

Litterscheid, Sina; Szalay, Dezsö

Conference Paper

Sequential, multidimensional screening

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Contract Theory, No. A12-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Litterscheid, Sina; Szalay, Dezsö (2014) : Sequential, multidimensional screening, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Contract Theory, No. A12-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100621>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

exterior and the interior design. The landowner's ideas for the exterior design are relatively standard, but her tastes for interior designs are quite particular. As a result, the constructor knows the costs of completing the exterior parts, but he only has a vague idea about the costs relating to the interior parts. However, he will learn these costs as time goes by.

This is a natural situation, in particular in large scale procurement environments, but the situation arises equally naturally in the context of price discrimination. Consumers thinking about buying a (new generation) smart phone know quite well how they value the services and applications they have already been consuming on their old phones. However, they may only have a vague idea about their valuation for new applications. Broadly speaking, a consumer switching from a standard mobile phone to a smart phone knows how many calls he needs to make but only time will tell how much data he will download with the phone. Moreover, casual evidence suggests that firms respond to this information structure.¹

In this paper, we would like to advance our understanding of contracting solutions in these types of environments. We study a model where a principal contracts with an agent to trade a *bundle* of services. Moreover, the agent has private information about the costs of producing one item in the bundle from the outset and privately learns the cost of producing the other item later on. Optimal contracting is dynamic; principal and agent get together both at the outset as well as later on when more information is available. At the outset, the agent decides whether or not he will eventually deliver the bundle of services, but the precise terms of the contract may still be left open at this time. At the second get-together, the remaining details of the contract are specified. The services are produced when all information is available - or, earlier, if the principal should prefer this - and the agent is paid when all services have been produced.

The literature has analyzed problems that share some, but not all the ingredients of our problem. In a nutshell, our problem is a convex combination of a tractable multidimensional problem of screening à la Armstrong and Rochet (1999) and a sequential screening problem

¹E.g., service provider Orange UK offers a choice of pay as you go services and monthly plans. Moreover, selecting into one of these plans limits the options to choose from later on. In particular, conditional on the selected plan a consumer has the possibility to buy one out of a given variety of additional bundles of services.

Specific examples of such options include plans Dolphin and Monkey <http://www.best-mobile-contracts.co.uk/networks/orange.html>. Dolphin offers different bundles of classical services with modern services, i.e. texts, minutes and data volume. Monkey offered bundles of free music and texts.

The key properties for our purposes are that a *bundle* of services is traded and that the *choice* from options is made *sequentially*.

à la Courty and Li (2000). To the best of our knowledge, this is the first paper that takes this approach. We describe the relevant literature in much greater depth below.

We raise and answer the following basic questions. What constraints does incentive compatibility impose in this environment; put differently, what is the set of implementable allocations in the present context? Moreover, what are the qualitative properties of an optimal allocation from the principal's perspective? Is it natural to expect the classical downward distortions in economic activity due to asymmetric information? What are the implications of different timings of production; in particular, how do optimal allocations and utility differ when the principal needs to have construction works on her house start before all information is available?

1.2 Main findings

To answer the first question, we begin with a detailed analysis of binding incentive and participation constraints in our framework. It is instructive to analyze the principal's design problem in two steps. In the first step, the allocation is taken as given and we search for the least cost way of implementing the given allocation. In the second step, we optimize over the allocations.

The problem of implementing given allocations is rich. In particular, one needs to understand systematically which deviations are most tempting for the agent as a function of the allocation that the principal wishes to implement. As is well known, in multidimensional problems the set of binding constraints at the optimum changes with the allocation. However, the timing of the agent's learning process simplifies our problem. Assuming that the agent's cost parameters are positively correlated, we can identify from the outset - for any incentive compatible allocation - two constraints that must be binding in any optimal contract. In particular, the agent with a high cost of constructing the exterior parts of the house is indifferent between participating and not participating; the agent with the low cost of exterior construction is indifferent between reporting this parameter truthfully or not and moreover obtains a rent. The level of this rent depends on what the agent with an initial low cost realization would report in the second round of communication, had he falsely reported his cost of constructing the exterior parts as high. Depending on how sensitive the allocation variables respond to information that arrives late, the agent's best deviation features truthtelling or lying after a false report in the first round of communication.²

²It is important to point out that an appropriate version of the revelation principle (Myerson (1986))

Once the pattern of binding constraints for given allocations is known, it is straightforward to optimize with respect to the allocation the principal wishes to implement. The solution depends crucially on the nature and strength of interactions between the items in the bundle in the principal's payoff function. For mild complements, optimal allocations induce truthtelling in the second period regardless of first period reports and display the classical downward distortions relative to the first-best allocation. For mild substitutes, truthtelling after a false first period report constitutes a binding constraint and the optimal allocation displays both up and downward distortions. For strong complements or strong substitutes, the optimal allocation triggers a second period lie after a first period lie. Moreover, the allocation can feature both up and downward distortions even in the very case where goods are strong complements for the principal.

Which of these cases are economically most relevant? What are reasonable assumptions on the strength of interactions between the items in the bundle the principal consumes? We obtain guidance from the comparative statics properties of the first-best allocation if we are willing to impose that changes in marginal costs of producing one item have more of an impact on the level of that item rather than the other one. If moreover the support of second period information is at least as wide as the support of first period information, then only the case of mild complements and substitutes is relevant. For applications in which these assumptions make sense, we provide a strikingly simple cook-book recipe: optimal contracts can be found by imposing truth-telling constraints *on and off* equilibrium path and the procedure picks up the optimum *even if the truthtelling constraints off path are binding*. This latter result strikes us as pretty surprising, because -as cannot be stressed enough - it has nothing to do with the revelation principle but rather emerges from the solution of the overall maximization problem.

It is essentially a corollary of our results that beginning construction works before the constructor has all information comes at a loss to the principal, unless the principal values each item in the bundle independently of the other item. That is, there is a strictly positive

applies in our environment. However, the principle implies only that the agent finds it optimal to announce both parameters truthfully, which implies in particular, that he will report the second parameter truthfully once he has been truthful about the first parameter. The dynamic literature has termed this behavior truthfulness on equilibrium path. The revelation principle has no implications whatsoever on what the agent does off equilibrium path, that is after a first period lie. Clearly, the agent has by definition an incentive to remain on equilibrium path. However, the utility loss to the principal to ensure such remaining on path depends on the agent's best alternative to truthtelling. Therefore, the off-path behavior becomes a crucial ingredient to the analysis.

option value of waiting if the items in the bundle are strict complements or substitutes for the principal. Given complementarity or substitutability, the efficient level of production of both items depends on the costs of producing both items. Even though waiting enables the agent to deviate in more complicated ways, the added flexibility is always valuable to the principal.

1.3 Literature

Our analysis builds on two strands of the literature: multidimensional screening on the one hand and sequential screening on the other hand. Armstrong and Rochet (1999) analyze a tractable model of two-dimensional screening. In their approach, the agent is endowed with all his information from the outset; our agent starts out with only one piece of information and obtains the remaining information later on. Moreover, screening is (optimally) static in their approach, while there are two rounds of contracting in our problem.³ In Courty and Li (2000) an agent is initially endowed with a vague idea about his onedimensional preference parameter and later on refines this information to a particular realization that he privately observes. Moreover, the principal and the agent get together twice - each time new information arrives - as in the present problem. The main difference is that there is only one payoff relevant parameter in Courty and Li (2000), while there are two in the present context. Moreover, there are also two allocation choices in the present context instead of one in their problem. So, our assumptions capture a different set of problems; Courty and Li (2000) analyze problems where information about one payoff relevant parameter is refined while we are interested in collecting multiple informational parameters over time to design a multidimensional object.

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 to consumers whose tastes follow a Markov process. Pavan et al. (2013) 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. In each period, the agent's private information is captured by a onedimensional parameter.

³A second difference is that we allow for substitutability and complementarity between the goods but restrict attention to positive correlation between the informational parameters, while Armstrong and Rochet (1999) consider neutral goods but allow for arbitrary correlations. See also Dana (1993) and Severinov (2008) for further static models of screening in the 2x2-model.

This is the key difference to our problem, where there are two payoff relevant parameters that simultaneously affect the agent’s payoff. Under this assumption, we obtain a natural taxonomy of cases featuring binding constraints with respect to one-shot deviations or double, dynamic deviations, respectively. The latter case, by definition, fails to satisfy the version of the one-stage-deviation principle by Pavan et al. (2013), which applies precisely when the best deviation for the agent is to lie once and then to return to truthful reporting strategies forever after. So, ultimately the qualitative differences of our contracting solutions as compared to those in Pavan et al. (2013) are due to the one-stage-deviation principle applying or failing, respectively. In turn, multidimensionality provides a natural reason for the failure of the one-stage-deviation principle.

Complementary to this paper is contemporaneous 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 the dynamic screening problem separation may not be feasible even though it would be feasible in the static counterpart of the model. In particular, Battaglini and Lamba (2013) provide natural examples where locally optimal contracts fail to satisfy global incentive constraints. 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 is standard, but the dynamic 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, in particular, dynamic pooling: initial separation followed by pooling in later periods.

The literature on dynamic and sequential screening has analyzed problems that share some of our model’s features but not all. We are by no means the first ones to analyze sequentially optimal lying strategies; however, to the best of our knowledge, this is the first paper that bridges the gap between the multidimensional and the sequential screening literature, and in this context dynamic lying strategies arise naturally. Sequentially optimal lies are also analyzed in Esó 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 mechanism 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), Krämer and Strausz (2012,2013) and Li and Shi (2013); for a combined model of moral hazard and adverse selection, see Garrett and Pavan (2013). Bhaskar (2013) analyzes dynamic deviation strategies in the pure moral hazard model. A question related to our timing application is addressed in Krämer 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.

We clearly do not do any justice to the multidimensional literature in this short account. For an in depth survey, see Rochet and Stole (2003); for general approaches, see Armstrong (1996) and Rochet and Choné (1998).⁴

The paper is organized as follows. In section two, we present the model and state the buyer's problem. Section three presents and solves the buyer's problem. Section four 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 five gives an example that is outside this regular structure. In section six, we discuss the optimal timing of productive decisions in our model. The final section concludes. Lengthy proofs are gathered in the appendix.

2 The Model

2.1 Setup

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

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

⁴See also Pavan et al. (2013) for a much more extensive survey of the literature on dynamic mechanism design.

where θ and η are cost shifters.

Contracting is a sequential process. At date 1, the seller knows the realization of θ (but not of η) and the conditional distribution of η given θ , whereas the buyer only knows the joint distribution of types. The cost realizations are binary, so that $\theta \in \{\underline{\theta}, \bar{\theta}\}$ and $\eta \in \{\underline{\eta}, \bar{\eta}\}$, where $\bar{\theta} > \underline{\theta} > 0$ and $\bar{\eta} > \underline{\eta} > 0$. The joint distribution is completely characterized by $\Pr(\theta = \underline{\theta}) = \alpha$ and $\lambda(\theta) \equiv \Pr\{\eta = \underline{\eta} | \theta\}$. At date 2, η becomes known to the seller but not to the buyer. Also, goods are produced and traded in exchange for the transfer T at that date. The game and the information structure is common knowledge.⁵

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 $\lim_{x \rightarrow 0} V_1(x, y) = \infty$ for all y and $\lim_{y \rightarrow 0} V_2(x, y) = \infty$ for all x .⁶ Further assumptions will be discussed as we go along.

2.2 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, where the agent announces each piece of information when it arrives. Thus, the contracting game is dynamic and involves two rounds of communication. In the first round at date 1, the seller reports a value $\hat{\theta} \in \{\underline{\theta}, \bar{\theta}\}$; in the second round at date 2, the seller reports a value $\hat{\eta} \in \{\underline{\eta}, \bar{\eta}\}$. The seller is given incentives to announce these values truthfully. This implies in particular, that truthfulness about η is optimal after a truthful report about θ . To rule out all feasible deviations by the seller, we need to analyze also what the seller would announce about η off equilibrium path, that is, had he falsely reported θ in the first round of communication.⁷ Since the optimal behavior of the agent in the second round depends on the first round report, $\hat{\theta}$, the first round true type, θ , and the second round true type, η , we need to distinguish between the incremental information that arrives in round two and the agent's

⁵In our model there are two choice variables, x and y , that interact with two informational variables, θ and η . The essential difference to Courty and Li (2000) is that θx enters the agent's payoff function.

⁶Throughout the paper, $V_i(x, y)$ and $V_{ij}(x, y)$ for $i, j = 1, 2$ denote partial and cross derivatives of the function V with respect to its arguments.

⁷It is important to notice that the revelation principle does not have any implications on 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 deviation in the first round of communication, we need to consider the possibility that the optimal thing to do in the second round after a first round lie is to lie again.

private information. That is, in the second period, the agent privately knows which node, identified by the triple $(\theta, \hat{\theta}, \eta)$, in the game tree has been reached. We let $\hat{\eta}^*(\theta, \hat{\theta}, \eta)$ for $\hat{\theta} \neq \theta$ denote the optimal report at node $(\theta, \hat{\theta}, \eta)$ and treat these reports as choice variables (subject to incentive compatibility constraints) of the principal.

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 about 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:

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

s.t.

$$T(\bar{\theta}, \underline{\eta}) - \bar{\theta}x(\bar{\theta}, \underline{\eta}) - \underline{\eta}y(\bar{\theta}, \underline{\eta}) \geq T(\bar{\theta}, \bar{\eta}) - \bar{\theta}x(\bar{\theta}, \bar{\eta}) - \bar{\eta}y(\bar{\theta}, \bar{\eta}), \quad (2)$$

$$T(\bar{\theta}, \bar{\eta}) - \bar{\theta}x(\bar{\theta}, \bar{\eta}) - \bar{\eta}y(\bar{\theta}, \bar{\eta}) \geq T(\bar{\theta}, \underline{\eta}) - \bar{\theta}x(\bar{\theta}, \underline{\eta}) - \bar{\eta}y(\bar{\theta}, \underline{\eta}) \quad (3)$$

$$T(\underline{\theta}, \underline{\eta}) - \underline{\theta}x(\underline{\theta}, \underline{\eta}) - \underline{\eta}y(\underline{\theta}, \underline{\eta}) \geq T(\underline{\theta}, \bar{\eta}) - \underline{\theta}x(\underline{\theta}, \bar{\eta}) - \underline{\eta}y(\underline{\theta}, \bar{\eta}) \quad (4)$$

$$T(\underline{\theta}, \bar{\eta}) - \underline{\theta}x(\underline{\theta}, \bar{\eta}) - \bar{\eta}y(\underline{\theta}, \bar{\eta}) \geq T(\underline{\theta}, \underline{\eta}) - \underline{\theta}x(\underline{\theta}, \underline{\eta}) - \bar{\eta}y(\underline{\theta}, \underline{\eta}) \quad (5)$$

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

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

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

$$\mathbb{E}_{\eta|\bar{\theta}} [T(\bar{\theta}, \eta) - \bar{\theta}x(\bar{\theta}, \eta) - \eta y(\bar{\theta}, \eta)] \geq 0, \quad (9)$$

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

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

for all $\theta, \hat{\theta} \in \{\underline{\theta}, \bar{\theta}\}$ and $\eta \in \{\underline{\eta}, \bar{\eta}\}$.

Constraints (2) through (5) are the second period constraints after a truthful report in the first period: after such a truthful report in period one, the seller must find it optimal to be truthful about η as well. We term these constraints ‘‘on-path constraints’’ for the obvious

reason. (6) and (7) are the first period incentive constraints. As of date one, the seller anticipates that after having misreported θ in the first period, he chooses the second period report optimally, as captured by (10). Since nodes $(\theta, \hat{\theta}, \eta)$ for $\theta \neq \hat{\theta}$ are off the equilibrium path (on equilibrium path, the first report is truthful), we term constraints ensuring any particular behavior at such off-path nodes “off-path incentive constraints”. (8) and (9) are the participation constraints.

Before diving into the quite intricate analysis, it is useful to take a bird’s eye view of the problem. At the first time of contracting, there are only two possible types - because the seller only knows θ but not yet η . Moreover, the seller decides whether or not he wishes to participate at that date. He anticipates optimal behavior at date 2, so each report gives rise to a continuation value. Due to this structure, our model has much in common with the simple (static) binary model of screening, so much of the logic of that model will carry over. The essential complication relative to the static counterpart is that the continuation values are endogenous and there is no simple shortcut to determine these values.

It is instructive to understand the properties of the first-best allocation.

2.3 The first-best

If the buyer and the seller both know θ at the outset and both learn η at date two, then both (8) and (9) are binding at the optimum and the optimal allocation satisfies

$$V_1(x(\theta, \eta), y(\theta, \eta)) = \theta \tag{11}$$

and

$$V_2(x(\theta, \eta), y(\theta, \eta)) = \eta \tag{12}$$

for $\theta \in \{\underline{\theta}, \bar{\theta}\}$ and $\eta \in \{\underline{\eta}, \bar{\eta}\}$.

To sharpen our intuition for “relevant” cases in the second-best, it proves useful to ask how the first-best solution depends on the cost parameters. Obviously, x and y move in the same direction in response to changes in the parameters if x and y are complements and move in opposite directions if x and y are substitutes. Beyond that, it is important to understand how strongly these choices respond to information that is learned in period 2.

Lemma 1 *If x and y are complements ($V_{12}(x, y) \geq 0$ for all x, y), then the first-best allocation, defined by (11) and (12), satisfies*

$$y(\theta, \underline{\eta}) - y(\theta, \bar{\eta}) \geq (\leq) x(\theta, \underline{\eta}) - x(\theta, \bar{\eta})$$

for arbitrary $\bar{\eta} - \underline{\eta} > 0$ if and only if $V_{12}(x, y) \leq (\geq) -V_{11}(x, y)$ for all x, y .

If x and y are substitutes, ($V_{12}(x, y) \leq 0$ for all x, y), then the first-best allocation satisfies

$$y(\theta, \underline{\eta}) - y(\theta, \bar{\eta}) \geq (\leq) - (x(\theta, \underline{\eta}) - x(\theta, \bar{\eta}))$$

for arbitrary $\bar{\eta} - \underline{\eta} > 0$ if and only if $V_{12}(x, y) \geq (\leq) V_{11}(x, y)$ for all x, y .

If the utility function of the buyer features interactions that are not too strong, then a change in η has a stronger impact on y than on x . We believe this is the natural case, but other cases are possible.⁸

We now address the buyer's problem under asymmetric information.

3 Analysis

We assume that the low cost producer in the first period is better to the buyer than the high cost producer in the sense of a weakly positive correlation

Assumption 1: θ and η are weakly positively correlated, that is $\lambda(\underline{\theta}) \geq \lambda(\bar{\theta})$.

First-order stochastic dominance is a regularity condition that is commonly used in the sequential screening literature. Assumption 1 implies the following Lemma:

Lemma 2 *If $\lambda(\underline{\theta}) \geq \lambda(\bar{\theta})$, 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 2 we can deduce that constraint (9) is binding at the optimum. Likewise, at least one of the first period 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

⁸In the context of strategic interactions, the natural case would generate stability of a system of best replies; see, e.g., Tirole (1988) for a discussion. Note moreover that joint concavity with respect to x and y requires that $V_{11}V_{22} - V_{12}^2 \geq 0$, so a concave function cannot be "irregular" both with respect to changes in η and θ .

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 3 $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \underline{\eta}$ for $x(\bar{\theta}, \underline{\eta}) \geq x(\bar{\theta}, \bar{\eta})$ and $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \bar{\eta}$ for $x(\underline{\theta}, \bar{\eta}) \geq x(\underline{\theta}, \bar{\eta})$. Likewise, $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \bar{\eta}$ for $x(\bar{\theta}, \underline{\eta}) \leq x(\bar{\theta}, \bar{\eta})$ and $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \underline{\eta}) = \underline{\eta}$ for $x(\underline{\theta}, \bar{\eta}) \leq x(\underline{\theta}, \bar{\eta})$.

The on-path constraints have some, however limited, implications for the optimal reports off path. In particular, it is *never* the case that the agent finds it optimal to lie at *all* off-path nodes in the second period. Depending on the monotonicity properties of the x -allocation, there are always some nodes at which truthtelling about the second period incremental information is automatically - by implication of the on-path constraints - induced. The intuition is quite simple. E.g., the on-path constraints of type $(\bar{\theta}, \underline{\eta})$ make reporting $\hat{\eta} = \underline{\eta}$ optimal for that type. At node $(\underline{\theta}, \bar{\theta}, \underline{\eta})$, the agent has an even stronger incentive to report $\hat{\eta} = \underline{\eta}$ if $x(\bar{\theta}, \underline{\eta}) \geq x(\bar{\theta}, \bar{\eta})$, because that boosts his extra rent from having exaggerated θ in round one.

The difficulty at this stage is of course that the monotonicity of the x -allocation with respect to η is not known and endogenous. Our solution strategy is as follows. Building on the insights from static models of screening, we aim for a reduced problem, where constraints (9) and (6) hold as equalities, while (7) (in addition to (8)) is slack. We solve this reduced problem and provide sufficient conditions such that its solution satisfies the neglected constraint (7). In turn, the reduced problem is tackled 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 first step problem we simultaneously optimize over payments and off-path reports.

3.1 The reduced problem

If the agent with first period cost $\bar{\theta}$ is indifferent between participating and not, the agent with first period type $\underline{\theta}$ is indifferent between being truthful and lying about θ , and the remaining first period constraints are slack, then the principal faces the standard trade-off between the efficiency of the allocation and the rent that needs to be given to ex ante type $\underline{\theta}$. Denote this rent as Δ . It is useful to split the principal's problem into two steps. In the first step, we take the allocation as given and determine optimal payments that implement the

allocation. Implementability of the allocation includes that the incentive constraint of the ex ante type $\bar{\theta}$ needs to be satisfied as well. Formally, letting Ω denote the expected profit ex ante type $\bar{\theta}$ can make by mimicking type $\underline{\theta}$, we require that $\Omega \leq 0$. Once the optimal payments are known, we maximize with respect to the allocation that the principal wishes to implement.

Identifying the minimal payments, while straightforward in the static model, is pretty involved in the present context. The reason is that implementation is much more flexible in the multidimensional context and so the pattern of binding constraints is not obvious. The reader who is not interested in the details of this step can skip subsection 3.1.1, consult Lemma 4 for the solution to the problem, and continue reading from subsection 3.1.2 onwards on a first go. However, to ultimately understand the structure of optimal allocations in sections 4 and 5 below, the reader needs to go back to subsection 3.1.1 to relate the pattern of binding constraints to the underlying model primitives.

3.1.1 Implementing given allocations at lowest cost

For a given allocation $(x, y)^9$, payments to types $(\bar{\theta}, \eta)$ and optimal off-path reporting at nodes $(\underline{\theta}, \bar{\theta}, \eta)$ for $\eta \in \{\underline{\eta}, \bar{\eta}\}$ solve the following problem:

$$\Delta \equiv \min_{\{T(\bar{\theta}, \eta), \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}}} \mathbb{E}_{\eta|\underline{\theta}} [T(\bar{\theta}, \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \eta)) - \underline{\theta}x(\bar{\theta}, \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \eta)) - \eta y(\bar{\theta}, \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \eta))] \quad (13)$$

s.t.

$$\begin{aligned} \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \eta) &\in \arg \max_{\hat{\eta}} T(\bar{\theta}, \hat{\eta}) - \underline{\theta}x(\bar{\theta}, \hat{\eta}) - \eta y(\bar{\theta}, \hat{\eta}) \text{ for } \eta \in \{\underline{\eta}, \bar{\eta}\} \\ \mathbb{E}_{\eta|\bar{\theta}} [T(\bar{\theta}, \eta) - \bar{\theta}x(\bar{\theta}, \eta) - \eta y(\bar{\theta}, \eta)] &= 0, \\ &(2), \text{ and } (3). \end{aligned}$$

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 η when he has misreported his parameter θ in the first period. However, if the buyer wishes to implement such a sequential lying strategy - 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.

⁹Throughout the paper we denote by (x, y) the allocation for all types $(\theta, \eta) \in \{\underline{\theta}, \bar{\theta}\} \times \{\underline{\eta}, \bar{\eta}\}$.

Once the solution to the first program is found, we can choose payments to types $(\underline{\theta}, \eta)$ and the optimal reporting at nodes $(\bar{\theta}, \underline{\theta}, \eta)$ for $\eta \in \{\underline{\eta}, \bar{\eta}\}$ to render constraint (7) as slack as can be. Formally, given the payments and reports $\{T(\bar{\theta}, \eta), \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}}$ that solve program (13), payments and reports $\{T(\underline{\theta}, \eta), \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}}$ solve the problem:

$$\Omega \equiv \min_{\{T(\underline{\theta}, \eta), \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}}} \mathbb{E}_{\eta|\bar{\theta}} [T(\underline{\theta}, \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \eta)) - \bar{\theta}x(\underline{\theta}, \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \eta)) - \eta y(\underline{\theta}, \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \eta))] \quad (14)$$

s.t.

$$\begin{aligned} \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \eta) &\in \arg \max_{\hat{\eta}} T(\underline{\theta}, \hat{\eta}) - \bar{\theta}x(\underline{\theta}, \hat{\eta}) - \eta y(\underline{\theta}, \hat{\eta}) \text{ for } \eta \in \{\underline{\eta}, \bar{\eta}\} \\ \mathbb{E}_{\eta|\underline{\theta}} [T(\underline{\theta}, \eta) - \underline{\theta}x(\underline{\theta}, \eta) - \eta y(\underline{\theta}, \eta)] &= \Delta, \\ &(4), \text{ and } (5), \end{aligned}$$

Notice that we solve problem (14) only after having solved problem (13). This procedure reflects our solution strategy that is based on reduced problems where constraint (7) is slack. As long as (7) is slack - formally, as long as $\Omega \leq 0$ - only the solution of problem (13) is directly payoff relevant. For this reason, we focus primarily on program (13) in the main text and relegate the solution to program (14) entirely to the appendix. We come back to these results only when we verify that the neglected constraint, (7), is indeed satisfied.

The solution to the programs depends on the allocation that the buyer wishes to implement. In particular, define $\Delta x(\theta) \equiv x(\theta, \underline{\eta}) - x(\theta, \bar{\eta})$, $\Delta y(\theta) \equiv y(\theta, \underline{\eta}) - y(\theta, \bar{\eta})$, and the following sets

$$\begin{aligned} \mathbb{X}_i(\theta) &\equiv \left\{ \{(x(\theta, \eta), y(\theta, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \mid (\bar{\eta} - \underline{\eta}) \Delta y(\theta) \geq (\bar{\theta} - \underline{\theta}) \Delta x(\theta) \geq 0 \right\}; \\ \mathbb{X}_{ii}(\theta) &\equiv \left\{ \{(x(\theta, \eta), y(\theta, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \mid (\bar{\eta} - \underline{\eta}) \Delta y(\theta) \geq -(\bar{\theta} - \underline{\theta}) \Delta x(\theta) \geq 0 \right\}; \\ \mathbb{X}_{iii}(\theta) &\equiv \left\{ \{(x(\theta, \eta), y(\theta, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \mid (\bar{\theta} - \underline{\theta}) \Delta x(\theta) \geq (\bar{\eta} - \underline{\eta}) \Delta y(\theta) \geq 0 \right\}; \\ \mathbb{X}_{iv}(\theta) &\equiv \left\{ \{(x(\theta, \eta), y(\theta, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \mid -(\bar{\theta} - \underline{\theta}) \Delta x(\theta) \geq (\bar{\eta} - \underline{\eta}) \Delta y(\theta) \geq 0 \right\}. \end{aligned}$$

For future reference, also define $\mathbb{X}_j^{int}(\theta)$ for $j = i, \dots, iv$ as these same sets when all the defining inequalities are strict, $\mathbb{X}_j \equiv \mathbb{X}_j(\underline{\theta}) \cup \mathbb{X}_j(\bar{\theta})$ for $j = i, \dots, iv$, and finally $\mathbb{X}(\theta) \equiv \bigcup_{j=i}^{iv} \mathbb{X}_j(\theta)$. These sets are depicted in the following graph:

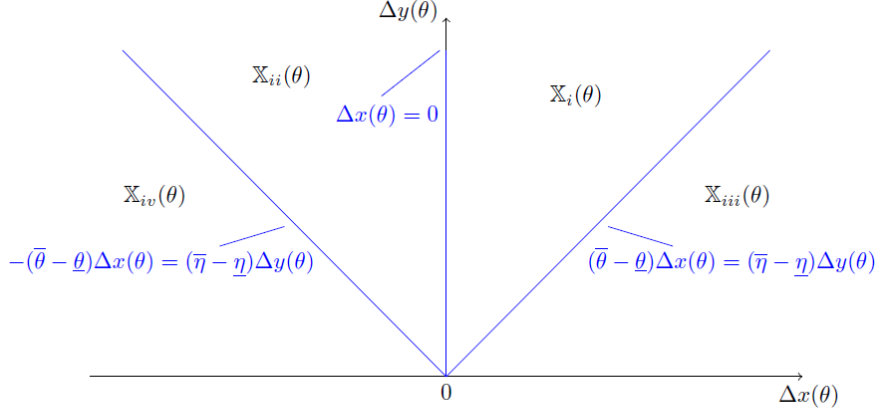


Figure 1. The space of implementable allocations is divided into four regions, i through iv, for each type $\theta \in \{\underline{\theta}, \bar{\theta}\}$. The cost minimizing payments that implement allocations within each regime depend on the regime itself.

Only y -allocations that are monotonic in η are incentive compatible. Hence, we only need to consider such allocations. From Lemma 3 we know that depending on the monotonicity of the x -allocation, truthful reporting is automatic at some nodes off path. Whether it is optimal to induce truthful reporting at the remaining nodes off path depends on which of the sets $\mathbb{X}_j(\theta)$ for $\theta \in \{\underline{\theta}, \bar{\theta}\}$ and $j = i, \dots, iv$ contain the allocation x, y that is implemented. Note that the sets $\mathbb{X}_j(\theta)$ for $j = i, \dots, iv$ are defined for both values of θ . The solution to program (13) depends on which set $\mathbb{X}_j(\bar{\theta})$ contains the allocation offered to ex ante type $\bar{\theta}$; the solution to (14) depends on $\mathbb{X}_j(\underline{\theta})$. Very conveniently, the dividing lines between the sets have isomorphic representations. The complete set of implementable allocations is thus given by $\mathbb{X}_j(\underline{\theta}) \times \mathbb{X}_k(\bar{\theta})$ for $j, k = i, ii, iii, iv$, leaving us with 16 possibilities. However, it turns out that under very natural conditions, the solution of the overall problem has the property that $j = k$, so there are only 4 cases *economically* relevant in our model. Therefore, to economize on space, we just present our result anticipating this result:

Lemma 4 For $(x, y) \in \mathbb{X}_i$

$$\Delta = \Delta_i \equiv \mathbb{E}_{\eta|\underline{\theta}} [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \eta)] + (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})$$

incentive compatibility constraints (2) and (5) are binding and all types report truthfully off path;

for $(x, y) \in \mathbb{X}_{ii}$

$$\Delta = \Delta_{ii} \equiv \mathbb{E}_{\eta|\bar{\theta}} [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \eta)] + (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})$$

all types report truthfully off path and the agent is indifferent between truthfully reporting and lying at nodes $(\underline{\theta}, \bar{\theta}, \underline{\eta})$ and $(\bar{\theta}, \underline{\theta}, \bar{\eta})$;

for $(x, y) \in \mathbb{X}_{iii}$

$$\Delta = \Delta_{iii} \equiv (\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \underline{\eta}) + (1 - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta}) - (1 - \lambda(\underline{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \underline{\eta})$$

incentive compatibility constraints (2) and (5) are binding and $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \underline{\eta}$ and $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \underline{\eta}) = \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \bar{\eta}$;

for $(x, y) \in \mathbb{X}_{iv}$

$$\Delta = \Delta_{iv} \equiv (\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \bar{\eta}) + \lambda(\underline{\theta}) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta}) - \lambda(\bar{\theta}) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \underline{\eta})$$

incentive compatibility constraints (3) and (4) are binding and $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \bar{\eta}$ and $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \hat{\eta}^*(\bar{\theta}, \underline{\theta}, \underline{\eta}) = \underline{\eta}$;

Allocations in \mathbb{X}_i induce truthtelling off path automatically in the sense that we can naïvely assume truthtelling off path. Solving program (13) under this hypothesis, we find that payments are minimized if constraint (2) is binding. In turn, for these payments, it is straightforward to verify that $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \bar{\eta}$. Taken together with Lemma 3, this implies the result. Intuitively, suppose that x is independent of η . In this case, truthtelling about η is simply a question of monotonicity of y in η . This intuition generalizes to all allocations in \mathbb{X}_i that share the property that x and y move in the same direction and y is more sensitive to changes in η than x is.

For allocations in \mathbb{X}_{ii} conjecturing truthtelling naïvely 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 at nodes $(\underline{\theta}, \bar{\theta}, \bar{\eta})$ and $(\bar{\theta}, \underline{\theta}, \underline{\eta})$ is indeed to tell the truth, this needs to be ensured explicitly with the appropriate constraints at nodes $(\underline{\theta}, \bar{\theta}, \underline{\eta})$ and $(\bar{\theta}, \underline{\theta}, \bar{\eta})$. Moreover, these constraints are binding at the optimum. Finally, when the dependency of the x -allocation on information η becomes strong, it becomes too costly to insist on truthtelling at all nodes off path. Instead, the cheapest way to implement any given allocation in sets \mathbb{X}_{iii} and \mathbb{X}_{iv} induces some type to lie off path. Intuitively, take again the extreme cases within sets \mathbb{X}_{iii} and \mathbb{X}_{iv} and suppose y is independent of η . Clearly, there is no way to induce truthtelling about η in period 2 in this case. Instead, the seller chooses the report that maximizes his rent from being able to produce x at lower cost, so he chooses the report that maximizes $x(\theta, \hat{\eta})$. E.g., for $(x, y) \in \mathbb{X}_{iii}$ the seller always reports $\hat{\eta} = \underline{\eta}$.

The functional form of the minimal rents that the agent with first period type $\underline{\theta}$ depends on which of the regimes i through iv prevails. The cases are ordered by increasing complexity.

Case i (where $(x, y) \in \mathbb{X}_i$) is the standard one, where the agent announces η truthfully in period two regardless of the report about θ . The expected rent of an agent with parameter $\theta = \underline{\theta}$ consists of two parts. First, the agent has a lower cost of producing x than the agent with $\theta = \bar{\theta}$. The expected cost advantage is $\mathbb{E}_{\eta|\underline{\theta}} [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \eta)]$, because η (which will be announced truthfully) is not yet known when θ is announced. Secondly, type $\underline{\theta}$ has a higher probability of being relatively more efficient at producing y . Moreover, in period two, an agent with a parameter $\eta = \underline{\eta}$ receives a higher utility than an agent with parameter $\bar{\eta}$, because this agent could always overstate his cost of producing y . At the optimum, the agent of type $(\bar{\theta}, \underline{\eta})$ is exactly indifferent between reporting the truth and mimicking type $(\bar{\theta}, \bar{\eta})$, so the difference between type $(\bar{\theta}, \underline{\eta})$'s and type $(\bar{\theta}, \bar{\eta})$'s utility is exactly $(\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})$. The expected additional gain - due to having a low type θ in period one - from this rent arising from having a low rather than a high value of η is exactly equal to $(\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})$.¹⁰

Case ii still features truthtelling on and off equilibrium path, but the agent must be kept indifferent between reporting honestly and exaggerating his η parameter at two nodes; in particular at node $(\underline{\theta}, \bar{\theta}, \underline{\eta})$. We can again split the agent's expected rent as of the first period into the expected direct gain from misreporting θ , $\mathbb{E}_{\eta|\underline{\theta}} [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \eta)]$, and the expected additional gain arising from the combined facts that an agent with a low parameter η obtains a rent in period two and that an agent with parameter $\theta = \underline{\theta}$ forms expectations based on a more favorable distribution of η than an agent with parameter $\theta = \bar{\theta}$ does. Formally, the difference in utilities between types $(\bar{\theta}, \underline{\eta})$ and $(\bar{\theta}, \bar{\eta})$ is set so as to keep the agent who has exaggerated θ in the first period from exaggerating η in the second period and so this difference equals $(\bar{\theta} - \underline{\theta}) (x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta})) + (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})$. Multiplying this term by the difference in distributions, $(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))$, and simplifying, we obtain Δ_{ii} .

In cases iii and iv , the allocation and the associated cost minimizing payments induce the agent to lie off equilibrium path. To avoid repetition, we focus on case iii only. In this case, the most tempting deviation to the agent with $\theta = \underline{\theta}$ is to exaggerate θ and to underreport η at node $(\underline{\theta}, \bar{\theta}, \bar{\eta})$. Put differently, a double deviation involving both parameters is strictly better to the agent with type $\underline{\theta}$ than a single deviation. As a result, the expected cost advantage due to having a low rather than a high value of θ is simply equal to $(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \underline{\eta})$, because the agent reports $\hat{\eta} = \underline{\eta}$ for both realizations of η . Moreover, he obtains the utility level

¹⁰Pavan et al. (2013) have termed the latter expression an "impulse response function", because the term measures the impact of the agent's current information on future allocation choices.

In our problem, both x and y are determined at date two, so things are slightly different, but the intuition is similar.

that type $(\bar{\theta}, \underline{\eta})$ obtains, $(1 - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})$, minus the expected loss in case he has a higher η realization in period two than the type he imitates, $(1 - \lambda(\underline{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \underline{\eta})$.

The optimal payments and the value of the second minimization problem can be found in the appendix. The reason we do not state these things in the main text is that we do not need these results for the discussion of the reduced problem that we now solve.

3.1.2 Optimal allocations in the reduced problem

We can now turn to the design of the optimal allocations in the reduced problem(s). Since we are neglecting constraint (7), we allow for any $\{(x(\underline{\theta}, \eta), y(\underline{\theta}, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}(\underline{\theta})$. Formally, the reduced problem for each constraint set is

$$W_j \equiv \max_{\substack{\{(x(\bar{\theta}, \eta), y(\bar{\theta}, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_j(\bar{\theta}) \\ \{(x(\underline{\theta}, \eta), y(\underline{\theta}, \eta))\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}(\underline{\theta})}} \mathbb{E}_{\theta} \mathbb{E}_{\eta|\theta} [V(x(\theta, \eta), y(\theta, \eta)) - \theta x(\theta, \eta) - \eta y(\theta, \eta)] - \alpha \Delta_j, \quad (\text{P}_j)$$

where Δ_j is defined in Lemma 4. The principal faces a classical trade-off between efficiency and rent extraction, with the complication that the functional form of the rent expression depends on the qualitative features of the allocation that is being implemented, as explained in Lemma 4.

The overall optimum for the buyer is

$$W = \max \{W_i, W_{ii}, W_{iii}, W_{iv}\}.$$

The solution has the following simple structure:

Proposition 1 *Suppose that either $V_{12}(x, y) \geq 0$ for all x, y or $V_{12}(x, y) \leq 0$ for all x, y . If in addition $V_{12}(x, y) \in \left[V_{11}(x, y) \frac{\bar{\eta} - \underline{\eta}}{\bar{\theta} - \underline{\theta}}, -V_{11}(x, y) \frac{\bar{\eta} - \underline{\eta}}{\bar{\theta} - \underline{\theta}} \right]$ for all x, y , then $W = \max \{W_i, W_{ii}\}$. Moreover, $W_i \geq W_{ii}$ if $V_{12} \geq 0$ for all x, y and $W_{ii} > W_i$ if $V_{12} < 0$ for all x, y .*

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

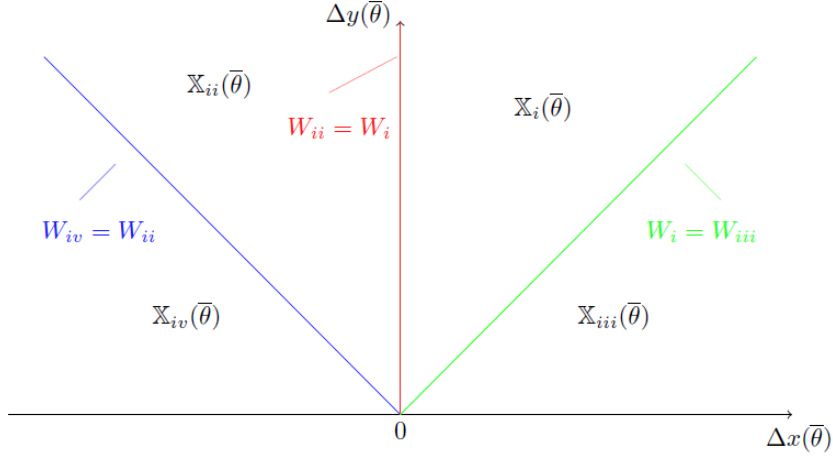


Figure 2. Along the dividing line between any two regimes, payoffs from adjacent programs are equal.

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 i and region ii , the payoffs from programs P_i and P_{ii} are identical for a given allocation. Formally, we have $W_i = W_{ii}$ for a given allocation that satisfies $x(\bar{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta})$. Moreover, none of the programs P_j 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 P_i satisfies $x(\bar{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta})$, whereas the solution to program P_{ii} does not. Since, the allocation that maximizes program P_i is feasible also under program P_{ii} , but is not chosen, it follows by strict concavity of the problem that the value of the objective under program P_{ii} is strictly higher. Likewise, for $V_{12} \geq 0$, the solution to program P_{ii} satisfies $x(\bar{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta})$, so the same argument can be made. However, the subtle difference in this case is that the optimal allocation under program P_i might also lie on the feasibility constraint $x(\bar{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta})$. However, since programs P_i and P_{ii} are identical on the feasibility constraint $x(\bar{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta})$, we have payoff dominance in the weak sense. Essentially the same arguments can be used to compare payoffs from programs P_i and P_{iii} and between programs P_{ii} and P_{iv} .

Complements versus substitutes are enough to determine whether program P_i or program P_{ii} gives a higher payoff. The reason is as follows. We know from Lemma 4 that the rent expression Δ_j is identical in cases i and ii except for differences in $\lambda(\underline{\theta})$ and $\lambda(\bar{\theta})$. Thus, the question is simply which set $X_j(\bar{\theta})$ matches better with the complementarity/substitutability

between the goods. If x and y are complements, then both x and y should be allowed to move in the same direction in response to changes in η . If x and y are substitutes, then x and y should be allowed to vary in opposite directions in response to changes in η . Moreover, the proposition not only compares payoffs between programs P_i and P_{ii} but between all programs P_j . Recall the conditions from Lemma 1 that make y more responsive to changes in η than x . The conditions in Proposition 1 simply adjust the earlier conditions for differences in the supports of early and late information. In particular, the conditions are *identical* if the supports of the parameters have the same width. A pattern of relatively larger variation in y than in x matches better with the sets $\mathbb{X}_i(\bar{\theta})$ and $\mathbb{X}_{ii}(\bar{\theta})$ than with the sets $\mathbb{X}_{iii}(\bar{\theta})$ and $\mathbb{X}_{iv}(\bar{\theta})$. Consequently, under the condition given in the proposition, the maximum of the reduced problem is attained either by problem P_i or P_{ii} .

The reader may verify that the sufficient condition in the proposition 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, the condition is satisfied for $V_{12}(x, y) = \delta \in \left[-\frac{\bar{\eta}-\eta}{\theta-\underline{\theta}}, \frac{\bar{\eta}-\eta}{\theta-\underline{\theta}}\right]$. Note that the utility function is jointly concave in x and y for $\delta \in [-1, 1]$. Thus, for $\frac{\bar{\eta}-\eta}{\theta-\underline{\theta}} \geq 1$, the set of parameter values that *violate* the condition becomes empty. Conversely, there is always a nonempty set of parameter values that generate a concave buyer problem and satisfy the sufficient condition even if $\frac{\bar{\eta}-\eta}{\theta-\underline{\theta}} < 1$. In this sense - at least in this example - the sufficient condition isolates the important case rather than the pathological one. Therefore, we impose henceforth

Assumption 2: either $V_{12}(x, y) \geq 0$ for all x, y or $V_{12}(x, y) \leq 0$ for all x, y and in addition $V_{12}(x, y) \in \left[V_{11}(x, y) \frac{\bar{\eta}-\eta}{\theta-\underline{\theta}}, -V_{11}(x, y) \frac{\bar{\eta}-\eta}{\theta-\underline{\theta}}\right]$ for all x, y .

We solve the full problem under this assumption. However, there are clearly cases that violate Assumption 2. For that reason, we discuss a particular case that violates Assumption 2 in section 5 below.

3.2 The solution to the full 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 the set $\mathbb{X}_j(\underline{\theta})$ that contains the allocation offered to the ex-ante type $\underline{\theta}$, $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\eta, \bar{\eta}\}}$. Note that this allocation corresponds simply

to the first-best allocation for that type. Since intuition and proof for the following result are essentially the same as for Lemma 1, we state without further discussion:

Lemma 5 *The first-best allocation defined by (11) and (12) satisfies $(x, y) \in$*

$$\begin{aligned} \mathbb{X}_i & \quad \text{if } 0 \leq V_{12}(x, y) \leq -\frac{\bar{\eta}-\eta}{\bar{\theta}-\theta}V_{11}(x, y) \text{ for all } x, y; \\ \mathbb{X}_{ii} & \quad \text{if } 0 \geq V_{12}(x, y) \geq \frac{\bar{\eta}-\eta}{\bar{\theta}-\theta}V_{11}(x, y) \text{ for all } x, y; \\ \mathbb{X}_{iii} & \quad \text{if } V_{12}(x, y) \geq -\frac{\bar{\eta}-\eta}{\bar{\theta}-\theta}V_{11}(x, y) \text{ for all } x, y; \\ \mathbb{X}_{iv} & \quad \text{if } V_{12}(x, y) \leq \frac{\bar{\eta}-\eta}{\bar{\theta}-\theta}V_{11}(x, y) \text{ for all } x, y. \end{aligned}$$

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

Combining Proposition 1 and Lemma 5, we observe that the solution to the reduced problem satisfies $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_j(\underline{\theta})$ if and only if $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_j(\bar{\theta})$ for $j = i, ii$. Obviously, this was the reason to economize on space in Lemma 4 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) goods are independent ($V_{12}(x, y) = 0$ for all x, y) or

II)

$$\max_{x, y} \left| \frac{V_{12}}{V_{11}V_{22} - V_{12}^2}(x, y) \right| \leq \frac{\bar{\theta} - \theta}{\bar{\eta} - \eta} \min_{x, y} \left| \frac{V_{22}}{V_{11}V_{22} - V_{12}^2}(x, y) \right|, \quad (15)$$

and either

a) $(x, y) \in \mathbb{X}_i^{int}$ (which holds in particular for α small enough if x, y are strict complements for all x, y), or

b) $(x, y) \in \mathbb{X}_{ii}^{int}$ (which holds in particular for α small enough if x, y are strict substitutes for all x, y) and in addition $\lambda(\underline{\theta}) = \lambda(\bar{\theta})$.

For independent goods, Assumption 2 is enough to guarantee that the reduced problem picks up the overall optimum. For strict complements or substitutes we need to impose additional structure. For strict complements (substitutes) and α sufficiently small, the entire allocation is an element of \mathbb{X}_i^{int} (\mathbb{X}_{ii}^{int}). The reason is that in the limit where α tends to zero, the second best allocation converges to the first-best allocation, whose properties we have described in Lemma 5. Building on this insight, we can go back to Lemma 4 (to be precise

to the proof of Lemma 4 in the appendix) and check the precise functional form of the neglected constraint, (7), and verify whether it is true that $\Omega \leq 0$. Indeed, we have $\Omega \leq 0$ for complements if condition (15) holds; we have $\Omega \leq 0$ for substitutes if condition (15) holds and on top of this the cost parameters are independent, $\lambda(\underline{\theta}) = \lambda(\bar{\theta})$.¹¹

Condition (15) restricts V_{12} relative to V_{22} . The new condition is imposed because Ω depends on the allocation offered to both ex ante types, not just the one offered to one particular type. The condition is satisfied in our example for $\delta \in \left[-\frac{\bar{\theta}-\underline{\theta}}{\bar{\eta}-\underline{\eta}}, \frac{\bar{\theta}-\underline{\theta}}{\bar{\eta}-\underline{\eta}}\right]$. The set of parameters that satisfy both Assumption 2 and condition (15) is always nonempty. If $\frac{\bar{\theta}-\underline{\theta}}{\bar{\eta}-\underline{\eta}} = 1$, then the conditions are identical; otherwise, one set is a strict subset of the other.

4 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 $\bar{\theta}$ is an allocation in $\mathbb{X}_i^{int}(\bar{\theta})$ and $\mathbb{X}_{ii}^{int}(\bar{\theta})$. In this case, the optimal allocation for ex ante type $\bar{\theta}$ satisfies

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) = \bar{\theta} + \frac{\alpha}{(1-\alpha)} \frac{\lambda_j}{\lambda(\bar{\theta})} (\bar{\theta} - \underline{\theta})$$

$$V_2(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) = \underline{\eta}$$

and

$$V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) = \bar{\theta} + \frac{\alpha}{(1-\alpha)} \frac{(1-\lambda_j)}{(1-\lambda(\bar{\theta}))} (\bar{\theta} - \underline{\theta})$$

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

where $j = i, ii$ and by convention $\lambda_i = \lambda(\underline{\theta})$ and $\lambda_{ii} = \lambda(\bar{\theta})$. The optimal allocation for ex ante type $\underline{\theta}$ is given by (11) and (12).

For the case of complements, the optimal allocation for ex ante type $\bar{\theta}$ displays the standard downward distortions relative to the first-best. For strictly positive complementarities,

¹¹Note that the conditions in part II of the proposition are far from necessary. E.g., one can also derive sufficient conditions for the case of substitutes and strictly positive correlation.

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, $x(\bar{\theta}, \underline{\eta})$ is distorted downwards and as a result, $y(\bar{\theta}, \underline{\eta})$ is distorted upwards.

Building on the discussion following Lemma 4, the intuition for the first-order conditions is straightforward. The first-order conditions for $x(\bar{\theta}, \underline{\eta})$ display the trade-off between efficiency and extraction of rents due to lower costs of producing x . $(1 - \alpha)\lambda(\bar{\theta})$ and $(1 - \alpha)(1 - \lambda(\bar{\theta}))$, respectively, are the probabilities that the cost realizations equal $(\bar{\theta}, \underline{\eta})$ and $(\bar{\theta}, \bar{\eta})$, respectively. These are the weights attached to the efficiency motive. On the other hand, a change in the x allocation affects the agent's expected rent by $\lambda_j(\bar{\theta} - \underline{\theta})$ and $(1 - \lambda_j)(\bar{\theta} - \underline{\theta})$, respectively. A change in the $y(\bar{\theta}, \underline{\eta})$ allocation does not affect the agent's rent, whereas a change in the $y(\bar{\theta}, \bar{\eta})$ allocation affects the agent's expected rent by $(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\theta} - \underline{\theta})$. These effects are weighted by α , the probability that $\theta = \underline{\theta}$.

5 The case of strong interactions

So far, we have characterized optimal allocations for regular cases, where the strength of interactions between the goods is relatively mild. If the ratio $\frac{\bar{\eta} - \underline{\eta}}{\bar{\theta} - \underline{\theta}}$ is relatively large, then “most” utility functions will display relatively mild interactions between the goods in this sense. This loose 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 $\mathbb{X}_i \cup \mathbb{X}_{ii}$ becomes optimal. Specifically, we have the following result:

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

For $\delta \in \left(\frac{\bar{\eta} - \underline{\eta}}{\bar{\theta} - \underline{\theta}}, 1\right)$, it follows from Lemma 5 that the solution of the reduced problem is an element of \mathbb{X}_{iii}^{int} for α close to zero. Moreover, it is straightforward to verify that $\Omega \leq 0$ in the example. Hence, we have shown that it can be strictly optimal to induce lying off equilibrium path.

Consider now the structure of the optimal allocation for the case where $(x, y) \in \mathbb{X}_{iii}^{int}$. The first-order conditions for the allocation offered to ex ante type $\bar{\theta}$ are as follows:

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

and

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

This allocation displays upwards distortions in the quantity $y(\bar{\theta}, \underline{\eta})$, for given quantity $x(\bar{\theta}, \underline{\eta})$. Since we are considering complements, this upwards distortion does not arise simply as a compensating effect due to a downward distortion in $x(\bar{\theta}, \underline{\eta})$, but rather reflects the particular structure of binding incentive constraints for this particular case. Recall from Lemma 4 that the best deviation of an agent with $\theta = \underline{\theta}$ is to report $\hat{\theta} = \bar{\theta}$ in period one and $\hat{\eta} = \underline{\eta}$ in period two, regardless of the actual realization of η . Hence, the reduction in $x(\bar{\theta}, \underline{\eta})$ reflects the fact that the conditional probability of receiving report $\hat{\eta} = \underline{\eta}$ is one; vice versa, there is no rent reduction motive when choosing $x(\bar{\theta}, \bar{\eta})$ at all, because the agent with $\theta = \underline{\theta}$ is never going to imitate type $(\bar{\theta}, \bar{\eta})$. Recall moreover, that in addition to the rents from producing x more efficiently, the agent with type $\theta = \underline{\theta}$ obtains rents from producing y more efficiently; in particular, the agent would obtain (when deviating to $\hat{\theta} = \bar{\theta}$), the utility level that type $(\bar{\theta}, \underline{\eta})$ obtains, $(1-\lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta})$, minus the expected loss in case he has a higher η realization in period two than the type he imitates, $(1-\lambda(\underline{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \underline{\eta})$. As a result, all else equal (that is, for a given x -allocation), the principal reduces $y(\bar{\theta}, \bar{\eta})$ below the first-best level and increases $y(\bar{\theta}, \underline{\eta})$ beyond the first-best level. Whether the overall production levels are above or below first-best depends on the specific utility function.

6 Sequential screening and the value of waiting

What if x needs to be determined already in period one? We can obtain the optimal mechanism with sequential production from our problem if we add the requirement that

$$x(\theta, \underline{\eta}) = x(\theta, \bar{\eta}) \text{ for } \theta \in \{\underline{\theta}, \bar{\theta}\}. \quad (16)$$

Technically, (16) 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 $\theta x(\hat{\theta})$ is sunk by the time the report about η 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 η . 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, (16), 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 behavior).

Solving the transfer minimization problems (13) and (14) for given allocation choices x and y , under the consistency condition (16) 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

$$\begin{aligned} \max_{x(\theta), y(\theta, \eta)} \quad & \mathbb{E}_\theta \mathbb{E}_{\eta|\theta} [V(x(\theta), y(\theta, \eta)) - \theta x(\theta) - \eta y(\theta, \eta)] \\ & - \alpha [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}) + (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta})] \end{aligned}$$

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

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

The following proposition is now obvious:

Proposition 4 *Delayed and early production achieve the same payoff only for independent goods. If $V_{12}(x, y) > (<) 0$ 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 $\bar{\theta}$ satisfies the first-order conditions

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

$$V_2(x(\bar{\theta}), y(\bar{\theta}, \underline{\eta})) = \underline{\eta},$$

and

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

The expected marginal benefit of $x(\bar{\theta})$ is equal to $\bar{\theta} + \frac{\alpha}{1-\alpha} (\bar{\theta} - \underline{\theta})$. For given allocation $y(\bar{\theta}, \eta)$, this corresponds to the standard result that $x(\bar{\theta})$ is distorted downwards relative to the first-best. Likewise, for given allocation $x(\bar{\theta})$, $y(\bar{\theta}, \underline{\eta})$ is set efficiently, while $y(\bar{\theta}, \bar{\eta})$ 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 truth-telling 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.

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.¹² 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 $\mathbb{X}_{ii}^{int}(\bar{\theta})$, we have that $y(\bar{\theta}, \underline{\eta}) > y(\bar{\theta}, \bar{\eta})$ and $x(\bar{\theta}, \bar{\eta}) > x(\bar{\theta}, \underline{\eta})$. If the x -allocation is now forced to take the common value $x(\bar{\theta})$, then, heuristically, $x(\bar{\theta}, \bar{\eta})$ is reduced while $x(\bar{\theta}, \underline{\eta})$ is increased. Since the marginal utility of consuming y still must take on the same value, the y -allocation has to respond more to η 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 .

¹²In the case of complements, the virtual marginal cost of $x(\bar{\theta}, \underline{\eta})$ is increased while the virtual marginal cost of $x(\bar{\theta}, \bar{\eta})$ is decreased for given level of y .

Thus, if the buyer has a choice, then starting production before all information is available is never strictly better than waiting until all information is available. In other words, our model features a nonnegative option value of waiting. The timing of production is irrelevant only in the case where the buyer's utility is additively separable in the utilities from consuming x and y .

7 Conclusion

This paper solves a tractable two-dimensional model of screening where 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. In this case, upward distortions may arise even in the case of complements.

As a simple by-product of our work we compare our solution to the literature by varying the timing of production in our model. It is always desirable to postpone all decisions until all information is available, if that is feasible. The comparison to the static model of two-dimensional screening is done in companion work. An interesting set of questions that we do not address in this work relates to repeating the interaction between the buyer and the seller.

Clearly, multidimensional problems are more complex than onedimensional ones. However, the timing of the information process imposes quite some structure on our problem. If on top of this, we are willing to impose assumptions on the buyer's utility function that we would be happy to impose already on problems under complete information, then the plethora of candidate solutions shrinks to only two possibilities. For mild complements, we obtain classical results, for mild substitutes, the results are close to but not quite equal to

the classical ones. However, perhaps more importantly, we can find the solution imposing simple truth-telling constraints on and off equilibrium path, a procedure that has no grounds whatsoever in the revelation principle per se, but rather emerges from implementing given allocations at lowest cost.

8 Appendix

Proof of Lemma 1. The first-best allocation $x(\theta, \eta), y(\theta, \eta)$ satisfies the system of first-order conditions (11) and (12) for $\theta \in \{\underline{\theta}, \bar{\theta}\}$ and $\eta \in \{\underline{\eta}, \bar{\eta}\}$. Define a new, artificial system of equations by

$$\begin{aligned} V_1(x(\theta, \eta), y(\theta, \eta)) &= \theta \\ V_2(x(\theta, \eta), y(\theta, \eta)) &= \eta \end{aligned} \tag{17}$$

for $\theta, \eta \in [\underline{\theta}, \bar{\theta}] \times [\underline{\eta}, \bar{\eta}]$. Note that the domain of the artificial system is obtained by a convexification of the original domain of definition; hence, by construction, the extreme points in the convexified domain are the cost types in the model. However, on the convexified domain, we can use calculus to determine differences between allocation choices. We prove the claims by direct evaluation of the differences. We focus on claim (i); the proof of claim (ii) uses the same methods and is therefore omitted.

Proof of claim (i): Since (17) is defined on a convex domain, we can write (by the fundamental theorem of calculus)

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

Totally differentiating the system (17), we have

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

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

By Cramer's rule

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

so

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

So, for any $\theta \in [\underline{\theta}, \bar{\theta}]$, $x(\theta, \bar{\eta}) \leq x(\theta, \underline{\eta})$ for $V_{12} \geq 0$ and $x(\theta, \bar{\eta}) > x(\theta, \underline{\eta})$ for $V_{12} < 0$. Thus, these inequalities hold in particular for $\theta \in \{\underline{\theta}, \bar{\theta}\}$.

Again using (17), the fundamental theorem, and Cramer's rule, we obtain

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

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

Combining these arguments, we have, for any θ ,

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

iff $V_{12} \geq 0$ and

$$0 \geq \int_{\underline{\eta}}^{\bar{\eta}} \frac{V_{11} + V_{12}}{V_{11}V_{22} - V_{12}^2}(\theta, \eta) d\eta,$$

which is satisfied if $V_{12} < -V_{11}$ for all (x, y) . Hence, these inequalities hold in particular for $\theta \in \{\underline{\theta}, \bar{\theta}\}$. Likewise, we have

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

for $\theta \in \{\underline{\theta}, \bar{\theta}\}$ if $V_{12} > -V_{11}$ for all (x, y) .

Proof of claim (ii): Similarly, one shows that for $V_{12} < 0$ we have

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

if $V_{12} \geq V_{11}$ for all (x, y) and

$$-(x(\theta, \underline{\eta}) - x(\theta, \bar{\eta})) \geq (y(\theta, \underline{\eta}) - y(\theta, \bar{\eta}))$$

for $V_{12} \leq V_{11}$ for all (x, y) . ■

Proof of Lemma 2. 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 $\lambda(\underline{\theta}) \geq \lambda(\bar{\theta})$.

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

From (6), we have

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

On the other hand

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

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

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

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

Hence, from (6), we have that

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

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

$$u(\bar{\theta}, \underline{\eta}) \geq u(\bar{\theta}, \bar{\eta}) + (\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}),$$

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

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

(9) written in terms of equilibrium utilities amounts to

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

which proves the claim. ■

Proof of Lemma 3. 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 (θ, η) .

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

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

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

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

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

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

From the on equilibrium path constraint 3, we know that

$$u(\bar{\theta}, \bar{\eta}) \geq u(\bar{\theta}, \underline{\eta}) - (\bar{\eta} - \underline{\eta})y(\bar{\theta}, \underline{\eta}).$$

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

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

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

$$x(\bar{\theta}, \bar{\eta}) \geq x(\bar{\theta}, \underline{\eta}).$$

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

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

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

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

Proof of lemma 4. We split the proof into two cases, depending on whether $x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})$ is nonnegative or nonpositive. For both cases, we first prove the part concerning the allocations of type $\bar{\theta}$. 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:

$$T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta}) \geq \bar{\theta}(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \underline{\eta}(y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) \quad (18)$$

$$T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta}) \leq \bar{\theta} (x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \bar{\eta} (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) \quad (19)$$

$$T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta}) \geq \underline{\theta} (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \underline{\eta} (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})) \quad (20)$$

$$T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta}) \leq \underline{\theta} (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \bar{\eta} (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})) \quad (21)$$

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

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

$$T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta}) \leq \underline{\theta} (x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \bar{\eta} (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})). \quad (22)$$

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

$$T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta}) \geq \underline{\theta} (x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \bar{\eta} (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})). \quad (23)$$

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

$$T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta}) \geq \underline{\theta} (x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \underline{\eta} (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) \quad (24)$$

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

$$T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta}) \leq \underline{\theta} (x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \underline{\eta} (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})). \quad (25)$$

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

$$T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta}) \geq \bar{\theta} (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \underline{\eta} (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})) \quad (26)$$

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

$$T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta}) \leq \bar{\theta} (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \underline{\eta} (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})). \quad (27)$$

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

$$T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta}) \leq \bar{\theta} (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \bar{\eta} (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})) \quad (28)$$

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

$$T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta}) \geq \bar{\theta} (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \bar{\eta} (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})). \quad (29)$$

Now we are ready to begin with the proof of the Lemma.

Suppose that $(x(\underline{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) \geq 0$. By Lemma 3 this implies that $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \underline{\eta}$. Adding the expected utility of the high type (which is zero by (9)) to the objective, we obtain the following problem:

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

subject to (18), (19), and
either (22) if $\hat{\eta}^(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \bar{\eta}$*
or (23) if $\hat{\eta}^(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \underline{\eta}$.*

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

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

subject to the constraints (18), (19), and (22). Note that (19) is automatically satisfied if (22) is. There exists a solution to the problem only if the constraint set is non-empty, that is, if the right-hand side of (22) is weakly larger than the right-hand side of (18). This is the case for $\{x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\bar{\theta})$. In this case (18) is binding. Using (9) and (18) to solve for the optimal payments, we have

$$\begin{pmatrix} T(\bar{\theta}, \underline{\eta}) \\ T(\bar{\theta}, \bar{\eta}) \end{pmatrix} = \begin{pmatrix} \bar{\theta}x(\bar{\theta}, \underline{\eta}) + \underline{\eta}y(\bar{\theta}, \underline{\eta}) + (1 - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}) \\ \bar{\theta}x(\bar{\theta}, \bar{\eta}) + \lambda(\bar{\theta})\underline{\eta}y(\bar{\theta}, \bar{\eta}) + (1 - \lambda(\bar{\theta}))\bar{\eta}y(\bar{\theta}, \bar{\eta}) \end{pmatrix}. \quad (30)$$

Substituting back into the objective we have obtain

$$\Delta_i = (\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}) + \mathbb{E}_{\eta|\underline{\theta}}[(\bar{\theta} - \underline{\theta})x(\bar{\theta}, \underline{\eta})].$$

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

$$\begin{aligned} & \min_{T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta})} (1 - \lambda(\bar{\theta})) \{T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta})\} \\ & + (\bar{\theta} - \underline{\theta})x(\bar{\theta}, \underline{\eta}) - (1 - \lambda(\underline{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \underline{\eta}) \\ & - (1 - \lambda(\bar{\theta}))(\bar{\theta}(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \underline{\eta}y(\bar{\theta}, \underline{\eta}) - \bar{\eta}y(\bar{\theta}, \bar{\eta})) \end{aligned}$$

subject to the constraints (18), (19), and (23). The right-hand side of (18) is weakly larger than the right-hand side of (23) for $\{x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iii}(\bar{\theta})$ and the reverse is

true for $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\bar{\theta})$. Clearly, the right-hand side of (19) is always larger than the right-hand side of (18). Therefore, for $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iii}(\bar{\theta})$, at the solution of the problem, constraint (18) holds as an equality. It follows that for $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \underline{\eta}$ and $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iii}(\bar{\theta})$, the transfers can be taken from (30) so that the objective takes value

$$\Delta_{iii} = (\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \underline{\eta}) - (1 - \lambda(\underline{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \underline{\eta}) + (1 - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \bar{\eta}).$$

For $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\bar{\theta})$ the off-path constraint (23) is binding. Substituting for the transfers implies that in this case

$$\hat{\Delta}_i = (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\bar{\theta}, \underline{\eta}) + \mathbb{E}_{\eta|\bar{\theta}} [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \eta)].$$

Since $\hat{\Delta}_i \geq \Delta_i$ if and only if $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\bar{\theta})$, for a given allocation $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\bar{\theta})$ the optimal payments are given by (30), $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \bar{\eta}$ and the information rent by Δ_i .

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

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

subject to

(20), (21), and either

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

(27) if $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \underline{\eta}) = \bar{\eta}$,

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

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

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

subject to the constraints (20), (21) and (26). The right-hand side of (20) is weakly smaller than the right-hand side of (26). Hence, the constraint set is nonempty if the right hand side of (26) is weakly smaller than the right-hand side of (21), which is exactly true for $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\underline{\theta})$. So, in this case, (21) is binding at the solution to the problem. Solving for the transfers from (21) and (6), we obtain

$$\begin{pmatrix} T(\underline{\theta}, \underline{\eta}) \\ T(\underline{\theta}, \bar{\eta}) \end{pmatrix} = \begin{pmatrix} \Delta + \underline{\theta}x(\underline{\theta}, \underline{\eta}) + (\lambda(\underline{\theta})\underline{\eta} + (1 - \lambda(\underline{\theta}))\bar{\eta})y(\underline{\theta}, \underline{\eta}) \\ \Delta + \underline{\theta}x(\underline{\theta}, \bar{\eta}) + \bar{\eta}y(\underline{\theta}, \bar{\eta}) - \lambda(\underline{\theta})(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \underline{\eta}) \end{pmatrix}. \quad (31)$$

Substituting these transfers back into the objective, we obtain

$$\Omega_i = -\lambda(\bar{\theta})(\bar{\theta} - \underline{\theta})x(\underline{\theta}, \underline{\eta}) - (\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \underline{\eta}) - (1 - \lambda(\bar{\theta}))(\bar{\theta} - \underline{\theta})x(\underline{\theta}, \bar{\eta}) + \Delta.$$

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

$$\Omega_i = \mathbb{E}_{\eta|\underline{\theta}}[(\bar{\theta} - \underline{\theta})x(\bar{\theta}, \eta)] - \mathbb{E}_{\eta|\bar{\theta}}[(\bar{\theta} - \underline{\theta})x(\underline{\theta}, \eta)] - (\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})(y(\underline{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})).$$

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

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

subject to the constraints (20), (21) and (27). The right-hand side of (27) is larger than the right-hand side of (21) for $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iii}(\underline{\theta})$. In this case, the feasible set is nonempty and at the solution (21) is binding; hence the transfers are given by (31) and the objective takes value

$$\begin{aligned} \Omega_{iii} &= -\lambda(\underline{\theta})(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \underline{\eta}) - (\bar{\theta} - \underline{\theta})x(\underline{\theta}, \bar{\eta}) + \lambda(\bar{\theta})(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \bar{\eta}) \\ & + \lambda(\underline{\theta})\underline{\theta}(x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + \Delta. \end{aligned}$$

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

$$\begin{aligned} \Omega_{iii} &= -\lambda(\underline{\theta})(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \underline{\eta}) + \lambda(\bar{\theta})(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \bar{\eta}) \\ & + \lambda(\underline{\theta})\underline{\theta}(x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) - (1 - \lambda(\underline{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \underline{\eta}) + (1 - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}). \end{aligned}$$

For $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\underline{\theta})$, the right-hand side of (27) is smaller than the right-hand side of (21). Moreover, the feasible set is always nonempty and thus at the solution constraint (27) is binding. Hence, we can substitute

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

into the objective and obtain

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

We have $\Omega_i \leq \hat{\Omega}_i$ for $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_i(\underline{\theta})$, so $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \underline{\eta}) = \underline{\eta}$ is cheaper to implement in that case.

Next consider the case where $(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) \leq 0$. By Lemma 3, this implies that $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \bar{\eta}) = \bar{\eta}$. Adding and subtracting the expected utility of type $\bar{\theta}$, we can write the objective as

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

subject to

(18), (19) and either

(24) if $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \underline{\eta}$, or

(25) if $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \bar{\eta}$.

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

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

subject to the constraints (18), (19), and (24). The right-hand side of (24) is always at least as large as the right-hand side of (18) (by the fact that $(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) \leq 0$). Hence, the constraint set is nonempty if the right-hand side of (19) is at least as large as the right-hand side of (24), which is precisely the case for $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{ii}(\bar{\theta})$. Since the objective is increasing in $T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta})$ and we are minimizing Δ , $T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta})$ is

set as small as possible, implying that (24) is binding. We can compute the transfers from (24) and (9). We obtain

$$\begin{pmatrix} T(\bar{\theta}, \underline{\eta}) \\ T(\bar{\theta}, \bar{\eta}) \end{pmatrix} = \begin{pmatrix} -\lambda(\bar{\theta})[(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}) - (\bar{\theta} - \underline{\theta})(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}))] + \bar{\theta}x(\bar{\theta}, \bar{\eta}) + \bar{\eta}y(\bar{\theta}, \bar{\eta}) + \underline{\theta}(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) + \underline{\eta}(y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) \\ -\lambda(\bar{\theta})[(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}) - (\bar{\theta} - \underline{\theta})(x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}))] + \bar{\theta}x(\bar{\theta}, \bar{\eta}) + \bar{\eta}y(\bar{\theta}, \bar{\eta}) \end{pmatrix}. \quad (32)$$

Substituting these transfers back into the objective, we obtain

$$\Delta_{ii} \equiv (\lambda(\underline{\theta}) - \lambda(\bar{\theta}))((\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta})) + \mathbb{E}_{\eta|\bar{\theta}}[(\bar{\theta} - \underline{\theta})x(\bar{\theta}, \eta)].$$

For $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iv}(\bar{\theta})$, no solution with $\hat{\eta}^*(\underline{\theta}, \bar{\theta}, \underline{\eta}) = \underline{\eta}$ exists.

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

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

subject to (18), (19), and (25). The right-hand side of (19) is weakly smaller than the right-hand side of (25) for $(x, y) \in \mathbb{X}_{iv}$. Since the objective is decreasing in $T(\bar{\theta}, \underline{\eta}) - T(\bar{\theta}, \bar{\eta})$ and we seek to minimize the objective function, at the optimum (19) must be binding. Thus, we can compute the optimal transfers from (19) and (9). We obtain

$$\begin{aligned} T(\bar{\theta}, \underline{\eta}) &= \bar{\theta}x(\bar{\theta}, \underline{\eta}) + \mathbb{E}_{\eta|\bar{\theta}}\eta y(\bar{\theta}, \underline{\eta}) \\ T(\bar{\theta}, \bar{\eta}) &= \bar{\theta}x(\bar{\theta}, \bar{\eta}) + \mathbb{E}_{\eta|\bar{\theta}}\eta y(\bar{\theta}, \bar{\eta}) - \bar{\eta}(y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) \end{aligned} \quad (33)$$

Substituting these transfers back into the objective, we obtain

$$\Delta_{iv} \equiv -\lambda(\bar{\theta})(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \underline{\eta}) + (\bar{\theta} - \underline{\theta})x(\bar{\theta}, \bar{\eta}) + \lambda(\underline{\theta})(\bar{\eta} - \underline{\eta})y(\bar{\theta}, \bar{\eta}).$$

For $\{x(\bar{\theta}, \eta), y(\bar{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iv}(\bar{\theta})$, the right-hand side of (25) is weakly smaller than the right-hand side of (19). Thus, (19) is slack. The right-hand side of (18) is smaller than the right-hand side of (25), which implies that constraint (25) must be binding and we obtain rent Δ_{ii} .

Consider next the second problem in case where $x(\underline{\theta}, \underline{\eta}) \leq x(\underline{\theta}, \bar{\eta})$. By lemma 3, this

implies that $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \underline{\eta}) = \underline{\eta}$. The objective then becomes

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

subject to

(21), (20), and either

(28) if $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \bar{\eta}$, or

(29) if $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \underline{\eta}$.

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 $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \bar{\eta}$. In that case the problem becomes

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

subject to (21), (20) and (28).

The right-hand side of (28) is always weakly smaller than the right-hand side of (21). Hence, (21) cannot become binding at the optimum. Moreover, the constraint set is non-empty exactly for $\{x(\underline{\theta}, \underline{\eta}), y(\underline{\theta}, \underline{\eta})\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{ii}(\underline{\theta})$. Since the objective is decreasing in $T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta})$, at the optimum, (28) is binding and we can compute the transfers from (28) and (6):

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

Since (28) is binding, we can substitute for

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

into the objective and obtain

$$\Omega_{ii} = -\mathbb{E}_{\eta|\underline{\theta}} [(\bar{\theta} - \underline{\theta})x(\underline{\theta}, \underline{\eta})] + \Delta - (\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})y(\underline{\theta}, \underline{\eta})$$

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

$$\Omega_{ii} = \mathbb{E}_{\eta|\bar{\theta}} [(\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \eta)] - \mathbb{E}_{\eta|\underline{\theta}} [(\bar{\theta} - \underline{\theta}) x(\underline{\theta}, \eta)] - (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) (y(\underline{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})).$$

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

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

subject to (21), (20), and (29).

The right-hand side of (20) is weakly larger than the right-hand side of (29) exactly for $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{iv}(\underline{\theta})$. Moreover, for such allocations, the constraint set is nonempty, and at the solution of the problem $T(\underline{\theta}, \underline{\eta}) - T(\underline{\theta}, \bar{\eta})$ reaches its lower bound, so (20) is binding. The transfers can then be computed from (6) and (20) :

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

Since (20) is binding, we can substitute

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

into the objective and obtain

$$\Omega_{iv} = (1 - \lambda(\underline{\theta})) (\bar{\eta} - \underline{\eta}) y(\underline{\theta}, \bar{\eta}) - (1 - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\underline{\theta}, \underline{\eta}) - (\bar{\theta} - \underline{\theta}) x(\underline{\theta}, \underline{\eta}) + \Delta$$

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

$$\begin{aligned} \Omega_{iv} &= -\lambda(\underline{\theta}) (\bar{\eta} - \underline{\eta}) (y(\underline{\theta}, \bar{\eta}) - y(\bar{\theta}, \bar{\eta})) - (\bar{\theta} - \underline{\theta}) (x(\underline{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) \\ &\quad - \lambda(\bar{\theta}) (\bar{\eta} - \underline{\eta}) (y(\bar{\theta}, \underline{\eta}) - y(\underline{\theta}, \underline{\eta})) - (\bar{\eta} - \underline{\eta}) (y(\underline{\theta}, \underline{\eta}) - y(\underline{\theta}, \bar{\eta})). \end{aligned}$$

For $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{ii}(\underline{\theta})$, the right-hand side of (29) is weakly larger than the right-hand side of (20). Moreover, since the right-hand side of (29) is smaller than the right-hand side of (21), the constraint set is nonempty. At the solution, (29) is binding, so we can substitute for

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

into the objective and obtain

$$\hat{\Omega}_{ii} = (1 - \lambda(\underline{\theta})) (\bar{\theta} - \underline{\theta}) (x(\underline{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) - (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) y(\underline{\theta}, \underline{\eta}) - (\bar{\theta} - \underline{\theta}) x(\underline{\theta}, \underline{\eta}) + \Delta$$

Since $\Omega_{ii} \leq \hat{\Omega}_{ii}$ for any $\{x(\underline{\theta}, \eta), y(\underline{\theta}, \eta)\}_{\eta \in \{\underline{\eta}, \bar{\eta}\}} \in \mathbb{X}_{ii}(\underline{\theta})$, implementing $\hat{\eta}^*(\bar{\theta}, \underline{\theta}, \bar{\eta}) = \underline{\eta}$, 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 P_{ii} ; part II does likewise for program P_i ; finally, part III compares the value of the objectives. The proof of the fact that $W_{iv} \leq W_{ii}$ and $W_{iii} \leq W_i$ is not given here but is available upon request from the authors; it uses essentially the same arguments.

Part I) Consider program P_{ii} . Up to a constant, the Lagrangian of program P_{ii} can be written as

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

The conditions of optimality are

$$((1 - \alpha) \lambda(\bar{\theta}) (V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \bar{\theta}) - \alpha \lambda(\bar{\theta}) (\bar{\theta} - \underline{\theta}) - \phi + \mu (\bar{\theta} - \underline{\theta})) = 0 \quad (34)$$

$$((1 - \alpha) (1 - \lambda(\bar{\theta})) (V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\theta}) - \alpha (1 - \lambda(\bar{\theta})) (\bar{\theta} - \underline{\theta}) + \phi - \mu (\bar{\theta} - \underline{\theta})) = 0 \quad (35)$$

$$((1 - \alpha) \lambda(\bar{\theta}) (V_2(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \underline{\eta}) + \mu (\bar{\eta} - \underline{\eta})) = 0 \quad (36)$$

$$((1 - \alpha) (1 - \lambda(\bar{\theta})) (V_2(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\eta}) - \alpha (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) - \mu (\bar{\eta} - \underline{\eta})) = 0. \quad (37)$$

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

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

$$(V_1(x(\bar{\theta}), y(\bar{\theta})) - \bar{\theta}) - \frac{\alpha}{1 - \alpha} (\bar{\theta} - \underline{\theta}) = 0 \quad (38)$$

and

$$\left(\begin{array}{c} V_2(x(\bar{\theta}), y(\bar{\theta})) \\ -\lambda(\bar{\theta}) \underline{\eta} - (1 - \lambda(\bar{\theta})) \bar{\eta} - \frac{\alpha}{1 - \alpha} (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) \end{array} \right) = 0. \quad (39)$$

Using (36), the Kuhn-Tucker-first-order-optimality-condition for $y(\bar{\theta}, \underline{\eta})$ and substituting (39) we have for $\mu \neq 0$

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

which simplifies to

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

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 *ii*.

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 $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) \geq 0$ or constraint $(\bar{\eta} - \underline{\eta}) (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) - (\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) \geq 0$ binds at the optimum of program *P_{ii}*.

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

$$\begin{aligned} V_2(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \underline{\eta} &= 0 \\ V_2(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\eta} - \frac{\alpha (\lambda(\underline{\theta}) - \lambda(\bar{\theta}))}{(1 - \alpha) (1 - \lambda(\bar{\theta}))} (\bar{\eta} - \underline{\eta}) &= 0. \end{aligned}$$

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

$$\begin{aligned} V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \bar{\theta} - \frac{\alpha}{1 - \alpha} (\bar{\theta} - \underline{\theta}) &= 0 \\ V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\theta} - \frac{\alpha}{1 - \alpha} (\bar{\theta} - \underline{\theta}) &= 0 \end{aligned}$$

which imply

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) = V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})). \quad (40)$$

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

$$V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) \quad (41)$$

Together conditions (40) and (41) imply

$$V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})). \quad (42)$$

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

$$V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) \leq V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \underline{\eta})) \quad (43)$$

which contradicts (42).

It follows that at least one constraint must be binding at the optimum of program P_{ii} .

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

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

Suppose, contrary to our claim,

$$(\bar{\eta} - \underline{\eta})(y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) - (\bar{\theta} - \underline{\theta})(x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta})) = 0$$

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

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

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

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

implying that

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta}))$$

if and only if

$$\frac{\alpha\lambda(\bar{\theta})(\bar{\theta} - \underline{\theta}) - \mu(\bar{\theta} - \underline{\theta})}{(1 - \alpha)\lambda(\bar{\theta})} < \frac{\alpha(1 - \lambda(\bar{\theta}))(\bar{\theta} - \underline{\theta}) + \mu(\bar{\theta} - \underline{\theta})}{(1 - \alpha)(1 - \lambda(\bar{\theta}))}$$

$$\iff$$

$$0 < \mu.$$

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

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

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) > V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})).$$

By $V_{12} \geq 0$

$$V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \underline{\eta})) \geq V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta}))$$

Together these imply that

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) > V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})),$$

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 $x(\bar{\theta}, \bar{\eta}) = x(\bar{\theta}, \underline{\eta})$.

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

If $V_{12} < 0$, then the solution to program P_{ii} satisfies $x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) > 0$.

Suppose not. We know that $\phi, \mu > 0$ is not possible. So, if $x(\bar{\theta}, \bar{\eta}) = x(\bar{\theta}, \underline{\eta})$, this would have to imply that $\mu = 0$. So, we would have $x(\bar{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta}) = x(\bar{\theta})$, $y(\bar{\theta}, \bar{\eta}) < y(\bar{\theta}, \underline{\eta})$ and $\mu = 0$. Adding up of conditions (34) and (35), the first-order conditions for $x(\bar{\theta}, \underline{\eta})$ and $x(\bar{\theta}, \bar{\eta})$, gives

$$\left(\begin{array}{c} \lambda(\bar{\theta}) (V_1(x(\bar{\theta}), y(\bar{\theta}, \underline{\eta})) - \bar{\theta}) + (1 - \lambda(\bar{\theta})) (V_1(x(\bar{\theta}), y(\bar{\theta}, \bar{\eta})) - \bar{\theta}) \\ - \frac{\alpha}{1-\alpha} (\bar{\theta} - \underline{\theta}) \end{array} \right) = 0. \quad (44)$$

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

$$V_1(x(\bar{\theta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}), y(\bar{\theta}, \bar{\eta})).$$

Together with (44), this implies that

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

Plugging the first of these inequalities into (34), we obtain

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

Plugging the latter of the inequalities into (35), we obtain

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

which simplifies to

$$\mu (\bar{\theta} - \underline{\theta}) > \phi.$$

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

Part II) Consider program P_i . Up to a constant, the Lagrangian of program P_i can be written as

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

The conditions of optimality are given by

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

$$(1 - \alpha)\lambda(\bar{\theta})(V_2(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \underline{\eta}) + \nu(\bar{\eta} - \underline{\eta}) = 0 \quad (46)$$

$$\left(\begin{array}{l} (1 - \alpha)(1 - \lambda(\bar{\theta}))(V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\theta}) \\ - \alpha(1 - \lambda(\underline{\theta}))(\bar{\theta} - \underline{\theta}) - \xi + \nu(\bar{\theta} - \underline{\theta}) \end{array} \right) = 0 \quad (47)$$

$$\left(\begin{array}{l} (1 - \alpha)(1 - \lambda(\bar{\theta}))(V_2(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\eta}) \\ - \alpha(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta}) - \nu(\bar{\eta} - \underline{\eta}) \end{array} \right) = 0 \quad (48)$$

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

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

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

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

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

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

and

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

Using the first-order condition with respect to $y(\bar{\theta}, \underline{\eta})$, (46), gives

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

Substituting for $V_2(x(\bar{\theta}), y(\bar{\theta}))$ gives

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

which simplifies to

$$-(1 - \lambda(\bar{\theta})) (1 - \alpha)\lambda(\bar{\theta}) - \alpha(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))\lambda(\bar{\theta}) = \nu,$$

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, we show that if $V_{12} < 0$, then $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) = 0$ at the solution to program P_i .

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

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

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

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

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

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) > V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})). \quad (49)$$

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

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \bar{\eta})). \quad (50)$$

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

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \bar{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})). \quad (51)$$

Taken together (50) and (51) imply that

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})).$$

which contradicts (49) derived previously from the first-order conditions.

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

Suppose that contrary to our claim, that the solution of program P_i satisfies

$(\bar{\eta} - \underline{\eta}) [y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})] = (\bar{\theta} - \underline{\theta}) [x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})] > 0$, $\xi = 0$ and $\nu > 0$. Adding up (45) and (47) gives

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

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

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta}))$$

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

$$V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})).$$

Taken together, we have

$$V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) < V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})).$$

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

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

Substituting into (45), using $\xi = 0$, and simplifying, we have

$$\alpha ((\lambda(\bar{\theta}) - \lambda(\underline{\theta})) (\bar{\theta} - \underline{\theta})) > \nu (\bar{\theta} - \underline{\theta}),$$

which would imply that $\nu < 0$, a contradiction.

It follows that for $V_{12} < 0$, the optimum of program P_i features $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) = 0$.

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

If $V_{12} \leq -V_{11} \frac{(\bar{\eta} - \underline{\eta})}{(\bar{\theta} - \underline{\theta})}$, then the optimum of program P_i features $(\bar{\eta} - \underline{\eta}) [y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})] > (\bar{\theta} - \underline{\theta}) [x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})]$.

We know that both constraints cannot bind simultaneously.

Hence, if $(\bar{\eta} - \underline{\eta}) [y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})] = (\bar{\theta} - \underline{\theta}) [x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})]$, then necessarily $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) > 0$. Suppose this is the case, so $\xi = 0$ and $\nu > 0$. Define $Y(\bar{\theta}, \bar{\eta}) = y(\bar{\theta}, \underline{\eta}) -$

$\frac{(\bar{\theta}-\underline{\theta})}{(\bar{\eta}-\underline{\eta})} [x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})]$. For $\nu \neq 0$ the first-order conditions with respect to x are given by

$$\begin{pmatrix} (1-\alpha)\lambda(\bar{\theta})(V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \bar{\theta}) - \alpha\lambda(\underline{\theta})(\bar{\theta} - \underline{\theta}) \\ -\frac{(\bar{\theta}-\underline{\theta})}{(\bar{\eta}-\underline{\eta})} ((1-\alpha)(1-\lambda(\bar{\theta}))(V_2(x(\bar{\theta}, \bar{\eta}), Y(\bar{\theta}, \bar{\eta})) - \bar{\eta}) - \alpha(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})) \end{pmatrix} = 0$$

and

$$\begin{pmatrix} (1-\alpha)(1-\lambda(\bar{\theta}))(V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\theta}) - \alpha(1-\lambda(\underline{\theta}))(\bar{\theta} - \underline{\theta}) \\ +\frac{(\bar{\theta}-\underline{\theta})}{(\bar{\eta}-\underline{\eta})} ((1-\alpha)(1-\lambda(\bar{\theta}))(V_2(x(\bar{\theta}, \bar{\eta}), Y(\bar{\theta}, \bar{\eta})) - \bar{\eta}) - \alpha(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))(\bar{\eta} - \underline{\eta})) \end{pmatrix} = 0.$$

These conditions imply

$$\begin{pmatrix} (1-\alpha)\lambda(\bar{\theta})(V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \bar{\theta}) - \alpha\lambda(\underline{\theta})(\bar{\theta} - \underline{\theta}) \\ + (1-\alpha)(1-\lambda(\bar{\theta}))(V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta})) - \bar{\theta}) - \alpha(1-\lambda(\underline{\theta}))(\bar{\theta} - \underline{\theta}) \end{pmatrix} = 0. \quad (52)$$

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

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

and

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

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

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

and $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) > 0$. $V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) \leq V_1(x(\bar{\theta}, \bar{\eta}), y(\bar{\theta}, \bar{\eta}))$ implies by (52) that

$$(V_1(x(\bar{\theta}, \underline{\eta}), y(\bar{\theta}, \underline{\eta})) - \bar{\theta}) \leq \frac{\alpha}{(1-\alpha)} (\bar{\theta} - \underline{\theta}) \quad (53)$$

By (53) into (45)

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

which is equivalent to

$$-\alpha(\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) \geq \nu (\bar{\eta} - \underline{\eta})$$

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

Part III) Comparison between the programs.

As a preliminary argument, note that if $x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) = 0$, then the objectives of programs P_i and P_{ii} become identical. To see this, note that the objectives are identical up to the costs of implementation, Δ . Moreover, it is easy to verify from Lemma 3 that $\Delta_{ii} - \Delta_i = 0$ for $x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) = 0$.

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

Likewise, for $V_{12} \geq 0$ the optimum of program P_{ii} satisfies $x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) = 0$. Hence, the solution is feasible under program P_i . If the solution of program P_i is on the line $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) = 0$, then the objectives are identical and hence the solutions of programs P_i and P_{ii} are identical. If the solution of program P_i is off the line $x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) = 0$, then, by revealed preference, the solution of program P_i yields a strictly higher expected payoff than the solution of program P_{ii} . Taken together, this implies weak payoff dominance of program P_i . ■

Proof of Lemma 5. The result is essentially a corollary to Lemma 1. By the same arguments as given there, we have

$$(\bar{\eta} - \underline{\eta}) (y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) > (\bar{\theta} - \underline{\theta}) (x(\bar{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta})) > 0$$

iff $V_{12} > 0$ and

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

which is satisfied if $V_{12} < -\frac{(\bar{\eta} - \eta)}{(\bar{\theta} - \theta)} V_{11}$ for all (x, y) .

The proof of the remaining statements uses identical arguments and is therefore omitted.

■

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}_j for $j = i$ or $j = ii$, respectively) satisfies

$$\begin{aligned} V_1(x, y)(\underline{\theta}, \eta) &= \underline{\theta} \\ V_2(x, y)(\underline{\theta}, \eta) &= \eta \end{aligned} \tag{54}$$

for $\eta \in \{\underline{\eta}, \bar{\eta}\}$ and

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

$$V_2(x, y)(\bar{\theta}, \underline{\eta}) = \underline{\eta}$$

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

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

where $j = i, ii$ and by convention $\lambda_i = \lambda(\underline{\theta})$ and $\lambda_{ii} = \lambda(\bar{\theta})$. Define the following artificial systems of equations for $\theta, \eta \in [\underline{\theta}, \bar{\theta}] \times [\underline{\eta}, \bar{\eta}]$:

$$V_1(x, y)(\theta, \eta) = \theta + \frac{\alpha}{(1 - \alpha)} \frac{\lambda_j}{\lambda(\bar{\theta})} (\theta - \underline{\theta}) \tag{57}$$

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

and

$$\begin{aligned}
V_1(x, y)(\theta, \eta) &= \theta + \frac{\alpha}{(1-\alpha)}(\theta - \underline{\theta}) \\
V_2(x, y)(\theta, \eta) &= \eta + \frac{\alpha}{(1-\alpha)} \frac{(\lambda(\underline{\theta}) - \lambda(\bar{\theta}))}{(1-\lambda(\bar{\theta}))}(\eta - \underline{\eta}).
\end{aligned} \tag{58}$$

Note that these systems are defined on convex domains. Moreover, the solution to (57) for $\theta = \underline{\theta}$ corresponds to the solution of (54), and for $\theta = \bar{\theta}$ and $\eta = \underline{\eta}$, the solution to (57) corresponds to the solution of (55). Likewise, for $\theta = \underline{\theta}$ and $\eta = \underline{\eta}$, the solution to (58) corresponds to the solution to (54); for $\theta = \bar{\theta}$ and $\eta = \underline{\eta}$, the solution to (58) corresponds to the solution of (55) for $\lambda_j = \lambda_{ii} = \lambda(\bar{\theta})$; and for $\theta = \bar{\theta}$ and $\eta = \bar{\eta}$, the solution to (58) corresponds to the solution to (56) for $\lambda_j = \lambda_{ii} = \lambda(\bar{\theta})$.

So, systems (57) and (58) are defined on convex domains. Moreover, the solutions to the systems at extreme points of the domain correspond to the economically meaningful solutions of (54), (55), and (56), respectively. Hence, we can conveniently apply calculus to the artificial system (57) and (58) to determine differences between allocation choices.

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

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

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

Sufficient conditions for the neglected constraint to hold are

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

and

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

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

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

and

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

for $\eta \in \{\underline{\eta}, \bar{\eta}\}$,

$$(1 - \alpha) [V_1(x(\bar{\theta})) - \bar{\theta}] = \alpha(\bar{\theta} - \underline{\theta}),$$

and finally

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

and

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

It is easy to see (by concavity of V), that $x(\underline{\theta}, \underline{\eta}) = x(\underline{\theta}, \bar{\eta}) > x(\bar{\theta})$, so

$(\mathbb{E}_{\eta|\bar{\theta}}[x(\underline{\theta}, \eta)] - \mathbb{E}_{\eta|\underline{\theta}}[x(\bar{\theta}, \eta)]) \geq 0$. is satisfied. By the same argument, we also have $y(\underline{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta}) \geq 0$.

Part II)

1. The case of complements.

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

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

Sufficient conditions for the neglected constraint to hold are

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

and

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

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

We can write

$$y(\underline{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta}) = y(\underline{\theta}, \underline{\eta}) - y(\bar{\theta}, \underline{\eta}) + y(\bar{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta}).$$

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

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

$$\min_{\eta \in \{\underline{\eta}, \bar{\eta}\}} x(\underline{\theta}, \eta) \geq \max_{\eta \in \{\underline{\eta}, \bar{\eta}\}} x(\bar{\theta}, \eta),$$

which in turn holds if

$$x(\underline{\theta}, \underline{\eta}) \geq x(\underline{\theta}, \bar{\eta}) \geq x(\bar{\theta}, \underline{\eta}) \geq x(\bar{\theta}, \bar{\eta}).$$

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

$$x(\underline{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) = x(\underline{\theta}, \bar{\eta}) - x(\underline{\theta}, \underline{\eta}) + x(\underline{\theta}, \underline{\eta}) - x(\bar{\theta}, \underline{\eta}).$$

The differences on the right-hand side of this equation can be conveniently computed from (57), since we argued above that the types on the right-hand side correspond to extreme points in the domain of definition of (57). Differentiating the system of equations (57), we obtain

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

where the first equality follows from setting $\theta = \underline{\theta}$ in (57) and applying Cramer's rule and the second equality from the mean value theorem, for some $\hat{\eta} \in [\underline{\eta}, \bar{\eta}]$. Likewise, by setting $\eta = \underline{\eta}$ in (57) and $j = i$ so that $\lambda_j = \lambda(\underline{\theta})$, and applying Cramer's rule, we have

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

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

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

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

In turn, this condition is satisfied if

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

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

2. The case of substitutes:

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

$$\left(\begin{aligned} & (\lambda(\underline{\theta}) - \lambda(\bar{\theta})) (\bar{\eta} - \underline{\eta}) (y(\underline{\theta}, \underline{\eta}) - y(\bar{\theta}, \bar{\eta})) \\ & + (\bar{\theta} - \underline{\theta}) x(\underline{\theta}, \underline{\eta}) - (\lambda(\bar{\theta}) (\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \underline{\eta}) + (1 - \lambda(\bar{\theta})) (\bar{\theta} - \underline{\theta}) x(\bar{\theta}, \bar{\eta})) \end{aligned} \right) \geq 0.$$

Equivalently, this can be written as

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

Recall that for $(x, y) \in X_{ii}$, we have

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

so the third term on the left-hand side is nonnegative. For the case where $\lambda(\underline{\theta}) = \lambda(\bar{\theta})$, the first term is zero and we only need to show that

$$x(\underline{\theta}, \underline{\eta}) - x(\bar{\theta}, \bar{\eta}) \geq 0.$$

We can write

$$x(\bar{\theta}, \bar{\eta}) - x(\underline{\theta}, \underline{\eta}) = x(\bar{\theta}, \bar{\eta}) - x(\bar{\theta}, \underline{\eta}) + x(\bar{\theta}, \underline{\eta}) - x(\underline{\theta}, \underline{\eta}).$$

The types on the right-hand side correspond to extreme points of the domain of definition of (58). Therefore, we obtain - by the same arguments as used for the complements case -

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

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

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

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

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

Finally, we need to show that the optimal allocations that solve the reduced problems P_i and P_{ii} , respectively, are elements of \mathbb{X}_i^{int} or \mathbb{X}_{ii}^{int} , respectively. Recall from Lemma 4 that the first-best allocation is an element of \mathbb{X}_i^{int} or \mathbb{X}_{ii}^{int} , respectively, precisely under the conditions that make either program P_i or P_{ii} generate a higher value to the principal. Now consider, for $j = i, ii, iii, iv$, the problems

$$\max_{(x,y) \in \cup_j \mathbb{X}_j} P_j$$

The solution to each of these problems converges uniformly to the first-best allocation as α goes to zero. It follows that the solution of program P_i is in \mathbb{X}_i^{int} for α close enough to zero if $0 < V_{12} < -\frac{\bar{\eta} - \underline{\eta}}{\bar{\theta} - \underline{\theta}} V_{11}$ and that the solution of program P_{ii} is in \mathbb{X}_{ii}^{int} for α close enough to zero if $\frac{\bar{\eta} - \underline{\eta}}{\bar{\theta} - \underline{\theta}} V_{11} < V_{12} < 0$. ■

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 α goes to zero, the allocations that achieve the maxima W_j are in \mathbb{X}_j^{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 β sufficiently large to generate interior solutions, the first-best allocation is given by

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

The neglected constraint for $(x, y) \in \mathbb{X}_{iii}$ takes the form

$$\begin{aligned} 0 & \geq (\bar{\theta} - \underline{\theta}) (x(\bar{\theta}, \underline{\eta}) - x(\underline{\theta}, \bar{\eta})) + (\bar{\eta} - \underline{\eta}) (y(\bar{\theta}, \bar{\eta}) - y(\bar{\theta}, \underline{\eta})) \\ & \quad + \lambda(\underline{\theta}) (\bar{\eta} - \underline{\eta}) (y(\bar{\theta}, \underline{\eta}) - y(\underline{\theta}, \underline{\eta})) + \lambda(\bar{\theta}) (\bar{\eta} - \underline{\eta}) (y(\underline{\theta}, \bar{\eta}) - y(\bar{\theta}, \bar{\eta})). \end{aligned}$$

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

$$0 \geq (\bar{\theta} - \underline{\theta}) \left(\frac{1}{1 - \delta^2} (- (\bar{\theta} - \underline{\theta}) + \delta (\bar{\eta} - \underline{\eta})) \right) + (\bar{\eta} - \underline{\eta}) \left(\frac{1}{1 - \delta^2} (- (\bar{\eta} - \underline{\eta})) \right) \\ + \lambda (\underline{\theta}) (\bar{\eta} - \underline{\eta}) \left(\frac{1}{1 - \delta^2} (- (\bar{\theta} - \underline{\theta}) \delta) \right) + \lambda (\bar{\theta}) (\bar{\eta} - \underline{\eta}) \left(\frac{1}{1 - \delta^2} ((\bar{\theta} - \underline{\theta}) \delta) \right),$$

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

9 References

- Armstrong, Mark and Rochet, Jean-Charles (1999): "Multi-dimensional screening: A user's guide", *European Economic Review*, Vol. 43, No. 4, 959-979.
- Baron, David P and Besanko, David (1984): "Regulation and information in a continuing relationship", *Information Economics and Policy*, Vol. 1, No. 3, 267-302.
- Battaglini, Marco (2005): "Long-term contracting with Markovian consumers", *American Economic Review*, Vol. 95, No. 3, 637-658.
- Battaglini, Marco and Rohit Lamba (2012): "Optimal Dynamic Contracting", unpublished.
- Bhaskar, V. (2013): "The Ratchet Effect Re-examined: a learning perspective", unpublished.
- Boleslavsky, Raphael and Said, Maher (2012): "Progressive Screening: Long-Term Contracting with a Privately Known Stochastic Process", *The Review of Economic Studies*, Vol. 80, No. 1, 1-34.
- Courty, Pascal and Hao, Li (2000): "Sequential screening", *The Review of Economic Studies*, Vol. 67, No. 4, 697-717.
- Dana, Jr., James D. (1993): "The Organization and Scope of Agents: Regulating Multiproduct Industries", *Journal of Economic Theory*, Vol. 59, No. 2, 288 - 310.

- Eső, Peter and Szentes, Balázs (2007a): "Optimal Information Disclosure in Auctions and the Handicap Auction", *The Review of Economic Studies*, Vol. 74, No. 3, 705-731.
- Eső, Péter and Szentes, Balázs (2007b): "The Price of Advice", *The RAND Journal of Economics*, Vol. 38, No. 4, 863-880.
- Eső, Péter and Szentes, Balázs (2013): "Dynamic Contracting with Adverse Selection: An Irrelevance Result", unpublished.
- Garrett, Daniel F. and Alessandro Pavan (2013): "Dynamic Managerial Compensation: On the Optimality of Seniority-based Schemes", 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, 323-358.
- Pavan, Alessandro, Ilya Segal and Juuso Toikka (2013): "Dynamic Mechanism Design: A Myersonian Approach", unpublished.
- Rochet, J.-C. and P. Choné (1998): "Ironing, sweeping, and multidimensional screening", *Econometrica*, Vol. 66, No. 4, 783-826.
- 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.
- Severinov, Sergej (2008): "The Value of Information and Optimal Organization", *The RAND Journal of Economics*, Vol. 39, No. 1, 238 – 265.
- Tirole, Jean (1988): "The Theory of Industrial Organization", MIT Press.