer minimizes the rent that needs to be given to the seller with ex ante type $\underline{\theta}$, taking into account that the optimal reporting strategy of this type in period two can be to misreport his parameter $\eta$ when he has misreported his parameter $\theta$ in the first period. However, if the buyer wishes to implement such a sequential lying strategy - simply because expected payments can be reduced this way - then he needs to explicitly make sure that the strategy is optimal from the seller's perspective as well.

Once the solution to the first program is found, we can choose payments to types $(\underline{\theta}, \eta)$ and the optimal reporting by types $\left(\overline{\theta}, \underline{\theta}, \eta\right)$ for $\eta \in \left\{\underline{\eta}, \overline{\eta}\right\}$ to render constraint (7) as slack as can be. Formally, for the payments and reports $T\left(\overline{\theta}, \eta\right), \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)$ that solve the first program, payments and reports $T\left(\underline{\theta}, \eta\right), \hat{\eta}^*\left(\overline{\theta}, \underline{\theta}, \eta\right)$ solve the problem:

$$\Omega \equiv \min_{\left\{T\left(\underline{\theta},\eta\right),\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right\}_{\eta\in\left\{\underline{\eta},\overline{\eta}\right\}}} \mathbb{E}_{\eta|\overline{\theta}}\left[T\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right) - \overline{\theta}x\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right) - \eta y\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right)\right] \tag{12}$$

$$s.t.$$

$$\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right) \in \arg\max_{\hat{\eta}} T\left(\underline{\theta},\hat{\eta}\right) - \overline{\theta}x\left(\underline{\theta},\hat{\eta}\right) - \eta y\left(\underline{\theta},\hat{\eta}\right) \text{ for } \eta \in \left\{\underline{\eta},\overline{\eta}\right\}$$

$$\mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\underline{\theta},\eta\right) - \underline{\theta}x\left(\underline{\theta},\eta\right) - \eta y\left(\underline{\theta},\eta\right)\right] = \Delta,$$

$$4, \text{ and } (5)\,,$$

The solution to these two programs depends obviously on the allocation that the buyer wishes to implement. In particular, define the following sets

$$\begin{aligned}
\mathbb{X}_a &\equiv \left\{(x,y) \mid \left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right) \geq \left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right) \geq 0\right\};\\
\mathbb{X}_b &\equiv \left\{(x,y) \mid \left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right) \geq \left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right) \geq 0\right\};\\
\mathbb{X}_c &\equiv \left\{(x,y) \mid -\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right) \geq \left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right) \geq 0\right\};\\
\mathbb{X}_d &\equiv \left\{(x,y) \mid \left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right) \geq -\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right) \geq 0\right\}.
\end{aligned}$$

For future reference, also define $\mathbb{X}_i^{int}$ as these same sets when all the defining inequalities are strict. These sets are depicted in the following graph:

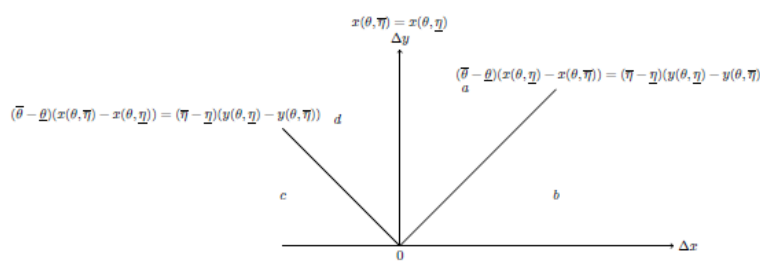


Figure 1: The space of implementable allocations is divided into four regions, a through d. The cost minimizing payments that implement allocations within each regime depend on the regime itself.

Only $y$-allocations that are monotonic in $\eta$ are incentive compatible. Hence, we only need to consider such allocations. From Lemma 2 we know that depending on the monotonicity of the $x$-allocation, one type reports automatically truthful also off path. Finally, whether it is optimal to have the other type report truthfully depends both on the $x-$ and on the $y-$ allocation. We state the important properties of the solution in the following Lemma. In particular, these are the values of the optima of the implementation problems stated above.

Moreover, for completeness we also state the optimal off-path reporting strategies. We have the following results:

**Lemma 3** *For* $(x,y)\in\mathbb{X}_a$

$$\begin{aligned}\Delta &= \Delta_a \equiv \mathbb{E}_{\eta|\underline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\eta\right)\right]+\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)\\ \Omega &= \Omega_a \equiv \left(\overline{\theta}-\underline{\theta}\right)\left(\mathbb{E}_{\eta|\underline{\theta}}\left[x\left(\overline{\theta},\eta\right)\right]-\mathbb{E}_{\eta|\overline{\theta}}\left[x\left(\underline{\theta},\eta\right)\right]\right)-\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)\end{aligned}$$

*and all types report truthfully off path;*
*for* $(x,y)\in\mathbb{X}_b$

$$\begin{aligned}\Delta &= \Delta_b \equiv \left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)+\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)-\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right)\\ \Omega &= \Omega_b \equiv \lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\overline{\eta}\right)-\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)-\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\overline{\eta}\right)+\Delta_b\end{aligned}$$

*and* $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)=\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\underline{\eta}$ *and* $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\overline{\eta}$*;*
*for* $(x,y)\in\mathbb{X}_c$

$$\begin{aligned}\Delta &= \Delta_c \equiv \left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)+\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)-\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right)\\ \Omega &= \Omega_c \equiv \left\{-\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)-\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right)+\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\overline{\eta}\right)+\Delta_c\right\}\end{aligned}$$

*and* $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)=\overline{\eta}$ *and* $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\underline{\eta}$*;*
*for* $(x,y)\in\mathbb{X}_d$

$$\begin{aligned}\Delta &= \Delta_d \equiv \left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)+\mathbb{E}_{\eta|\overline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\eta\right)\right]\\ \Omega &= \Omega_d \equiv -\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)-\mathbb{E}_{\eta|\underline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\eta\right)\right]+\Delta_d\end{aligned}$$

*and all types report truthfully off path and off path types* $\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)$ *and* $\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)$*.are indifferent between truthfully reporting and lying off path.*

The intuition is straightforward and can best be explained with the help of the graph. Allocations in $\mathbb{X}_a$ induce truthtelling off path automatically in the sense that we can naïvely assume truthtelling off path, that is, neglect any off path constraints altogether and simply impose truthtelling off path. For allocations in $\mathbb{X}_d$ such a naïve conjecture would prove to be false; the seller would not report truthfully off path if we simply took such behavior as given. While the optimal report off path by types $\left(\underline{\theta}, \overline{\theta}, \overline{\eta}\right)$ and $\left(\overline{\theta}, \underline{\theta}, \underline{\eta}\right)$ is indeed to tell the truth, this needs to be ensured explicitly with the appropriate constraints for types $\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)$ and $\left(\overline{\theta}, \underline{\theta}, \overline{\eta}\right)$. Moreover, these constraints are binding at the optimum. Finally, when dependencies of the $x-$allocation on information $\eta$ becomes strong, it becomes too costly to insist on truthtelling by all types off path. Recall that there is no reason at all to believe in truthtelling per se off equilibrium path. Instead, the cheapest way to implement any given allocation in sets $\mathbb{X}_b$ and $\mathbb{X}_c$ induces some type to lie off path.

There is one subtlety that we deemphasize but wish to mention nevertheless to avoid confusion. The sets $\mathbb{X}_i$ for $i = a, b, c, d$ are defined for both $\left(x\left(\underline{\theta}, \eta\right), y\left(\underline{\theta}, \eta\right)\right)$ and $\left(x\left(\overline{\theta}, \eta\right), y\left(\overline{\theta}, \eta\right)\right)$. The reason is that the dividing lines between the sets have isomorphic representations. To get a complete description of the implementable set of allocations one would have to define sets $\mathbb{X}_i\left(\theta\right)$; the complete set of implementable allocations is then given by $\mathbb{X}_i\left(\underline{\theta}\right) \times \mathbb{X}_j\left(\overline{\theta}\right)$ for $i, j = a, b, c, d$. However, it turns out that the solutions of the overall problem have the property that $i = j$. Therefore, to economize on space, we just present our result anticipating this result.

# 5 Optimal allocations in the reduced problem

We can now turn to the design of the optimal allocations. We address this question first in the reduced problems where we neglect constraint (7) altogether. Since the agent's rent depends on the allocation that is chosen, we need to perform this step of optimization separately for each set $\mathbb{X}_i\left(\overline{\theta}\right)$. Since we are neglecting constraint $(7)$, we allow for any $\left(x\left(\underline{\theta}, \eta\right), y\left(\underline{\theta}, \eta\right)\right) \in \mathbb{X}\left(\underline{\theta}\right)$, where $\mathbb{X}\left(\underline{\theta}\right) \equiv \left\{x\left(\underline{\theta}, \eta\right), y\left(\underline{\theta}, \eta\right) \middle| \, y\left(\underline{\theta}, \underline{\eta}\right) \geq y\left(\underline{\theta}, \overline{\eta}\right)\right\}$ is the union of all sets $\mathbb{X}_i\left(\underline{\theta}\right)$ for $i = a, b, c, d$. Formally, the reduced problem for each constraint set is

$$W_i \equiv \max_{\substack{\left(x\left(\overline{\theta}, \eta\right), y\left(\overline{\theta}, \eta\right)\right) \in \mathbb{X}_i\left(\overline{\theta}\right) \\ \left(x\left(\underline{\theta}, \eta\right), y\left(\underline{\theta}, \eta\right)\right) \in \mathbb{X}\left(\underline{\theta}\right)}} \mathbb{E}_\theta \mathbb{E}_{\eta|\theta}\left[V\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) - \theta x\left(\theta, \eta\right) - \eta y\left(\theta, \eta\right)\right] - \alpha \Delta_i. \qquad (\mathrm{P}_i)$$

The overall optimum for the buyer is

$$W = \max\left\{W_a, W_b, W_c, W_d\right\}.$$

The solution has the following simple structure:

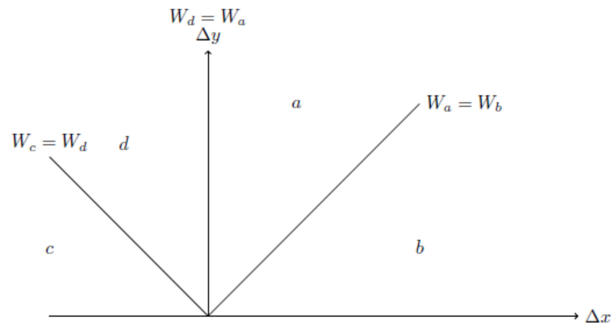


Figure 2: Along the dividing line between any two regimes, payoffs from adjacent programs are equal.

**Proposition 1** *If* $-V_{11}(x,y)\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}} \geq V_{12}(x,y) \geq V_{11}(x,y)\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}$ *for all* $x, y$*, then* $W = \max\{W_a, W_d\}$. *Moreover,* $W_d > W_a$ *if* $V_{12} < 0$ *for all* $x, y$ *and* $W_a \geq W_d$ *if* $V_{12} \geq 0$ *for all* $x, y$.

The intuition is straightforward and easiest to understand with the help of figure 2.

The idea to prove the results is as follows. The payoffs in the various regimes have a continuity structure that is displayed in the figure. For allocations that are feasible in two regions, say region a and region d, the payoffs from programs $\mathrm{P}_a$ and $\mathrm{P}_d$ are identical for a given allocation. Formally, we have $W_a = W_d$ for allocations that satisfy $x(\overline{\theta}, \underline{\eta}) = x(\overline{\theta}, \overline{\eta})$. Moreover, none of the programs $\mathrm{P}_i$ is ever so constrained that an allocation in the origin of the diagram is implemented. Hence, we can use simple revealed preference arguments to prove payoff dominance in the cases described in the proposition. For $V_{12} < 0$, the solution to program $\mathrm{P}_a$ actually does satisfy $x(\overline{\theta}, \underline{\eta}) = x(\overline{\theta}, \overline{\eta})$, whereas the solution to program $\mathrm{P}_d$ does not. Since, the allocation that maximizes program $\mathrm{P}_a$ is feasible also under program $\mathrm{P}_d$, but is not chosen, it follows by strict concavity of the problem that the value of the objective under program $\mathrm{P}_d$ is strictly higher. Likewise, for $V_{12} \geq 0$, the solution to program $\mathrm{P}_d$ satisfies $x(\overline{\theta}, \underline{\eta}) = x(\overline{\theta}, \overline{\eta})$, so the same argument can be made. However, the subtle difference in this case is that the optimal allocation under program $\mathrm{P}_a$ might also lie on the feasibility constraint $x(\overline{\theta}, \underline{\eta}) = x(\overline{\theta}, \overline{\eta})$. Hence, we can only establish payoff dominance in the weak sense.

The proposition not only compares payoffs between programs $\mathrm{P}_a$ and $\mathrm{P}_d$ but between all programs $\mathrm{P}_i$. Complements versus substitutes are enough to determine whether program $\mathrm{P}_a$ or program $\mathrm{P}_d$ gives a higher payoff. To rule out that the optimum is attained by programs $\mathrm{P}_c$ and $\mathrm{P}_b$, we need to place bounds on the strength of interactions between the goods. In the case of substitutes, if the substitutability is not too strong, in the sense that $V_{12}(x,y) \geq V_{11}(x,y)\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}$ for all $x, y$, then the solution to program $\mathrm{P}_c$ is on the dividing line between $\mathbb{X}_c\left(\overline{\theta}\right)$ and $\mathbb{X}_d\left(\overline{\theta}\right)$. However, on the dividing line we have $W_c = W_d$, so the payoff attained by program $\mathrm{P}_c$ could also be obtained by program $\mathrm{P}_d$. However, the solution to program $\mathrm{P}_d$ is not necessarily on the dividing line between $\mathbb{X}_c\left(\overline{\theta}\right)$ and $\mathbb{X}_d\left(\overline{\theta}\right)$. If the solution to program $\mathrm{P}_d$ is in the strict interior of $\mathbb{X}_d\left(\overline{\theta}\right)$, then $W_d > W_c$. By a similar argument, we can rule out that the principal ever wishes to implement an allocation that is in the set $\mathbb{X}_b\left(\overline{\theta}\right)$ if the complementarity between the goods is not too strong.

Conceptually, there is nothing special in the cases where the interaction between goods becomes strong. One can use the exact same methods to determine the optimum in these cases as well. However, the conditions needed to ensure that the reduced problem solves the overall problem become quite strong. So, we content ourselves showing that the optimum lies in the particular region for a special case of our model below. For the time being, we impose:

**Assumption 2:** $-V_{11}(x,y)\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}} \geq V_{12}(x,y) \geq V_{11}(x,y)\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}$ for all $x, y$.

The reader may verify that Assumption 2 captures a relevant parameter restriction with the help of the following example:

**Example 1** $V(x,y) = \beta^2 - \frac{1}{2}(x-\beta)^2 - \frac{1}{2}(y-\beta)^2 + \delta xy.$

In the example, Assumption 2 is satisfied for $\delta \in \left[-\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}, \frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}\right]$. Note that the utility function is jointly concave in $x$ and $y$ for $\delta \in (-1,1)$. Thus, depending on the relative variation in $\eta$ relative to $\theta$, the set of parameter values that *violate* Assumption 2 becomes empty. Conversely, there is always a nonempty set of parameter values that generate a concave buyer problem and satisfy Assumption 2. In this sense - at least in this example - Assumption 2 isolates the important case rather than the pathological one.

### 5.1 The solution to the full problem

Our results so far refer to the solution of a reduced problem where constraint (7) is dropped from the problem. Obviously, the reduced problem is of interest only if it solves the overall problem; that is, if the solution of the reduced problem satisfies the neglected constraint, (7).

Checking the neglected constraint requires knowing which off path behavior is implemented, which in turn requires knowing which type of allocation the principal wishes to implement. It is convenient to address this question by first asking in which of the sets the first-best allocation lies. This is useful for two reasons. First, the allocation for ex ante type $\underline{\theta}$ that maximizes the reduced problem corresponds simply to the first-best allocation for that type. Second, as $\alpha$ goes to zero, the second best allocation converges to the first-best allocation. Hence, by continuity, the second best allocation is close to the first-best allocation for $\alpha$ small.

The first-best allocation satisfies

$$V_1\left(x\left(\theta,\eta\right), y\left(\theta,\eta\right)\right) = \theta \tag{13}$$

and

$$V_2\left(x\left(\underline{\theta},\eta\right), y\left(\underline{\theta},\eta\right)\right) = \eta \tag{14}$$

for $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ and $\eta \in \left\{\underline{\eta}, \overline{\eta}\right\}$.

**Lemma 4** *The first-best allocation defined by* (13) *and* (14) *satisfies* $(x, y) \in$

$$\begin{array}{ll}
\mathbb{X}_a & if\ 0 \leq V_{12}\left(x,y\right) \leq -\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}} V_{11}\left(x,y\right)\ \ for\ all\ x, y; \\
\mathbb{X}_b & if\ V_{12}\left(x,y\right) \geq -\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}} V_{11}\left(x,y\right)\ \ for\ all\ x, y \\
\mathbb{X}_d & if\ 0 \geq V_{12} \geq \frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)} V_{11}\ for\ all\ x, y; \\
\mathbb{X}_c & if\ V_{12} \leq \frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)} V_{11}\ for\ all\ x, y.
\end{array}$$

*Moreover, the first-best allocation is in the interior of these sets if the defining inequalities are strict.*

It is now straightforward to check whether the neglected constraint is satisfied, by combining lemmas 3 and 4. Notice in particular that the solution to the reduced problem satisfies

$\{x(\underline{\theta},\eta), y(\underline{\theta},\eta)\}_{\eta\in\{\underline{\eta},\overline{\eta}\}} \in \mathbb{X}_i(\underline{\theta})$ for $i = a, d$ if and only if $\{x(\overline{\theta},\eta), y(\overline{\theta},\eta)\}_{\eta\in\{\underline{\eta},\overline{\eta}\}} \in \mathbb{X}_i(\overline{\theta})$ for $i = a, d$. Obviously, this was the reason to economize on space in Lemma 3 in the first place.

The reduced problem picks up the overall optimum under natural conditions.

**Proposition 2** *The solution to the reduced problem solves the overall problem under Assumption 2 if in addition either*
*i)* $V_{12} = 0$ *or*
*ii)* $V_{12} \neq 0$ *for all* $(x, y)$, $V_{12}$ *does not change sign,*

$$\max_{x,y}\left|\frac{V_{12}}{V_{11}\dot{V}_{22}-V_{12}^2}(x,y)\right| \leq \frac{\overline{\theta}-\underline{\theta}}{\overline{\eta}-\underline{\eta}}\min_{x,y}\left|\frac{V_{22}}{V_{11}\dot{V}_{22}-V_{12}^2}(x,y)\right|,$$

*and either*
*a)* $x, y \in \mathbb{X}_a^{int}$, *or*
*b) or* $x, y \in \mathbb{X}_d^{int}$ *and in addition* $\lambda(\underline{\theta}) = \lambda(\overline{\theta})$.

If the complementarity/subsitutability between $x$ and $y$ is not too strong relative to the concavity of the problem, then the reduced problem picks up the overall optimum under natural conditions. Without further conditions, this is the case if the goods are literally independent. For the case of complements and substitutes, we need to impose further conditions. The proof works as follows. Suppose the reduced problem has a solution that is either in $\mathbb{X}_a^{int}$ (for the case of complements) or in $\mathbb{X}_d^{int}$ (for the case of substitutes). Then, the neglected constraint that we identified in Lemma 3 is satisfied under the conditions given in part ii of the proposition. For the case of substitutes, we need to assume independent types in addition to that. The maximizers of problems $\mathrm{P}_a$ and $\mathrm{P}_d$, respectively are indeed in the interior of their allowed sets under the assumptions that render the value of these programs higher than the value of the other program, respectively and the additional assumption that $\alpha$ is sufficiently small. The reason is that in the limit where $\alpha$ tends to zero, the second best allocation converges to the first-best allocation, whose properties we have described in Lemma 4. The condition in the proposition is satisfied in our example for $\delta \in \left[-\frac{\overline{\theta}-\underline{\theta}}{\overline{\eta}-\underline{\eta}}, \frac{\overline{\theta}-\underline{\theta}}{\overline{\eta}-\underline{\eta}}\right]$. This is consistent with Assumption 2 for any differences between types. If $\frac{\overline{\theta}-\underline{\theta}}{\overline{\eta}-\underline{\eta}} = 1$, then the conditions are identical; otherwise, one set is a strict subset of the other.

The complete description of the optimum obviously includes a characterization of the optimal payments as well. These payments can be found in the proof to Lemma 3. We are now ready to discuss the structure of optimal allocations and the optimal timing of production.

## 6 The structure of optimal allocations

We can now investigate how the optimal allocation depends qualitatively on the interaction between goods in the buyer's utility function. To discuss this question in the simplest possible case, we simply state the result for the case where the optimum for ex ante type $\overline{\theta}$ is an allocation in $\mathbb{X}_a^{int}\left(\overline{\theta}\right)$ and $\mathbb{X}_d^{int}\left(\overline{\theta}\right)$. In this case, the optimal allocation for ex ante type $\overline{\theta}$ satisfies

$$\begin{aligned} V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) &= \overline{\theta}+\frac{\alpha}{(1-\alpha)}\frac{\lambda_i}{\lambda\left(\overline{\theta}\right)}\left(\overline{\theta}-\underline{\theta}\right) \\ V_2\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) &= \underline{\eta} \end{aligned}$$

and

$$\begin{aligned} V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right) &= \overline{\theta}+\frac{\alpha}{(1-\alpha)}\frac{(1-\lambda_i)}{\left(1-\lambda\left(\overline{\theta}\right)\right)}\left(\overline{\theta}-\underline{\theta}\right) \\ V_2\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right) &= \overline{\eta}+\frac{\alpha}{(1-\alpha)}\frac{\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)}{\left(1-\lambda\left(\overline{\theta}\right)\right)}\left(\overline{\eta}-\underline{\eta}\right), \end{aligned}$$

where $\lambda_a=\lambda\left(\underline{\theta}\right)$ and $\lambda_d=\lambda\left(\overline{\theta}\right)$, while the optimal allocation for ex ante type $\underline{\theta}$ is given by (13) and (14).

For the case of complements, the optimal allocation for ex ante type $\overline{\theta}$ displays the standard downward distortions relative to the first best. For strictly positive complementarities, all allocation variables are strictly below the first best optimal levels. This is quite different for the case of substitutes, which displays both upward and downward distortions. In particular, $x\left(\overline{\theta},\underline{\eta}\right)$ is distorted downwards and as result, $y\left(\overline{\theta},\underline{\eta}\right)$ is distorded upwards.

## 7 The case of strong interactions

So far, we have characterized optimal allocations for regular cases, where the strenght of interactions between the goods is relatively mild. If the ratio $\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}$ is relatively large, then "most" utility functions will display relatively mild interactions between the goods in this sense. This lose statement can be given a very precise meaning in the concrete example of negative quadratic utility. For that case, all concave utility functions satisfy Assumption 2 if the support of second period information is wider than the support of first period information. On the other hand, if the reverse is true, then one can give natural examples, where an allocation outside the sets $\mathbb{X}_a \cup \mathbb{X}_d$ becomes optimal. Specifically, we have the following result:

**Proposition 3** *Suppose that* $\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}} < 1$ *and consider the quadratic utility function of Example 1 with* $\delta \in \left(\frac{\overline{\eta}-\underline{\eta}}{\overline{\theta}-\underline{\theta}}, 1\right)$. *For that utility function, for* $\alpha$ *sufficiently close to zero, the overall optimal allocation satisfies* $x, y \in \mathbb{X}_b$.

The idea is simple. Assume the conditions of Lemma 4 hold so that the first-best is in the interior of one of the sets $\mathbb{X}_i$ for $i = a, b, c, d$. Now consider an allocation that maximizes each of the problems $\mathrm{P}_i$ for $i = a, b, c, d$, where the set of allowed allocations is $\mathbb{X}_i$. Now let $\alpha$ go to zero. The solutions to the problems $\mathrm{P}_i$ converge uniformly to the first-best allocation as $\alpha$ tends to zero. Therefore, for $\alpha$ positive but small, if the first best allocation is an element of $\mathbb{X}_i$, then the solution to $\mathrm{P}_i$ is still an element of $\mathbb{X}_i$. Moreover, we show in the proof of the proposition, that the neglected constraint is satisfied for the example. Hence, we have shown that there are natural conditions such that an allocation in, say, $\mathbb{X}_b$ becomes optimal. Thus, it can be strictly optimal to induce lying off equilibrium path.

The optimal allocation for the case where $x, y \in \mathbb{X}_b^{int}$ has interesting features. The first-order conditions for the allocation offered to ex ante type $\overline{\theta}$ are as follows:

$$\begin{aligned} V_1\left(x\left(\overline{\theta},\underline{\eta}\right), y\left(\overline{\theta},\underline{\eta}\right)\right) &= \overline{\theta} + \frac{\alpha}{(1-\alpha)}\frac{1}{\lambda\left(\overline{\theta}\right)}\left(\overline{\theta}-\underline{\theta}\right) \\ V_2\left(x\left(\overline{\theta},\underline{\eta}\right), y\left(\overline{\theta},\underline{\eta}\right)\right) &= \underline{\eta} - \frac{\alpha}{(1-\alpha)}\frac{\left(1-\lambda\left(\underline{\theta}\right)\right)}{\lambda\left(\overline{\theta}\right)}\left(\overline{\eta}-\underline{\eta}\right) \end{aligned}$$

and

$$\begin{aligned} V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right) &= \overline{\theta} \\ V_2\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right) &= \overline{\eta}+\frac{\alpha}{(1-\alpha)}\left(\overline{\eta}-\underline{\eta}\right) \end{aligned}$$

This allocation displays upwards distortions in the quantity $y\left(\overline{\theta},\underline{\eta}\right)$, for given quantity $x\left(\overline{\theta},\underline{\eta}\right)$. Since we are considering complements, this upwards distortion does not arise simply as a compensating effect due to a downard distortion in $x\left(\overline{\theta},\underline{\eta}\right)$, but rather reflects the particular structure of binding incentive constraints for this particular case.

# 8 Sequential screening and the value of waiting

Our formulation departs from the standard setup in sequential screening problems as this literature typically studies problems where in each period a given choice has to be made. The problems in the sequential and dynamic mechanism design literature correspond in our model to the case where $x$ has to be chosen right away after $\theta$ is reported and $y$ is chosen only later when $\eta$ is reported. We now show that, if given the option of waiting, it is always better to keep the flexibility to adjust both allocation choices to both informational parameters, except for the case where the buyer's utility is separable in the two goods.

We can obtain the optimal mechanism with sequential production from our problem if we add the consistency requirement that

$$x\left(\theta,\underline{\eta}\right)=x\left(\theta,\overline{\eta}\right)\text{ for }\theta\in\left\{\underline{\theta},\overline{\theta}\right\}. \tag{15}$$

Technically, (15) is a consistency requirement in the sense that the level of $x$ can only depend on information that is available when the level of $x$ is chosen.

It is straightforward to see that off-path lies are not an issue under this constraint. The reason is that $\theta x\left(\hat{\theta}\right)$ is sunk by the time the report about $\eta$ needs to be made and moreover enters the seller's profit in an additively separable way. So, seller types who have lied in the past correspond to types with different fixed costs of producing the $y$ good. However, fixed costs do not change the seller's incentive to report about $\eta$. So, the on-path incentive constraints automatically ensure that reporting is truthful also off path.

It is also obvious that sequential production cannot do better than delaying production of both goods until all information is there. The reason is that we are simply adding another constraint, (15) , to the buyer's problem and thereby eliminate some flexibility off equilibrium path (precisely because the on-path constraints automatically imply a particular off-path behaviour).

Solving the transfer minimization problems (11) and (12) for given allocation choices $x$ and $y$, under the consistency condition (15) and its implication of truthfulness of path, we find that at the solutions to these problems constraints (2) and (9) and (6) and (5) are binding. Using the optimal payments, the buyer's problem of finding an optimal allocation can be written as

$$\max_{x(\theta),y(\theta,\eta)} \mathbb{E}_{\theta}\mathbb{E}_{\eta|\theta}\left[V\left(x\left(\theta\right),y\left(\theta,\eta\right)\right)-\theta x\left(\theta\right)-\eta y\left(\theta,\eta\right)\right]$$
$$-\alpha\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta}\right)+\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)\right]$$

Moreover, the neglected incentive constraint (7) is equivalent to

$$\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)\geq\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta}\right)-x\left(\underline{\theta}\right)\right).$$

The following proposition is now obvious:

**Proposition 4** *Delayed and early production achieve the same payoff only for independent goods. For $V_{12}\left(x,y\right)\neq 0$ for all $x,y$, delayed production is strictly better than early production.*

The proof of the statement follows from the discussion in an obvious way and is therefore omitted. The logic is simply that the allocation under sequential production is always feasible under delayed production of both goods but is not chosen at the optimum, except for the case of independent goods.

It is instructive to take a closer look into the losses associated to sequential production. The allocation offered to ex ante type $\underline{\theta}$ is first-best efficient; that is, there is no distortion at the top. The allocation offered to ex ante type $\overline{\theta}$ satisfies the first-order conditions

$$\mathbb{E}_{\eta|\overline{\theta}}\left[V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\eta\right)\right)\right]=\overline{\theta}+\frac{\alpha}{1-\alpha}\left(\overline{\theta}-\underline{\theta}\right),$$

$$V_2\left(x\left(\overline{\theta}\right), y\left(\overline{\theta}, \underline{\eta}\right)\right) = \underline{\eta},$$

and

$$V_2\left(x\left(\overline{\theta}\right), y\left(\overline{\theta}, \overline{\eta}\right)\right) = \overline{\eta} + \frac{\alpha}{1-\alpha} \frac{\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)}{1 - \lambda\left(\overline{\theta}\right)} \left(\overline{\eta} - \underline{\eta}\right).$$

The expected marginal benefit of $x\left(\overline{\theta}\right)$ is equal to $\overline{\theta} + \frac{\alpha}{1-\alpha}\left(\overline{\theta} - \underline{\theta}\right)$. For given allocation $y\left(\overline{\theta}, \eta\right)$, this corresponds to the standard result that $x\left(\overline{\theta}\right)$ is distorted downwards relative to the first-best. Likewise, for given allocation $x\left(\overline{\theta}\right)$, $y\left(\overline{\theta}, \underline{\eta}\right)$ is set efficiently, while $y\left(\overline{\theta}, \overline{\eta}\right)$ is distorted downwards. Whether the entire allocation is higher or lower than first-best depends on the nature of interactions between the goods. For the case of independent goods, the overall allocation relates exactly as stated to the first-best allocation.

For nonzero interactions between the goods, there are two sources of losses for the principal due to choosing $x$ early on. Firstly, it is simply the case that both allocation choices should be adjusted to both cost conditions. Secondly, as we have explained at great lengths, it is sometimes not optimal to insist on truthtelling off path when both $x$ and $y$ are chosen late. Intuitively, it becomes easier to screen the information in the second round of reporting when the principal has more screening instruments available.

It is interesting to note that in the case of weak substitutes in the sense of Proposition 2, the first-order conditions differ *only* in that the marginal utilities interact with each other; the virtual cost expressions on the right hand side are identical for both timing configurations.[2] It is then straightforward to see how the optimal allocations differ from each other in the more flexible regime with delayed production and in the regime with early production of $x$. For an allocation in the regime with delayed production in $\mathbb{X}_d^{int}\left(\overline{\theta}\right)$, we have that $y\left(\overline{\theta}, \underline{\eta}\right) > y\left(\overline{\theta}, \overline{\eta}\right)$ and $x\left(\overline{\theta}, \overline{\eta}\right) > x\left(\overline{\theta}, \underline{\eta}\right)$. If the $x-$allocation is now forced to take the common value $x\left(\overline{\theta}\right)$, then, heuristically, $x\left(\overline{\theta}, \overline{\eta}\right)$ is reduced while $x\left(\overline{\theta}, \underline{\eta}\right)$ is increased. Since the marginal utility of consuming $y$ still must take on the same value, the $y-$allocation has to respond more to $\eta$ than it does in the flexible regime. Hence, the variation in the level of $y$ is increased in response to the reduction in the variation in the level of $x$.

We find these observations interesting; however, it should be stressed that the optimal timing of decisions is an issue only if the production process is flexible in this regard. The

[2] In the case of complements, the virtual marginal cost of $x\left(\overline{\theta}, \underline{\eta}\right)$ is increased while the virtual marginal cost of $x\left(\overline{\theta}, \overline{\eta}\right)$ is decreased for given level of $y$.

literature we are aware of so far works under the assumption that there is no flexibility; in each period a production decision has to be taken. We show that, if there is flexibility and absent discounting, then the designer wants to postpone all production decisions until all information is present.

# 9 Conclusion

This paper solves a tractable two-dimensional model of screening when the agent produces two goods, knows one cost parameter from the outset, and learns a second one at some later date. Depending on whether the goods are complements or substitutes and on how strongly the goods interact, a different pattern of binding constraints arises at the optimum. For weak complements, we obtain a standard solution, where the principal only needs to worry about single deviations. As a result, the solution to the full problem could also be obtained by a naive procedure that simply imposes truthtelling at all nodes of the game, even at those that are not reached if the agent is truthful early on in the game. For weak substitutes, it is still true that the solution can be obtained by imposing truthtelling on and off equilibrium path. However, now a truthtelling constraint off equilibrium path is binding at the optimum. As a result, the solution displays both upward and downward distortions. Finally, in the case of strong interactions between the goods, it may become optimal for the principal to give up on truthtelling off path and let the agent lie again after a first lie.

# 10 Appendix

**Proof of Lemma 1.** Note first that at least one participation constraint must be binding; otherwise all payments could be reduced by the same amount, resulting in higher buyer surplus. To prove the statement, it suffices to show the standard result that (9) together with (6) imply (8). This is true if $\lambda(\underline{\theta}) \geq \lambda(\overline{\theta})$.

Let $u(\theta, \eta)$ denote equilibrium utility.

From $(6)$, we have

$$\begin{aligned}
& \mathbb{E}_{\eta|\underline{\theta}}\left[u\left(\underline{\theta}, \eta\right)\right] \\
= & \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\underline{\theta}, \eta\right) - \underline{\theta} x\left(\underline{\theta}, \eta\right) - \eta y\left(\underline{\theta}, \eta\right)\right] \\
\geq & \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)\right) - \underline{\theta} x\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)\right) - \eta y\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)\right)\right]
\end{aligned}$$

On the other hand

$$\begin{aligned}
& \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)\right) - \underline{\theta} x\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)\right) - \eta y\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \eta\right)\right)\right] \\
\geq & \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\overline{\theta}, \eta\right) - \underline{\theta} x\left(\overline{\theta}, \eta\right) - \eta y\left(\overline{\theta}, \eta\right)\right]
\end{aligned}$$

since $\hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)$ and $\hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \overline{\eta}\right)$ are chosen optimally. Moreover,

$$\begin{aligned}
& \mathbb{E}_{\eta|\underline{\theta}}\left[T\left(\overline{\theta}, \eta\right) - \underline{\theta} x\left(\overline{\theta}, \eta\right) - \eta y\left(\overline{\theta}, \eta\right)\right] \\
= & \mathbb{E}_{\eta|\underline{\theta}}\left[u\left(\overline{\theta}, \eta\right) + \left(\overline{\theta} - \underline{\theta}\right) x\left(\overline{\theta}, \eta\right)\right] \\
\geq & \mathbb{E}_{\eta|\underline{\theta}}\left[u\left(\overline{\theta}, \eta\right)\right],
\end{aligned}$$

where the last inequality follows since production is non-negative.

Hence, from $(6)$, we have that

$$\mathbb{E}_{\eta|\underline{\theta}}\left[u\left(\underline{\theta}, \eta\right)\right] \geq \mathbb{E}_{\eta|\underline{\theta}}\left[u\left(\overline{\theta}, \eta\right)\right].$$

Now, from (2) it is straightforward to see that

$$u\left(\overline{\theta}, \underline{\eta}\right) \geq u\left(\overline{\theta}, \overline{\eta}\right) + \left(\overline{\eta} - \underline{\eta}\right) y\left(\overline{\theta}, \overline{\eta}\right),$$

and thus $u\left(\overline{\theta}, \underline{\eta}\right) \geq u\left(\overline{\theta}, \overline{\eta}\right)$. Using $\lambda(\underline{\theta}) \geq \lambda(\overline{\theta})$, we have moreover that

$$\mathbb{E}_{\eta|\underline{\theta}}\left[u\left(\overline{\theta}, \eta\right)\right] \geq \mathbb{E}_{\eta|\overline{\theta}}\left[u\left(\overline{\theta}, \eta\right)\right].$$

(9) written in terms of equilibrium utilities amounts to

$$\mathbb{E}_{\eta|\overline{\theta}}\left[u\left(\overline{\theta}, \eta\right)\right] \geq 0,$$

which proves the claim. ■

**Proof of Lemma 2.** The proof is by direct inspection. We consider all four off-path types in sequence.

Recall that $u(\theta,\eta)$ denotes the equilibrium utility of type $(\theta,\eta)$.

Consider type $\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)$, that is an agent with preference parameters $\underline{\theta},\overline{\eta}$ who has sent a first period report $\hat{\theta}=\overline{\theta}$. By reporting $\hat{\eta}=\overline{\eta}$, he obtains utility

$$T\left(\overline{\theta},\overline{\eta}\right)-\underline{\theta}x\left(\overline{\theta},\overline{\eta}\right)-\overline{\eta}y\left(\overline{\theta},\overline{\eta}\right)=u\left(\overline{\theta},\overline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)$$

If he reports $\hat{\eta}=\underline{\eta}$, then he obtains utility

$$T\left(\overline{\theta},\underline{\eta}\right)-\underline{\theta}x\left(\overline{\theta},\underline{\eta}\right)-\overline{\eta}y\left(\overline{\theta},\underline{\eta}\right)=u\left(\overline{\theta},\underline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)-\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right).$$

Type $\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)$ prefers to report $\hat{\eta}=\overline{\eta}$ if

$$u\left(\overline{\theta},\overline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)\geq u\left(\overline{\theta},\underline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)-\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right)$$

From the on equilibrium path constraint 3, we know that

$$u\left(\overline{\theta},\overline{\eta}\right)\geq u\left(\overline{\theta},\underline{\eta}\right)-\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right).$$

adding $\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)$ to both sides we get

$$u\left(\overline{\theta},\overline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)\geq u\left(\overline{\theta},\underline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)-\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right),$$

which implies that $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)=\overline{\eta}$ if

$$x\left(\overline{\theta},\overline{\eta}\right)\geq x\left(\overline{\theta},\underline{\eta}\right).$$

It is easy to demonstrate the other results by the exact same procedure. In particular:

$\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\underline{\eta}$ follows from the on-path constraint (2) if $x\left(\overline{\theta},\underline{\eta}\right)\geq x\left(\overline{\theta},\overline{\eta}\right)$;

$\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\underline{\eta}$ follows from the on-path constraint (4) if $x\left(\underline{\theta},\overline{\eta}\right)\geq x\left(\underline{\theta},\underline{\eta}\right)$; and

$\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\overline{\eta}$ follows from the on-path constraint (5) if $x\left(\underline{\theta},\underline{\eta}\right)\geq x\left(\underline{\theta},\overline{\eta}\right)$. ■

**Proof of lemma 3.** We split the proof in two cases, depending on whether $x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)$ is nonnegative or nonpositive. For both cases, we first prove the part concerning the allocations of type $\overline{\theta}$. Afterwards we turn to the allocation for type $\underline{\theta}$.

Preliminaries:

For convenience, note that the on path constraints (2) – (5) can be rewritten as follows:

$$T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\geq\overline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)\tag{16}$$

$$T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\leq\overline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\overline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)\tag{17}$$

$$T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\geq\underline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\overline{\eta}\right)\right)\tag{18}$$

$$T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\leq\underline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\overline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\overline{\eta}\right)\right)\tag{19}$$

Likewise, the off path-constraints take the following form:

Type $\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)$ prefers to report $\hat{\eta}=\overline{\eta}$ if

$$T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\leq\underline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\overline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right).\tag{20}$$

and prefers to report $\hat{\eta}=\underline{\eta}$ if

$$T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\geq\underline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\overline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right).\tag{21}$$

Type $\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)$ prefers to report $\hat{\eta}=\underline{\eta}$ if

$$T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\geq\underline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)\tag{22}$$

and prefers to report $\hat{\eta}=\overline{\eta}$ if

$$T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\leq\underline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right).\tag{23}$$

Type $\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)$ prefers to report $\hat{\eta}=\underline{\eta}$ if

$$T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\geq\overline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\overline{\eta}\right)\right)\tag{24}$$

and prefers to report $\hat{\eta}=\overline{\eta}$ if

$$T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\leq\overline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\overline{\eta}\right)\right).\tag{25}$$

Type $\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)$ prefers to report $\hat{\eta}=\overline{\eta}$ if

$$T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\leq\overline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\overline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\overline{\eta}\right)\right)\tag{26}$$

and prefers to report $\hat{\eta} = \underline{\eta}$ if

$$T\left(\underline{\theta}, \underline{\eta}\right) - T\left(\underline{\theta}, \overline{\eta}\right) \geq \overline{\theta}\left(x\left(\underline{\theta}, \underline{\eta}\right) - x\left(\underline{\theta}, \overline{\eta}\right)\right) + \overline{\eta}\left(y\left(\underline{\theta}, \underline{\eta}\right) - y\left(\underline{\theta}, \overline{\eta}\right)\right). \tag{27}$$

Suppose that $\left(x\left(\overline{\theta}, \underline{\eta}\right) - x\left(\overline{\theta}, \overline{\eta}\right)\right) \geq 0$. By Lemma 2 this implies that $\hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right) = \underline{\eta}$. Adding the expected utility of the high type (which is zero by (9)) to the objective, we obtain the following problem:

$$\Delta \equiv \min_{\left\{T\left(\overline{\theta},\eta\right)\right\}_{\eta\in\{\underline{\eta},\overline{\eta}\}},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)} \left\{ \begin{array}{c} \left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left(T\left(\overline{\theta},\underline{\eta}\right)-\overline{\theta}x\left(\overline{\theta},\underline{\eta}\right)-\underline{\eta}y\left(\overline{\theta},\underline{\eta}\right)\right)+\lambda(\underline{\theta})\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right) \\ +(1-\lambda(\underline{\theta}))\left(T\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)\right)-\underline{\theta}x\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)\right)-\overline{\eta}y\left(\overline{\theta},\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)\right)\right) \\ -\left(1-\lambda\left(\overline{\theta}\right)\right)\left[T\left(\overline{\theta},\overline{\eta}\right)-\overline{\theta}x\left(\overline{\theta},\overline{\eta}\right)-\overline{\eta}y\left(\overline{\theta},\overline{\eta}\right)\right] \end{array} \right\}$$

$$\begin{gathered} \textit{subject to } (16), (17), \textit{ and} \\ \textit{either } (20) \textit{ if } \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \overline{\eta}\right) = \overline{\eta} \\ \textit{or } (21) \textit{ if } \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \overline{\eta}\right) = \underline{\eta}. \end{gathered}$$

Consider now both possible off path reports. If $\hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \overline{\eta}\right) = \overline{\eta}$, then the objective is

$$\min_{T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)} \left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)\left[T\left(\overline{\theta}, \underline{\eta}\right) - \overline{\theta}x\left(\overline{\theta}, \underline{\eta}\right) - \underline{\eta}y\left(\overline{\theta}, \underline{\eta}\right)\right] + \lambda\left(\underline{\theta}\right)\left(\overline{\theta} - \underline{\theta}\right)x\left(\overline{\theta}, \underline{\eta}\right)$$
$$- \left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)\left[T\left(\overline{\theta}, \overline{\eta}\right) - \overline{\theta}x\left(\overline{\theta}, \overline{\eta}\right) - \overline{\eta}y\left(\overline{\theta}, \overline{\eta}\right)\right] + \left(1 - \lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta} - \underline{\theta}\right)x\left(\overline{\theta}, \overline{\eta}\right)$$

subject to the constraints $(16), (17)$, and $(20)$. Note that (17) is automatically satisfied if (20) is. There exists a solution to the problem only if the constraint set is non-empty, that is, if the right-hand side of (20) is weakly larger than the right-hand side of $(16)$. This is the case for $(x, y) \in \mathbb{X}_a\left(\overline{\theta}\right)$. In this case (16) is binding. Using (9) and (16) to solve for the optimal payments, we have

$$\begin{pmatrix} T\left(\overline{\theta}, \underline{\eta}\right) \\ T\left(\overline{\theta}, \overline{\eta}\right) \end{pmatrix} = \begin{pmatrix} \overline{\theta}x\left(\overline{\theta}, \underline{\eta}\right) + \underline{\eta}y\left(\overline{\theta}, \underline{\eta}\right) + \left(1 - \lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta} - \underline{\eta}\right)y\left(\overline{\theta}, \overline{\eta}\right) \\ \overline{\theta}x\left(\overline{\theta}, \overline{\eta}\right) + \lambda\left(\overline{\theta}\right)\underline{\eta}y\left(\overline{\theta}, \overline{\eta}\right) + \left(1 - \lambda\left(\overline{\theta}\right)\right)\overline{\eta}y\left(\overline{\theta}, \overline{\eta}\right) \end{pmatrix}. \tag{28}$$

Substituting back into the objective we have obtain

$$\Delta_a = \left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta} - \underline{\eta}\right)y\left(\overline{\theta}, \overline{\eta}\right) + \mathbb{E}_{\eta|\underline{\theta}}\left[\left(\overline{\theta} - \underline{\theta}\right)x\left(\overline{\theta}, \eta\right)\right].$$

On the other hand, if $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)=\underline{\eta}$, then the problem is

$$\begin{gathered}\min\nolimits_{T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)}\left(1-\lambda\left(\overline{\theta}\right)\right)\left\{T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)\right\}\\ +\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)-\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right)\\ -\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\underline{\eta}y\left(\overline{\theta},\underline{\eta}\right)-\overline{\eta}y\left(\overline{\theta},\overline{\eta}\right)\right)\end{gathered}$$

subject to the constraints $(16)$, $(17)$, and $(21)$. The right-hand side of $(16)$ is weakly larger than the right-hand side of (21) for $(x,y)\in\mathbb{X}_b\left(\overline{\theta}\right)$ and the reverse is true for $(x,y)\in\mathbb{X}_a\left(\overline{\theta}\right)$. Clearly, the right-hand side of (17) is always larger than the right-hand side of $(16)$. Therefore, for $(x,y)\in\mathbb{X}_b\left(\overline{\theta}\right)$, at the solution of the problem, constraint (16) holds as an equality. It follows that for $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\overline{\eta}\right)=\underline{\eta}$ and $(x,y)\in\mathbb{X}_b\left(\overline{\theta}\right)$, the transfers can be taken from (28) so that the objective takes value

$$\Delta_b=\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)-\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right)+\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right).$$

For $(x,y)\in\mathbb{X}_a\left(\overline{\theta}\right)$ the right-hand side of (17) becomes smaller than the right-hand side of $(21)$, so the feasible set becomes empty. Hence, no solution exists.

Next consider the second problem for the case where $x\left(\underline{\theta},\underline{\eta}\right)\geq x\left(\underline{\theta},\overline{\eta}\right)$. By lemma 2 this implies that $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\overline{\eta}$.So, the problem can be written as

$$\Omega=\min_{\{T(\underline{\theta},\eta)\}_{\eta\in\{\underline{\eta},\overline{\eta}\}},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)}\left\{\begin{array}{c}\lambda\left(\overline{\theta}\right)\left[T\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right)-\overline{\theta}x\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right)-\underline{\eta}y\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\eta\right)\right)\right]\\ +\left(1-\lambda\left(\overline{\theta}\right)\right)\left[T(\underline{\theta},\overline{\eta})-\overline{\theta}x(\underline{\theta},\overline{\eta})-\overline{\eta}y(\underline{\theta},\overline{\eta})\right]\\ -\mathbb{E}_{\eta|\underline{\theta}}[T(\underline{\theta},\eta)-\underline{\theta}x(\underline{\theta},\eta)-\eta y(\underline{\theta},\eta)]+\Delta\end{array}\right\}$$

$$\begin{gathered}\text{subject to}\\ (18),(19),\ \text{and either}\\ (24)\ \text{if}\ \hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\underline{\eta},\ \text{or}\\ (25)\ \text{if}\ \hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\overline{\eta},\end{gathered}$$

where the objective is obtained from substituting the constraint (6) as an equality into the objective.

Consider first the case where $\hat{\eta}^*\left(\overline{\theta}, \underline{\theta}, \underline{\eta}\right) = \underline{\eta}$. In this case, the problem is

$$\begin{aligned}
&\min_{T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)} -\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left\{T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\right\} \\
&+\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left[\overline{\theta}x\left(\underline{\theta},\underline{\eta}\right)+\underline{\eta}y\left(\underline{\theta},\underline{\eta}\right)-\underline{\theta}x\left(\underline{\theta},\overline{\eta}\right)-\overline{\eta}y\left(\underline{\theta},\overline{\eta}\right)\right] \\
&-\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\overline{\eta}\right)-\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right)+\Delta.
\end{aligned}$$

subject to the constraints $(18), (19)$ and $(24)$. The right-hand side of (18) is weakly smaller than the right-hand side of $(24)$. Hence, the constraint set is nonempty if the right hand side of (24) is weakly smaller than the right-hand side of $(19)$, which is exactly true for $(x, y) \in \mathbb{X}_a\left(\underline{\theta}\right)$. So, in this case, (19) is binding at the solution to the problem. Solving for the transfers from (19) and $(6)$, we obtain

$$\begin{pmatrix} T\left(\underline{\theta},\underline{\eta}\right) \\ T\left(\underline{\theta},\overline{\eta}\right) \end{pmatrix} = \begin{pmatrix} \Delta+\underline{\theta}x\left(\underline{\theta},\underline{\eta}\right)+\left(\lambda\left(\underline{\theta}\right)\underline{\eta}+\left(1-\lambda\left(\underline{\theta}\right)\right)\overline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right) \\ \Delta+\underline{\theta}x\left(\underline{\theta},\overline{\eta}\right)+\overline{\eta}y\left(\underline{\theta},\overline{\eta}\right)-\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right) \end{pmatrix}. \tag{29}$$

Substituting these transfers back into the objective, we obtain

$$\Omega_a = -\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right)-\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)-\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\overline{\eta}\right)+\Delta.$$

For future reference, we note that for $(x, y) \in \mathbb{X}_a$, this can be written as

$$\Omega_a = \mathbb{E}_{\eta|\underline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\eta\right)\right]-\mathbb{E}_{\eta|\overline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\eta\right)\right]-\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right).$$

Consider next the case where $\hat{\eta}^*\left(\overline{\theta}, \underline{\theta}, \underline{\eta}\right) = \overline{\eta}$. In this case, the problem becomes

$$\begin{aligned}
&\min_{T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)} -\lambda\left(\underline{\theta}\right)\left\{T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)\right\} \\
&-\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\overline{\eta}\right)+\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\overline{\eta}\right) \\
&+\lambda\left(\underline{\theta}\right)\left[\underline{\theta}x\left(\underline{\theta},\underline{\eta}\right)-\underline{\theta}x\left(\underline{\theta},\overline{\eta}\right)+\underline{\eta}y\left(\underline{\theta},\underline{\eta}\right)-\overline{\eta}y\left(\underline{\theta},\overline{\eta}\right)\right]+\Delta.
\end{aligned}$$

subject to the constraints $(18), (19)$ and $(25)$. The right-hand side of (25) is larger than the right-hand side of (19) for $(x, y) \in \mathbb{X}_b\left(\underline{\theta}\right)$. In this case, the feasible set is nonempty and at the solution (19) is binding; hence the transfers are given by (29) and the objective takes value

$$\begin{aligned}
\Omega_b = {} & -\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)-\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\overline{\eta}\right)+\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\overline{\eta}\right) \\
&+\lambda\left(\underline{\theta}\right)\underline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\Delta.
\end{aligned}$$

Again, for future reference, if $(x, y) \in \mathbb{X}_b$, then we can write

$$\begin{aligned} \Omega_b &= -\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right) y\left(\underline{\theta}, \underline{\eta}\right)+\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right) y\left(\underline{\theta}, \overline{\eta}\right) \\ &\quad +\lambda\left(\underline{\theta}\right) \underline{\theta}\left(x\left(\underline{\theta}, \underline{\eta}\right)-x\left(\underline{\theta}, \overline{\eta}\right)\right)-\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right) y\left(\overline{\theta}, \underline{\eta}\right)+\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right) y\left(\overline{\theta}, \overline{\eta}\right). \end{aligned}$$

For $(x, y) \in \mathbb{X}_a\left(\underline{\theta}\right)$, the right-hand side of (25) is smaller than the right-hand side of (19). Moreover, the feasible set is always nonempty and thus at the solution constraint (25) is binding. Hence, we can substitute

$$T\left(\underline{\theta}, \underline{\eta}\right)-T\left(\underline{\theta}, \overline{\eta}\right)=\overline{\theta}\left(x\left(\underline{\theta}, \underline{\eta}\right)-x\left(\underline{\theta}, \overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\underline{\theta}, \underline{\eta}\right)-y\left(\underline{\theta}, \overline{\eta}\right)\right)$$

into the objective and obtain

$$\begin{aligned} \hat{\Omega}_a &= -\lambda\left(\underline{\theta}\right)\left\{\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\underline{\theta}, \underline{\eta}\right)-x\left(\underline{\theta}, \overline{\eta}\right)\right)+\left(\overline{\eta}-\underline{\eta}\right) y\left(\underline{\theta}, \overline{\eta}\right)\right\} \\ &\quad -\left(\overline{\theta}-\underline{\theta}\right) x\left(\underline{\theta}, \overline{\eta}\right)+\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right) y\left(\underline{\theta}, \overline{\eta}\right)+\Delta. \end{aligned}$$

We have $\Omega_a \leq \hat{\Omega}_a$ for $(x, y) \in \mathbb{X}_a\left(\underline{\theta}\right)$, so $\hat{\eta}^*\left(\overline{\theta}, \underline{\theta}, \underline{\eta}\right)=\underline{\eta}$ is cheaper to implement in that case.

Next consider the case where $\left(x\left(\overline{\theta}, \underline{\eta}\right)-x\left(\overline{\theta}, \overline{\eta}\right)\right) \leq 0$. By Lemma 2, this implies that $\hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \overline{\eta}\right)=\overline{\eta}$. Adding and subtracting the expected utility of type $\overline{\theta}$, we can write the objective as

$$\Delta \equiv \min_{\left\{T\left(\overline{\theta}, \eta\right)\right\}_{\eta \in\left\{\underline{\eta}, \overline{\eta}\right\}}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)}\left\{\begin{array}{c} \lambda(\underline{\theta})\left(T\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)\right)-\underline{\theta} x\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)\right)-\underline{\eta} y\left(\overline{\theta}, \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)\right)\right) \\ -\left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left(T\left(\overline{\theta}, \overline{\eta}\right)-\overline{\theta} x\left(\overline{\theta}, \overline{\eta}\right)-\overline{\eta} y\left(\overline{\theta}, \overline{\eta}\right)\right)+(1-\lambda(\underline{\theta}))\left(\overline{\theta}-\underline{\theta}\right) x\left(\overline{\theta}, \overline{\eta}\right) \\ -\lambda\left(\overline{\theta}\right)\left[T\left(\overline{\theta}, \underline{\eta}\right)-\overline{\theta} x\left(\overline{\theta}, \underline{\eta}\right)-\underline{\eta} y\left(\overline{\theta}, \underline{\eta}\right)\right] \end{array}\right\}$$

$$\begin{gathered} \textit{subject to} \\ (16),(17) \text{ and either} \\ (22) \text{ if } \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)=\underline{\eta}, \text{ or} \\ (23) \text{ if } \hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)=\overline{\eta}. \end{gathered}$$

Consider first the case where the off-path report is $\hat{\eta}^*\left(\underline{\theta}, \overline{\theta}, \underline{\eta}\right)=\underline{\eta}$. In this case, the objective is

$$\Delta \equiv \min_{T\left(\overline{\theta}, \underline{\eta}\right)-T\left(\overline{\theta}, \overline{\eta}\right)}\left\{\begin{array}{c} \left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left(T\left(\overline{\theta}, \underline{\eta}\right)-\overline{\theta} x\left(\overline{\theta}, \underline{\eta}\right)-\underline{\eta} y\left(\overline{\theta}, \underline{\eta}\right)\right)+\lambda(\underline{\theta})\left(\overline{\theta}-\underline{\theta}\right) x\left(\overline{\theta}, \underline{\eta}\right) \\ -\left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left(T\left(\overline{\theta}, \overline{\eta}\right)-\overline{\theta} x\left(\overline{\theta}, \overline{\eta}\right)-\overline{\eta} y\left(\overline{\theta}, \overline{\eta}\right)\right)+(1-\lambda(\underline{\theta}))\left(\overline{\theta}-\underline{\theta}\right) x\left(\overline{\theta}, \overline{\eta}\right) \end{array}\right\}$$

subject to the constraints $(16), (17)$, and $(22)$. The right-hand side of (22) is always at least as large as the right-hand side of (16) (by the fact that $\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\leq 0$). Hence, the constraint set is nonempty if the right-hand side of (17) is at least as large as the right-hand side of $(22)$, which is precisely the case for $(x,y)\in\mathbb{X}_d\left(\overline{\theta}\right)$. Since the objective is increasing in $T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)$ and we are minimizing $\Delta$, $T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)$ is set as small as possible, implying that (22) is binding. We can compute the transfers from (22) and $(9)$. We obtain

$$\begin{pmatrix} T\left(\overline{\theta},\underline{\eta}\right) \\ T\left(\overline{\theta},\overline{\eta}\right) \end{pmatrix} = \begin{pmatrix} -\lambda\left(\overline{\theta}\right)\left[\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\right]+\overline{\theta}x\left(\overline{\theta},\overline{\eta}\right)+\overline{\eta}y\left(\overline{\theta},\overline{\eta}\right)+\underline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)+\underline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right) \\ -\lambda\left(\overline{\theta}\right)\left[\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\right]+\overline{\theta}x\left(\overline{\theta},\overline{\eta}\right)+\overline{\eta}y\left(\overline{\theta},\overline{\eta}\right) \end{pmatrix}. \tag{30}$$

Substituting these transfers back into the objective, we obtain

$$\Delta_d \equiv \left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)\right)+\mathbb{E}_{\eta|\overline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\eta\right)\right].$$

For $(x,y)\in\mathbb{X}_c\left(\overline{\theta}\right)$, no solution with $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\underline{\eta}$ exists.

Suppose thus that $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\overline{\eta}$. In this case, the objective is

$$\Delta\equiv\min_{T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)}\left\{\lambda\left(\overline{\theta}\right)\left(T\left(\overline{\theta},\overline{\eta}\right)-\overline{\theta}x\left(\overline{\theta},\overline{\eta}\right)-\overline{\eta}y\left(\overline{\theta},\overline{\eta}\right)-\left[T\left(\overline{\theta},\underline{\eta}\right)-\overline{\theta}x\left(\overline{\theta},\underline{\eta}\right)-\underline{\eta}y\left(\overline{\theta},\underline{\eta}\right)\right]\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)+\lambda(\underline{\theta})\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)\right\}$$

subject to $(16), (17)$, and $(23)$. The right-hand side of (17) is weakly smaller than the right-hand side of (23) for $(x,y)\in\mathbb{X}_c$. Since the objective is decreasing in $T\left(\overline{\theta},\underline{\eta}\right)-T\left(\overline{\theta},\overline{\eta}\right)$ and we seek to minimize the objective function, at the optimum (17) mus be binding. Thus, we can compute the optimal transfers from (17) and $(9)$. We obtain

$$\begin{pmatrix} T\left(\overline{\theta},\underline{\eta}\right) \\ T\left(\overline{\theta},\overline{\eta}\right) \end{pmatrix} = \begin{pmatrix} \overline{\theta}x\left(\overline{\theta},\overline{\eta}\right)+\mathbb{E}_{\eta|\overline{\theta}}\eta y\left(\overline{\theta},\eta\right)+\lambda\left(\overline{\theta}\right)\left[\overline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\right] \\ \overline{\theta}x\left(\overline{\theta},\overline{\eta}\right)+\mathbb{E}_{\eta|\overline{\theta}}\eta y\left(\overline{\theta},\eta\right)-\overline{\eta}\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)-\left(1-\lambda\left(\overline{\theta}\right)\right)\left[\overline{\theta}\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\right] \end{pmatrix}. \tag{31}$$

Substituting these transfers back into the objective, we obtain

$$\Delta_c \equiv -\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\underline{\eta}\right)+\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)+\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right).$$

For $(x,y)\in\mathbb{X}_d\left(\overline{\theta}\right)$, the right-hand side of (23) is weakly smaller than the right-hand side of (17) and moreover, the constraint set is empty. Hence, $\hat{\eta}^*\left(\underline{\theta},\overline{\theta},\underline{\eta}\right)=\overline{\eta}$ cannot be implemented for $(x,y)\in\mathbb{X}_d\left(\overline{\theta}\right)$.

Consider next the second problem in case where $x\left(\underline{\theta},\underline{\eta}\right) \leq x\left(\underline{\theta},\overline{\eta}\right)$. By lemma 2, this implies that $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\underline{\eta}\right)=\underline{\eta}$. The objective then becomes

$$\Omega = \min_{\{T(\underline{\theta},\eta)\}_{\eta\in\{\underline{\eta},\overline{\eta}\}},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)} \begin{pmatrix} -\left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left[T\left(\underline{\theta},\underline{\eta}\right)-\underline{\theta}x\left(\underline{\theta},\underline{\eta}\right)-\underline{\eta}y\left(\underline{\theta},\underline{\eta}\right)\right]-\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right) \\ +\left(1-\lambda\left(\overline{\theta}\right)\right)\left(T\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)\right)-\overline{\theta}x\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)\right)-\overline{\eta}y\left(\underline{\theta},\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)\right)\right) \\ -(1-\lambda(\underline{\theta}))[T(\underline{\theta},\overline{\eta})-\underline{\theta}x(\underline{\theta},\overline{\eta})-\overline{\eta}y(\underline{\theta},\overline{\eta})]+\Delta \end{pmatrix}$$

subject to

(19), (18), and either

(26) if $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\overline{\eta}$, or

(27) if $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\underline{\eta}$.

where we have added the difference between the right- and the left-hand side of (6), which is zero by the fact that this constraint binds.

Consider first the possibility that $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right)=\overline{\eta}$. In that case the problem becomes

$$\Omega = \min_{T\left(\underline{\theta},\underline{\eta}\right)-T(\underline{\theta},\overline{\eta})} \begin{pmatrix} -\left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left[T\left(\underline{\theta},\underline{\eta}\right)-\underline{\theta}x\left(\underline{\theta},\underline{\eta}\right)-\underline{\eta}y\left(\underline{\theta},\underline{\eta}\right)\right] \\ +\left(\lambda(\underline{\theta})-\lambda\left(\overline{\theta}\right)\right)\left(T(\underline{\theta},\overline{\eta})-\overline{\theta}x(\underline{\theta},\overline{\eta})-\overline{\eta}y(\underline{\theta},\overline{\eta})\right) \\ -\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right)-(1-\lambda(\underline{\theta}))\left(\overline{\theta}-\underline{\theta}\right)x(\underline{\theta},\overline{\eta})+\Delta \end{pmatrix}$$

subject to (19), (18) and (26).

The right-hand side of (26) is always weakly smaller than the right-hand side of (19). Hence, (19) cannot become binding at the optimum. Moreover, the constraint set is non-empty exactly for $(x,y)\in\mathbb{X}_d\left(\underline{\theta}\right)$. Since the objective is decreasing in $T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)$, at the optimum, (26) is binding and we can compute the transfers from (26) and (6):

$$\begin{pmatrix} T\left(\underline{\theta},\underline{\eta}\right) \\ T(\underline{\theta},\overline{\eta}) \end{pmatrix} = \begin{pmatrix} \Delta+(1-\lambda(\underline{\theta}))\left[\overline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x(\underline{\theta},\overline{\eta})\right)+\overline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y(\underline{\theta},\overline{\eta})\right)\right]+\lambda(\underline{\theta})\left(\underline{\theta}x\left(\underline{\theta},\underline{\eta}\right)+\underline{\eta}y\left(\underline{\theta},\underline{\eta}\right)\right)+(1-\lambda(\underline{\theta}))[\underline{\theta}x(\underline{\theta},\overline{\eta})+\overline{\eta}y(\underline{\theta},\overline{\eta})] \\ \Delta-\lambda(\underline{\theta})\left[\overline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x(\underline{\theta},\overline{\eta})\right)+\overline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y(\underline{\theta},\overline{\eta})\right)\right]+\lambda(\underline{\theta})\left(\underline{\theta}x\left(\underline{\theta},\underline{\eta}\right)+\underline{\eta}y\left(\underline{\theta},\underline{\eta}\right)\right)+(1-\lambda(\underline{\theta}))[\underline{\theta}x(\underline{\theta},\overline{\eta})+\overline{\eta}y(\underline{\theta},\overline{\eta})] \end{pmatrix}$$

Since (26) is binding, we can substitute for

$$T\left(\underline{\theta},\underline{\eta}\right)-T\left(\underline{\theta},\overline{\eta}\right)=\overline{\theta}\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\overline{\eta}\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\overline{\eta}\right)\right)$$

into the objective and obtain

$$\Omega_d=-\mathbb{E}_{\eta|\underline{\theta}}\left[\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\eta\right)\right]+\Delta-\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)$$

If $(x, y) \in \mathbb{X}_d$, then this can be written as

$$\Omega_d = \mathbb{E}_{\eta|\overline{\theta}} \left[ \left( \overline{\theta} - \underline{\theta} \right) x \left( \overline{\theta}, \eta \right) \right] - \mathbb{E}_{\eta|\underline{\theta}} \left[ \left( \overline{\theta} - \underline{\theta} \right) x \left( \underline{\theta}, \eta \right) \right] - \left( \lambda \left( \underline{\theta} \right) - \lambda \left( \overline{\theta} \right) \right) \left( \overline{\eta} - \underline{\eta} \right) \left( y \left( \underline{\theta}, \underline{\eta} \right) - y \left( \overline{\theta}, \overline{\eta} \right) \right).$$

Consider finally the possibility that $\hat{\eta}^* \left( \overline{\theta}, \underline{\theta}, \overline{\eta} \right) = \underline{\eta}$. In that case the problem becomes

$$\Omega = \min_{T\left(\underline{\theta},\underline{\eta}\right) - T\left(\underline{\theta},\overline{\eta}\right)} \begin{pmatrix} (1-\lambda(\underline{\theta}))\left\{T\left(\underline{\theta},\underline{\eta}\right) - \underline{\theta} x\left(\underline{\theta},\underline{\eta}\right) - \underline{\eta} y\left(\underline{\theta},\underline{\eta}\right) - \left[T\left(\underline{\theta},\overline{\eta}\right) - \underline{\theta} x\left(\underline{\theta},\overline{\eta}\right) - \overline{\eta} y\left(\underline{\theta},\overline{\eta}\right)\right]\right\} \\ -\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right) y\left(\underline{\theta},\underline{\eta}\right) - \left(\overline{\theta}-\underline{\theta}\right) x\left(\underline{\theta},\underline{\eta}\right) + \Delta \end{pmatrix}$$

subject to (19), (18), and (27).

The right-hand side of (18) is weakly larger than the right-hand side of (27) exactly for $(x, y) \in \mathbb{X}_c(\underline{\theta})$. Moreover, for such allocations, the constraint set is nonempty, and at the solution of the problem $T\left(\underline{\theta}, \underline{\eta}\right) - T\left(\underline{\theta}, \overline{\eta}\right)$ reaches its lower bound, so (18) is binding. The transfers can then be computed from (6) and (18):

$$\begin{pmatrix} T\left(\underline{\theta}, \underline{\eta}\right) \\ T\left(\underline{\theta}, \overline{\eta}\right) \end{pmatrix} = \begin{pmatrix} \Delta + (1 - \lambda(\underline{\theta}))\left(\overline{\eta} - \underline{\eta}\right) y\left(\underline{\theta}, \overline{\eta}\right) + \underline{\theta} x\left(\underline{\theta}, \underline{\eta}\right) + \underline{\eta} y\left(\underline{\theta}, \underline{\eta}\right) \\ \Delta + \underline{\theta} x\left(\underline{\theta}, \overline{\eta}\right) + \left(\lambda(\underline{\theta})\, \underline{\eta} + (1 - \lambda(\underline{\theta}))\, \overline{\eta}\right) y\left(\underline{\theta}, \overline{\eta}\right) \end{pmatrix}$$

Since (18) is binding, we can substitute

$$T\left(\underline{\theta}, \underline{\eta}\right) - T\left(\underline{\theta}, \overline{\eta}\right) = \underline{\theta}\left(x\left(\underline{\theta}, \underline{\eta}\right) - x\left(\underline{\theta}, \overline{\eta}\right)\right) + \underline{\eta}\left(y\left(\underline{\theta}, \underline{\eta}\right) - y\left(\underline{\theta}, \overline{\eta}\right)\right)$$

into the objective and obtain

$$\Omega_c = (1 - \lambda(\underline{\theta}))\left(\overline{\eta} - \underline{\eta}\right) y\left(\underline{\theta}, \overline{\eta}\right) - \left(1 - \lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta} - \underline{\eta}\right) y\left(\underline{\theta}, \underline{\eta}\right) - \left(\overline{\theta} - \underline{\theta}\right) x\left(\underline{\theta}, \underline{\eta}\right) + \Delta$$

For future reference, if $(x, y) \in \mathbb{X}_c$, then we can write

$$\begin{aligned} \Omega_c \;=\; & -\lambda(\underline{\theta})\left(\overline{\eta} - \underline{\eta}\right)\left(y\left(\underline{\theta}, \overline{\eta}\right) - y\left(\overline{\theta}, \overline{\eta}\right)\right) - \left(\overline{\theta} - \underline{\theta}\right)\left(x\left(\underline{\theta}, \underline{\eta}\right) - x\left(\overline{\theta}, \overline{\eta}\right)\right) \\ & -\lambda\left(\overline{\theta}\right)\left(\overline{\eta} - \underline{\eta}\right)\left(y\left(\overline{\theta}, \underline{\eta}\right) - y\left(\underline{\theta}, \underline{\eta}\right)\right) - \left(\overline{\eta} - \underline{\eta}\right)\left(y\left(\underline{\theta}, \underline{\eta}\right) - y\left(\underline{\theta}, \overline{\eta}\right)\right). \end{aligned}$$

For $(x, y) \in \mathbb{X}_d(\underline{\theta})$, the right-hand side of (27) is weakly larger than the right-hand side of (18). Moreover, since the right-hand side of (27) is smaller than the right-hand side of (19), the constraint set is nonempty. At the solution, (27) is binding, so we can substitute for

$$T\left(\underline{\theta}, \underline{\eta}\right) - T\left(\underline{\theta}, \overline{\eta}\right) = \overline{\theta}\left(x\left(\underline{\theta}, \underline{\eta}\right) - x\left(\underline{\theta}, \overline{\eta}\right)\right) + \overline{\eta}\left(y\left(\underline{\theta}, \underline{\eta}\right) - y\left(\underline{\theta}, \overline{\eta}\right)\right)$$

into the objective and obtain

$$\hat{\Omega}_d = (1-\lambda(\underline{\theta}))\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)-\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\underline{\theta},\underline{\eta}\right)-\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right)+\Delta$$

Since $\Omega_d \leq \hat{\Omega}_d$ for any $(x,y) \in \mathbb{X}_d\left(\underline{\theta}\right)$, implementing $\hat{\eta}^*\left(\overline{\theta},\underline{\theta},\overline{\eta}\right) = \underline{\eta}$, the principal cannot gain by implementing this report. ■

**Proof of Proposition 1.** The proof of the first statements is given in three parts. Part i establishes properties of the solution of program d; part ii does likewise for program a; finally, part iii compares the value of the objectives. The proof of the fact that $W_c \leq W_d$ and $W_b \leq W_a$ is not given here but is available upon request from the authors; it uses essentially the same arguments.

Part i) Consider program d. Up to a constant program d can be written as

$$\begin{aligned}
&(1-\alpha)\,\mathbb{E}_{\eta|\overline{\theta}}\left[V\left(x\left(\overline{\theta},\eta\right),y\left(\overline{\theta},\eta\right)\right)-\overline{\theta}x\left(\overline{\theta},\eta\right)-\eta y\left(\overline{\theta},\eta\right)\right]\\
&-\alpha\left\{\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)+\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)+\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)\right\}\\
&+\phi\left[x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)\right]+\mu\left\{\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)\right)\right\}
\end{aligned}$$

The conditions of optimality are

$$\left((1-\alpha)\lambda\left(\overline{\theta}\right)\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)-\phi+\mu\left(\overline{\theta}-\underline{\theta}\right)\right)=0 \tag{32}$$

$$\left((1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)+\phi-\mu\left(\overline{\theta}-\underline{\theta}\right)\right)=0 \tag{33}$$

$$\left((1-\alpha)\lambda\left(\overline{\theta}\right)\left(V_2\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\underline{\eta}\right)+\mu\left(\overline{\eta}-\underline{\eta}\right)\right)=0 \tag{34}$$

$$\left((1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_2\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\eta}\right)-\alpha\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)-\mu\left(\overline{\eta}-\underline{\eta}\right)\right)=0. \tag{35}$$

We show by contradiction that at most one constraint binds at the optimum of program d.

Suppose both constraints bind. If $\phi,\mu>0$, then $x\left(\overline{\theta},\underline{\eta}\right)=x\left(\overline{\theta},\overline{\eta}\right)=x\left(\overline{\theta}\right)$ and $y\left(\overline{\theta},\underline{\eta}\right)=y\left(\overline{\theta},\overline{\eta}\right)=y\left(\overline{\theta}\right)$ and the conditions of optimality imply that

$$\left(V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta}\right)\right)-\overline{\theta}\right)-\frac{\alpha}{1-\alpha}\left(\overline{\theta}-\underline{\theta}\right)=0 \tag{36}$$

and

$$\begin{pmatrix} V_2\left(x\left(\overline{\theta}\right),y\left(\overline{\theta}\right)\right)\\ -\lambda\left(\overline{\theta}\right)\underline{\eta}-\left(1-\lambda\left(\overline{\theta}\right)\right)\overline{\eta}-\frac{\alpha}{1-\alpha}\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\end{pmatrix}=0. \tag{37}$$

Using $(34)$, the Kuhn-Tucker-first-order-optimality-condition for $y\left(\overline{\theta},\underline{\eta}\right)$ and substituting (37) we have for $\mu \neq 0$

$$\left((1-\alpha)\,\lambda\left(\overline{\theta}\right)\left(\lambda\left(\overline{\theta}\right)\underline{\eta}+\left(1-\lambda\left(\overline{\theta}\right)\right)\overline{\eta}+\frac{\alpha}{1-\alpha}\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)-\underline{\eta}\right)+\mu\left(\overline{\eta}-\underline{\eta}\right)\right)=0$$

which simplifies to

$$(1-\alpha)\,\lambda\left(\overline{\theta}\right)\left(\left(1-\lambda\left(\overline{\theta}\right)\right)+\frac{\alpha}{1-\alpha}\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\right)=-\mu$$

This implies $\mu < 0$ which contradicts the supposition that both constraints bind at the optimum. It follows that at most one constraint binds at the optimum of program d.

Further results require a case distinction between $V_{12} < 0$ and $V_{12} \geq 0$.

Case i) $V_{12} \geq 0$.

First, we show that if $V_{12} \geq 0$, then either constraint $x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\geq 0$ or constraint $\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)-\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)\geq 0$ binds at the optimum of program d.

Suppose no constraint binds. Then the first-order conditions with respect to $y$ are given by

$$V_2\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\underline{\eta}=0$$

$$V_2\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\eta}-\frac{\alpha\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)}{(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)}\left(\overline{\eta}-\underline{\eta}\right)=0.$$

The first-order conditions with respect to $x$ are given by

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}-\frac{\alpha}{1-\alpha}\left(\overline{\theta}-\underline{\theta}\right)=0$$

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}-\frac{\alpha}{1-\alpha}\left(\overline{\theta}-\underline{\theta}\right)=0$$

which imply

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)=V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right). \tag{38}$$

By concavity, $V_{11} < 0$, and $x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)>0$, we have

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)<V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) \tag{39}$$

Together conditions (38) and (39) imply

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)<V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right). \tag{40}$$

By complementarity, $V_{12}\geq 0$, and $y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)>0$, we have

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)\leq V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) \tag{41}$$

which contradicts (40).

It follows that at least one constraint must be binding at the optimum of program d.

Next, we show that if $V_{12}\geq 0$, then the optimal allocation satisfies

$$\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)\right)>0.$$

Suppose, contrary to our claim,

$$\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)\right)=0$$

and moreover $\phi=0$ and $\mu>0$.

The first-order conditions with respect to $x$ are given by

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}-\frac{\alpha\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)-\mu\left(\overline{\theta}-\underline{\theta}\right)}{(1-\alpha)\lambda\left(\overline{\theta}\right)}=0$$

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}-\frac{\alpha\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)+\mu\left(\overline{\theta}-\underline{\theta}\right)}{(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)}=0$$

implying that

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)<V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)$$

if and only if

$$\begin{array}{ccc}\dfrac{\alpha\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)-\mu\left(\overline{\theta}-\underline{\theta}\right)}{(1-\alpha)\lambda\left(\overline{\theta}\right)} & < & \dfrac{\alpha\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)+\mu\left(\overline{\theta}-\underline{\theta}\right)}{(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)} \\ & \Longleftrightarrow & \\ 0 & < & \mu.\end{array}$$

However, we must have $V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)\geq V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)$.

To see this, note that by $V_{11} < 0$

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) > V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right).$$

By $V_{12} \geq 0$

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) \geq V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)$$

Together these imply that

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) > V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right),$$

so that the conditions above would imply that $\mu < 0$, a contradiction.

It follows from these arguments that the optimal allocation for $V_{12} \geq 0$ satisfies $x\left(\overline{\theta},\overline{\eta}\right) = x\left(\overline{\theta},\underline{\eta}\right)$.

Case ii) $V_{12} < 0$.

If $V_{12} < 0$, then the solution to program d satisfies $x\left(\overline{\theta},\overline{\eta}\right) - x\left(\overline{\theta},\underline{\eta}\right) > 0$.

Suppose not. We know that $\phi, \mu > 0$ is not possible. So, if $x\left(\overline{\theta},\overline{\eta}\right) = x\left(\overline{\theta},\underline{\eta}\right)$, this would have to imply that $\mu = 0$. So, we would have $x\left(\overline{\theta},\underline{\eta}\right) = x\left(\overline{\theta},\overline{\eta}\right) = x\left(\overline{\theta}\right)$, $y\left(\overline{\theta},\overline{\eta}\right) < y\left(\overline{\theta},\underline{\eta}\right)$ and $\mu = 0$. Adding up of conditions (32) and (33), the first-order conditions for $x\left(\overline{\theta},\underline{\eta}\right)$ and $x\left(\overline{\theta},\overline{\eta}\right)$, gives

$$\left(\begin{array}{c} \lambda\left(\overline{\theta}\right)\left(V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)+\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right) \\ -\frac{\alpha}{1-\alpha}\left(\overline{\theta}-\underline{\theta}\right) \end{array}\right) = 0. \quad (42)$$

$V_{12} < 0$ and $y\left(\overline{\theta},\underline{\eta}\right) - y\left(\overline{\theta},\overline{\eta}\right) > 0$ imply that

$$V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right).$$

Together with (42), this implies that

$$V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < \overline{\theta} + \frac{\alpha}{(1-\alpha)}\left(\overline{\theta}-\underline{\theta}\right) < V_1\left(x\left(\overline{\theta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right).$$

Plugging the first of these inequalities into (32), we obtain

$$\left((1-\alpha)\lambda\left(\overline{\theta}\right)\frac{\alpha}{(1-\alpha)}\left(\overline{\theta}-\underline{\theta}\right) - \alpha\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right) - \phi + \mu\left(\overline{\theta}-\underline{\theta}\right)\right) > 0.$$

Plugging the latter of the inequalities into $(33)$, we obtain

$$\left((1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\frac{\alpha}{(1-\alpha)}\left(\overline{\theta}-\underline{\theta}\right)-\alpha\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)+\phi-\mu\left(\overline{\theta}-\underline{\theta}\right)\right)<0,$$

which simplifies to

$$\mu\left(\overline{\theta}-\underline{\theta}\right)>\phi.$$

For $\mu=0$ this implies $\phi<0$. Hence, $V_{12}<0$ implies that $x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)>0$.

Part ii: Consider program a. Up to a constant program a can be written as

$$\begin{array}{l}(1-\alpha)\,\mathbb{E}_{\eta|\overline{\theta}}\left[V\left(x\left(\overline{\theta},\eta\right),y\left(\overline{\theta},\eta\right)\right)-\overline{\theta}x\left(\overline{\theta},\eta\right)-\eta y\left(\overline{\theta},\eta\right)\right]\\ -\alpha\left\{\begin{array}{c}\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)+\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)\\ +\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)y\left(\overline{\theta},\overline{\eta}\right)\end{array}\right\}\\ +\xi\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right]+\nu\left[\begin{array}{c}\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)\\ -\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\end{array}\right]\end{array}$$

The conditions of optimality are given by

$$\left(\begin{array}{c}(1-\alpha)\,\lambda\left(\overline{\theta}\right)\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)\\ -\alpha\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)+\xi-\nu\left(\overline{\theta}-\underline{\theta}\right)\end{array}\right)=0 \tag{43}$$

$$(1-\alpha)\,\lambda\left(\overline{\theta}\right)\left(V_2\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\underline{\eta}\right)+\nu\left(\overline{\eta}-\underline{\eta}\right)=0 \tag{44}$$

$$\left(\begin{array}{c}(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right)\\ -\alpha\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)-\xi+\nu\left(\overline{\theta}-\underline{\theta}\right)\end{array}\right)=0 \tag{45}$$

$$\left(\begin{array}{c}(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_2\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\eta}\right)\\ -\alpha\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)-\nu\left(\overline{\eta}-\underline{\eta}\right)\end{array}\right)=0 \tag{46}$$

$$\xi,\nu\geq 0$$

$$\xi\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right]=0$$

$$\nu\left[\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\right]=0$$

First, we show by contradiction that at most one constraint binds at the optimum of program a.

So suppose both constraints bind at the optimum, i.e. $\xi, \nu > 0$.´If $\xi, \nu > 0$, then $y\left(\overline{\theta}, \underline{\eta}\right) = y\left(\overline{\theta}, \overline{\eta}\right) = y\left(\overline{\theta}\right)$ and $x\left(\overline{\theta}, \underline{\eta}\right) = x\left(\overline{\theta}, \overline{\eta}\right) = x\left(\overline{\theta}\right)$. Then

$$V_1\left(x\left(\overline{\theta}\right), y\left(\overline{\theta}\right)\right) = \overline{\theta} + \frac{\alpha}{(1-\alpha)}\left(\overline{\theta} - \underline{\theta}\right)$$

and

$$V_2\left(x\left(\overline{\theta}\right), y\left(\overline{\theta}\right)\right) = \lambda\left(\overline{\theta}\right)\underline{\eta} + \left(1 - \lambda\left(\overline{\theta}\right)\right)\overline{\eta} + \frac{\alpha\left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)}{(1-\alpha)}\left(\overline{\eta} - \underline{\eta}\right).$$

Using the first-order condition with respect to $y\left(\overline{\theta}, \underline{\eta}\right)$, (44), gives

$$V_2\left(x\left(\overline{\theta}\right), y\left(\overline{\theta}\right)\right) = \underline{\eta} - \frac{\nu}{(1-\alpha)\lambda\left(\overline{\theta}\right)}\left(\overline{\eta} - \underline{\eta}\right).$$

Substituting for $V_2\left(x\left(\overline{\theta}\right), y\left(\overline{\theta}\right)\right)$ gives

$$\lambda\left(\overline{\theta}\right)\underline{\eta} + \left(1 - \lambda\left(\overline{\theta}\right)\right)\overline{\eta} + \frac{\alpha\left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)}{(1-\alpha)}\left(\overline{\eta} - \underline{\eta}\right) = \underline{\eta} - \frac{\nu}{(1-\alpha)\lambda\left(\overline{\theta}\right)}\left(\overline{\eta} - \underline{\eta}\right)$$

which simplifies to

$$-\left(1 - \lambda\left(\overline{\theta}\right)\right)(1-\alpha)\lambda\left(\overline{\theta}\right) - \alpha\left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)\lambda\left(\overline{\theta}\right) = \nu,$$

implying that $\nu < 0$. It follows that at the optimum $\xi, \nu > 0$ is not true.

Further results require a case distinction between $V_{12} < 0$ and $V_{12} \geq 0$.

Case i) $V_{12} < 0$.

First, whe show that if $V_{12} < 0$, then $x\left(\overline{\theta}, \underline{\eta}\right) - x\left(\overline{\theta}, \overline{\eta}\right) = 0$ at the solution to program a.

To show this, we establish first that $V_{12} < 0$ implies that at least one constraint binds. Moreover, we show that $\left(\overline{\eta} - \underline{\eta}\right)\left[y\left(\overline{\theta}, \underline{\eta}\right) - y\left(\overline{\theta}, \overline{\eta}\right)\right] > \left(\overline{\theta} - \underline{\theta}\right)\left[x\left(\overline{\theta}, \underline{\eta}\right) - x\left(\overline{\theta}, \overline{\eta}\right)\right]$ at the optimum of program a.

Suppose no constraint binds at the optimum, i.e. $\left(\overline{\eta} - \underline{\eta}\right)\left[y\left(\overline{\theta}, \underline{\eta}\right) - y\left(\overline{\theta}, \overline{\eta}\right)\right] > \left(\overline{\theta} - \underline{\theta}\right)\left[x\left(\overline{\theta}, \underline{\eta}\right) - x\left(\overline{\theta}, \overline{\eta}\right)\right] > 0$. Then $\xi = \nu = 0$. The first-order conditions with respect to $x$ are given by

$$V_1\left(x\left(\overline{\theta}, \underline{\eta}\right), y\left(\overline{\theta}, \underline{\eta}\right)\right) - \overline{\theta} - \frac{\alpha\lambda\left(\underline{\theta}\right)}{(1-\alpha)\lambda\left(\overline{\theta}\right)}\left(\overline{\theta} - \underline{\theta}\right) = 0$$

$$V_1\left(x\left(\overline{\theta}, \overline{\eta}\right), y\left(\overline{\theta}, \overline{\eta}\right)\right) - \overline{\theta} - \frac{\alpha\left(1 - \lambda\left(\underline{\theta}\right)\right)}{(1-\alpha)\left(1 - \lambda\left(\overline{\theta}\right)\right)}\left(\overline{\theta} - \underline{\theta}\right) = 0$$

Since $\frac{\lambda(\underline{\theta})}{\lambda(\overline{\theta})} > \frac{(1-\lambda(\underline{\theta}))}{(1-\lambda(\overline{\theta}))}$, these conditions imply that

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) > V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right). \tag{47}$$

However, by $V_{12} < 0$ and $y\left(\overline{\theta},\underline{\eta}\right) - y\left(\overline{\theta},\overline{\eta}\right) > 0$

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right). \tag{48}$$

By $V_{11} < 0$ and $x\left(\overline{\theta},\underline{\eta}\right) - x\left(\overline{\theta},\overline{\eta}\right) > 0$

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right). \tag{49}$$

Taken together (48) and (49) imply that

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right).$$

which contradicts (47) derived previously from the first order-conditions.

It follows that at least one constraint must bind at the optimum of program a if $V_{12} < 0$.

Suppose that contrary to our claim, that the solution of program a satisfies $\left(\overline{\eta}-\underline{\eta}\right)\left[y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right]$ $=\left(\overline{\theta}-\underline{\theta}\right)\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right] > 0$, $\xi = 0$ and $\nu > 0$. Adding up (43) and (45) gives

$$\begin{pmatrix} (1-\alpha)\lambda\left(\overline{\theta}\right)\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right) \\ +(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right) \end{pmatrix} = 0.$$

By $V_{11} < 0$ and $x\left(\overline{\theta},\underline{\eta}\right) - x\left(\overline{\theta},\overline{\eta}\right) > 0$

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)$$

By $V_{12} < 0$ and $y\left(\overline{\theta},\underline{\eta}\right) - y\left(\overline{\theta},\overline{\eta}\right) > 0$

$$V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right).$$

Taken together, we have

$$V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right) < V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right).$$

Combining with the implications of the first-order conditions with respect to $x$ we obtain

$$\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right) < \frac{\alpha}{(1-\alpha)}\left(\overline{\theta}-\underline{\theta}\right) < \left(V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right).$$

Substituting into $(43)$, using $\xi = 0$, and simplifying, we have

$$\alpha\left(\left(\lambda\left(\overline{\theta}\right)-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)\right)>\nu\left(\overline{\theta}-\underline{\theta}\right),$$

which would imply that $v < 0$, a contradiction.

It follows that for $V_{12} < 0$, the optimum of program a displays $x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)=0$.

Case ii): $V_{12} \geq 0$.

If $V_{12} \leq -V_{11}\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}$, then the optimum of program a displays $\left(\overline{\eta}-\underline{\eta}\right)\left[y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right]>\left(\overline{\theta}-\underline{\theta}\right)\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right]$.

We know that both constraints cannot bind simultaneously. Hence, if $\left(\overline{\eta}-\underline{\eta}\right)\left[y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right]=\left(\overline{\theta}-\underline{\theta}\right)\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right]$, then necessarily $x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)>0$. Suppose this is the case, so $\xi = 0$ and $\nu > 0$. Define $Y\left(\overline{\theta},\overline{\eta}\right)=y\left(\overline{\theta},\underline{\eta}\right)-\frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)}\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right]$. For $\nu \neq 0$ the first order conditions with respect to $x$ are given by

$$\begin{aligned}&\left(\begin{array}{c}(1-\alpha)\lambda\left(\overline{\theta}\right)\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)\\ -\frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)}\left((1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_2\left(x\left(\overline{\theta},\overline{\eta}\right),Y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\eta}\right)-\alpha\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\right)\end{array}\right)\\ =\ &0\end{aligned}$$

and

$$\begin{aligned}&\left(\begin{array}{c}(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)\\ +\frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)}\left((1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_2\left(x\left(\overline{\theta},\overline{\eta}\right),Y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\eta}\right)-\alpha\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\right)\end{array}\right)\\ =\ &0.\end{aligned}$$

These conditions imply

$$\left(\begin{array}{c}(1-\alpha)\lambda\left(\overline{\theta}\right)\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)\\ +(1-\alpha)\left(1-\lambda\left(\overline{\theta}\right)\right)\left(V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)-\overline{\theta}\right)-\alpha\left(1-\lambda\left(\underline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)\end{array}\right)=0. \quad (50)$$

Define $s$ such that $s > 0$, $s=y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)$ and $x\left(\overline{\theta},\underline{\eta}\right)=x\left(\overline{\theta},\overline{\eta}\right)+\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}s$. Then by

$$\begin{aligned}&V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)\\ =\ &V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)+\int_0^s\frac{\partial V_1\left(x\left(\overline{\theta},\overline{\eta}\right)+\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}k,y\left(\overline{\theta},\overline{\eta}\right)+k\right)}{\partial k}dk\end{aligned}$$

and

$$\frac{\partial V_1\left(x\left(\overline{\theta},\overline{\eta}\right)+\frac{(\overline{\eta}-\underline{\eta})}{(\overline{\theta}-\underline{\theta})}k,y\left(\overline{\theta},\overline{\eta}\right)+k\right)}{\partial k}=\left(\begin{array}{c}\frac{(\overline{\eta}-\underline{\eta})}{(\overline{\theta}-\underline{\theta})}V_{11}\left(x\left(\overline{\theta},\overline{\eta}\right)+\frac{(\overline{\eta}-\underline{\eta})}{(\overline{\theta}-\underline{\theta})}k,y\left(\overline{\theta},\overline{\eta}\right)+k\right)\\+V_{12}\left(x\left(\overline{\theta},\overline{\eta}\right)+\frac{(\overline{\eta}-\underline{\eta})}{(\overline{\theta}-\underline{\theta})}k,y\left(\overline{\theta},\overline{\eta}\right)+k\right)\end{array}\right)$$

we have $V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)\leq V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)$ since $V_{12}\leq -V_{11}\frac{(\overline{\eta}-\underline{\eta})}{(\overline{\theta}-\underline{\theta})}$ implies

$$\frac{\partial V_1\left(x\left(\overline{\theta},\overline{\eta}\right)+\frac{(\overline{\eta}-\underline{\eta})}{(\overline{\theta}-\underline{\theta})}k,y\left(\overline{\theta},\overline{\eta}\right)+k\right)}{\partial k}\leq 0 \text{ for all } k>0$$

and $x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)>0$. $V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)\leq V_1\left(x\left(\overline{\theta},\overline{\eta}\right),y\left(\overline{\theta},\overline{\eta}\right)\right)$ implies by (50) that

$$\left(V_1\left(x\left(\overline{\theta},\underline{\eta}\right),y\left(\overline{\theta},\underline{\eta}\right)\right)-\overline{\theta}\right)\leq\frac{\alpha}{(1-\alpha)}\left(\overline{\theta}-\underline{\theta}\right) \tag{51}$$

By (51) into (43)

$$\left(\begin{array}{c}(1-\alpha)\lambda\left(\overline{\theta}\right)\frac{\alpha}{(1-\alpha)}\left(\overline{\theta}-\underline{\theta}\right)\\-\alpha\lambda\left(\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)+\xi-\nu\left(\overline{\theta}-\underline{\theta}\right)\end{array}\right)\geq 0$$

which is equivalent to

$$-\alpha\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\geq\nu\left(\overline{\eta}-\underline{\eta}\right)$$

which is true only if $\nu<0$ since $\xi=0$. Hence, we get a contradiction to $\nu>0$ contradicting that $\left(\overline{\eta}-\underline{\eta}\right)\left[y\left(\overline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right]\geq\left(\overline{\theta}-\underline{\theta}\right)\left[x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right]$ binds in singularity at the optimum of program a.

Part iii) Comparison between the programs.

As a preliminary argument, note that if $x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)=0$, then the value of the objectives of programs a and d become identical. To see this, note that the objectives are identical up to the costs of implementation, $\Delta$. Moreover, it is easy to verify from Lemma 3 that $\Delta_d-\Delta_a$ for $x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)=0$.

For $V_{12}<0$, the maximum of program a satisfies $x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)=0$, whereas the maximum of program d satisfies $x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)>0$. Hence, the solution of program a is feasible but not chosen. By revealed preference, this implies that the solution to program d is preferred.

Likewise, for $V_{12} \geq 0$ the optimum of program d satisfies $x\left(\overline{\theta}, \overline{\eta}\right) - x\left(\overline{\theta}, \underline{\eta}\right) = 0$. Hence, the solution is feasible under program a. Therefore, by revealed preference, the solution of program a yields a weakly higher expected payoff than the solution of program d. ■

**Proof of Lemma 4.** The first-best allocation satisfies the system of conditions

$$V_1\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) = \theta$$

$$V_2\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) = \eta$$

for $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ and $\eta \in \left\{\underline{\eta}, \overline{\eta}\right\}$.

By the fundamental theorem of calculus

$$x\left(\theta, \overline{\eta}\right) - x\left(\theta, \underline{\eta}\right) = \int_{\underline{\eta}}^{\overline{\eta}} \frac{\partial}{\partial \eta} x\left(\theta, \eta\right) d\eta.$$

Totally differentiating the system of first-order conditions, we have

$$V_{11}\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) dx + V_{12}\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) dy = 0$$

$$V_{21}\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) dx + V_{22}\left(x\left(\theta, \eta\right), y\left(\theta, \eta\right)\right) dy = d\eta$$

By Cramer's rule

$$\frac{dx}{d\eta} = \frac{-V_{12}}{V_{11}V_{22} - V_{12}^2}$$

so

$$x\left(\theta, \overline{\eta}\right) - x\left(\theta, \underline{\eta}\right) = \int_{\underline{\eta}}^{\overline{\eta}} \frac{-V_{12}}{V_{11}V_{22} - V_{12}^2}\left(\theta, \eta\right) d\eta.$$

So $x\left(\theta, \overline{\eta}\right) \leq x\left(\theta, \underline{\eta}\right)$ for $V_{12} \geq 0$ and $x\left(\theta, \overline{\eta}\right) > x\left(\theta, \underline{\eta}\right)$ for $V_{12} < 0$.

Again by the fundamental theorem and by Cramer's rule

$$y\left(\theta, \overline{\eta}\right) - y\left(\theta, \underline{\eta}\right) = \int_{\underline{\eta}}^{\overline{\eta}} \frac{V_{11}}{V_{11}V_{22} - V_{12}^2}\left(\theta, \eta\right) d\eta.$$

By concavity, we have $y\left(\theta, \overline{\eta}\right) - y\left(\theta, \underline{\eta}\right) < 0$.

Combining these arguments, we have

$$\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right)>\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right)>0$$

iff $V_{12}>0$ and

$$0>\int_{\underline{\eta}}^{\overline{\eta}}\frac{V_{11}+\frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)}V_{12}}{V_{11}V_{22}-V_{12}^{2}}\left(\theta,\eta\right)d\eta,$$

which is satisfied if $V_{12}<-\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}V_{11}$ for all $\left(x,y\right)$. Likewise, we have

$$\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right)>\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right)>0$$

if $V_{12}>-\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}V_{11}$ for all $\left(x,y\right)$.

Similarly, one shows that for $V_{12}<0$ we have

$$\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right)>-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right)>0$$

if $V_{12}>\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}V_{11}$ for all $\left(x,y\right)$ and

$$-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right)>\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right)$$

for $V_{12}<\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}V_{11}$ for all $\left(x,y\right)$. ■

**Proof of Proposition 2.** We show that the neglected constraint is satisfied under the assumptions.

Preliminaries:

For convenience, recall that the unconstrained solution (in the sense of unconstrained by the implementation sets $\mathbb{X}_i$ for $i=a$ or $i=d$, respectively) satisfies

$$\begin{aligned}V_1\left(x,y\right)\left(\underline{\theta},\eta\right)&=\underline{\theta}\\V_2\left(x,y\right)\left(\underline{\theta},\eta\right)&=\eta\end{aligned}\qquad(52)$$

for $\eta\in\left\{\underline{\eta},\overline{\eta}\right\}$ and

$$\begin{aligned}V_1\left(x,y\right)\left(\overline{\theta},\underline{\eta}\right)&=\overline{\theta}+\frac{\alpha}{\left(1-\alpha\right)}\frac{\lambda_i}{\lambda\left(\overline{\theta}\right)}\left(\overline{\theta}-\underline{\theta}\right)\\V_2\left(x,y\right)\left(\overline{\theta},\underline{\eta}\right)&=\underline{\eta}\end{aligned}\qquad(53)$$

$$V_1(x,y)(\overline{\theta},\overline{\eta}) = \overline{\theta} + \frac{\alpha}{(1-\alpha)}\frac{(1-\lambda_i)}{(1-\lambda(\overline{\theta}))}(\overline{\theta}-\underline{\theta}) \tag{54}$$

$$V_2(x,y)(\overline{\theta},\overline{\eta}) = \overline{\eta} + \frac{\alpha}{(1-\alpha)}\frac{(\lambda(\underline{\theta})-\lambda(\overline{\theta}))}{(1-\lambda(\overline{\theta}))}(\overline{\eta}-\underline{\eta}),$$

where $\lambda_a = \lambda(\underline{\theta})$ and $\lambda_d = \lambda(\overline{\theta})$. Notice, that (52) and (53) have the common representation

$$V_1(x,y)(\theta,\underline{\eta}) = \theta + \frac{\alpha}{(1-\alpha)}\frac{\lambda_i}{\lambda(\overline{\theta})}(\theta-\underline{\theta}) \tag{55}$$

$$V_2(x,y)(\theta,\eta) = \eta$$

so that for $\theta = \underline{\theta}$, (55) is equivalent to (52) and for $\theta = \overline{\theta}$ and $\eta = \overline{\eta}$, (55) is equivalent to (53). Likewise, note that for $\lambda_d = \lambda(\overline{\theta})$, (52), (53), and (54) have the common representation

$$V_1(x,y)(\theta,\eta) = \theta + \frac{\alpha}{(1-\alpha)}(\theta-\underline{\theta}) \tag{56}$$

$$V_2(x,y)(\theta,\eta) = \eta + \frac{\alpha}{(1-\alpha)}\frac{(\lambda(\underline{\theta})-\lambda(\overline{\theta}))}{(1-\lambda(\overline{\theta}))}(\eta-\underline{\eta}),$$

so that for $\theta = \underline{\theta}$ and $\eta \in \eta \in \{\underline{\eta},\overline{\eta}\}$, (56) is equivalent to (52), for $\theta = \overline{\theta}$ and $\eta = \underline{\eta}$, (56) is equivalent to (53), and for $\theta = \overline{\theta}$ and $\eta = \overline{\eta}$, (56) is equivalent to (55).

Part i):The case of independent goods: $V_{12} = 0$.

From Proposition 1 we know that program $\mathrm{P}_a$ solves the reduced problem for $V_{12} = 0$. Hence, the neglected constraint takes the form

$$\begin{aligned}&(\lambda(\underline{\theta})-\lambda(\overline{\theta}))(\overline{\eta}-\underline{\eta})(y(\underline{\theta},\underline{\eta})-y(\overline{\theta},\overline{\eta}))\\&+(\overline{\theta}-\underline{\theta})\left(\mathbb{E}_{\eta|\overline{\theta}}[x(\underline{\theta},\eta)]-\mathbb{E}_{\eta|\underline{\theta}}[x(\overline{\theta},\eta)]\right)\geq 0.\end{aligned}$$

Sufficient conditions for the neglected constraint to hold are

$$y(\underline{\theta},\underline{\eta}) - y(\overline{\theta},\overline{\eta}) \geq 0$$

and

$$\left(\mathbb{E}_{\eta|\overline{\theta}}[x(\underline{\theta},\eta)]-\mathbb{E}_{\eta|\underline{\theta}}[x(\overline{\theta},\eta)]\right)\geq 0.$$

Moreover, we know again from Proposition 1 that $x\left(\overline{\theta},\underline{\eta}\right)=x\left(\overline{\theta},\overline{\eta}\right)=x\left(\overline{\theta}\right)$ at the solution. So, the relevant first-order conditions describing the optimum simplify to

$$V_1\left(x\left(\underline{\theta},\eta\right)\right)=\underline{\theta}$$

$$V_2\left(y\left(\underline{\theta},\eta\right)\right)=\eta$$

for $\eta\in\left\{\underline{\eta},\overline{\eta}\right\}$ and moreover

$$(1-\alpha)\,\mathbb{E}_{\eta|\overline{\theta}}\left[V_1\left(x\left(\overline{\theta}\right)\right)-\overline{\theta}\right]=\alpha\left(\overline{\theta}-\underline{\theta}\right)$$

$$V_2\left(y\left(\overline{\theta},\overline{\eta}\right)\right)=\overline{\eta}+\frac{\alpha}{(1-\alpha)}\frac{\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)}{\left(1-\lambda\left(\overline{\theta}\right)\right)}\left(\overline{\eta}-\underline{\eta}\right).$$

It is easy to see (by concavity of $V$), that $x\left(\underline{\theta},\underline{\eta}\right)=x\left(\underline{\theta},\overline{\eta}\right)>x\left(\overline{\theta}\right)$, so $\left(\mathbb{E}_{\eta|\overline{\theta}}\left[x\left(\underline{\theta},\eta\right)\right]-\mathbb{E}_{\eta|\underline{\theta}}\left[x\left(\overline{\theta},\eta\right)\right]\right)\geq 0$. is satisfied. By the same argument, we also have $y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\geq 0$.

Part ii)

1. The case of complements.

For the case of complements with $0\leq V_{12}<-V_{11}\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}$ for all $x,y,$, by Lemmas 3 and 4, the neglected constraint (7) is equivalent to

$$\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right)$$
$$+\left(\overline{\theta}-\underline{\theta}\right)\left(\mathbb{E}_{\eta|\overline{\theta}}\left[x\left(\underline{\theta},\eta\right)\right]-\mathbb{E}_{\eta|\underline{\theta}}\left[x\left(\overline{\theta},\eta\right)\right]\right)\geq 0.$$

Sufficient conditions for the neglected constraint to hold are

$$y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\geq 0$$

and

$$\left(\mathbb{E}_{\eta|\overline{\theta}}\left[x\left(\underline{\theta},\eta\right)\right]-\mathbb{E}_{\eta|\underline{\theta}}\left[x\left(\overline{\theta},\eta\right)\right]\right)\geq 0.$$

We now provide sufficient conditions such that the unconstrained solution satisfies these monotonicity restrictions.

Incentive compatibility with respect to $\eta$ alone requires that $y\left(\overline{\theta},\underline{\eta}\right)\geq y\left(\overline{\theta},\overline{\eta}\right)$. Hence, a sufficient condition for $y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\geq 0$ is that $\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\underline{\eta}\right)\right)\geq 0$. In turn, this follows trivially from the fact that $\theta + \frac{\alpha}{(1-\alpha)}\frac{\lambda_i}{\lambda(\overline{\theta})}(\theta - \underline{\theta})$ is increasing in $\theta$ and thus that an increase in $\theta$ reduces $x$, which by complementarity reduces $y$.

A sufficient condition for $\left(\mathbb{E}_{\eta|\overline{\theta}}\left[x(\underline{\theta},\eta)\right] - \mathbb{E}_{\eta|\underline{\theta}}\left[x\left(\overline{\theta},\eta\right)\right]\right) \geq 0$ is that

$$\min_{\eta\in\left\{\underline{\eta},\overline{\eta}\right\}} x(\underline{\theta},\eta) \geq \max_{\eta\in\left\{\underline{\eta},\overline{\eta}\right\}} x\left(\overline{\theta},\eta\right),$$

which in turn holds if

$$x\left(\underline{\theta},\underline{\eta}\right) \geq x(\underline{\theta},\overline{\eta}) \geq x\left(\overline{\theta},\underline{\eta}\right) \geq x\left(\overline{\theta},\overline{\eta}\right).$$

It is straightforward to see that $x\left(\underline{\theta},\underline{\eta}\right) \geq x(\underline{\theta},\overline{\eta})$, since $x$ and $y$ are complements. Similarly, $x\left(\overline{\theta},\underline{\eta}\right) \geq x\left(\overline{\theta},\overline{\eta}\right)$ follows from the fact that $\frac{\lambda(\underline{\theta})}{\lambda(\overline{\theta})} \geq \frac{(1-\lambda(\underline{\theta}))}{\left(1-\lambda\left(\overline{\theta}\right)\right)}$ and that $x$ and $y$ are complements. So, we need to show that $x(\underline{\theta},\overline{\eta}) \geq x\left(\overline{\theta},\underline{\eta}\right)$. We can write

$$x(\underline{\theta},\overline{\eta}) - x\left(\overline{\theta},\underline{\eta}\right) = x(\underline{\theta},\overline{\eta}) - x\left(\underline{\theta},\underline{\eta}\right) + x\left(\underline{\theta},\underline{\eta}\right) - x\left(\overline{\theta},\underline{\eta}\right).$$

Let $(x,y)$ be determined by (55). Differentiating the system of equations (55), we can write

$$x(\underline{\theta},\overline{\eta}) - x\left(\underline{\theta},\underline{\eta}\right) = \int_{\underline{\eta}}^{\overline{\eta}} \frac{-V_{12}}{V_{11}V_{22} - V_{12}^2}(\underline{\theta},\eta)\,d\eta = \left(\overline{\eta} - \underline{\eta}\right)\frac{-V_{12}}{V_{11}V_{22} - V_{12}^2}(\underline{\theta},\hat{\eta}).$$

where the first equality follows from Lemma 3 and the second equality from the mean value theorem, for some $\hat{\eta} \in \left[\underline{\eta},\overline{\eta}\right]$. Likewise, we have

$$\begin{aligned} x\left(\underline{\theta},\underline{\eta}\right) - x\left(\overline{\theta},\underline{\eta}\right) &= \int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial x\left(\theta,\underline{\eta}\right)}{\partial\theta}d\theta = \left(1 + \frac{\alpha}{(1-\alpha)}\frac{\lambda(\underline{\theta})}{\lambda\left(\overline{\theta}\right)}\right)\int_{\overline{\theta}}^{\underline{\theta}} \frac{V_{22}}{V_{11}V_{22} - V_{12}^2}d\theta \\ &= -\left(\overline{\theta} - \underline{\theta}\right)\left(1 + \frac{\alpha}{(1-\alpha)}\frac{\lambda(\underline{\theta})}{\lambda\left(\overline{\theta}\right)}\right)\frac{V_{22}}{V_{11}V_{22} - V_{12}^2}\left(\hat{\theta},\underline{\eta}\right). \end{aligned}$$

for some $\hat{\theta} \in \left[\underline{\theta},\overline{\theta}\right]$, where the last equality follows again by the mean value theorem.

So, we have $x(\underline{\theta},\overline{\eta}) \geq x\left(\overline{\theta},\underline{\eta}\right)$ iff

$$\left(\overline{\eta} - \underline{\eta}\right)\frac{-V_{12}}{V_{11}V_{22} - V_{12}^2}(\underline{\theta},\hat{\eta}) - \left(\overline{\theta} - \underline{\theta}\right)\left(1 + \frac{\alpha}{(1-\alpha)}\frac{\lambda(\underline{\theta})}{\lambda\left(\overline{\theta}\right)}\right)\frac{V_{22}}{V_{11}V_{22} - V_{12}^2}\left(\hat{\theta},\underline{\eta}\right) \geq 0.$$

In turn, this condition is satisfied if

$$\begin{aligned} & \frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)}\left(1+\frac{\alpha}{(1-\alpha)}\frac{\lambda\left(\underline{\theta}\right)}{\lambda\left(\overline{\theta}\right)}\right)\min_{x,y}\frac{-V_{22}}{V_{11}V_{22}-V_{12}^2}(x,y) \\ \geq\ & \max_{x,y}\frac{V_{12}}{V_{11}V_{22}-V_{12}^2}(x,y). \end{aligned}$$

Since the left-hand side is increasing in $\alpha$, the condition is hardest to satisfy for $\alpha = 0$, which is the condition given in the proposition.

2. The case of substitutes:

For $0 > V_{12} > V_{11}\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}$ for all $x, y$, the neglected constraint is equivalent to

$$\left(\begin{array}{c} \left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right) \\ +\left(\overline{\theta}-\underline{\theta}\right)x\left(\underline{\theta},\underline{\eta}\right)-\left(\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\underline{\eta}\right)+\left(1-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\theta}-\underline{\theta}\right)x\left(\overline{\theta},\overline{\eta}\right)\right) \end{array}\right)\geq 0.$$

Equivalently, this can be written as

$$\left(\begin{array}{c} \left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\underline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right) \\ +\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\underline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)-\left(\lambda\left(\overline{\theta}\right)\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\right)\right) \end{array}\right)\geq 0.$$

Recall that for $x, y \in X_d$, we have

$$\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\theta,\underline{\eta}\right)-y\left(\theta,\overline{\eta}\right)\right)\geq-\left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\theta,\underline{\eta}\right)-x\left(\theta,\overline{\eta}\right)\right)\geq 0.$$

For the case where $\lambda\left(\underline{\theta}\right)=\lambda\left(\overline{\theta}\right)$, we only need to show that

$$x\left(\underline{\theta},\underline{\eta}\right)-x\left(\overline{\theta},\overline{\eta}\right)\geq 0.$$

We can write

$$x\left(\overline{\theta},\overline{\eta}\right)-x\left(\underline{\theta},\underline{\eta}\right)=x\left(\overline{\theta},\overline{\eta}\right)-x\left(\overline{\theta},\underline{\eta}\right)+x\left(\overline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\underline{\eta}\right).$$

we can write

$$\begin{aligned} x\left(\overline{\theta},\overline{\eta}\right)-x\left(\underline{\theta},\underline{\eta}\right) &= \left(\overline{\eta}-\underline{\eta}\right)\left(1+\frac{\alpha}{(1-\alpha)}\frac{\left(\lambda\left(\underline{\theta}\right)-\lambda\left(\overline{\theta}\right)\right)}{\left(1-\lambda\left(\overline{\theta}\right)\right)}\right)\frac{-V_{12}}{V_{11}V_{22}-V_{12}^2}\left(\overline{\theta},\hat{\eta}\right) \\ &\quad +\left(\overline{\theta}-\underline{\theta}\right)\frac{1}{1-\alpha}\frac{V_{22}}{V_{11}V_{22}-V_{12}^2}\left(\hat{\theta},\underline{\eta}\right) \end{aligned}$$

for some values $\hat{\theta} \in \left[\underline{\theta}, \overline{\theta}\right]$ and $\hat{\eta} \in \left[\underline{\eta}, \overline{\eta}\right]$. Hence, we have $x\left(\underline{\theta}, \underline{\eta}\right) - x\left(\overline{\theta}, \overline{\eta}\right) \geq 0$, if

$$\begin{aligned} & \frac{\overline{\eta} - \underline{\eta}}{\overline{\theta} - \underline{\theta}} \left( 1 - \alpha + \alpha \frac{\left(\lambda\left(\underline{\theta}\right) - \lambda\left(\overline{\theta}\right)\right)}{\left(1 - \lambda\left(\overline{\theta}\right)\right)} \right) \min_{x,y} \frac{V_{12}}{V_{11}V_{22} - V_{12}^2} \\ \geq \ & \max_{x,y} \frac{V_{22}}{V_{11}V_{22} - V_{12}^2} \end{aligned}$$

Since $\frac{\left(\lambda(\underline{\theta}) - \lambda(\overline{\theta})\right)}{\left(1 - \lambda(\overline{\theta})\right)} < 1$, the expression on the left-hand side of the inequality is smallest for $\alpha = 0$, so the condition is satisfied if

$$\frac{\overline{\eta} - \underline{\eta}}{\overline{\theta} - \underline{\theta}} \min_{x,y} \frac{V_{12}}{V_{11}V_{22} - V_{12}^2} \geq \max_{x,y} \frac{V_{22}}{V_{11}V_{22} - V_{12}^2}.$$

Finally, we need to show that the optimal allocations that solve the reduced problems $\mathrm{P}_a$ and $\mathrm{P}_d$, respectively, are elements of $\mathbb{X}_a^{int}$ or $\mathbb{X}_d^{int}$, respectively. Recall from Lemma 4 that the first best allocation is an element of $\mathbb{X}_a^{int}$ or $\mathbb{X}_d^{int}$, respectively, precisely under the conditions that make either program $\mathrm{P}_a$ or $\mathrm{P}_d$ generate a higher value to the principal. Now consider, for $i = a.b.c.d$, the problems

$$\max_{x,y \in \cup_i \mathbb{X}_i} \mathrm{P}_i$$

The solution to each of these problems converges uniformly to the first-best allocation as $\alpha$ goes to zero. It follows that the solution of program $\mathrm{P}_a$ is in $\mathbb{X}_a^{int}$ for $\alpha$ close enough to zero if $0 < V_{12} < -\frac{\overline{\eta} - \eta}{\overline{\theta} - \underline{\theta}} V_{11}$ and that the solution of program $\mathrm{P}_d$ is in $\mathbb{X}_d^{int}$ for $\alpha$ close enough to zero if $\frac{\overline{\eta} - \eta}{\overline{\theta} - \underline{\theta}} V_{11} < V_{12} < 0$. ■

**Proof of Proposition 3.** From Lemma 4, we have conditions such that the first-best allocation is in $\mathbb{X}_i^{int}$. Hence, in the limit as $\alpha$ goes to zero, the allocations that achieve the maxima $W_i$ are in $\mathbb{X}_i^{int}$. So, we need to show that these maximizers satisfy the neglected constraint. We focus on the case of strong complements. Exactly the same argument can be given for strong substitutes.

For the example, for $\delta \in (-1, 1)$ and $\beta$ sufficiently large to generate interior solutions, the first-best allocation is given by

$$\begin{aligned} x\left(\theta, \eta\right) &= \tfrac{1}{1-\delta^2}\left(\beta\left(1+\delta\right) - \theta - \delta\eta\right) \\ y\left(\theta, \eta\right) &= \tfrac{1}{1-\delta^2}\left(\beta\left(1+\delta\right) - \eta - \theta\delta\right) \end{aligned}$$

The neglected constraint for $x, y \in \mathbb{X}_b$ takes the form

$$\begin{aligned} 0 \;\geq\;& \left(\overline{\theta}-\underline{\theta}\right)\left(x\left(\overline{\theta},\underline{\eta}\right)-x\left(\underline{\theta},\overline{\eta}\right)\right)+\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\overline{\eta}\right)-y\left(\overline{\theta},\underline{\eta}\right)\right) \\ &+\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\overline{\theta},\underline{\eta}\right)-y\left(\underline{\theta},\underline{\eta}\right)\right)+\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)\left(y\left(\underline{\theta},\overline{\eta}\right)-y\left(\overline{\theta},\overline{\eta}\right)\right). \end{aligned}$$

The first-best allocation is in $\mathbb{X}_b$ for $\delta > \frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)}$. The buyer's problem remains concave for $\delta < 1$. Both conditions are satisfied for a nonempty set of parameters only if $\frac{\left(\overline{\eta}-\underline{\eta}\right)}{\left(\overline{\theta}-\underline{\theta}\right)} < 1$. For the example, the neglected constraint is equivalent to

$$\begin{aligned} 0 \;\geq\;& \left(\overline{\theta}-\underline{\theta}\right)\left(\frac{1}{1-\delta^2}\left(-\left(\overline{\theta}-\underline{\theta}\right)+\delta\left(\overline{\eta}-\underline{\eta}\right)\right)\right)+\left(\overline{\eta}-\underline{\eta}\right)\left(\frac{1}{1-\delta^2}\left(-\left(\overline{\eta}-\underline{\eta}\right)\right)\right) \\ &+\lambda\left(\underline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)\left(\frac{1}{1-\delta^2}\left(-\left(\overline{\theta}-\underline{\theta}\right)\delta\right)\right)+\lambda\left(\overline{\theta}\right)\left(\overline{\eta}-\underline{\eta}\right)\left(\frac{1}{1-\delta^2}\left(\left(\overline{\theta}-\underline{\theta}\right)\delta\right)\right), \end{aligned}$$

which is satisfied if $\delta \leq \frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)}$. Since $\frac{\left(\overline{\theta}-\underline{\theta}\right)}{\left(\overline{\eta}-\underline{\eta}\right)} > 1$, this condition is automatically satisfied. ■

# 11 References


Armstrong, Mark and Rochet, Jean-Charles (1999): "Multi-dimensional screening: A user's guide", European Economic Review, 43(4-6), 959-979.

Baron, David P and Besanko, David (1984): "Regulation and information in a continuing relationship", Information Economics and Policy, 1(3), 267–302.

Battaglini, Marco (2005): "Long-term contracting with Markovian consumers", American Economic Review, 95, 637-658.

Battaglini, Marco and Rohit Lamba (2012): "Optimal Dynamic Contracting", unpublished.

Boleslavsky, Raphael and Said, Maher (2012): "Progressive Screening: Long-Term Contracting with a Privately Known Stochastic Process", The Review of Economic Studies, 80(1), 1-34.

Courty, Pascal and Hao, Li (2000): "Sequential screening", The Review of Economic Studies, 67(4), 697–717.

Eső, Peter and Szentes, Balázs (2007a): "Optimal Information Disclosure in Auctions and the Handicap Auction", The Review of Economic Studies, 74(3), 705-731.

Eső, Péter and Szentes, Balázs (2007b): "The Price of Advice", The RAND Journal of Economics, 38(4), 863-880.

Eső, Péter and Szentes, Balázs (2013): "Dynamic Contracting with Adverse Selection: An Irrelevance Result", unpublished.

Krähmer, Daniel and Roland Strausz (2008): "Ex post private information and monopolistic screening", The B.E. Journal of Theoretical Economics: Vol. 8: Iss. 1 (Topics), Article 25

Krähmer, Daniel and Roland Strausz (2012): "The Benefits of Sequential Screening", unpublished.

Krähmer, Daniel and Roland Strausz (2013): "Ex post information rents and disclosure in sequential screening", unpublished.

Li, Hao and Xianwen Shi (2013): “Discriminatory Information Disclosure” unpublished

Myerson, R. (1986): “Multistage Games with Communication”, Econometrica, Vol. 54, No. 2, pp. 323-358

Pavan, Alessandro, Ilya Segal and Juuso Toikka (2012): "Dynamic Mechanism Design", unpublished.

Rochet, J.-C. and L. A. Stole, (2003) “The Economics of Multidimensional Screening” Chapter 5 in Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, Vol. I edited by M. Dewatripont, L. P. Hansen, and S. Turnovsky, Cambridge University Press, Cambridge, 2003, 150-197.