

Bilateral Trade with Justification*

Matthias Lang**

February 2019

Abstract

With cheap talk, there are clear limits to incentive-compatible communication as demonstrated by Myerson and Satterthwaite (1983). In natural language, we often justify our statements to make our claims more convincing. By justification, I refer to a message that transmits new information to the recipient and that is partially verifiable by the recipient. The idea is that the sender provides a number of reasons and the recipient derives the message's meaning from the reasons provided. The recipient knows for one reason whether it is appropriate. Hence, she becomes aware of some deviations. I apply justification to the canonical model of bilateral trade. I show that justification allows for efficient trade in contrast to previous impossibility results.

JEL classifications: D81, D82, D86

Keywords: Mechanism Design, Bilateral Trade, Justification, Communication, Information

*This draft is made available to stimulate discussion and to gather feedback. It has not benefited from a formal peer review. Please do not (re)post this paper. Please contact the author for the latest draft before citing. All comments and suggestions are greatly appreciated. I thank Joachim Henkel and the participants of the Micro Seminar at University of Konstanz, MSM Research Seminar at University of Duisburg-Essen, and BEWIP Seminar at TU Munich for helpful comments. Financial support by the German Research Foundation (DFG) through the Collaborative Research Centre (CRC TRR 190) is gratefully acknowledged.

**University of Munich (LMU), Geschwister-Scholl-Platz 1, 80539 Munich (Germany). CESifo, Munich (Germany) Email: matthias.lang@econ.lmu.de

1 Introduction

Economic analysis often simplifies or neglects communication and information although these are important components of almost any economic activity. Commonly, there is a dichotomy with cheap talk on the one hand side and disclosure of hard information on the other hand side. I try to bridge these two sides and provide some middle ground. We observe the same in natural language with people justifying their statements to make their claims more convincing.

The methodological contribution of this paper is to model justification. By justification, I refer to messages that transmit new information to the recipient and that are partially verifiable by the recipient. Instead of making a direct statement, e.g., that costs are high, the sender provides a number of reasons, arguments and points to support her statement. Indeed, the recipient derives the sender's statement from the reasons provided. If the sender wants to change her statement she must change several reasons. The recipient knows for one reason whether it is appropriate for the sender's statement. If the sender deviates from equilibrium play, the recipient becomes aware of such a deviation with some probability. Therefore, justification makes the sender's claims more convincing in the sense that the recipient notices some deviations.

I apply justification to the canonical setting of bilateral trade by Myerson and Satterthwaite (1983). Justification applies to bilateral trade in the following way: The seller learns all necessary inputs. Each input costs 1. The buyer only knows the distribution of inputs. By reading product reviews or observing the product, the buyer learns for one input whether that input is necessary. This shared signal, however, is uninformative about the seller's costs. The seller reports the necessary inputs, but her message is not necessarily truthful. The buyer replies with an unverifiable message about the shared signal. As the messages are the only third-party enforceable information, the mechanism just depends on these messages. To show the strength and power of justification, I limit the buyer's information as far as possible. Nevertheless, the buyer can to some extent verify the seller's message although the seller's costs and the buyer's information are stochastically independent and, hence, uncorrelated.

I do not assume an exogenous verification technology, type-dependent message spaces, or that messages are verifiable by a third party. All messages are unverifiable, but contractible. Hence, a third party cannot tell whether a message is truthful. The mechanism uses the fact that the seller and the buyer share an observation of the environment. This shared observation is very limited. First, it has mass zero with respect to the seller's information. Second, the buyer cannot infer anything about the seller's costs from this information. Giving more information to the buyer makes it eas-

ier for him to detect distorted reports by the seller. The seller recalls all observations to justify her costs. By providing this additional information, she makes herself vulnerable to scrutiny. If she were to distort the reported costs, she has to lie about several inputs. No matter how she distorts the reported costs, there is a strictly positive probability that the buyer becomes aware of any distortion that increases the price. The reason is that the seller does not know which inputs the buyer has learned.

My model of justification is much more general and not limited to bilateral trade. Justification applies more generally to mechanism design, moral hazard with subjective evaluations, hold-up and many more settings whenever contracting parties interact and share some information. The mechanism allows the better-informed party to provide justification as defined above.

Related Literature

There are different approaches in the literature concerning partial verifiability. First, agents disclose hard information that does not fully reveal the state of the world. For example, in Shin (1994), two senders present hard information to an arbitrator who wants to match the state of the world with her action. The hard information takes the form that the state lies above or below certain privately-known thresholds. The senders only report evidence that supports their position because they have conflicting preferences. Dziuda (2011) considers a more general model but with only one sender. Again full revelation is impossible. Alternatively, Fishman and Hagerty (1990) assume that the sender learns a finite number of signal about the state of the world. The sender, however, can only disclose one of these signals to the recipient. They determine the optimal amount of discretion for the sender's choice.

Second, agents' message spaces depend on their types going back to Milgrom (1981) and Green and Laffont (1986). Notice that this second approach is more general than the first one. Green and Laffont (1986) focus on direct mechanisms and provide conditions on how message spaces depend on types for the revelation principle to be valid. Deneckere and Severinov (2008) show that focusing on direct mechanisms is restrictive and instead sequential and password mechanisms should be considered. Okuno-Fujiwara et al. (1990) provide conditions on a game's payoffs for complete revelation of types. Lipman and Seppi (1995) provide conditions on how message spaces depend on types for complete revelation of types. Glazer and Rubinstein (2006) and Sher (2014) characterize optimal reporting strategies and show that randomization and commitment by the recipient are unnecessary. Returning to the cheap talk setting of Crawford and Sobel (1982), Seidmann and Winter (1997) and Mathis (2008) provide conditions

on how message spaces depend on types for existence of fully separating equilibria. Correspondingly, Giovannoni and Seidmann (2007) provide conditions on preferences for existence of separating equilibria. Lang (2019) considers moral hazard with partial verifiability. Finally, Bull and Watson (2004), Ben-Porath and Lipman (2012), and Kartik and Tercieux (2012) characterize implementable social choice functions in complete information environments.¹

Third, the recipient could choose to verify some aspects of the sender's message. Glazer and Rubinstein (2004) consider a sender and a receiver. The receiver takes a binary action. The receiver's optimal action depends on the state of the world that is two dimensional. The sender always prefers one action and sends a message about the state of the world. Depending on this message, the receiver can verify the one dimension of the message. Glazer and Rubinstein (2004) show that finding an optimal mechanism is equivalent to solving a linear programming problem. In addition, there is often a deterministic optimal mechanism.

Fourth, restrictions of the contracting language make it ex-ante impossible to describe some events. In Al-Najjar et al. (2006) and Anderlini and Felli (1994), these restrictions make incomplete contracts optimal. In my paper, there are no restrictions on the mechanism designer. Yet, the state of the world is private information and needs to be communicated ex-post. This communication can be supplemented by justification that makes messages partly verifiable. Although I use a similar representation of the states of the world as an infinite binary sequence, their approach is conceptually and technically different from what I do here.

Now turn to Myerson and Satterthwaite (1983) who study the bilateral trade setting to which I apply justification. They show that without justification it is impossible to find a mechanism that is individually rational, ex-post efficient and requires no outside subsidies. Hence, it is impossible to implement efficient trade. Recently, a couple of papers re-examine this impossibility for different preferences. For example, Wolitzky (2016) considers ambiguity-averse agents and provides conditions so that efficient trade is possible. Garratt and Pycia (2016) consider concave utilities and provides conditions that make efficient trade possible. Benkert (2017) considers loss-averse agents and confirms previous impossibility results.

This paper is organized as follows. Section 2 introduces the canonical bilateral trade setting. In Section 3, I convey the basic intuition of justification in a simplified setting. Then, Section 4 provides the details of a justification and proves the power of justification. Finally, Section 5 considers strongly budget-balanced mechanisms. Section 6

¹Kartik and Tercieux (2012) use an even more general approach of partial verifiability with the costs of sending certain messages depending on players' types.

contains the concluding remarks. All formal proofs are relegated to an appendix.

2 Trade Model

There is a seller (she) who has an object to sell to a single buyer (he). The buyer's valuation, v , is distributed according to the distribution F on $[0, \bar{v}]$ with a $\bar{v} > 0$. The buyer's valuation is his private information. The seller's costs c , are distributed independently according to the distribution G on $[0, \bar{c}]$ with a $\bar{c} > 0$. The seller's costs are her private information. If trade occurs, the seller's payoff is $p - c$ and the buyer's payoff is $v - p$; otherwise, both have a payoff of zero. Both the seller and the buyer are expected payoff maximizers. The following properties are common objectives for mechanisms:

First, the mechanism should be efficient.

Definition 1 (Ex-post efficiency). A mechanism is ex-post efficient if there is trade with probability one if and only if $v \geq c$.

Thus, any ex-post efficient mechanism creates a surplus of

$$\int_0^{\bar{v}} \int_0^v v - cdG(c)dF(v).$$

Second, both seller and buyer should be willing to enter the mechanism as participation is voluntary.

Definition 2 (Individual Rationality). A mechanism is individually rational if both seller and buyer expect nonnegative utilities.

Third, the mechanism should be budget-balanced, so that it requires no outside subsidies.² I return to the issue of budget balance in Section 5 which discusses strong budget balance, so that any outside payments are prohibited.

Definition 3 (Budget Balance). A mechanism is budget-balanced if it requires no outside subsidies.

It turns out, however, that it is impossible to find a mechanism for bilateral trade with these three properties.

²Outside subsidies make efficient implementation by Vickrey-Clarke-Groves mechanisms possible as demonstrated by Vickrey (1961), Clarke (1971), and Groves (1973).

Proposition 1. [Myerson and Satterthwaite (1983)] Assume that there are continuous and positive densities $f(v)$ and $g(c)$ for all $v, c \in [0, \min\{\bar{c}, \bar{v}\}]$. Then there is no mechanism that is individually rational, ex-post efficient and budget-balanced.

The assumption about the densities basically requires that there is sufficiently much uncertainty about costs and valuations. The way I state the assumption is the classical one by Myerson and Satterthwaite (1983). The result is more general, however. This result is based on the revelation principle that allows to focus on truthful and direct mechanisms. Such a mechanism specifies a probability of trade t and a price p conditional on the messages the parties send. Denote the buyer's message by \mathcal{M}_b and the seller's message by \mathcal{M}_s . A truthful mechanism requires incentive compatibility. Thus, the mechanism must incentivize both seller and buyer to report their valuations truthfully. Hence, we can write the optimization as:

$$\begin{aligned} & \max \int_0^{\bar{v}} \int_0^{\bar{v}} p(v, c) - t(v, c)c + t(v, c)v - p(v, c) dG(c) dF(v) & (A) \\ \text{subject to} & \int_0^{\bar{v}} p(v, c) - t(v, c)c - (p(v, \mathcal{M}_s) - t(v, \mathcal{M}_s)c) dF(v) \geq 0 & \forall \mathcal{M}_s \quad (IC_S) \\ & \int_0^{\bar{c}} t(v, c)v - p(v, c) - (t(\mathcal{M}_b, c)v - p(\mathcal{M}_b, c)) dG(c) \geq 0 & \forall \mathcal{M}_b \quad (IC_B) \\ & \int_0^{\bar{v}} p(v, c) - t(v, c)c dF(v) \geq 0 & \forall c \in [0, \bar{c}] \quad (IR_S) \\ & \int_0^{\bar{c}} t(v, c)v - p(v, c) dG(c) \geq 0 & \forall v \in [0, \bar{v}]. \quad (IR_B) \end{aligned}$$

It is easy to see that the prices cancel out in the objective so that the integrand is simply $(v - c)t(v, c)$. The conditions (IC_S) and (IC_B) ensure that truth-telling is optimal for seller and buyer. The conditions (IR_S) and (IR_B) guarantee that seller and buyer are willing to participate in the mechanism. The way I write down the program ensures budget balance.³

Before introducing the general model of justification, I discuss the basic intuition in a simplified setting with discrete valuations.

3 Justification: Basic Intuition

Assume that the production of the good potentially requires up to k_s inputs. Each input has costs of 1. The seller knows all necessary inputs and her costs are $c = \sum_{i=1}^{k_s} c_i$. The buyer only knows the distribution of inputs. Hence, $\text{Prob}(c_i = 1) = r_c \in (0, 1)$ and

³In general, there could be transfers τ_s for the seller and τ_b for the buyer instead of a price p .

- In period 1, nature determines preferences, technology and samples.
- In period 2, seller and buyer learn their private information.
- In period 3, seller and buyer can agree to the mechanism.
- In period 4, seller and buyer report their information.
- In period 5, the mechanism implements trade and payments.

Figure 1: Timing of the Model

$\text{Corr}(c_i, c_j) = \rho \in (-1, 1)$ for $i \neq j$. By inspection of the good or reading some reviews, the buyer learns for one input i^* whether it is required. The buyer's sample is drawn with probability $h^b(\cdot)$ with full support on $\mathbb{N}_{k_s} = \{1, 2, \dots, k_s\}$.

Vice versa, the buyer appreciates up to k_b features of the product. Each feature yields a utility of 1. The buyer knows his preferences and, hence, which feature of the product he appreciates. Thus, his valuation equals $v = \sum_{i=1}^{k_b} v_i$. The seller only knows the distribution of preferences. Hence, $\text{Prob}(v_i = 1) = r_v \in (0, 1)$, and $\text{Corr}(v_i, v_j) \in (-1, 1)$ for $i \neq j$. By analyzing social media or the cookies in the buyer's browser, the seller learns for one feature f^* whether it is appreciated by the buyer. The seller's sample is drawn with probability $h^s(\cdot)$ with full support on $\mathbb{N}_{k_b} = \{1, 2, \dots, k_b\}$. Denote $h_{min}^s = \min_{f \in \mathbb{N}_{k_s}} h^s(f)$ and $h_{min}^b = \min_{i \in \mathbb{N}_{k_b}} h^b(i)$. By assumption, h_{min}^s and h_{min}^b are positive. Payoffs are the same as in the trade model of Section 2. Figure 1 summarizes the timing.

Proposition 2. *There is a mechanism that is individually rational, ex-post efficient and budget-balanced.*

The interpretation of the mechanism becomes easier after introducing some notation. Begin with the messages. The buyer can send messages

$$\mathcal{M}_b = (m_b, M_b) \in \{0, 1, \dots, k_s\} \times 2^{\{1, 2, \dots, k_b\}}.^4$$

In equilibrium, he reports the features he appreciates $M_b = \{i \in \mathbb{N}_{k_b} | v_i = 1\}$ and his sample of the seller's inputs. In particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = 0$ otherwise. Vice versa, the seller can send messages

$$\mathcal{M}_s = (m_s, M_s) \in \{0, 1, \dots, k_s\} \times 2^{\{1, 2, \dots, k_s\}}.$$

In equilibrium, she reports the necessary inputs $M_s = \{i \in \mathbb{N}_{k_s} | c_i = 1\}$ and her sample

⁴Here 2^S denotes the power set of the set S , i.e., the set of all subsets of the set S .

of the buyer's preferences. In particular, she reports $m_s = f^*$ if $v_{f^*} = 1$ and $m_s = 0$ otherwise.

After learning their sample of the inputs and the features, both buyer and seller update their beliefs. Denote the seller's posterior about the buyer's valuation by $\tilde{F}(v)$ and the buyer's posterior about the seller's costs by $\tilde{G}(c)$ with the associated probability measures $\tilde{f}(v)$ and $\tilde{g}(c)$.

Finally, turn to payments. In addition to the price, there is a payment π_s by the seller to a third party. Denote $p^*(M_b, M_s) = \chi|M_b| + (1 - \chi)|M_s|$ with some relative bargaining power $\chi \in [0, 1]$ of the seller. Consider the following mechanism:

$$\left\{ \begin{array}{ll} t = 0 & p = 0 \quad \pi_s = 0 & \text{if } |M_b| < |M_s| \\ t = 1 & p = p^*(M_b, M_s) \quad \pi_s = 0 & \text{if } |M_b| \geq |M_s|, m_b \notin M_s \text{ and } m_s \notin (\mathbb{N}_{k_b} \setminus M_b) \\ t = 1 & p = p^*(M_b, M_s) + \frac{\chi}{h_{min}^s} \quad \pi_s = \frac{\chi}{h_{min}^s} & \text{if } |M_b| \geq |M_s|, m_b \notin M_s \text{ and } m_s \in (\mathbb{N}_{k_b} \setminus M_b) \\ t = 1 & p = p^*(M_b, M_s) \quad \pi_s = \frac{1-\chi}{h_{min}^b} & \text{if } |M_b| \geq |M_s|, m_b \in M_s \text{ and } m_s \notin (\mathbb{N}_{k_b} \setminus M_b) \\ t = 0 & p = \frac{\chi}{h_{min}^s} \quad \pi_s = \frac{\chi}{h_{min}^s} + \frac{1-\chi}{h_{min}^b} & \text{otherwise.} \end{array} \right. \quad (1)$$

It is easy to see that the mechanism (1) is budget-balanced. Above reporting strategies ensure that $m_b \notin M_s$ and $m_s \notin (\mathbb{N}_{k_b} \setminus M_b)$ in equilibrium. Hence, the mechanism (1) is also ex-post efficient if seller and buyer follow above reporting strategies. Individual rationality and incentive compatibility of the above strategies are not that obvious. Thus, let us study these in turn.

To study the buyer's utilities, suppose the seller reports as specified above. Then, the buyer expects payoffs of

$$\sum_{i=1}^v [v - p^*(v, i)] \tilde{g}(i) = (1 - \chi) \sum_{i=1}^v (v - i) \tilde{g}(i) \geq 0$$

if he follows above reporting strategy. These expected payoffs are clearly nonnegative so that the mechanism (1) is individually rational for the buyer.

Turning to incentive compatibility, consider deviations by the buyer. Deviating to a message with $|M_b| \geq v$ at least weakly increases the price and, hence, reduces the buyer's utility. Deviating to a message with $|M_b| < v$ has three effects: the price decreases, the probability of trade decreases, and a price surcharge of χ/h_{min}^s could be triggered. The first effect benefits the buyer while the second and third effect hurt him.

In particular, such a deviation yields an expected utility of at most

$$\sum_{i=1}^{|M_b|} \left[v - (\chi|M_b| + (1 - \chi)i) - \text{Prob}(m_s \in (\mathbb{N}_{k_b} \setminus M_b)) \frac{\chi}{h_{min}^s} \right] \tilde{g}(i).$$

Hence, such a deviation is unprofitable if

$$(1 - \chi) \sum_{i=|M_b|+1}^v (v - i) \tilde{g}(i) \geq \sum_{i=1}^{|M_b|} \left[\chi(v - |M_b|) - \text{Prob}(m_s \in (\mathbb{N}_{k_b} \setminus M_b)) \frac{\chi}{h_{min}^s} \right] \tilde{g}(i)$$

The left-hand side of this inequality is nonnegative, while the right-hand side is non-positive, because the right-hand side equals

$$\chi \left[v - |M_b| - \frac{\text{Prob}(m_s \in (\mathbb{N}_{k_b} \setminus M_b))}{h_{min}^s} \right] \sum_{i=1}^{|M_b|} \tilde{g}(i) \leq \chi(v - |M_b|)[1 - 1] \sum_{i=1}^{|M_b|} \tilde{g}(i) = 0.$$

The reason is that $\text{Prob}(m_s \in (\mathbb{N}_{k_b} \setminus M_b)) \geq h_{min}^s(v - |M_b|)$. Therefore, any deviation in M_b is unprofitable. Following above reporting strategy with respect to m_b could yield the first three cases in the mechanism (1). A deviation in m_b has no effect in the first case. If the second case applies, a deviation in m_b could make the fourth case apply — not affecting the buyer. If the third case applies, a deviation in m_b could make the fifth case apply — decreasing the buyer's payoffs while not affecting the buyer if the third case still applies following the deviation. Therefore, any deviation in m_b is unprofitable and, in addition, any joint deviation in M_b and m_b is unprofitable. Consequently, it is optimal for the buyer to follow above reporting strategy. I postpone the remainder of the proof to the appendix.

As the buyer's information is correlated with the seller's costs and the seller's information is correlated with the buyer's valuation, buyer and seller update their beliefs about each other's valuation. At the same time, this correlation might prima facie remind of ideas by McAfee and Reny (1992) and Crémer and McLean (1988) to exploit correlation in agents' beliefs.⁵ It is obvious from the mechanism that justification works differently and works even if their conditions are violated. To strengthen these differences, there is no such correlation in the main model. Indeed, the buyer's information and the seller's costs will be stochastically independent and vice versa.

⁵Also Gresik (1991) uses correlation to achieve efficient trade but for binary valuations.

4 Justification: Main Model

The buyer's valuation $v \in [0, \bar{v}]$ and the seller's costs $c \in [0, \bar{c}]$ are drawn from distributions $F(v)$ and $G(c)$ with continuous densities. Instead of k_s inputs and k_b features, there is a continuum of features and inputs now. In particular, each input $i \in (0, 1]$ is necessary with probability r_c , while each feature $f \in (0, 1]$ is appreciated with probability r_v . The inputs and features are essentially pairwise independent as defined by Sun (2006, Definition 2.7).

Lemma 1. *There is a probability space that satisfies these assumptions and guarantees a law of large numbers.*

The seller's costs are $\bar{c} \int_0^1 c_i d\lambda(i)$ with an extension λ of the Lebesgue measure on $[0, 1]$ as introduced by Sun and Zhang (2009, Theorem 1). The probabilities r_c and r_v are determined as follows $r_c = c/\bar{c}$ and $r_v = v/\bar{v}$. Hence, the seller's costs equal c with probability one. The buyer's valuation is $\bar{v} \int_0^1 v_f d\lambda(f)$ and equals v with probability one. Again the buyer learns for one input i^* whether it is necessary. That input i^* is drawn uniformly. Similarly, the feature f^* the seller observes is drawn uniformly.

To describe the messages, I need some additional notation: Denote by

$$\mathcal{N} = \{X \subseteq (0, 1] \mid X \text{ is } \lambda\text{-measurable}\}.$$

Then the buyer reports $\mathcal{M}_b = (m_b, M_b) \in [0, 1] \times \mathcal{N}_2$ and the seller reports $\mathcal{M}_s = (m_s, M_s) \in [0, 1] \times \mathcal{N}_2$.

Theorem 1. *There is a mechanism that is individually rational, ex-post efficient and budget-balanced.*

In equilibrium, the buyer reports the features he appreciates $M_b = \{f \in (0, 1] \mid v_f = 1\}$ and his sample i^* of the seller's inputs. In particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = 0$ otherwise. Vice versa, in equilibrium, the seller reports the necessary inputs $M_s = \{i \in (0, 1] \mid c_i = 1\}$ and her sample f^* of the buyer's preferences. In particular, she reports $m_s = f^*$ if $v_{f^*} = 1$ and $m_s = 0$ otherwise.

Adjust the definition of p^* to $p^*(M_b, M_s) = \chi \bar{v} \lambda(M_b) + (1 - \chi) \bar{c} \lambda(M_s)$ and consider the following mechanism:

$$\left\{ \begin{array}{ll} t = 0, & p = 0, \quad \pi_s = 0 & \text{if } \bar{v}\lambda(M_b) < \bar{c}\lambda(M_s) \\ t = 1, & p = p^*(M_b, M_s), \quad \pi_s = 0 & \text{if } \bar{v}\lambda(M_b) \geq \bar{c}\lambda(M_s), m_b \notin M_s \text{ and } m_s \notin ((0, 1] \setminus M_b) \\ t = 1, & p = p^*(M_b, M_s) + \bar{v}\chi, \quad \pi_s = \bar{v}\chi & \text{if } \bar{v}\lambda(M_b) \geq \bar{c}\lambda(M_s), m_b \notin M_s \text{ and } m_s \in ((0, 1] \setminus M_b) \\ t = 1, & p = p^*(M_b, M_s), \quad \pi_s = \bar{c}(1 - \chi) & \text{if } \bar{v}\lambda(M_b) \geq \bar{c}\lambda(M_s), m_b \in M_s \text{ and } m_s \notin ((0, 1] \setminus M_b) \\ t = 0, & p = \bar{v}\chi, \quad \pi_s = \bar{v}\chi + \bar{c}(1 - \chi) & \text{otherwise.} \end{array} \right. \quad (2)$$

It is easy to see that the mechanism (2) is budget-balanced. Above reporting strategies ensure that $m_b \notin M_s$ and $m_s \notin ((0, 1] \setminus M_b)$ in equilibrium. Hence, the mechanism (2) is also ex-post efficient if seller and buyer follow above reporting strategies. Individual rationality and incentive compatibility of the above strategies are not that obvious. Thus, let us study these in turn.

To study the seller's payoffs, suppose the buyer reports as specified above. Denote the conditional expectation of the buyer's valuation by $v^*(c) = \mathbb{E}(v|v \geq c)$. If the seller follows above reporting strategy, she expects payoffs of

$$\text{Prob}(\text{trade})(\mathbb{E}(p|\text{trade}) - c) = (1 - F(c))(v^*(c) - c)\chi = \chi \int_c^{\bar{v}} \tilde{v} - c dF(\tilde{v})$$

which is nonnegative. Therefore, the mechanism (2) is individually rational for the seller.

Turning to incentive compatibility, consider deviations by the seller. If the seller deviates in M_s , there are two cases to distinguish. Any deviation with $\bar{c}\lambda(M_s) \leq c$ makes the seller worse off by at least weakly reducing the price. A deviation with $\bar{c}\lambda(M_s) > c$ has three effects: the price increases, the probability of trade decreases, and additional payments $\pi_s = \bar{c}(1 - \chi)$ could be triggered. The first effect benefits the seller while the second and third effect hurt her. In particular, such a deviation implies payoffs of at most:

$$\begin{aligned} & (1 - F(\bar{c}\lambda(M_s))) \left[\chi v^*(\bar{c}\lambda(M_s)) + (1 - \chi)\bar{c}\lambda(M_s) - c - \left(\lambda(M_s) - \frac{c}{\bar{c}}\right)\bar{c}(1 - \chi) \right] = \\ & = (1 - F(\bar{c}\lambda(M_s)))(v^*(\bar{c}\lambda(M_s)) - c)\chi = \chi \int_{\bar{c}\lambda(M_s)}^{\bar{v}} \tilde{v} - c dF(\tilde{v}) \end{aligned}$$

The reason is that $\text{Prob}(m_b \in M_s) \geq \lambda(M_s) - c/\bar{c}$. Hence, such a deviation is unprofitable if

$$\chi \int_c^{\bar{v}} \tilde{v} - c dF(\tilde{v}) \geq \chi \int_{\bar{c}\lambda(M_s)}^{\bar{v}} \tilde{v} - c dF(\tilde{v})$$

$$\Leftrightarrow \int_c^{\bar{c}\lambda(M_s)} \tilde{v} - c dF(\tilde{v}) \geq 0 \quad \Leftrightarrow \quad \mathbb{E}(v | \bar{c}\lambda(M_s) \geq v \geq c) \geq c.$$

By definition of a conditional expectation, the last inequality is valid. Therefore, any deviation in M_s is unprofitable. Following above reporting strategy with respect to m_s could yield the first, the second or the fourth case in the mechanism (2). A deviation in m_s has no effect in the first case. If the second case applies, a deviation in m_s could make the third case apply — not affecting the seller. If the fourth case applies, a deviation in m_s could make the fifth case apply — decreasing the seller’s payoffs while not affecting the seller if the fourth case still applies following the deviation. Therefore, any deviation in m_s is unprofitable and, in addition, any joint deviation in M_s and m_s is unprofitable. Consequently, following above reporting strategy is optimal for the seller. I postpone the remainder of the proof to the appendix.

This section shows the full power of justification limiting the amount of information of the counterpart to a minimum. The mechanism does not require outside subsidies and there are no outside payments on the equilibrium path. The mechanism requires payments to third parties off the equilibrium path, however. Many results in the literature are in line with this property. For example, Kojima and Yamashita (2017, p. 1399) “do not regard [strong] budget balance to be indispensable as long as the mechanism runs no budget deficit.” Nevertheless, the question arises: Is it possible to find a mechanism without such payments?

5 Strong Budget Balance

Begin by defining strong budget balance.

Definition 4 (Strong Budget Balance). A mechanism is strongly budget-balanced if there are no outside payments.

The preceding mechanisms are budget-balanced, but not strongly budget-balanced. It turns out to be difficult to make them strongly budget-balanced. For a finite grid of values for costs and valuation, however, there are strongly budget-balanced mechanisms. For this purpose, suppose $v \in \mathcal{V} = \{\frac{1}{n}\bar{v}, \frac{2}{n}\bar{v}, \dots, \bar{v}\}$ and $c \in \{0, \frac{2}{n}\bar{v}, \frac{4}{n}\bar{v}, \dots, \frac{n-1}{n}\bar{v}\}$ with an odd $n \geq 3$. For low values of n , the step size between different values and, hence, the minimal deviation is sufficiently large so that there are efficient and strongly budget-balanced mechanisms for any distribution of valuations.

Proposition 3. *There is a mechanism that is individually rational, ex-post efficient and strongly budget-balanced if $n < 7$.*

The proof shows that the mechanism

$$\left\{ \begin{array}{l} t = 1 \quad p = \bar{v} \frac{\lambda(M_b) + \lambda(M_s)}{2} \\ \quad \text{if } \lambda(M_b) \geq \lambda(M_s), \bar{v}\lambda(M_b) \in \mathcal{V}, \bar{v}(1 - \lambda(M_s)) \in \mathcal{V}, m_b \notin M_s \text{ and } m_s \notin ((0, 1] \setminus M_b) \\ t = 0 \quad p = 0 \quad \text{otherwise.} \end{array} \right. \quad (3)$$

is individually rational, ex-post efficient and strongly budget-balanced for sufficiently large step size. The step size unfortunately also determines the number of possible realizations of seller's costs and buyer's valuation. Nevertheless, even the case with binary valuations which corresponds to $n = 3$ in my model is widely used in applications, such as Schweizer (1989), Chatterjee and Samuelson (1987), and Fudenberg and Tirole (1983). Matsuo (1989) considers efficient mechanisms for binary valuations though their conditions are quite restrictive. Notice that my Proposition 3 is valid even if the conditions of the theorem in Matsuo (1989) are violated.

For distributions with larger supports and $n \geq 7$, the possibility result requires additional conditions on the distributions of seller's costs and buyer's valuation.

Theorem 2. *There is a mechanism that is incentive compatible, individually rational, ex-post efficient and strongly budget-balanced if $n \geq 7$, $f(\bar{v} \frac{n-2}{n}) > 2f(\bar{v})$, and $g(\bar{v} \frac{2}{n}) > 2g(0)$ as well as $f(\bar{v} \frac{2k-1}{n}) > 3f(\bar{v} \frac{2k+1}{n})$ and $g(\bar{v} \frac{2k+2}{n}) > 3g(\bar{v} \frac{2k}{n})$ for all $k \in \{1, 2, \dots, \frac{n-3}{2}\}$.*

The conditions in Theorem 2 require that high costs of the seller are more likely than low costs. Vice versa, low values of the buyer are more likely than high values. Notice that such distributions are also the relevant ones. In the opposite case, when expected costs are low and expected values are high, it is easy for buyer and seller to agree on trade. It is exactly in the case of Theorem 2 that it is difficult for buyer and seller to agree on trade because on average it is optimal not to trade. Nevertheless, there could be large gains from trade in some states of the world. Justification helps to realize these gains from trade and it helps exactly when it is most difficult to reach agreement.

6 Conclusions

I consider justification. Justification enriches the message space to make messages partially verifiable. The idea is that the sender provides a number of reasons, arguments and points to make her claim. The recipient derives the sender's claim from the provided reasons. Therefore, changing her claim requires changing several arguments. Indeed,

the larger the change, the more arguments have to be adjusted. The recipient knows for one reason whether it is appropriate for the sender's claim. Hence, if the sender deviates from equilibrium play the recipient notices such a deviation with some probability. Justification works even if the recipient cannot infer anything about the sender's type or equilibrium claim from the one reason she knows.

I apply justification to mechanism design and choose the canonical setting of bilateral trade for this purpose. I show that justification allows for ex-post efficient mechanisms that are individually rational and budget-balanced. Consequently, justification is powerful enough to overturn previous impossibility results, e.g., by Myerson and Satterthwaite (1983).

...

I assume simultaneous reporting of all messages. The mechanisms considered in this paper, however, are robust to different timing of the messages. In particular, sequential messages are possible as long as justifications are provided first or at least other messages are not made public before the sender provides justification. In the model, this assumption requires that the buyer does not observe the seller's message m_s before justifying his valuation. Vice versa, the seller shall not observe the buyer's message m_b before justifying her costs. ...

A Appendix

Proof of Proposition 1: See Myerson and Satterthwaite (1983, Corollary 1). \square

Proof of Proposition 2: In the main text, I state a budget-balanced mechanism (1) and corresponding reporting strategies. Following the specified reporting strategies guarantees that $|M_b| = v$ and $|M_s| = c$ such that there is trade if and only if $v \geq c$. Therefore, the mechanism (1) is ex-post efficient. In addition, I show that the reporting strategy is optimal for the buyer and he participates voluntarily because the mechanism is individually rational for him.

Now turn to the seller. For this purpose, suppose the buyer follows the specified reporting strategy. Then the seller expects payoffs of

$$\sum_{i=c}^{k_b} [p^*(i, c) - c] \tilde{f}(i) = \chi \sum_{i=c}^{k_b} (i - c) \tilde{f}(i) \geq 0$$

if she follows above reporting strategy. These expected payoffs are clearly nonnegative so that the mechanism is individually rational for the seller.

Turning to incentive compatibility, consider deviations by the seller. Deviating

to a message with $|M_s| \leq c$ at least weakly lowers the price and, hence, the seller's payoffs. Deviating to a message with $|M_s| > c$ has three effects: the price increases, the probability of trade decreases, and additional payments of $\pi_s = (1 - \chi)/h_{min}^b$ could be triggered. The first effect benefits the seller while the second and third effect hurt her. In particular, such a deviation yields expected payoffs of at most

$$\sum_{i=|M_s|}^{k_b} \left[\chi i + (1 - \chi)|M_s| - c - \text{Prob}(m_b \in M_s) \frac{1 - \chi}{h_{min}^b} \right] \tilde{f}(i).$$

Hence, any deviation is unprofitable if

$$\chi \sum_{i=c}^{|M_s|-1} (i - c) \tilde{f}(i) \geq \sum_{i=|M_s|}^{k_b} \left[(1 - \chi)(|M_s| - c) - \text{Prob}(m_b \in M_s) \frac{1 - \chi}{h_{min}^b} \right] \tilde{f}(i)$$

The left-hand side of the inequality is nonnegative, while the right-hand side is nonpositive, because the right-hand side equals

$$\begin{aligned} (1 - \chi) \left[|M_s| - c - \text{Prob}(m_b \in M_s) \frac{1}{h_{min}^b} \right] \sum_{i=|M_s|}^{k_b} \tilde{f}(i) &\leq \\ &\leq (1 - \chi)(|M_s| - c)[1 - 1] \sum_{i=|M_s|}^{k_b} \tilde{f}(i) = 0. \end{aligned}$$

The reason is that $\text{Prob}(m_b \in M_s) \geq h_{min}^b(|M_s| - c)$. Therefore, any deviation in M_s is unprofitable. Following above reporting strategy with respect to m_s could yield the first, the second or the fourth case in the mechanism (1). A deviation in m_s has no effect in the first case. If the second case applies, a deviation in m_s could make the third case apply — not affecting the seller. If the fourth case applies, a deviation in m_s could make the fifth case apply — decreasing the seller's payoffs while not affecting the seller if the fourth case still applies following the deviation. Therefore, any deviation in m_s is unprofitable and, in addition, any joint deviation in M_s and m_s is unprofitable. Consequently, it is optimal for the seller to follow the specified reporting strategy.

To sum up, there is a mechanism that is individually rational, ex-post efficient and budget-balanced. \square

Proof of Lemma 1: A probability space for the random variable v (c , resp.) is constructed in the common way using the Borel σ -algebra. For a given v and r_v (c and r_c , resp.), I require a suitable probability space for the c_i and v_f . For this purpose, I follow the approach by Sun (2006). Hence, I consider a Fubini extension instead of

the usual continuum product based on the Kolmogorov construction. Sun and Zhang (2009, Theorem 1 and Corollary 2) prove that there exist a set Ω , a probability space on Ω , an extension λ of the Lebesgue measure $\bar{\lambda}$ on $[0, 1]$, a Fubini extension on $[0, 1] \times \Omega$ and a process $z: [0, 1] \times \Omega \mapsto \mathbb{R}$, such that the random variables $z(i, \cdot)$ are essentially pairwise independent with the required distribution: $\Pr(\{\omega \in \Omega | z(i, \omega) = 1\}) = r_v$ (r_c , resp.) for $i \in [0, 1]$ almost surely. By definition of a Fubini extension, the integral $\int_{[0,1]} z(i, \omega) d\lambda(i)$ is well defined for all $\omega \in \Omega$. In addition, Sun (2006, Theorem 2.8) proves that the integral equals r_v (r_c , resp.) almost surely. \square

Proof of Theorem 1: In the main text, I state a budget-balanced mechanism (2) and corresponding reporting strategies. Following the specified reporting strategies guarantees that $\bar{v}\lambda(M_b) = v$ and $\bar{c}\lambda(M_s) = c$ such that there is trade if and only if $v \geq c$. Therefore, the mechanism (2) is ex-post efficient. In addition, I show that the reporting strategy is optimal for the seller and she participates voluntarily because the mechanism (2) is individually rational for her.

Now turn to the buyer. For this purpose, denote the conditional expectation of the seller's costs by $c^*(v) = \mathbb{E}(c | v \geq c)$ and suppose the seller follows the specified reporting strategy. Then the buyer expects payoffs of

$$\text{Prob}(\text{trade})(v - \mathbb{E}(p | \text{trade})) = G(v)(v - c^*(v))(1 - \chi) = (1 - \chi) \int_0^v v - \tilde{c} dG(\tilde{c}).$$

if he follows above reporting strategy. These expected payoffs are clearly nonnegative so that the mechanism (2) is individually rational for the buyer.

Turning to incentive compatibility, consider deviations by the buyer. If the buyer deviates in M_b , there are two cases to distinguish. Any deviation with $\bar{v}\lambda(M_b) \geq v$ makes the buyer worse off by increasing the price and potentially triggering trade at a price above his valuation while there are no benefits. A deviation with $\bar{v}\lambda(M_b) < v$ has three effects: the price decreases, the probability of trade decreases, and a price surcharge of $\bar{v}\chi$ could be triggered. The first effect benefits the buyer while the second and third effect hurt him. In particular, this deviation implies a utility of at most:

$$\begin{aligned} G(\bar{v}\lambda(M_b)) \left[v - (\chi \bar{v}\lambda(M_b) + (1 - \chi)c^*(\bar{v}\lambda(M_b))) - \left(\frac{v}{\bar{v}} - \lambda(M_b)\right)\bar{v}\chi \right] = \\ = G(\bar{v}\lambda(M_b))(v - c^*(\bar{v}\lambda(M_b)))(1 - \chi) = (1 - \chi) \int_0^{\bar{v}\lambda(M_b)} v - \tilde{c} dG(\tilde{c}) \end{aligned}$$

The reason is that $\text{Prob}(m_s \in ((0, 1] \setminus M_b)) \geq v/\bar{v} - \lambda(M_b)$. Hence, such a deviation is

unprofitable if

$$\begin{aligned}
(1 - \chi) \int_0^v v - \tilde{c} dG(\tilde{c}) &\geq (1 - \chi) \int_0^{\bar{v}\lambda(M_b)} v - \tilde{c} dG(\tilde{c}) \\
&\Leftrightarrow \int_{\bar{v}\lambda(M_b)}^v v - \tilde{c} dG(\tilde{c}) \geq 0 \\
&\Leftrightarrow \mathbb{E}(c|v \geq c \geq \bar{v}\lambda(M_b)) \leq v.
\end{aligned}$$

By definition of a conditional expectation, the last inequality is valid. Therefore, any deviation in M_b is unprofitable. Following above reporting strategy with respect to m_b could yield the first three cases in the mechanism (2). A deviation in m_b has no effect in the first case. If the second case applies, a deviation in m_b could make the fourth case apply — not affecting the buyer. If the third case applies, a deviation in m_b could make the fifth case apply — decreasing the buyer's payoffs while not affecting the buyer if the third case still applies following the deviation. Therefore, any deviation in m_b is unprofitable and, in addition, any joint deviation in M_b and m_b is unprofitable. Consequently, it is optimal for the buyer to follow the specified reporting strategy.

To sum up, there is a mechanism that is individually rational, ex-post efficient and budget-balanced. \square

Proof of Proposition 3: Consider the same reporting strategies as in Theorem 1 and the mechanism:

$$\left\{ \begin{array}{l} t = 1 \quad p = \bar{v} \frac{\lambda(M_b) + \lambda(M_s)}{2} \\ \quad \text{if } \lambda(M_b) \geq \lambda(M_s), \bar{v}\lambda(M_b) \in \mathcal{V}, \bar{v}(1 - \lambda(M_s)) \in \mathcal{V}, m_b \notin M_s \text{ and } m_s \notin ((0, 1] \setminus M_b) \\ t = 0 \quad p = 0 \quad \text{otherwise.} \end{array} \right.$$

It is easy to see that the mechanism is strongly budget-balanced. The reporting strategies ensure that $m_b \notin M_s$ and $m_s \notin ((0, 1] \setminus M_b)$ as well as $\bar{v}\lambda(M_b) = v \in \mathcal{V}$ and $\bar{v}\lambda(M_s) = c$, thus, $\bar{v}(1 - \lambda(M_s)) \in \mathcal{V}$ in equilibrium. Hence, there is trade if and only if $v \geq c$. Therefore, the mechanism is ex-post efficient if seller and buyer follow their reporting strategies. Individual rationality and incentive compatibility of the above strategies are not that obvious. Thus, let us study these in turn.

In a slight abuse of notation denote by $f(v)$ and $g(c)$ the probability measures for the buyer's valuation and the seller's costs. To study the buyer's utilities, suppose the

seller reports as specified in Theorem 1. Then, the buyer expects payoffs of

$$\sum_{i=0}^{\frac{nv-\bar{v}}{2\bar{v}}} \left[v - \frac{v + i\frac{2}{n}\bar{v}}{2} \right] g\left(i\frac{2}{n}\bar{v}\right) = \frac{1}{2} \sum_{i=0}^{\frac{nv-\bar{v}}{2\bar{v}}} (v - i\frac{2}{n}\bar{v})g\left(i\frac{2}{n}\bar{v}\right) \geq 0$$

if he follows his reporting strategy. These expected payoffs are nonnegative so that above mechanism is individually rational for the buyer.

Turning to incentive compatibility, consider deviations by the buyer. For this purpose, suppose the seller follows the specified reporting strategy. Deviating to a message with $\bar{v}\lambda(M_b) \geq v$ at least weakly increases the price and, hence, reduces the buyer's utility. Any deviation with $\bar{v}\lambda(M_b) \notin \mathcal{V}$ is clearly unprofitable. Deviating to a message with $\bar{v}\lambda(M_b) \in \mathcal{V}$ and $\bar{v}\lambda(M_b) < v$ has two effects: the price decreases and the probability of trade decreases. The first effect benefits the buyer while the second effect hurts him. In particular, such a deviation yields an expected utility of at most

$$\begin{aligned} & \text{Prob}(m_s \notin ((0, 1] \setminus M_b)) \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left[v - \frac{\bar{v}}{2}(\lambda(M_b) + i\frac{2}{n}) \right] g\left(i\frac{2}{n}\bar{v}\right) \leq \\ & \leq \left(1 - \frac{v}{\bar{v}} + \lambda(M_b)\right) \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left[v - \frac{\lambda(M_b)\bar{v}}{2} - i\frac{\bar{v}}{n} \right] g\left(i\frac{2}{n}\bar{v}\right). \end{aligned}$$

The reason is that $\text{Prob}(m_s \in ((0, 1] \setminus M_b)) \geq \frac{v}{\bar{v}} - \lambda(M_b)$ and that the sum is nonnegative. Such a deviation is unprofitable if

$$\begin{aligned} & \frac{1}{2} \sum_{i=0}^{\frac{nv-\bar{v}}{2\bar{v}}} (v - i\frac{2}{n}\bar{v})g\left(i\frac{2}{n}\bar{v}\right) \geq \left(1 - \frac{v}{\bar{v}} + \lambda(M_b)\right) \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left[v - \frac{\lambda(M_b)\bar{v}}{2} - i\frac{\bar{v}}{n} \right] g\left(i\frac{2}{n}\bar{v}\right) \\ \Leftrightarrow & \frac{1}{2} \sum_{i=\frac{n\lambda(M_b)+1}{2}}^{\frac{nv-\bar{v}}{2\bar{v}}} (v - i\frac{2}{n}\bar{v})g\left(i\frac{2}{n}\bar{v}\right) \geq \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left\{ \left[v - \frac{\lambda(M_b)\bar{v}}{2} - i\frac{\bar{v}}{n} \right] \left(1 - \frac{v}{\bar{v}} + \lambda(M_b)\right) - \frac{1}{2}(v - i\frac{2}{n}\bar{v}) \right\} g\left(i\frac{2}{n}\bar{v}\right) \end{aligned} \quad (4)$$

The right-hand side of this inequality equals

$$\begin{aligned} & \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left\{ \left[v - \frac{\lambda(M_b)\bar{v}}{2} - i\frac{\bar{v}}{n} - \frac{v}{2} + i\frac{\bar{v}}{n} \right] - \left[v - \frac{\lambda(M_b)\bar{v}}{2} - i\frac{\bar{v}}{n} \right] \left(\frac{v}{\bar{v}} - \lambda(M_b) \right) \right\} g\left(i\frac{2}{n}\bar{v}\right) = \\ & = \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left\{ \frac{1}{2}(v - \lambda(M_b)\bar{v}) - \left[\frac{v}{\bar{v}} - \frac{\lambda(M_b)}{2} - \frac{i}{n} \right] (v - \lambda(M_b)\bar{v}) \right\} g\left(i\frac{2}{n}\bar{v}\right) = \end{aligned}$$

$$= (v - \lambda(M_b)\bar{v}) \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left[\frac{1}{2} - \frac{v}{\bar{v}} + \frac{\lambda(M_b)}{2} + \frac{i}{n} \right] g\left(i\frac{2}{n}\bar{v}\right)$$

For $n < 7$, $\lambda(M_b) \in \{1/3, 1/5, 3/5\}$ as $\lambda(M_b) < v/\bar{v} \leq 1$. Hence, the sum goes at most up to $i = 1$. The case $i = 1$ is only possible for $n = 5$, $\lambda(M_b) = 3/5$ and $v = \bar{v}$. Then the term in square brackets equals $1/2 - 1 + 3/10 + 1/5 = 0$. For $i = 0$, the term in square brackets equals $1/2 - v/\bar{v} + \lambda(M_b)/2 \leq \max\{\lambda(M_b)/2 - 1/2, 1/2 - 3/5 + 1/10\} \leq 0$. Therefore, the right-hand side of inequality (4) is nonpositive, while the left-hand side is nonnegative. Consequently, any deviation in M_b is unprofitable.

Following above reporting strategy with respect to m_b could yield both cases in the mechanism. A deviation in m_b has no effect in the second case. If the first case applies, a deviation in m_b could make the second case apply — decreasing the buyer's payoffs while not affecting the buyer if the first case still applies following the deviation. Therefore, any deviation in m_b is unprofitable and, in addition, any joint deviation in M_b and m_b is unprofitable. Consequently, it is optimal for the buyer to follow the specified reporting strategy.

To study the seller's payoffs, suppose the buyer reports as specified above. If the seller follows above reporting strategy, she expects payoffs of

$$\sum_{i=\frac{nc}{2\bar{v}}}^{\frac{n-1}{2}} \left[\frac{2i+1}{n}\bar{v} + c - c \right] f\left(\frac{2i+1}{n}\bar{v}\right) = \frac{1}{2} \sum_{i=\frac{nc}{2\bar{v}}}^{\frac{n-1}{2}} \left(\frac{2i+1}{n}\bar{v} - c \right) f\left(\frac{2i+1}{n}\bar{v}\right) \geq 0$$

which are nonnegative. Therefore, the mechanism is individually rational for the seller.

Turning to incentive compatibility, consider deviations by the seller. If the seller deviates in M_s , there are two cases to distinguish. Any deviation with $\bar{c}\lambda(M_s) \leq c$ makes the seller worse off by at least weakly reducing the price. Any deviation with $\bar{v}(1 - \lambda(M_s)) \notin \mathcal{V}$ is clearly unprofitable. A deviation with $\bar{v}(1 - \lambda(M_s)) \in \mathcal{V}$ and $\bar{c}\lambda(M_s) > c$ has two effects: the price increases and the probability of trade decreases. The first effect benefits the seller while the second effect hurts her. In particular, such a deviation implies payoffs of at most:

$$\begin{aligned} & \text{Prob}(m_b \notin M_s) \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left[\frac{\bar{v}}{2} \left(\lambda(M_s) + \frac{2i+1}{n} \right) - c \right] f\left(\frac{2i+1}{n}\bar{v}\right) \leq \\ & \leq \left(1 + \frac{c}{\bar{v}} - \lambda(M_s) \right) \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left[\frac{\lambda(M_s)\bar{v}}{2} + \bar{v} \frac{2i+1}{2n} - c \right] f\left(\frac{2i+1}{n}\bar{v}\right). \end{aligned}$$

The reason is that $\text{Prob}(m_b \in M_s) \geq \lambda(M_s) - \frac{c}{\bar{v}}$ and that the sum is nonnegative. Such a deviation is unprofitable if

$$\begin{aligned}
& \frac{1}{2} \sum_{i=\frac{nc}{2\bar{v}}}^{\frac{n-1}{2}} \left(\frac{2i+1}{n} \bar{v} - c \right) f\left(\frac{2i+1}{n} \bar{v} \right) \geq \left(1 + \frac{c}{\bar{v}} - \lambda(M_s) \right) \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left[\frac{\lambda(M_s)\bar{v}}{2} + \bar{v} \frac{2i+1}{2n} - c \right] f\left(\frac{2i+1}{n} \bar{v} \right) \\
& \Leftrightarrow \frac{1}{2} \sum_{i=\frac{nc}{2\bar{v}}}^{\frac{n\lambda(M_s)}{2}-1} \left(\frac{2i+1}{n} \bar{v} - c \right) f\left(\frac{2i+1}{n} \bar{v} \right) \geq \\
& \geq \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left\{ \left[\frac{\lambda(M_s)\bar{v}}{2} + \bar{v} \frac{2i+1}{2n} - c \right] \left(1 + \frac{c}{\bar{v}} - \lambda(M_s) \right) - \frac{1}{2} \left(\frac{2i+1}{n} \bar{v} - c \right) \right\} f\left(\frac{2i+1}{n} \bar{v} \right) \quad (5)
\end{aligned}$$

The right-hand side of this inequality equals

$$\begin{aligned}
& \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left\{ \left[\frac{\lambda(M_s)\bar{v}}{2} + \bar{v} \frac{2i+1}{2n} - \frac{c}{2} - \bar{v} \frac{2i+1}{2n} \right] - \left[\frac{\lambda(M_s)\bar{v}}{2} + \bar{v} \frac{2i+1}{2n} - c \right] \left(\lambda(M_s) - \frac{c}{\bar{v}} \right) \right\} f\left(\frac{2i+1}{n} \bar{v} \right) = \\
& = \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left\{ \frac{1}{2} (\lambda(M_s)\bar{v} - c) - \left[\frac{\lambda(M_s)}{2} + \frac{2i+1}{2n} - \frac{c}{\bar{v}} \right] (\lambda(M_s)\bar{v} - c) \right\} f\left(\frac{2i+1}{n} \bar{v} \right) = \\
& = (\lambda(M_s)\bar{v} - c) \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left[\frac{1}{2} - \frac{\lambda(M_s)}{2} - \frac{2i+1}{2n} + \frac{c}{\bar{v}} \right] f\left(\frac{2i+1}{n} \bar{v} \right)
\end{aligned}$$

For $n < 7$, $\lambda(M_s) \geq 2/n$ as $\lambda(M_s) > c/\bar{v} \geq 0$. Hence, the sum begins at least at $i = 1$. The case $i = 2$ is only possible for $n = 5$. Then the term in square brackets equals $1/2 - \frac{\lambda(M_s)}{2} - 1/2 + \frac{c}{\bar{v}} = \frac{c}{\bar{v}} - \frac{\lambda(M_s)}{2}$. However, $\frac{c}{\bar{v}} - \frac{\lambda(M_s)}{2} < 0$ for $c = 0$ and $\frac{c}{\bar{v}} - \frac{\lambda(M_s)}{2} = 0$ for $c = 2/5$ and $\lambda(M_s) = 4/5$. The case $i = 1$ is only possible for $c = 0$. Then the term in square brackets equals $1/2 - \lambda(M_s)/2 - 3/(2n) \leq 1/2 - 1/n - 3/(2n) = (n-5)/(2n) \leq 0$. Therefore, the right-hand side of inequality (5) is nonpositive, while the left-hand side is nonnegative. Consequently, any deviation in M_s is unprofitable.

Following above reporting strategy with respect to m_s could yield both cases in the mechanism. A deviation in m_s has no effect in the second case. If the first case applies, a deviation in m_s could make the second case apply — decreasing the seller's payoffs while not affecting the seller if the first case still applies following the deviation. Therefore, any deviation in m_s is unprofitable and, in addition, any joint deviation in M_s and m_s is unprofitable. Consequently, it is optimal for the seller to follow the specified reporting strategy. To sum up, the mechanism (3) is individually rational,

ex-post efficient and strongly budget-balanced. \square

Proof of Theorem 2: Return to the mechanism (3). The proof of Proposition 3 shows that the mechanism is individually rational and strongly budget-balanced. The reporting strategies in Theorem 1 ensure that $m_b \notin M_s$ and $m_s \notin ((0, 1] \setminus M_b)$ as well as $\bar{v}\lambda(M_b) = v \in \mathcal{V}$ and $\bar{v}\lambda(M_s) = c$, thus, $\bar{v}(1 - \lambda(M_s)) \in \mathcal{V}$ in equilibrium. Hence, there is trade if and only if $v \geq c$. Therefore, the mechanism is ex-post efficient if seller and buyer follow their reporting strategies. Incentive compatibility of the reporting strategies is not that obvious.

Begin with the buyer. For this purpose, suppose the seller follows the specified reporting strategy. The proof of Proposition 3 shows that it is optimal for the buyer to follow the reporting strategy if inequality (4) is satisfied. Simplifying the right-hand side of that inequality as in the proof of Proposition 3 yields the buyer's incentive compatibility:

$$\frac{1}{2} \sum_{i=\frac{n\lambda(M_b)+1}{2}}^{\frac{nv-\bar{v}}{2\bar{v}}} (v - i\frac{2}{n}\bar{v})g(i\frac{2}{n}\bar{v}) \geq (v - \lambda(M_b)\bar{v}) \sum_{i=0}^{\frac{n\lambda(M_b)-1}{2}} \left(\frac{1}{2} - \frac{v}{\bar{v}} + \frac{\lambda(M_b)}{2} + \frac{i}{n} \right) g(i\frac{2}{n}\bar{v}) \quad (6)$$

Any deviation with $\lambda(M_b) = 1/n$ is unprofitable if

$$\begin{aligned} \frac{1}{2}(v - \frac{2}{n}\bar{v})g(\frac{2}{n}\bar{v}) &\geq (v - \frac{1}{n}\bar{v}) \left(\frac{1}{2} - \frac{v}{\bar{v}} + \frac{1}{2n} \right) g(0) \\ \Leftrightarrow g(\frac{2}{n}\bar{v})/g(0) &\geq \frac{v - \frac{1}{n}\bar{v}}{v - \frac{2}{n}\bar{v}} \left(1 - 2\frac{v}{\bar{v}} + \frac{1}{n} \right) \end{aligned}$$

The right-hand side decreases in $v \geq 3\bar{v}/n$. Therefore such a deviation is unprofitable if the last inequality is satisfied for the smallest $v > \lambda(M_b)\bar{v}$, namely $v = 3\bar{v}/n$. Then the right-hand side equals $2(1 - 5/n)$. Consequently, any deviation with $\lambda(M_b) = 1/n$ is unprofitable if $g(\frac{2}{n}\bar{v}) \geq 2g(0)$.

Any deviation with $\lambda(M_b) = k/n$ and odd $k \geq 3$ is unprofitable if

$$(v - \frac{k+1}{n}\bar{v})g(\frac{k+1}{n}\bar{v}) \geq (v - \frac{k}{n}\bar{v}) \sum_{i=0}^{\frac{k-1}{2}} \left(1 - 2\frac{v}{\bar{v}} + \frac{k}{n} + 2\frac{i}{n} \right) g(i\frac{2}{n}\bar{v}) \quad (7)$$

Inserting $g(\bar{v}\frac{2}{n}) > 2g(0)$ and $g(\bar{v}\frac{2i+2}{n}) > 3g(\bar{v}\frac{2i}{n})$ for all $i \in \{1, 2, \dots, \frac{k-3}{2}\}$, the right-

hand side of this inequality is smaller than

$$\left(v - \frac{k}{n}\bar{v}\right)g\left(\frac{k-1}{n}\bar{v}\right) \left(\left(\frac{1}{3}\right)^{\frac{k-3}{2}} \frac{1}{2} \left(1 - 2\frac{v}{\bar{v}} + \frac{k}{n}\right) + \sum_{i=1}^{\frac{k-1}{2}} \left(\frac{1}{3}\right)^{\frac{k-1}{2}-i} \left(1 - 2\frac{v}{\bar{v}} + \frac{k}{n} + 2\frac{i}{n}\right) \right)$$

Hence, inequality (7) is satisfied if

$$\begin{aligned} & g\left(\frac{k+1}{n}\bar{v}\right)/g\left(\frac{k-1}{n}\bar{v}\right) \geq \\ & \geq \frac{v - \frac{k}{n}\bar{v}}{v - \frac{k+1}{n}\bar{v}} \left(\left(\frac{1}{3}\right)^{\frac{k-3}{2}} \frac{1}{2} \left(1 - 2\frac{v}{\bar{v}} + \frac{k}{n}\right) + \sum_{i=1}^{\frac{k-1}{2}} \left(\frac{1}{3}\right)^{\frac{k-1}{2}-i} \left(1 - 2\frac{v}{\bar{v}} + \frac{k}{n} + 2\frac{i}{n}\right) \right). \end{aligned}$$

The right-hand side decreases in $v \geq (k+2)\bar{v}/n$. Therefore such a deviation is unprofitable if the last inequality is satisfied for the smallest $v > \lambda(M_b)\bar{v}$, namely $v = (k+2)\bar{v}/n$. Then the right-hand side equals

$$\begin{aligned} & 2 \left(\left(\frac{1}{3}\right)^{\frac{k-3}{2}} \frac{1}{2} \left(1 - \frac{k+4}{n}\right) + \sum_{i=1}^{\frac{k-1}{2}} \left(\frac{1}{3}\right)^{\frac{k-1}{2}-i} \left(1 - \frac{k+4-2i}{n}\right) \right) \leq \\ & \leq \frac{1}{3^{\frac{k-3}{2}}} + 2 \sum_{i=1}^{\frac{k-1}{2}} \left(\frac{1}{3}\right)^{\frac{k-1}{2}-i} = \frac{1}{3^{\frac{k-3}{2}}} + 2 \sum_{i=0}^{\frac{k-3}{2}} \frac{1}{3^i} = \frac{1}{3^{\frac{k-3}{2}}} + 2 \frac{1 - \left(\frac{1}{3}\right)^{\frac{k-1}{2}}}{1 - \frac{1}{3}} = \\ & = \frac{1}{3^{\frac{k-3}{2}}} + 3 \left(1 - \left(\frac{1}{3}\right)^{\frac{k-1}{2}}\right) = 3. \end{aligned}$$

because $k+4-2i > 0$ for all $i \leq \frac{k-1}{2}$ and the common formula for a geometric series applies. Consequently, any deviation with $\lambda(M_b) = k/n$ and odd $k \geq 3$ is unprofitable if $g(\bar{v}^{\frac{2}{n}}) > 2g(0)$ and $g(\bar{v}^{\frac{2i+2}{n}}) > 3g(\bar{v}^{\frac{2i}{n}})$ for all $i \in \{1, 2, \dots, \frac{n-3}{2}\}$.

Continue with the seller. For this purpose, suppose the buyer follows the specified reporting strategy. The proof of Proposition 3 shows that it is optimal for the seller to follow the reporting strategy if inequality (5) is satisfied. Simplifying the right-hand side of that inequality as in the proof of Proposition 3 yields the buyer's incentive compatibility:

$$\begin{aligned} & \frac{1}{2} \sum_{i=\frac{nc}{2\bar{v}}}^{\frac{n\lambda(M_s)}{2}-1} \left(\frac{2i+1}{n}\bar{v} - c\right) f\left(\frac{2i+1}{n}\bar{v}\right) \geq \\ & \geq (\lambda(M_s)\bar{v} - c) \sum_{i=\frac{n\lambda(M_s)}{2}}^{\frac{n-1}{2}} \left[\frac{1}{2} - \frac{\lambda(M_s)}{2} - \frac{2i+1}{2n} + \frac{c}{\bar{v}} \right] f\left(\frac{2i+1}{n}\bar{v}\right) \end{aligned} \quad (8)$$

Any deviation with $\lambda(M_s) = (n-1)/n$ is unprofitable if

$$\begin{aligned} \frac{1}{2} \left(\frac{n-2}{n} \bar{v} - c \right) f \left(\frac{n-2}{n} \bar{v} \right) &\geq \left(\frac{n-1}{n} \bar{v} - c \right) \left(\frac{1}{2} - \frac{n-1}{2n} - \frac{1}{2} + \frac{c}{\bar{v}} \right) f(\bar{v}) \\ \Leftrightarrow f \left(\frac{n-2}{n} \bar{v} \right) / f(\bar{v}) &\geq \frac{\frac{n-1}{n} \bar{v} - c}{\frac{n-2}{n} \bar{v} - c} \left(2 \frac{c}{\bar{v}} - \frac{n-1}{n} \right) \end{aligned}$$

The right-hand side increases in $c \leq (n-3)\bar{v}/n$. Therefore such a deviation is unprofitable if the last inequality is satisfied for the largest $c < \lambda(M_s)\bar{v}$, namely $c = (n-3)\bar{v}/n$. Then the right-hand side equals $2(1-5/n)$. Consequently, any deviation with $\lambda(M_s) = (n-1)/n$ is unprofitable if $f(\frac{n-2}{n}\bar{v}) \geq 2f(\bar{v})$.

Any deviation with $\lambda(M_s) = k/n$ and even $k \leq n-3$ is unprofitable if

$$\left(\frac{k-1}{n} \bar{v} - c \right) f \left(\frac{k-1}{n} \bar{v} \right) \geq \left(\frac{k}{n} \bar{v} - c \right) \sum_{i=k/2}^{\frac{n-1}{2}} \left(1 - \frac{k}{n} - \frac{2i+1}{n} + 2 \frac{c}{\bar{v}} \right) f \left(\frac{2i+1}{n} \bar{v} \right) \quad (9)$$

Inserting $f(\bar{v} \frac{n-2}{n}) > 2f(\bar{v})$ and $f(\bar{v} \frac{2i-1}{n}) > 3f(\bar{v} \frac{2i+1}{n})$ for all $i \in \{\frac{k}{2} + 1, \frac{k}{2} + 2, \dots, \frac{n-3}{2}\}$, the right-hand side of this inequality is smaller than

$$\left(\frac{k}{n} \bar{v} - c \right) f \left(\frac{k+1}{n} \bar{v} \right) \left(\left(\frac{1}{3} \right)^{\frac{n-3-k}{2}} \frac{1}{2} \left(2 \frac{c}{\bar{v}} - \frac{k}{n} \right) + \sum_{i=k/2}^{\frac{n-3}{2}} \left(\frac{1}{3} \right)^{i-\frac{k}{2}} \left(1 - \frac{k}{n} - \frac{2i+1}{n} + 2 \frac{c}{\bar{v}} \right) \right)$$

Hence, inequality (9) is satisfied if

$$\begin{aligned} f \left(\frac{k-1}{n} \bar{v} \right) / f \left(\frac{k+1}{n} \bar{v} \right) &\geq \\ &\geq \frac{\frac{k}{n} \bar{v} - c}{\frac{k-1}{n} \bar{v} - c} \left(\left(\frac{1}{3} \right)^{\frac{n-3-k}{2}} \frac{1}{2} \left(2 \frac{c}{\bar{v}} - \frac{k}{n} \right) + \sum_{i=k/2}^{\frac{n-3}{2}} \left(\frac{1}{3} \right)^{i-\frac{k}{2}} \left(1 - \frac{k}{n} - \frac{2i+1}{n} + 2 \frac{c}{\bar{v}} \right) \right). \end{aligned}$$

The right-hand side increases in $c \leq (k-2)\bar{v}/n$. Therefore such a deviation is unprofitable if the last inequality is satisfied for the largest $c < \lambda(M_s)\bar{v}$, namely $c = (k-2)\bar{v}/n$.

Then the right-hand side equals

$$\begin{aligned} &2 \left(\left(\frac{1}{3} \right)^{\frac{n-3-k}{2}} \frac{1}{2} \left(\frac{k-4}{n} \right) + \sum_{i=k/2}^{\frac{n-3}{2}} \left(\frac{1}{3} \right)^{i-\frac{k}{2}} \left(1 + \frac{k-4}{n} - \frac{2i+1}{n} \right) \right) \leq \\ &\leq \left(\frac{1}{3} \right)^{\frac{n-3-k}{2}} + 2 \sum_{i=k/2}^{\frac{n-3}{2}} \left(\frac{1}{3} \right)^{i-\frac{k}{2}} = \frac{1}{3^{\frac{n-3-k}{2}}} + 2 \sum_{i=0}^{\frac{n-k-3}{2}} \left(\frac{1}{3} \right)^i = \frac{1}{3^{\frac{n-3-k}{2}}} + 2 \frac{1 - \left(\frac{1}{3} \right)^{\frac{n-k-1}{2}}}{1 - \frac{1}{3}} = \end{aligned}$$

$$= \frac{1}{3^{\frac{n-3-k}{2}}} + 3\left(1 - \left(\frac{1}{3}\right)^{\frac{n-k-1}{2}}\right) = 3$$

because $k - 4 - (2i + 1) < 0$ for all $i \geq \frac{k}{2}$ and the common formula for a geometric series applies. Consequently, any deviation with $\lambda(M_s) = k/n$ and even $k \leq n - 3$ is unprofitable if $f(\bar{v}\frac{n-2}{n}) > 2f(\bar{v})$ and $f(\bar{v}\frac{2i-1}{n}) > 3f(\bar{v}\frac{2i+1}{n})$ for all $i \in \{1, 2, \dots, \frac{n-3}{2}\}$.

To sum up, the mechanism (3) is individually rational, ex-post efficient and strongly budget-balanced if the conditions in Theorem 2 are satisfied. \square

References

- Al-Najjar, N. I., Anderlini, L., and Felli, L. (2006). Undescribable Events. *Review of Economic Studies*, 73(4):849–868.
- Anderlini, L. and Felli, L. (1994). Incomplete Written Contracts: Undescribable States of Nature. *Quarterly Journal of Economics*, 109(4):1085–1124.
- Ben-Porath, E. and Lipman, B. L. (2012). Implementation with partial provability. *Journal of Economic Theory*, 147(5):1689–1724.
- Benkert, J.-M. (2017). Bilateral Trade with Loss-Averse Agents. *University of Zurich Working Paper*.
- Bull, J. and Watson, J. (2004). Evidence disclosure and verifiability. *Journal of Economic Theory*, 118(1):1–31.
- Chatterjee, K. and Samuelson, L. (1987). Bargaining with Two-sided Incomplete Information: An Infinite Horizon Model with Alternating Offers. *The Review of Economic Studies*, 54(2):175–192.
- Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, 11(1):17–33.
- Crawford, V. P. and Sobel, J. (1982). Strategic Information Transmission. *Econometrica*, 50(6):1431–1451.
- Cr mer, J. and McLean, R. P. (1988). Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions. *Econometrica*, 56(6):1247–1257.
- Deneckere, R. and Severinov, S. (2008). Mechanism design with partial state verifiability. *Games and Economic Behavior*, 64(2):487–513. Special Issue in Honor of Michael B. Maschler.
- Dziuda, W. (2011). Strategic Argumentation. *Journal of Economic Theory*, 146(4):1362–1397.
- Fishman, M. J. and Hagerty, K. M. (1990). The Optimal Amount of Discretion to Allow in Disclosure. *The Quarterly Journal of Economics*, 105(2):427–444.
- Fudenberg, D. and Tirole, J. (1983). Sequential Bargaining with Incomplete Information. *The Review of Economic Studies*, 50(2):221–247.

- Garratt, R. and Pycia, M. (2016). Efficient bilateral trade. *Working Paper*.
- Giovannoni, F. and Seidmann, D. J. (2007). Secrecy, two-sided bias and the value of evidence. *Games and Economic Behavior*, 59(2):296–315.
- Glazer, J. and Rubinstein, A. (2004). On Optimal Rules of Persuasion. *Econometrica*, 72(6):1715–1736.
- Glazer, J. and Rubinstein, A. (2006). A Study in the Pragmatics of Persuasion: a game theoretical Approach. *Theoretical Economics*, 1(4):395–410.
- Green, J. R. and Laffont, J.-J. (1986). Partially Verifiable Information and Mechanism Design. *Review of Economic Studies*, 53(3):447–456.
- Gresik, T. A. (1991). Efficient bilateral trade with statistically dependent beliefs. *Journal of Economic Theory*, 53(1):199–205.
- Groves, T. (1973). Incentives in teams. *Econometrica*, 41(4):617–631.
- Kartik, N. and Tercieux, O. (2012). Implementation with evidence. *Theoretical Economics*, 7(2):323–355.
- Kojima, F. and Yamashita, T. (2017). Double auction with interdependent values: Incentives and efficiency. *Theoretical Economics*, 12(3):1393–1438.
- Lang, M. (2019). Communicating Subjective Evaluations. *Journal of Economic Theory*, 179:163–199.
- Lipman, B. L. and Seppi, D. J. (1995). Robust Inference in Communication Games with Partial Provability. *Journal of Economic Theory*, 66(2):370–405.
- Mathis, J. (2008). Full revelation of information in Sender-Receiver games of persuasion. *Journal of Economic Theory*, 143(1):571–584.
- Matsuo, T. (1989). On incentive compatible, individually rational, and ex post efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 49(1):189–194.
- McAfee, R. P. and Reny, P. J. (1992). Correlated Information and Mechanism Design. *Econometrica*, 60(2):395–421.
- Milgrom, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics*, 12(2):380–391.
- Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281.
- Okuno-Fujiwara, M., Postlewaite, A., and Suzumura, K. (1990). Strategic Information Revelation. *The Review of Economic Studies*, 57(1):25–47.
- Schweizer, U. (1989). Litigation and Settlement under Two-Sided Incomplete Information. *The Review of Economic Studies*, 56(2):163–177.
- Seidmann, D. J. and Winter, E. (1997). Strategic Information Transmission with Verifiable Messages. *Econometrica*, 65(1):163–169.

- Sher, I. (2014). Persuasion and dynamic communication. *Theoretical Economics*, 9(1):99–136.
- Shin, H. S. (1994). The Burden of Proof in a Game of Persuasion. *Journal of Economic Theory*, 64(1):253–264.
- Sun, Y. (2006). The exact law of large numbers via Fubini extension and characterization of insurable risks. *Journal of Economic Theory*, 126(1):31–69.
- Sun, Y. and Zhang, Y. (2009). Individual risk and Lebesgue extension without aggregate uncertainty. *Journal of Economic Theory*, 144(1):432–443.
- Vickrey, W. (1961). Counterspeculation, Auctions, and Competitive Sealed Tenders. *The Journal of Finance*, 16(1):8–37.
- Wolitzky, A. (2016). Mechanism design with maxmin agents: Theory and an application to bilateral trade. *Theoretical Economics*, 11(3):971–1004.