

Lang, Matthias

Working Paper

Mechanism Design with Narratives

CESifo Working Paper, No. 8502

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Lang, Matthias (2020) : Mechanism Design with Narratives, CESifo Working Paper, No. 8502, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/223574>

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.

Mechanism Design with Narratives

Matthias Lang

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Mechanism Design with Narratives

Abstract

Individuals use narratives as rationales or justifications to make their claims more convincing. I provide a general framework for partial verifiability based on narratives. Narratives give many reasons and arguments. The receiver derives the message's meaning by aggregating these reasons; her private information tells her whether some potential reasons support the sender's claims. Therefore, the receiver detects misreports with positive probability. Narratives flexibly allow for different degrees of partial verifiability and allow using the revelation principle. Considering mechanism design as an example, I prove that narratives are so powerful to implement efficient trade in the canonical bilateral-trade setting.

JEL-Codes: D810, D820, D860.

Keywords: narrative, communication, partial verifiability, mechanism design, bilateral trade.

Matthias Lang
University of Munich (LMU)
Geschwister-Scholl-Platz 1
Germany – 80539 Munich
matthias.lang@econ.lmu.de

August 2020

This draft is made available to stimulate discussion and to gather feedback. All comments and suggestions are greatly appreciated. I thank Andreas Asseuer, Yves Breitmoser, Matthew Ellman, Agustin Faure, Joachim Henkel, Alida Johannsen, Yves Le Laouanq, Johannes Meier, Martin Pollrich, Roland Strausz, Gabriel Vollert, Jonas von Wangenheim, Leeat Yariv and the audiences at the CMID, ESEM, EARIE and the Japanese-German Workshop on Contracts and Incentives as well as the participants of the Micro Seminar at HU Berlin and at the 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.

1 Introduction

A lot of communication is between the polar cases of cheap talk and disclosing hard information. Start-ups present product ideas to venture capitalists or crowdfunders. Their products and services are still under development but adding project details and milestones makes projects more convincing. Companies supply earning figures to attract funding and to determine tax liabilities. Auditors and the SEC ask companies for many details, breakdowns and segmentations of revenues and costs to make earning figures reliable. Nonetheless, earnings still leave much flexibility to companies, not to mention, widely used, adjusted earnings. Companies talk to regulators about their costs, market power or consumer data to avoid regulation. Regulators require many details of market interactions. Most organizations interview job candidates to learn about their skills. Candidates tell how they acquired skills instead of just listing their skills. All these senders provide many reasons, arguments, points, details, segmentations and breakdowns to make and support their claims. I capture this kind of communication between the extremes of cheap talk and disclosing hard information by offering a general and flexible framework for partial verifiability; a goal the literature has pursued at least since Green and Laffont (1986).¹ I call the communication, that senders in above examples use to make their claims more convincing, narratives. This interpretation of “narratives as rationales or justifications” follows previous literature, like Bénabou et al. (2019, p. 1).² I then apply this framework to mechanism design. Demonstrating their possibilities, narratives implement efficient trade in the canonical bilateral-trade setting in contrast to famous impossibility results. Moreover, the framework satisfies the revelation principle making analyses very tractable. In particular, classical tools and methods of mechanism design are applicable to determine optimal contracts and study implementability.

It is helpful to contrast narratives with traditional modeling of communication. In traditional mechanism design, communication is brief and concise. Building on ideas in Holmström (1979) and Kim (1995), communication usually focuses on sufficient statistics of payoff-relevant types, like valuations or costs. Narratives in my framework are longer and more elaborate messages that spread out the payoff-relevant information in many pieces. Instead of making direct statements, e.g., firms do not tell regulators that ‘costs are high’, narratives provide many reasons, arguments and points to make and support the sender’s claim. The receiver derives the sender’s statement from those

¹See Glazer and Rubinstein (2006) and Carroll and Egorov (2019) for more recent contributions and the related literature section for detailed discussions.

²In contrast to Eliaz and Spiegler (2019), Bénabou et al. (2019) and I consider fully rational agents. Shiller (2017) considers a different interpretation of narratives as ‘a gem for conversation.’

reasons. The more the sender wants to change her statement, the more reasons she must change. The receiver knows whether some potential reasons are appropriate and support the sender's statement. If the sender misreports, the receiver detects deviations with some probability proportional to the deviation's size. Thus, narratives make the sender's claims more convincing because the receiver might notice misreporting. Senders trade off gains from misreporting and probabilities of detection. Senders can send any messages as they face no communication costs and no restrictions on messages spaces, messages are unverifiable and only the equilibrium establishes the messages' meaning. Therefore, the revelation principle is valid in my framework. Receivers, however, can use their information to verify senders' narratives partially.

I apply narratives to the canonical setting of bilateral trade by Myerson and Satterthwaite (1983) addressing a major problem of allocative efficiency: A seller and a buyer bargain over the terms of trade. The buyer's valuation for the good and the seller's costs are private information. The seller knows all inputs required for production. Each input has the same costs. The seller could directly tell the buyer her costs, i.e., the number of necessary inputs. This form of cheap talk is how communication is traditionally modeled. It cannot achieve all possible gains from trade. Thus, valuable opportunities for trade are lost. In a narrative, instead, the seller tells the buyer all necessary inputs to justify her costs. Her message is not necessarily truthful. By providing this additional information, she makes herself vulnerable to scrutiny. The buyer only knows the distribution of inputs. By reverse engineering, reading product reviews or inspecting the product, the buyer learns for a few inputs whether those inputs are necessary. This sample of the buyer is uninformative about the seller's costs: the seller's costs and the buyer's information are stochastically independent. The buyer reports the observed inputs in an unverifiable message. To show the possibilities of narratives, I limit the buyer's information as far as possible. Even in that limit and when observing only one input out of a continuum in his sample, the buyer can verify the seller's narrative to some extent. The buyer's information about his sample and the seller's narrative as a more elaborate message ensure that the buyer detects any misreports with positive probability –mirroring the communication in above examples. Moreover, the probability of detection is proportional to the gains of misreporting ensuring that mechanisms can deter small *and* large deviations with common contractual clauses. Narratives contain a lot of 'cheap talk' for all the inputs not contained in the buyer's sample but messages are still informative along this dimension. This can explain why we see so much exchange of unverifiable arguments in real life - which is difficult to reconcile with standard cheap talk. Narratives, thus, enable ex-post efficient trade in

contrast to previous impossibility results that the literature discusses extensively and that are part of the core curriculum in economics. Narratives work although the seller's costs and the buyer's information are stochastically independent and, hence, uncorrelated. Thus, narratives mitigate widely documented frictions caused by asymmetric information.

To demonstrate the flexibility of my framework, I introduce the strength of narratives. The narratives' strength – a measure of the support of the distribution of the buyer's sample – varies the efficiency of second-best contracts. This strength could relate to common cultural backgrounds, length and depth of contractual relationships between sender and receiver, the familiarity of the buyer with the product category or the familiarity of the seller with buyers' tastes. Stronger narratives allow the receiver to detect more deviations making incentive compatibility easier to satisfy and second-best contracts more efficient. When narratives are sufficiently strong, optimal contracts approximate first-best welfare arbitrarily close. These different degrees of partial verifiability make my framework very flexible.

This framework for partial verifiability based on narratives is general and not limited to bilateral trade. Narratives apply to all economic environments, including mechanism design, taxation, market regulation, finance, auditing and many other settings whenever contracting parties interact and communicate. Narratives make messages partially verifiable. I do *not* assume correlated valuations, exogenous verification technologies, type-dependent message spaces, or messages that are verifiable by third parties in contrast to other approaches for partial verifiability which I discuss next.

After reviewing the related literature, this paper is organized as follows. Section 2 considers bilateral trade. Section 3 conveys the basic intuition of narratives. Section 4 proves the power of narratives – ensuring efficiency in bilateral trade. Section 5 introduces the strength of narratives and calculates second-best contracts. Section 6 concludes and discusses more familiar indirect implementations. Most formal proofs are relegated to Appendix A.

Related Literature

The literature offers different approaches concerning partial verifiability. First, agents disclose hard information that does not fully reveal senders' types. For example, in Shin (1994), senders present hard information. Messages indicate whether the type is above or below privately-known thresholds. Dziuda (2011) considers more general disclosure but only one sender. Again full revelation is impossible. Alternatively, Fishman and Hagerty (1990) assume that senders learn finitely many signals about types. The

sender can disclose only one of these signals. They determine the optimal discretion for senders' choices. Schweighofer-Kodritsch and Strausz (2020) discuss the game-theoretic problems created by disclosure and why the revelation principle is usually violated in these settings – I will return to this problem throughout this paper.

Second, agents' message spaces depend on their types (Milgrom, 1981, Green and Laffont, 1986). This approach is more general than the first one. Green and Laffont (1986) focus on direct mechanisms allowing for any dependency of message spaces on types. Deneckere and Severinov (2008) show that sequential and password mechanisms are superior to direct mechanisms – violating the revelation principle. Okuno-Fujiwara et al. (1990) provide conditions on a game's payoffs for full revelation. Lipman and Seppi (1995) provide conditions on how message spaces depend on types for full revelation. Glazer and Rubinstein (2006) and Sher (2014) characterize optimal reporting strategies. Hart et al. (2017) show that commitment is irrelevant in certain classes of evidence games. Seidmann and Winter (1997) and Mathis (2008) provide conditions on how message spaces depend on types for full revelation in certain games. Correspondingly, Giovannoni and Seidmann (2007) provide conditions on preferences for full revelation. Lang (2019) considers moral hazard with partial verifiability. Hagenbach et al. (2014) study pre-play communication. Finally, Bull and Watson (2004), Ben-Porath and Lipman (2012), and Kartik and Tercieux (2012) characterize implementable social choice functions for complete information.

Third, senders' types are multidimensional and receivers can verify one dimension of senders' types. Payoffs are exogenous. With binary actions, Glazer and Rubinstein (2004) show that finding optimal mechanisms is equivalent to solving linear programming problems. Carroll and Egorov (2019) consider finite multidimensional type spaces and monotone payoffs. They prove that full revelation is possible for submodular payoffs but impossible for (strictly) supermodular payoffs.

Fourth, the contracting language is restricted, e.g., some events are ex-ante impossible to describe. In Al-Najjar et al. (2006) and Anderlini and Felli (1994), these restrictions make incomplete contracts optimal. In Jakobsen (2020), the principal can write complex contracts that agents do not fully understand.

Fifth, Caragiannis et al. (2012) and Ball and Kattwinkel (2019) consider probabilistic verification. Receivers have an exogenous verification technology to audit messages. These audits depend on agents' types. Narratives can be seen as a microfoundation of probabilistic verification – not relying on exogenous verification technologies.

Finally, I relate to Myerson and Satterthwaite (1983). They show that efficient trade is impossible without narratives. Recently, a couple of papers re-examine this

impossibility for different preferences. Ambiguity aversion (Wolitzky, 2016) and concave utilities (Garratt and Pycia, 2016) can sometimes enable efficient trade. Benkert (2017) proves impossibility for loss-averse agents.

2 Trade Model

Begin with the canonical model of bilateral trade. A seller (she) has an object to sell to a single buyer (he). The buyer's valuation, v , is distributed according to a distribution F on $[0, \nu)$ with $\nu > 0$. The buyer's valuation is his private information. The seller's costs c , are distributed independently according to a distribution G on $[0, \gamma)$ with $\gamma > 0$. The seller's costs are her private information. Both distributions have continuous and positive densities. If trade occurs at a price p , the seller's utility is $p - c$ and the buyer's utility is $v - p$; reservation utilities are zero. Both the seller and the buyer maximize expected payoffs. I also assume, in line with mechanism design traditions, that the designer can choose an equilibrium of the ensuing game. The following properties are common objectives in mechanism design:

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

Definition (Individual Rationality). A mechanism is individually rational if both seller and buyer participate voluntarily given their types.

Finally, the mechanism should be budget-balanced, so that it requires no outside subsidies.³ I return to the issue of budget balance in the conclusion.

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

It is impossible to find a mechanism for bilateral trade with these three properties.

Proposition 1. *[Myerson and Satterthwaite (1983)] No mechanism is ex-post efficient, individually rational, and budget-balanced.*

Proof: See Myerson and Satterthwaite (1983, Corollary 1). □

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

³Outside subsidies can implement efficient trade as demonstrated by Vickrey (1961), Clarke (1971), and Groves (1973).

3 Narratives: Basic Intuition

Assume that production requires up to $I \geq 2$ potential inputs. Denote $c_i = 1$ if input i is necessary and $c_i = 0$ if input i is unnecessary. Each input has the same costs. The seller knows all necessary inputs and her costs are $c = \sum_{i=1}^I c_i$. The buyer only knows the distribution of inputs: $\text{Prob}(c_i = 1) \in (0, 1)$ and $\text{Corr}(c_i, c_j) \in (-1, 1)$ for $i \neq j$. By inspection of the good, reverse engineering or reading some reviews, the buyer learns for one input i^* whether that input is necessary. That input i^* is drawn with positive probabilities $h(\cdot)$ from $\{1, 2, \dots, I\}$. Denote $h_{\min} = \min_{i \in \{1, \dots, I\}} h(i)$. The buyer's valuation v is drawn from a distribution $F(v)$ with finite or continuous support on \mathbb{R}_0^+ . To summarize, the seller's private information is the vector (c_1, \dots, c_I) while the buyer's private information is (v, i^*, c_{i^*}) . Payoffs are the same as in the trade model of Section 2. For ease of exposition, I study an asymmetric setting with only sellers using narratives about their costs. The cases of only buyers using narratives about their valuations or both sides using narratives are analogous, however. If buyers use narratives, their narratives contain, for example, all the features they appreciate instead of their overall valuation. Appendix B considers this symmetric setting.

After learning their information, seller and buyer report their information if they agree to participate in the mechanism. The mechanism then implements trade and payments. A mechanism specifies probabilities of trade t and payments conditional on the buyer's message \mathcal{M}_b and the seller's message \mathcal{M}_s . In addition to the buyer paying a price p to the seller, the seller pays π to a third party. I return to these payments in the conclusion and discuss common interpretations.

Proposition 2. *The following mechanism is ex-post efficient, individually rational, and budget-balanced: The buyer reports $\mathcal{M}_b = (m_b, M_b) \in \{0, 1, \dots, I\} \times \text{support of } F$. The seller reports⁴*

$$M_s \in 2^{\{1, 2, \dots, I\}}.$$

Payments and probabilities of trade are

$$\begin{cases} t = 0, & p = 0, & \pi = 0 & \text{if } M_b < |M_s| \\ t = 1, & p = |M_s|, & \pi = 0 & \text{if } M_b \geq |M_s| \text{ and } m_b \notin M_s \\ t = 1, & p = |M_s|, & \pi = 1/h_{\min} & \text{otherwise.} \end{cases}$$

Intuitively, the mechanism implements no trade if the number of reported necessary

⁴For any finite set M , let 2^M denote the power set (the set of all subsets of M) and let $|M|$ denote the cardinality of M .

inputs is above the buyer's reported valuation. Otherwise, trade takes places and the buyer pays a price equal to the number of reported necessary inputs. If the buyer learns about an unnecessary input and the seller includes this input in her narrative, the seller has to make an additional payment $1/h_{min}$. To see that the last case only occurs out of equilibrium, consider equilibrium strategies. In equilibrium, the seller reports the necessary inputs

$$M_s = \{i \in \{1, 2, \dots, I\} | c_i = 1\}.$$

Vice versa, the buyer, in equilibrium, reports his valuation $M_b = v$ and his sample of the seller's inputs by stating the name of the observed input if that input is unnecessary. In particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = 0$ otherwise. These strategies ensure that $m_b \notin M_s$. Hence, the mechanism is ex-post efficient in equilibrium. It is obvious that the mechanism is budget-balanced because it requires no outside subsidies.

To study the buyer's incentives, suppose the seller follows her strategy. In equilibrium, the buyer expects nonnegative payoffs making the mechanism individually rational for the buyer. The buyer receives the entire surplus in equilibrium so that he cannot gain by deviating: A deviation with $M_b > v$ lowers his utilities by making him buy the object at a price above his valuation. A deviation with $v > M_b$ lowers the buyer's utilities by not being able to buy the object at some prices below his valuation. Any other deviation by the buyer does not change his payoffs. Therefore, his strategy is optimal for the buyer.

Turning to the seller, suppose the buyer follows his strategy. In equilibrium, the seller expects payoffs

$$\sum_{i=c}^I [c - c] f(i) = 0$$

making the mechanism individually rational for the seller. A deviation with $|M_s| \leq c$ at least weakly lowers the price and, hence, cannot yield positive profits. A deviation with $|M_s| > c$ has three effects: the price increases, probabilities of trade decrease, and additional payments $\pi = 1/h_{min}$ could be triggered. The first effect benefits the seller while the second and third effect hurt her. In particular, expected payoffs of such a deviation are below

$$\sum_{i=|M_s|}^I \left[|M_s| - c - \text{Prob}(m_b \in M_s) \frac{1}{h_{min}} \right] f(i) \leq (|M_s| - c)[1 - 1] \sum_{i=|M_s|}^I f(i) = 0.$$

These deviation payoffs are nonpositive, because $\text{Prob}(m_b \in M_s) \geq h_{min}(|M_s| - c)$ as the probability that a given input is observed by the buyer is at least h_{min} and the seller

has to include at least $|M_s| - c$ unnecessary inputs in her report for such a deviation. Consequently, any deviation in M_s is unprofitable and her strategy is optimal for the seller.

To attain efficiency, all the surplus has to go to the buyer in this toy model. In general, narratives do not require this. In symmetric models, any split of the surplus between seller and buyer is possible. See Appendix B for an example. As the buyer's information is correlated with the seller's costs, the buyer updates his beliefs about the seller'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 narratives work differently. To emphasize these differences and formalize them, the main model has no such correlation and I will prove that the spanning conditions are violated. Indeed, the buyer's information and the seller's costs will be stochastically independent emphasizing the power of narratives. I now turn to the general model.

4 First-Best by Narratives

The buyer's valuation $v \in [0, \nu)$ and the seller's costs $c \in [0, \gamma)$ are drawn from distributions $F(v)$ and $G(c)$ with continuous and positive densities. Instead of I inputs, production requires potential inputs $i \in [0, \gamma)$. Each input i is necessary with probability c/γ . The inputs are essentially pairwise independent as defined by Sun (2006, Definition 2.7). For most readers, it will be convenient to neglect these technical details and think about identically and independently distributed inputs.⁶

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

$$\int c_i d\lambda(i) = c \quad \text{almost surely}$$

with an extension λ of the Lebesgue measure on $[0, 1]$ as introduced by Sun and Zhang (2009, Theorem 1).

Again the buyer learns for one input i^* whether it is necessary. That input i^* is drawn uniformly. Therefore, the buyer cannot infer anything about the seller's costs c from his sample i^* because the buyer's information is stochastically independent and

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

⁶See also the discussion about independence on a continuum in Ali and Bénabou (2020, Footnote 18).

uncorrelated with the seller's costs c . Nevertheless, narratives make efficient trade feasible.

Theorem 1. *The following mechanism is ex-post efficient, individually rational, and budget-balanced: The buyer reports*

$$(m_b, M_b) \in [0, \gamma] \times [0, \nu).$$

The seller reports

$$M_s \in \{X \subseteq [0, \gamma) | X \text{ is } \lambda\text{-measurable}\}.$$

*Payments and probabilities of trade are*⁷

$$\begin{cases} t = 0, & p = 0, & \pi = 0 & \text{if } M_b < \lambda(M_s) \\ t = 1, & p = \lambda(M_s), & \pi = 0 & \text{if } M_b \geq \lambda(M_s) \text{ and } m_b \notin M_s \\ t = 1, & p = \lambda(M_s), & \pi = \gamma & \text{otherwise.} \end{cases} \quad (1)$$

In equilibrium, the seller's narrative describes the necessary inputs

$$M_s = \{i \in [0, \gamma) | c_i = 1\};$$

the buyer reports his valuation $M_b = v$ and his sample i^* of the seller's inputs by stating the name of the observed input if that input is unnecessary. In particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = \gamma$ otherwise.

Trade and payments depend on the buyer's reported valuation M_b , the seller's costs as inferred from her narrative $\lambda(M_s)$, and whether the seller's narrative is consistent with the buyer's report. The seller's narrative is consistent with the buyer's report if the buyer reports a sample $m_b = i^*$ that is not contained in the necessary inputs as reported by the seller. The seller's narrative is also consistent with the buyer's report if the buyer reports to have observed a necessary input, $m_b = \gamma$. These strategies ensure that the seller's narrative is consistent with the buyer's report in equilibrium – making the mechanism ex-post efficient. The mechanism is also budget-balanced as I only consider payments made by the seller and the buyer. In the next steps, I investigate individual rationality and incentive compatibility.

To study the seller's incentives, suppose the buyer follows his strategy. In equilib-

⁷Remember that t is the probability of trade, the buyer pays the price p to the seller, and the seller pays π to a third party.

rium, the seller expects payoffs

$$\text{Prob}(\text{trade})(\mathbb{E}(p|\text{trade}) - c) = (1 - F(c))(c - c) = 0$$

making the mechanism individually rational for the seller. Turning to incentive compatibility, consider deviations by the seller. A deviation with $\lambda(M_s) \leq c$ makes the seller worse off by reducing the price. A deviation with $\lambda(M_s) > c$ has three effects: the price increases, probabilities of trade decrease, and additional payments π could be triggered. The first effect benefits the seller while the second and third effect hurt her. In particular, payoffs of such a deviation are below

$$\begin{aligned} & \left(1 - F(\lambda(M_s))\right) [\lambda(M_s) - c - \text{Prob}(m_b \in M_s)\gamma] \\ & \leq \left(1 - F(\lambda(M_s))\right) \left[\lambda(M_s) - c - \frac{1}{\gamma}(\lambda(M_s) - c)\gamma\right] = 0. \end{aligned}$$

As the seller's narrative must contain at least a mass of $\lambda(M_s) - c$ unnecessary inputs, the probability that the buyer observed one of these inputs is at least $(\lambda(M_s) - c)/\gamma$. Hence, $\text{Prob}(m_b \in M_s) \geq (\lambda(M_s) - c)/\gamma$ and such a deviation is unprofitable. Therefore, any deviation is unprofitable. Consequently, her strategy is optimal for the seller. As the buyer gets the entire surplus, the mechanism is also incentive compatible for him. I postpone the remainder of the proof to Appendix A.

This section limits the information of the counterpart to a minimum showing the possibilities of narratives. Indeed, the buyer's information has mass zero and is stochastically independent from the seller's costs. Narratives nevertheless allow the buyer to cross-check the seller's reported costs learning with some probability if the reported costs are incorrect. This kind of partial verifiability allows the mechanism designer to attain first best in many settings. Narratives differ from the ideas of McAfee and Reny (1992) and Crémer and McLean (1988). Crémer and McLean (1988) consider only finite type spaces. Hence, I focus on McAfee and Reny (1992). As the buyer's valuation and the seller's costs are independently distributed the difference seems suggestive. In addition, I implement efficient trade with bounded payments while McAfee and Reny (1992) can only approximate first best and require arbitrarily large payments the closer they approximate first best. Finally, I prove that my setting violates their spanning condition.

Lemma 2. *The setting violates the spanning condition of McAfee and Reny (1992, Theorem 2).*

Lemma 2 emphasizes the differences between the narratives proposed here and ear-

lier ideas in the literature to use correlated information shedding new light on the results of McAfee and Reny (1992) and Crémer and McLean (1988). Frequently, we are interested in partial verifiability that does not immediately yield first-best. Narratives allow doing so by restricting the receiver's sample.

5 The Strength of Narratives

The narratives' strength $\delta \in \mathbb{N}$ measures the familiarity of the buyer with the product or the depth of the relationship between buyer and seller. If a big car producer procures some well-known components, a high δ shows her ability to comprehend and cross-check the supplier's narratives. If a little experienced consumer buys some antique furniture or painting, a low δ shows her limited ability to cross-check a seller's narratives. Vice versa, if the buyer tells the narratives, the higher δ , the better the seller understands the value-generating aspects for the buyer. Thus, longer relationships between banks and borrowers or the availability of big data increase δ – making narratives more powerful and strengthening the seller's hand even if the additional information or knowledge does not affect the seller's prior about the buyer's valuation. Therefore, I call δ the narratives' strength.

For ease of exposition, I describe one of many possible specifications instead of aiming for maximum generality. Consider the same model as in Section 4 with these changes: To simplify the exposition, I set $\nu = \gamma = 1$. The necessary inputs are randomly located on a circle with circumference one in the following way: A starting point σ is drawn from a distribution with full support on a grid $\mathcal{G} = \{0, 1/\delta, 2/\delta, \dots, (\delta - 1)/\delta\}$. Beginning from σ , the next inputs with mass c are necessary. See Figure 1 for a graphical depiction. Both c and σ are private information of the seller. The buyer only knows their distribution. The buyer learns for one input i^* whether it is necessary. That input is drawn uniformly from \mathcal{G} . The buyer's sample is private information of the buyer. The seller only knows the distribution and, hence, its support. Again, the buyer cannot infer anything about the seller's costs c from his sample i^* . Indeed the buyer's information is stochastically independent and uncorrelated with the seller's costs c .

The seller's knowledge about the grid \mathcal{G} allows her to understate her costs to some extent without the buyer being able to detect such deviations. Therefore, it is impossible to achieve ex-post efficiency. Narratives still make simple mechanisms surprisingly

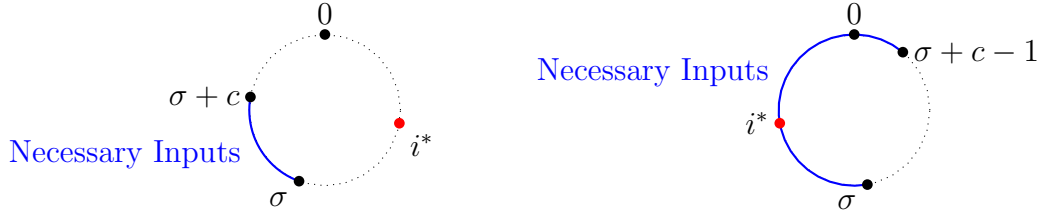


Figure 1: Two Examples of Necessary Inputs $[\sigma, \sigma + c]$ and Buyer's Sample i^*

efficient. For this purpose, define an agreement function

$$A(m_b, \tilde{\sigma}, \tilde{c}) = \begin{cases} 1 & \text{if } 0 > m_b - \tilde{\sigma} > \tilde{c} - 1 \text{ or } m_b > \tilde{c} + \tilde{\sigma} \text{ or } m_b = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Correspondingly, messages agree, $A = 1$, if the buyer reports to have observed a necessary input, $m_b = 1$. Messages also agree if the buyer reports a sample $m_b = i^*$ that is outside the set of necessary inputs as reported by the seller. Finally, denote $\mathbb{N}^k = \{1, 2, \dots, k\}$.

Proposition 3. *The following mechanism is budget-balanced, individually rational, and incentivizes buyer and seller to report truthfully as defined below. The buyer reports*

$$\mathcal{M}_b = (m_b, M_b) \in \{0, 1/\delta, 2/\delta, \dots, (\delta - 1)/\delta, 1\} \times [0, 1)$$

and the seller reports

$$\mathcal{M}_s = (m_s, M_s) \in \mathcal{G} \times [0, 1).$$

Probabilities of trade and payments are

$$\begin{cases} t = 0, & p = 0, & \pi = 0 & \text{if } M_b < p^*(M_s) \\ t = 1, & p = p^*(M_s), & \pi = 0 & \text{if } M_b \geq p^*(M_s) \text{ and } A(m_b, m_s, M_s) = 1 \\ t = 1, & p = p^*(M_s), & \pi = 1 & \text{otherwise} \end{cases} \quad (3)$$

with a price

$$p^*(c) = \frac{k}{\delta}$$

if $c \in [(k - 1)/\delta, k/\delta)$ for a $k \in \mathbb{N}^\delta$.

The seller, in equilibrium, reports the necessary inputs by sending messages $m_s = \sigma$ and $M_s = c$. The buyer reports his valuation $M_b = v$ and his sample i^* of the seller's inputs by stating the name of the observed input if that input is unnecessary. In

particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = 1$ otherwise. These strategies ensure that $A(m_b, m_s, M_s) = 1$. The mechanism is budget-balanced. Next, I investigate individual rationality and incentive compatibility.

To study the seller's incentives, suppose the buyer follows her strategy. In equilibrium, the seller expects payoffs

$$\text{Prob}(\text{trade})(\mathbb{E}(p|\text{trade}) - c) = (1 - F(p^*(c)))(p^*(c) - c) > 0$$

making the mechanism individually rational for the seller. A deviation with $M_s \leq c$ could make the seller worse off by trading at prices below costs and/or by triggering additional payments π . A deviation with $M_s > c$, $p^*(M_s) = p^*(c)$ and $m_s = \sigma$ does not change the seller's payoffs. A deviation with $M_s > c$ and $(p^*(M_s) \neq p^*(c) \text{ or } m_s \neq \sigma)$ has three effects: the price increases, probabilities of trade decrease, and additional payments π could be triggered. The first effect benefits the seller while the second and third effect hurt her. By deviating and claiming more necessary inputs, the price $p^*(M_s)$ increases by $1/\delta$ for each grid point that the seller additionally includes in her report $[m_s, m_s + M_s]$. At the same time, she has to make an additional payment of 1 with probability $1/\delta$ for each grid point that she additionally includes in her report. Hence, such a deviation is unprofitable. Therefore, any deviation in M_s and m_s is unprofitable. Consequently, her strategy is optimal for the seller. I postpone the remainder of the proof to Appendix A.

The mechanism (3) does not achieve efficiency. If narratives become stronger, however, welfare increases and converges to first-best.

Lemma 3. *Welfare generated by the mechanism in Proposition 3 converges to first best for sufficiently strong narratives, $\delta \rightarrow \infty$. In particular, for sufficiently strong narratives, this mechanism attains higher efficiency than any mechanism that does not use narratives.*

A natural follow-up question regards the characterization of a second-best mechanism in this setting. Previous models of partial verifiability often made it difficult to find second-best mechanisms as the revelation principle fails. For instance, Green and Laffont (1986) provide this example: An agent has one of three types, $\theta \in \{\theta_1, \theta_2, \theta_3\}$. Type θ_1 can report θ_1 or θ_2 . Type θ_2 can report θ_2 or θ_3 . Type θ_3 can report θ_3 . The principal takes a binary action, e.g., whether to procure a good from the agent. The agent prefers procurement independently of her type. The principal wants to procure the good from types θ_2 or θ_3 but not from type θ_1 . This outcome is implementable by the allocation rule: Procure the good if and only if the agent reports θ_3 . Then types θ_2

and θ_3 report θ_3 . It does not matter whether type θ_1 reports θ_1 or θ_2 because she will never get the procurement contract. This outcome function cannot be implemented truthfully, though. Whenever type θ_2 receives the contract upon reporting θ_2 , type θ_1 can ensure that she also gets the procurement contract by reporting θ_2 . Hence, the revelation principle fails in the framework of Green and Laffont (1986). Using narratives, messages are partially verifiable *and* the revelation principle is valid.

Proposition 4. *Any implementable outcome function can be implemented truthfully. Hence, the revelation principle is valid.*

Therefore, I can focus on direct mechanisms to find a second-best mechanism. It turns out that, as in the classical setting without narratives, additional assumptions on the distributions are necessary:

$$c + \frac{G(c) - G((k-1)/\delta)}{g(c)} \text{ increases in } c \in \left[\frac{k-1}{\delta}, \frac{k}{\delta} \right) \text{ for all } k \in \mathbb{N}^\delta, \quad (4)$$

$$\text{and } v - \frac{1 - F(v)}{f(v)} \text{ increases in } v. \quad (5)$$

For $\delta = 1$, these assumptions about increasing virtual values are common in mechanism design. See, e.g., Myerson (1981). To ease notation, introduce some definitions. Define $Z(c) = (k-1)/\delta$ if $c \in [\frac{k-1}{\delta}, \frac{k}{\delta})$ for a $k \in \mathbb{N}^\delta$. This function tells us which interval $[(k-1)/\delta, k/\delta)$ contains the value c . Furthermore, define a trading rule

$$Y(\tilde{c}, \tilde{v}, \alpha) = \begin{cases} 1 & \text{if } \tilde{c} + \alpha \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})} \leq \tilde{v} - \alpha \frac{1 - F(\tilde{v})}{f(\tilde{v})} \\ 0 & \text{otherwise} \end{cases}$$

and the value $\alpha \in (0, 1]$ so that

$$\int_0^1 \int_0^1 \left(\tilde{v} - \frac{1 - F(\tilde{v})}{f(\tilde{v})} - \tilde{c} - \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})} \right) Y(\tilde{c}, \tilde{v}, \alpha) dG(\tilde{c}) dF(\tilde{v}) = 0.$$

The value α governs the efficiency of the second-best mechanism. Lower α makes the second-best mechanism more efficient. Finally, define the values

$$\mathcal{T}_s(c) = \int_0^1 Y(c, \tilde{v}, \alpha) dF(\tilde{v}) \text{ and } \mathcal{T}_b(v) = \int_0^1 Y(\tilde{c}, v, \alpha) dG(\tilde{c})$$

for all $c, v \in [0, 1]$. The value $\mathcal{T}_b(M_b)$ equals the expected probability of trade from the buyer's point of view if he reports M_b . The value $\mathcal{T}_s(M_s)$ equals the expected probability of trade from the seller's point of view if she reports M_s .

Theorem 2. *If the distributions F and G satisfy conditions (4) and (5), the following mechanism is optimal: The buyer reports*

$$\mathcal{M}_b = (m_b, M_b) \in \{0, 1/\delta, 2/\delta, \dots, (\delta - 1)/\delta, 1\} \times [0, 1)$$

and the seller reports

$$\mathcal{M}_s = (m_s, M_s) \in \mathcal{G} \times [0, 1).$$

Probabilities of trade are $t = Y(M_s, M_b, \alpha)$. Payments equal $\pi = 2(1 - A(m_b, m_s, M_s))$ and

$$\begin{aligned} p = p^*(M_s, M_b) = & \mathcal{T}_s(M_s)M_s + \int_{M_s}^{Z(M_s)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} + \mathcal{T}_b(M_b)M_b - \int_0^{M_b} \mathcal{T}_b(\tilde{v})d\tilde{v} \\ & - \int_0^1 \left(\mathcal{T}_s(c')c' + \int_{c'}^{Z(c')+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} \right) dG(c') \end{aligned}$$

with the agreement function $A(m_b, m_s, M_s)$ defined in (2).

Narratives limit the seller's scope for deviations. Small deviations, however, are undetectable. Therefore, classical tools of mechanism design are useful in determining the solution in contrast to previous partial-verifiability settings like Green and Laffont (1986). It is impossible to attain efficiency in this setting — reinforcing the earlier result that the setting violates the spanning conditions of McAfee and Reny (1992) and Crémer and McLean (1988). Accordingly, there is no trade if the buyer's valuation is only marginally above the seller's costs. Narratives decrease this inefficiency and increase welfare. Thus, narratives should be used from a welfare perspective and communication should not be limited to summary statistics, like c and v in this example. Stronger narratives increase welfare. In particular, similarly to Lemma 3, welfare can approach ex-post efficiency arbitrarily close for sufficiently strong narratives.

6 Conclusions

I consider narratives as rationales or justifications following previous literature, e.g. Bénabou et al. (2019). I develop a general framework for partial verifiability based on narratives. Narratives enrich messages to make messages partially verifiable. Narratives provide a number of reasons, arguments and points to make the sender's claim instead of just stating the claim. The receiver derives the sender's claim from these reasons. Therefore, misreporting requires changing several arguments. Indeed, the larger the intended change, the more arguments the sender has to adjust. The receiver

knows whether some potential reasons are appropriate and support the sender's claim. Hence, if the sender deviates from equilibrium the receiver notices such a deviation with positive probability. Narratives work even if the receiver knows only one reason out of a continuum and cannot infer anything about the sender's valuation from the reasons she knows. I apply narratives to mechanism design. As an example, I consider the canonical setting of bilateral trade. I show that narratives allow for ex-post efficient trade and ensure individual rationality as well as budget balance. Consequently, narratives are so powerful to overturn the famous impossibility result by Myerson and Satterthwaite (1983). In addition, I demonstrate how to adjust the strength of narratives. Thus, narratives offer a powerful, general, and flexible framework to work with partial verifiability. With narratives, the classical revelation principle is still valid so that classical tools can be used to determine optimal contracts and study feasibility as well as implementability.

I consider direct mechanisms which are very tractable for analysis, but sometimes too abstract for direct applications. Indirect implementation makes contracts more familiar. Consider as an example the setting of Section 4: Buyer and seller participate in a contract with dispute procedures. First, the seller makes an offer M_s that contains her narrative and all necessary inputs reflecting the complexity of the object. Observing this offer, the buyer decides whether to buy the good at a price, $p = \lambda(M_s)$. The buyer has the possibility to dispute the seller's claim. If he disputes the seller's claim or claims loss of trust, the seller has to sell the good to the buyer for a price, $p = \lambda(M_s)$. In addition, the seller has to pay the costs of arbitration $\pi = \gamma$.⁸ It is easy to see that this contract is pay-off equivalent to the mechanism in Theorem 1. Consequently, this contract with dispute procedures is individually rational, ex-post efficient and budget-balanced.

Such dispute procedures can be implemented internally, by a mediator or by the legal system. These dispute procedures are quite common. See Fenn et al. (1997) for examples in the construction industry and the chemical process industry. Notice that this indirect implementation is also feasible if disputing a claim is costly by adjusting the contract to reimburse these costs. For this purpose, some of the payments by the seller are paid to the buyer to reimburse his costs for disputing the seller's claim. If the costs for disputing the seller's claim are not prohibitively high, this reimbursement is possible by diverting the seller's payments.

The mechanisms considered so far do not require outside subsidies and have no payments to third parties on the equilibrium path. The mechanisms require payments

⁸The buyer disputes the seller's claim if he wants to trade, i.e., $v \geq \lambda(M_s)$, $c_{i^*} = 0$, and $i^* \in M_s$.

to third parties off the equilibrium path, however. Many results in the literature, e.g., Wolitzky (2016), 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.” In the case of discrete type spaces and symmetric settings, it is possible to find conditions on the distributions of costs and valuations that ensure strong budget balance, i.e., no payments to third parties.

In line with the canonical bilateral-trade setting, I assume simultaneous reporting of all messages. The mechanisms of Theorems 1 and 2 as well as Propositions 2 and 3, however, are robust to different timings. In particular, sequential communication is possible as long as narratives are told first or other messages are not made public before the sender tells her narratives. Thus, in bilateral trade, the seller must not observe the buyer’s message m_b before telling the narrative about her costs (and vice versa for the buyer in symmetric settings).

A Appendix

Proof of Proposition 2: The proposition and the following discussion state a budget-balanced mechanism and corresponding strategies. These strategies guarantee that $M_b = v$ and $|M_s| = c$ such that there is trade if and only if $v \geq c$. Therefore, the mechanism is ex-post efficient. I show that these strategies are optimal for buyer and seller. Both participate voluntarily because the mechanism is individually rational for them. To sum up, the mechanism is ex-post efficient, individually rational, and budget-balanced. \square

Proof of Lemma 1: A probability space for the random variable c is constructed in the usual way using the Borel σ -algebra. For given c and probability c/γ , I require a suitable probability space for the inputs c_i . Following the approach by Sun (2006), 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 on $[0, \gamma]$, a Fubini extension on $[0, \gamma] \times \Omega$ and a measurable process $z: [0, \gamma] \times \Omega \mapsto \mathbb{R}$, such that the random variables $z(i, \cdot)$ are essentially pairwise independent with

$$\text{Prob}(\{\omega \in \Omega | z(i, \omega) = 1\}) = c/\gamma$$

for all $i \in [0, \gamma)$ almost surely. By definition of a Fubini extension, the integral $\int_{[0, \gamma)} z(i, \omega) d\lambda(i)$ is well defined for all $\omega \in \Omega$. Finally, Sun (2006, Theorem 2.8) proves that the integral equals $\gamma c/\gamma = c$ almost surely. \square

Proof of Theorem 1: The theorem and the following discussion state a budget-balanced mechanism and corresponding strategies. These strategies guarantee that $M_b = v$ and $\lambda(M_s) = c$ such that there is trade if and only if $v \geq c$. Therefore, the mechanism is ex-post efficient. In addition, for the seller, I show that the strategy is optimal and the mechanism is individually rational. The seller's strategy is also feasible because the suggested messages are measurable as realizations of a measurable process according to Lemma 1.

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

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

in equilibrium. These equilibrium payoffs are nonnegative so that the mechanism is individually rational for the buyer. Turning to incentive compatibility, consider deviations by the buyer. A deviation with $M_b \geq v$ makes the buyer worse off by potentially

triggering trade at a price above his valuation. Deviations with $M_b < v$ decrease probabilities of trade. In particular, this deviation implies utilities below

$$G(M_b) [v - c^*(M_b)] = \int_0^{M_b} v - \tilde{c} dG(\tilde{c})$$

Hence, such a deviation is unprofitable if

$$\int_0^v v - \tilde{c} dG(\tilde{c}) \geq \int_0^{M_b} v - \tilde{c} dG(\tilde{c}) \Leftrightarrow \int_{M_b}^v v - \tilde{c} dG(\tilde{c}) \geq 0$$

The last inequality is valid because $v - \tilde{c} > 0$ for all $\tilde{c} \in [M_b, v)$. Therefore, any deviation in M_b is unprofitable. A deviation in m_b does not affect the buyer. Therefore, any deviation in m_b is unprofitable and, in addition, any joint deviation in M_b and m_b is unprofitable. Consequently, his strategy is optimal for the buyer.

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

Proof of Lemma 2: Denote the seller's private information and, hence, her type by τ_1 with the type space T_1 and the buyer's private information by τ_2 with the type space T_2 . The seller's type is a profile of necessary inputs, $\tau_1 = (c_i)_{i \in [0, \gamma]}$. The buyer's type is his valuation and his sample of the seller's inputs, $\tau_2 = (v, i^*, s)$. I denote by h the conditional densities of a type conditional on the other player's type. According to McAfee and Reny (1992, Theorem 2) it is possible to use the correlation if and only if for every type $\tau_j^0 \in T_j$ and every probability measure μ on the type space T_j , $\mu(\{\tau_j^0\}) \neq 1$ implies $h(\cdot | \tau_j^0) \neq \int_{T_j} h(\cdot | \tau_j) d\mu(\tau_j)$.

Begin with the buyer and $j = 2$. The conditional density of τ_1 conditional on τ_2 is either 0 or equal to the unconditional density of $(c_i)_{i \in [0, \gamma] \setminus \{i^*\}}$ and, hence, determined by the distribution $G(c)$ and the essential independence. The first case occurs if $s \neq c_{i^*}$. The second case occurs if $s = c_{i^*}$. Pick any $\tau_2^0 = (v_0, i_0^*, s_0) \in T_2$ and denote by \cdot the dot product.⁹ Consider the measure μ_2 on T_2 determined by

$$\begin{aligned} \mu_2(\hat{T}_2) &= 0 \text{ with } \hat{T}_2 = \{\tau_2 = (v, i^*, s) \in T_2 | i^* \neq i_0^* \text{ or } s \neq s_0\} \text{ and} \\ \mu_2(\tilde{T}_2) &= \frac{1}{\nu} \lambda(\{(1, 0, 0) \cdot \tau_2 | \tau_2 \in \tilde{T}_2\}) \text{ with the Lebesgue measure } \lambda \end{aligned}$$

for all $\tilde{T}_2 \subseteq T_2 \setminus \hat{T}_2 = \{\tau_2 = (v, i^*, s) \in T_2 | i^* = i_0^* \text{ and } s = s_0\}$. The definition of μ_2 guarantees countable additivity. Thus, it is enough to verify that $\mu_2(T_2) = \mu_2(T_2 \setminus \hat{T}_2) = \nu/\nu = 1$ and, hence, $0 \leq \mu_2(T) \leq 1$ for all $T \subseteq T_2$ to ensure that μ_2 is a probability measure.

⁹Notice that $(1, 0, 0) \cdot \tau_2 = v \in (0, \nu]$ for any $\tau_2 = (v, i^*, s)$.

It is obvious that $\mu_2(\{\tau_2^0\}) = 0 < 1$ and $h(\tau_1|\tau_2^0) = h(\tau_1|\tau_2)$ for all $\tau_2 \in T_2 \setminus \hat{T}_2$. Hence,

$$h(\tau_1|\tau_2^0) = \int_{T_2} h(\tau_1|\tau_2) d\mu_2(\tau_2) \quad \forall \tau_1 \in T_1$$

violating the spanning condition of McAfee and Reny (1992, Theorem 2).

Continue with the seller and $j = 1$. The conditional density of τ_2 conditional on τ_1 is either 0 or equal to the unconditional density of (v, i^*) and, hence, determined by the distribution $F(v)$ and the uniform distribution of i^* on $[0, \gamma]$. The first case occurs if the third component of $\tau_2 = (v, i^*, s)$ equals the i^* component of τ_1 , namely $(0, 0, 1) \cdot \tau_2 = c_{i^*}$. The second case occurs if this equality is not satisfied and, hence, $(0, 0, 1) \cdot \tau_2 = s = 1 - c_{i^*}$. Pick any $\tau_1^0 = (c_i^0)_{i \in [0, \gamma]} \in T_1$. Consider the measure μ_1 on T_1 determined by

$$\begin{aligned} \mu_1(T_1 \setminus \hat{T}_1) &= 0 \text{ with } \hat{T}_1 = \{\tau_1 \in T_1 | \exists \hat{i}: c_i = c_i^0 \text{ for all } i \in ([0, \gamma] \setminus \{\hat{i}\})\} \text{ and} \\ \mu_1(\tilde{T}_1) &= \frac{1}{\gamma} \lambda(\{\hat{i} \in (0, \gamma] | \exists \tau_1 \in \tilde{T}_1: c_i = c_i^0 \text{ for all } i \in ([0, \gamma] \setminus \{\hat{i}\})\}) \text{ for all } \tilde{T}_1 \subseteq \hat{T}_1 \end{aligned}$$

with the Lebesgue measure λ . The definition of μ_1 guarantees countable additivity. Thus, it is enough to verify that $\mu_1(T_1) = \mu_1(\hat{T}_1) = \gamma/\gamma = 1$ and, hence, $0 \leq \mu_1(T) \leq 1$ for all $T \subseteq T_1$ to ensure that μ_1 is a probability measure.

It is obvious that $\mu_1(\{\tau_1^0\}) = 0 \neq 1$ and

$$h(\tau_2|\tau_1^0) = \int_{T_1} h(\tau_2|\tau_1) d\mu_1(\tau_1) \quad \forall \tau_2 \in T_2$$

violating the spanning condition of McAfee and Reny (1992, Theorem 2). Therefore, the mechanisms of McAfee and Reny (1992) and Crémer and McLean (1988) to achieve efficiency do not work here. \square

Proof of Proposition 3: The proposition and the following discussion state a budget-balanced mechanism and corresponding strategies. For the seller, I show that the strategy is optimal and the mechanism (3) is individually rational.

Now turn to the buyer. If there is no trade, his payoffs are zero. In addition, trade takes place only if his reported valuation is above the price p^* set in the mechanism. These expected payoffs are clearly nonnegative in equilibrium making the mechanism (3) individually rational for the buyer. A deviation with $M_b \geq v$ makes the buyer worse off by potentially triggering trade at a price above his valuation. Deviations with $M_b < v$ decrease probabilities of trade making him worse off. Therefore, any deviation in M_b is unprofitable. A deviation in m_b has no effect for the buyer. Therefore, any (joint) deviation in M_b and m_b is unprofitable. Consequently, his strategy is optimal for the buyer. \square

Proof of Lemma 3: For $\delta \rightarrow \infty$, $\lim p^*(c) = c$ and $p^*(c) \geq c$ for all $c \in [0, 1]$. Therefore, trade occurs in the mechanism (3) with probability 0 if $v < c$ and for $\delta \rightarrow \infty$ with probability 1 if $v \geq c$. Hence, the mechanism (3) is ex-post efficient for $\delta \rightarrow \infty$. Therefore, for sufficiently large δ , the mechanism generates a higher surplus than any fixed-price mechanism or any other classical mechanism that does not use narratives because these mechanisms cannot attain efficiency according to Myerson and Satterthwaite (1983, Corollary 1). \square

Proof of Proposition 4: The framework consists of well-defined type spaces, typical quasi-linear preferences and common priors. In addition, any communication is possible. Therefore, any implementable outcome function can be implemented truthfully according to Mas-Colell et al. (1995, Proposition 23.D.1). \square

Proof of Theorem 2: I begin by proving budget balance and incentive compatibility. Then I turn to individual rationality before considering optimality. Finally, I show existence of α .

Budget Balance and Incentive Compatibility

The mechanism in Theorem 2 is budget-balanced. In equilibrium, the buyer reports his valuation $M_b = v$ and his sample i^* of the seller's inputs by stating the name of the observed input if that input is unnecessary. In particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = 1$ otherwise. The seller, in equilibrium, reports the necessary inputs by sending messages $m_s = \sigma$ and $M_s = c$. These strategies ensure that $A(m_b, m_s, M_s) = 1$ in equilibrium.

Next, I consider incentive compatibility. To study the buyer's payoffs, suppose the seller follows her equilibrium strategy. Then the buyer's expected payoffs are

$$\begin{aligned} v\mathcal{T}_b(M_b) - \mathbb{E}(p^*(c, M_b)) &= v\mathcal{T}_b(M_b) - M_b\mathcal{T}_b(M_b) + \int_0^{M_b} \mathcal{T}_b(\tilde{v})d\tilde{v} \\ &\quad - \int_0^1 \left(\mathcal{T}_s(c')c' + \int_{c'}^{Z(c')+\frac{1}{\delta}} \mathcal{T}_s(\tilde{c})d\tilde{c} \right) dG(c') + \int_0^1 \left(\mathcal{T}_s(c')c' + \int_{c'}^{Z(c')+\frac{1}{\delta}} \mathcal{T}_s(\tilde{c})d\tilde{c} \right) dG(c') \\ &= (v - M_b)\mathcal{T}_b(M_b) + \int_0^{M_b} \mathcal{T}_b(\tilde{v})d\tilde{v} \end{aligned} \tag{6}$$

If the buyer reports $M_b = v$, his expected payoffs $U_b(v)$ are $\int_0^v \mathcal{T}_b(\tilde{v})d\tilde{v}$. Turning to incentive compatibility, consider deviations by the buyer. If the buyer deviates in M_b , his expected payoffs change by

$$\frac{\partial v\mathcal{T}_b(M_b) - \mathbb{E}(p^*(c, M_b))}{\partial M_b} = (v - M_b) \frac{\partial \mathcal{T}_b(M_b)}{\partial M_b} - \mathcal{T}_b(M_b) + \mathcal{T}_b(M_b) = \frac{\partial \mathcal{T}_b(M_b)}{\partial M_b} (v - M_b).$$

Assumptions (4) and (5) ensure that, for all $\tilde{\alpha} \in [0, 1]$, $c + \tilde{\alpha}(G(c) - G(Z(c)))/g(c)$ strictly increases in $c \in [(k-1)/\delta, k/\delta]$ for all $k \in \mathbb{N}^\delta$ and $v - \tilde{\alpha}(1 - F(v))/f(v)$ strictly increases in v . Hence, for all $\tilde{\alpha} \in [0, 1]$, $Y(\tilde{c}, \tilde{v}, \tilde{\alpha})$ decreases in $\tilde{c} \in [(k-1)/\delta, k/\delta]$ for all $k \in \mathbb{N}^\delta$ and increases in \tilde{v} . Therefore, $\mathcal{T}_s(c)$ decreases in c in each interval $[(k-1)/\delta, k/\delta]$ and $\mathcal{T}_b(v)$ increases in v . Hence, it is optimal to report $M_b = v$ because the positive derivative of $\mathcal{T}_b(v)$ ensures that the buyer's expected payoffs are maximised at $M_b = v$. Any deviation in m_b does not change the buyer's expected payoffs. Consequently, his strategy is optimal for the buyer.

Now turn to the seller. To study the seller's payoffs, suppose the buyer follows his equilibrium strategy. If the seller reports any messages (m_s, M_s) that guarantee an agreement, i.e., $A(m_b, m_s, M_s) = 1$ for all m_b consistent with the buyer's strategy and the seller's information, her expected payoffs are

$$\begin{aligned} \mathbb{E}(p^*(M_s, v)) - c\mathcal{T}_s(M_s) &= (M_s - c)\mathcal{T}_s(M_s) + \int_{M_s}^{Z(M_s)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} \\ &+ \int_0^1 \left(\mathcal{T}_b(v')v' - \int_0^{v'} \mathcal{T}_b(\tilde{v})d\tilde{v} \right) dF(v') - \int_0^1 \left(\mathcal{T}_s(c')c' + \int_{c'}^{Z(c')+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} \right) dG(c') \end{aligned} \quad (7)$$

The last two terms are constant in the seller's message. If the seller deviates in (m_s, M_s) still ensuring an agreement, her expected payoffs change by

$$\begin{aligned} \mathbb{E}(p^*(M_s, v)) - c\mathcal{T}_s(M_s) - \mathbb{E}(p^*(c, v)) + c\mathcal{T}_s(c) \\ = (M_s - c)\mathcal{T}_s(M_s) + \int_{M_s}^{Z(M_s)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} - \int_c^{Z(c)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c}. \end{aligned} \quad (8)$$

If $Z(c) = Z(M_s)$, it is optimal to report $M_s = c$ because the derivative of the seller's expected payoffs equals $(M_s - c)(\partial\mathcal{T}_s(M_s)/\partial M_s)$ and $\mathcal{T}_s(\cdot)$ is decreasing. If $Z(c) > Z(M_s)$, $c \geq Z(M_s) + 1/\delta > M_s$ and, hence, $(M_s - c)\mathcal{T}_s(M_s) < 0$. In addition, the last term in Eq. (8) is negative and the sum of the first two terms is also nonpositive as

$$\int_{M_s}^{Z(M_s)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} \leq (Z(M_s) + 1/\delta - M_s)\mathcal{T}_s(M_s) \leq (c - M_s)\mathcal{T}_s(M_s) = -(M_s - c)\mathcal{T}_s(M_s).$$

Therefore, expected payoffs decrease and deviations with $Z(c) > Z(M_s)$ are unprofitable. If $Z(c) < Z(M_s)$, any such deviations cannot ensure an agreement. In particular, the seller expects to pay π with probability of at least $Z(M_s) - Z(c)$ because there are at least $\delta Z(M_s)$ grid points contained in the set of inputs $[m_s, m_s + M_s]$ reported by the seller and the buyer observes each grid point with probability $1/\delta$. The mechanism charges the penalty only if trade occurs. Therefore, the gains of such a deviation are at most

$$(M_s - c)\mathcal{T}_s(M_s) + \int_{M_s}^{Z(M_s)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} - \int_c^{Z(c)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} - \pi(Z(M_s) - Z(c))$$

$$\begin{aligned}
&\leq (Z(M_s) + \frac{1}{\delta} - c - \pi(Z(M_s) - Z(c)))\mathcal{T}_s(M_s) - \int_c^{Z(c)+1/\delta} \mathcal{T}_s(\tilde{c})d\tilde{c} \\
&\leq (\frac{1}{\delta} - (\pi - 1)(Z(M_s) - Z(c)))\mathcal{T}_s(M_s) \leq (\frac{1}{\delta} - (\pi - 1)\frac{1}{\delta})\mathcal{T}_s(M_s) = 0.
\end{aligned}$$

The first inequality results from the fact that $\mathcal{T}_s(\cdot)$ is decreasing in $[M_s, Z(M_s) + 1/\delta)$. The second inequality results from $c \geq Z(c)$ and a positive integral in the second line. The last equality results from $\pi = 2$ and $Z(M_s) - Z(c) \geq 1/\delta$. Notice that $\pi = 1$ is also incentive compatible if $\lim_{c' \nearrow k/\delta} \mathcal{T}_s(c') \geq \mathcal{T}_s((k' - 1)/\delta)$ for all $k, k' \in \mathbb{N}^\delta$ with $k < k'$.¹⁰ Hence, it is optimal to report $M_s = c$. Any deviation in m_s weakly decreases the seller's expected payoffs. Therefore, any deviation in M_s and m_s is unprofitable. Consequently, her strategy is optimal for the seller.

Individually Rationality

Notice that the buyer's expected payoffs are zero for $v = 0$ according to Eq. (6) and they increase in his valuation v with derivative $\mathcal{T}_b(v) \geq 0$ in equilibrium. Hence, the mechanism in Theorem 2 is individually rational for the buyer. The seller's expected payoffs $U_s(c)$ decrease in her costs $c \in [(k - 1)/\delta, k/\delta)$ for all $k \in \mathbb{N}^\delta$ with derivative $-\mathcal{T}_s(c) \leq 0$. Denote $U_s^-(c') = \lim_{\tilde{c} \nearrow c'} U_s(\tilde{c})$. It remains to show that $U_s^-(k/\delta) \geq 0$ for all $k \in \mathbb{N}^\delta$ to ensure individually rationality. For this purpose, consider a more general mechanism with prices $p(c, v)$ and probabilities of trade $t(c, v)$ such that the buyer's expected utilities $U_b(v)$ increase in v and equal

$$U_b(v) = \int_0^1 vt(\tilde{c}, v) - p(\tilde{c}, v)dG(\tilde{c}) = U_b(0) + \int_0^v \int_0^1 t(\tilde{c}, \tilde{v})dG(\tilde{c})d\tilde{v} \quad (9)$$

and the seller's expected utilities $U_s(c)$ decrease in $c \in [(k - 1)/\delta, k/\delta)$ for all $k \in \mathbb{N}^\delta$ and equal

$$U_s(c) = \int_0^1 p(c, \tilde{v}) - ct(c, \tilde{v})dF(\tilde{v}) = U_s^-(\frac{k}{\delta}) + \int_c^{k/\delta} \int_0^1 t(\tilde{c}, \tilde{v})dF(\tilde{v})d\tilde{c} \quad (10)$$

for all $c \in [(k - 1)/\delta, k/\delta)$ and $k \in \mathbb{N}^\delta$. Notice that $\mathcal{T}_s(c) = \int_0^1 t(c, \tilde{v})dF(\tilde{v})$ and $\mathcal{T}_b(v) = \int_0^1 t(\tilde{c}, v)dG(\tilde{c})$ for the special case of the mechanism in Theorem 2, that satisfies these conditions. Denote $\Xi = \sum_{k=1}^\delta U_s^-(\frac{k}{\delta}) (G(\frac{k}{\delta}) - G(\frac{k-1}{\delta}))$. Hence, the seller's ex-ante expected utilities are

$$\begin{aligned}
&\int_0^1 U_s(c')dG(c') = \sum_{k=1}^\delta \int_{(k-1)/\delta}^{k/\delta} U_s(c')dG(c') \\
&\stackrel{\text{Eq. (10)}}{=} \sum_{k=1}^\delta \int_{(k-1)/\delta}^{k/\delta} \left(U_s^-(\frac{k}{\delta}) + \int_{c'}^{k/\delta} \int_0^1 t(\tilde{c}, \tilde{v})dF(\tilde{v})d\tilde{c} \right) dG(c')
\end{aligned} \quad (11)$$

¹⁰The reason is that this condition guarantees $\lim_{c' \nearrow Z(c)+1/\delta} \mathcal{T}_s(c') \geq \mathcal{T}_s(M_s)$ and the second line of above inequality is at most $(1/\delta - c + Z(c))\mathcal{T}_s(M_s) - (Z(c) + 1/\delta - c) \lim_{c' \nearrow Z(c)+1/\delta} \mathcal{T}_s(c')$ for $\pi = 1$.

$$\begin{aligned}
&= \sum_{k=1}^{\delta} \left[\int_{(k-1)/\delta}^{k/\delta} \int_{c'}^{k/\delta} \int_0^1 t(\tilde{c}, \tilde{v}) dF(\tilde{v}) d\tilde{c} dG(c') + U_s^- \left(\frac{k}{\delta} \right) \left(G \left(\frac{k}{\delta} \right) - G \left(\frac{k-1}{\delta} \right) \right) \right] \\
&\stackrel{\text{integration by parts}}{=} \sum_{k=1}^{\delta} \left[\int_{(k-1)/\delta}^{k/\delta} G(\tilde{c}) \int_0^1 t(\tilde{c}, \tilde{v}) dF(\tilde{v}) d\tilde{c} - G \left(\frac{k-1}{\delta} \right) \int_{(k-1)/\delta}^{k/\delta} \int_0^1 t(\tilde{c}, \tilde{v}) dF(\tilde{v}) d\tilde{c} \right] + \Xi \\
&= \int_0^1 \left(G(\tilde{c}) - G(Z(\tilde{c})) \right) \int_0^1 t(\tilde{c}, \tilde{v}) dF(\tilde{v}) d\tilde{c} + \Xi
\end{aligned}$$

Next, plug Eq. (11) into the total expected surplus.

$$\int_0^1 \int_0^1 (v' - c') t(c', v') dG(c') dF(v') = \int_0^1 U_b(v') dF(v') + \int_0^1 U_s(c') dG(c') \quad (12)$$

$$= U_b(0) + \int_0^1 \int_0^{v'} \int_0^1 t(\tilde{c}, \tilde{v}) dG(\tilde{c}) d\tilde{v} dF(v') + \int_0^1 U_s(c') dG(c')$$

$$\stackrel{\text{int. by parts \& Eq.(11)}}{=} U_b(0) + \int_0^1 (1 - F(\tilde{v})) \int_0^1 t(\tilde{c}, \tilde{v}) dG(\tilde{c}) d\tilde{v} + \int_0^1 \left(G(\tilde{c}) - G(Z(\tilde{c})) \right) \int_0^1 t(\tilde{c}, \tilde{v}) dF(\tilde{v}) d\tilde{c} + \Xi$$

$$= U_b(0) + \int_0^1 \int_0^1 t(\tilde{c}, \tilde{v}) \left(\left(G(\tilde{c}) - G(Z(\tilde{c})) \right) f(\tilde{v}) + (1 - F(\tilde{v})) g(\tilde{c}) \right) d\tilde{c} d\tilde{v} + \Xi$$

Combining the first and last line of Eq. (12) yields

$$\int_0^1 \int_0^1 \left(\tilde{v} - \frac{1 - F(\tilde{v})}{f(\tilde{v})} - \tilde{c} - \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})} \right) t(\tilde{c}, \tilde{v}) dG(\tilde{c}) dF(\tilde{v}) = U_b(0) + \Xi \quad (13)$$

The definition of α ensures that the left-hand side of this equation equals zero for the mechanism in Theorem 2. Therefore,

$$\Xi = \sum_{k=1}^{\delta} U_s^- \left(\frac{k}{\delta} \right) \left(G \left(\frac{k}{\delta} \right) - G \left(\frac{k-1}{\delta} \right) \right) = 0$$

in this mechanism because $U_b(0) = 0$. According to Eq. (7),

$$U_s^- \left(\frac{k}{\delta} \right) = \int_0^1 \left(\mathcal{T}_b(v') v' - \int_0^{v'} \mathcal{T}_b(\tilde{v}) d\tilde{v} \right) dF(v') - \int_0^1 \left(\mathcal{T}_s(c') c' + \int_{c'}^{Z(c') + \frac{1}{\delta}} \mathcal{T}_s(\tilde{c}) d\tilde{c} \right) dG(c')$$

is independent of k for all $k \in \mathbb{N}^\delta$. Hence, $U_s^-(k/\delta) = U_s^-(k'/\delta)$ for all $k, k' \in \mathbb{N}^\delta$ and $U_s^-(k/\delta) = 0$ for all $k \in \mathbb{N}^\delta$. Consequently, the mechanism in Theorem 2 is individually rational for the seller.

Optimality

The seller's costs c are her private information. For the moment, assume that the designer learns $c \in [(k-1)/\delta, k/\delta)$ for a $k \in \mathbb{N}^\delta$. In this case, the designer cannot infer anything about the seller's costs c and the necessary inputs even if the designer were to know the buyer's sample. Therefore, the narratives do not matter and this setting

is equivalent to the classical setting considered by Myerson and Satterthwaite (1983). Consider a mechanism with probabilities of trade $t(c, v)$ and price $p(c, v)$. Standard arguments show that any incentive compatible mechanism has to satisfy conditions (9) and (10). Therefore, Eq. (13) is valid in such a mechanism. Individual Rationality is satisfied if $U_b(0) \geq 0$ and $U_s^-(\frac{k}{\delta}) \geq 0$ for all $k \in \mathbb{N}^\delta$. Hence, any incentive compatible and individual rational mechanism has to ensure that the left-hand side of Eq. (13) is nonnegative. Therefore, optimal mechanisms are equivalent to probabilities of trade $t(c, v)$ that maximize

$$\int_0^1 \int_0^1 (v' - c') t(c', v') dG(c') dF(v')$$

subject to this constraint. Using κ as the Lagrange multiplier, the Lagrangian equals

$$(1 + \kappa) \int_0^1 \int_0^1 \left(\tilde{v} - \frac{\kappa}{1 + \kappa} \frac{1 - F(\tilde{v})}{f(\tilde{v})} - \tilde{c} - \frac{\kappa}{1 + \kappa} \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})} \right) t(\tilde{c}, \tilde{v}) dG(\tilde{c}) dF(\tilde{v}).$$

This Lagrangian is maximized by $t(\tilde{c}, \tilde{v}) = 1$ if $\tilde{v} - \frac{\kappa}{1 + \kappa} \frac{1 - F(\tilde{v})}{f(\tilde{v})} \geq \tilde{c} + \frac{\kappa}{1 + \kappa} \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})}$ and $t = 0$ otherwise. The mechanism in Theorem 2 implements such probabilities of trade and is incentive compatible as well as individually rational as shown above. Furthermore, the definition of α ensures that the left-hand side of Eq. (13) is zero and, hence, the constraint is satisfied with equality. Finally, the mechanism guarantees that $U_b(0) = 0$ and $U_s^-(k/\delta) = 0$ for all $k \in \mathbb{N}^\delta$. Consequently, the mechanism in Theorem 2 is optimal.

Existence of α

Define

$$I(\tilde{\alpha}, k) = \int_0^1 \int_{(k-1)/\delta}^{k/\delta} \left(\tilde{v} - \frac{1 - F(\tilde{v})}{f(\tilde{v})} - \tilde{c} - \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})} \right) Y(\tilde{c}, \tilde{v}, \tilde{\alpha}) dG(\tilde{c}) dF(\tilde{v})$$

for all $k \in \mathbb{N}^\delta$. Obviously, $I(1, k) \geq 0$ for all k because $Y(\tilde{c}, \tilde{v}, 1) = 0$ if the term in brackets is negative. Thus, $\sum_{k=1}^\delta I(1, k) \geq 0$. Next, $\sum_{k=1}^\delta I(0, k) < 0$ because by the steps above $\sum_{k=1}^\delta I(0, k) \geq 0$ would imply the existence of an ex-post efficient mechanism contradicting Myerson and Satterthwaite (1983, Corollary 1). Finally, it remains to prove continuity of $I(\tilde{\alpha}, k)$. Notice that $c + \tilde{\alpha} \frac{G(c) - G(Z(c))}{g(c)}$ strictly increases in $c \in [(k-1)/\delta, k/\delta]$ for all $\tilde{\alpha} \in [0, 1]$. Hence, $c + \tilde{\alpha} \frac{G(c) - G(Z(c))}{g(c)} = v - \tilde{\alpha} \frac{1 - F(v)}{f(v)}$ has at most one solution in c for given v and $\tilde{\alpha} \in [0, 1]$. This solution is continuous in v and $\tilde{\alpha}$. Therefore, there is a continuous function $\zeta^k(v, \tilde{\alpha}): [0, 1] \times [0, 1] \mapsto [(k-1)/\delta, k/\delta]$ so that

$$I(\tilde{\alpha}, k) = \int_0^1 \int_{(k-1)/\delta}^{\zeta^k(\tilde{v}, \tilde{\alpha})} \tilde{v} - \frac{1 - F(\tilde{v})}{f(\tilde{v})} - \tilde{c} - \frac{G(\tilde{c}) - G(Z(\tilde{c}))}{g(\tilde{c})} dG(\tilde{c}) dF(\tilde{v}).$$

Thus, $\sum_{k=1}^\delta I(\tilde{\alpha}, k)$ is continuous in $\tilde{\alpha}$. Consequently, Bolzano's theorem ensures that there is an $\alpha \in (0, 1]$ such that $\sum_{k=1}^\delta I(\alpha, k) = 0$. \square

B Symmetric Model

Return to the setting of Section 4. To achieve symmetry, assume that the buyer appreciates some features $f \in [0, \nu)$ of the product. Denote $v_f = 1$ if feature f is appreciated and $v_f = 0$ if feature f is not appreciated. Each feature $f \in [0, \nu)$ is appreciated with probability v/ν and yields the same utility. The features are essentially pairwise independent. The buyer knows his preferences and, hence, which features of the product he appreciates. His valuation v equals $\int_0^\nu v_f d\lambda(f)$ almost surely with an extension λ of the Lebesgue measure on $[0, \max\{\gamma, \nu\}]$. The seller only knows the distribution of preferences. By analyzing social media data or cookies in the buyer's browser, the seller learns for one feature f^* whether it is appreciated by the buyer. That feature f^* is drawn uniformly.

Equivalently to Theorem 1, Theorem 3 provides an ex-post efficient mechanisms with both seller and buyer using narratives. The mechanisms can implement any distribution of the surplus between buyer and seller in contrast to the asymmetric settings above where all the surplus goes to one side. For this purpose, define $p^*(M_b, M_s) = \chi\lambda(M_b) + (1 - \chi)\lambda(M_s)$ with the seller's bargaining power $\chi \in [0, 1]$.

Theorem 3. *The following mechanism implements any distribution of the surplus between buyer and seller. It is individually rational, ex-post efficient and budget-balanced: The buyer reports $\mathcal{M}_b = (m_b, M_b) \in [0, \gamma] \times \{X \subseteq [0, \nu) \mid X \text{ is } \lambda\text{-measurable}\}$ and the seller reports $\mathcal{M}_s = (m_s, M_s) \in [0, \nu] \times \{X \subseteq [0, \gamma] \mid X \text{ is } \lambda\text{-measurable}\}$. Payments and probabilities of trade are*

$$\left\{ \begin{array}{ll} t = 0, & p = 0, \quad \pi = 0 \quad \text{if } \lambda(M_b) < \lambda(M_s) \\ t = 1, & p = p^*(M_b, M_s), \quad \pi = 0 \quad \text{if } \lambda(M_b) \geq \lambda(M_s), m_b \notin M_s \text{ and } m_s \in (M_b \cup \{\nu\}) \\ t = 1, & p = p^*(M_b, M_s) + \nu\chi, \quad \pi = \nu\chi \quad \text{if } \lambda(M_b) \geq \lambda(M_s), m_b \notin M_s \text{ and } m_s \notin (M_b \cup \{\nu\}) \\ t = 1, & p = p^*(M_b, M_s), \quad \pi = \gamma(1 - \chi) \quad \text{if } \lambda(M_b) \geq \lambda(M_s), m_b \in M_s \text{ and } m_s \in (M_b \cup \{\nu\}) \\ t = 0, & p = \nu\chi, \quad \pi = \nu\chi + \gamma(1 - \chi) \quad \text{otherwise.} \end{array} \right.$$

In equilibrium, the buyer reports the features he appreciates $M_b = \{f \in [0, \nu) \mid v_f = 1\}$ and his sample i^* of the seller's inputs. In particular, he reports $m_b = i^*$ if $c_{i^*} = 0$ and $m_b = \gamma$ otherwise. Vice versa, in equilibrium, the seller reports the necessary inputs $M_s = \{i \in [0, \gamma) \mid c_i = 1\}$ and her sample f^* of the buyer's preferences. In particular, she reports $m_s = f^*$ if $v_{f^*} = 1$ and $m_s = \nu$ otherwise. These strategies ensure that $m_b \notin M_s$ and $m_s \in (M_b \cup \{0\})$. Hence, the mechanism is ex-post efficient in equilibrium. It is easy to see that the mechanism is budget-balanced. Next, study individual rationality and incentive compatibility of above strategies in turn.

To study the seller's payoffs, suppose the buyer follows his strategy. Denote the conditional expectation of the buyer's valuation by $v^*(c) = \mathbb{E}(v|v \geq c)$. In equilibrium, the seller expects payoffs

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

which are nonnegative making the mechanism individually rational for the seller. Turning to incentive compatibility, consider deviations by the seller. A deviation with $\lambda(M_s) \leq c$ makes the seller worse off by reducing the price. Her expected payoffs are below

$$\begin{aligned} & (1 - F(\lambda(M_s)))(\chi v^*(\lambda(M_s)) + (1 - \chi)\lambda(M_s) - c) < (1 - F(\lambda(M_s)))(v^*(\lambda(M_s)) - c)\chi \\ & = \chi \int_{\lambda(M_s)}^\nu \tilde{v} - c dF(\tilde{v}) \end{aligned}$$

which is below her equilibrium payoffs because $\tilde{v} - c < 0$ for all $\tilde{v} \in [\lambda(M_s), c)$. A deviation with $\lambda(M_s) > c$ has three effects: the price increases, probabilities of trade decrease, and additional payments $\pi = \gamma(1 - \chi)$ could be triggered. The first effect benefits the seller while the second and third effect hurt her. In particular, payoffs of such a deviation are below

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

As the seller's narrative must contain at least a mass of $\lambda(M_s) - c$ unnecessary inputs, the probability that the buyer observed one of these inputs is at least $(\lambda(M_s) - c)/\gamma$. Hence, $\text{Prob}(m_b \in M_s) \geq (\lambda(M_s) - c)/\gamma$ and such a deviation is unprofitable if

$$\chi \int_c^\nu \tilde{v} - c dF(\tilde{v}) \geq \chi \int_{\lambda(M_s)}^\nu \tilde{v} - c dF(\tilde{v}) \Leftarrow \int_c^{\lambda(M_s)} \tilde{v} - c dF(\tilde{v}) > 0$$

The last inequality is valid because $\tilde{v} - c > 0$ for all $\tilde{v} \in (c, \lambda(M_s)]$. A deviation in m_s has no effect for the seller. Therefore, any (joint) deviation in M_s and m_s is unprofitable. Consequently, her strategy is optimal for the seller.

Now turn to the buyer. Denote the conditional expectation of the seller's costs by $c^*(v) = \mathbb{E}(c|v \geq c)$ and suppose the seller follows her strategy. In equilibrium, the buyer expects payoffs

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

which are nonnegative making the mechanism individually rational for the buyer. Turning to incentive compatibility, consider deviations by the buyer. A deviation with $\lambda(M_b) \geq v$ makes the buyer worse off by increasing the price. His expected payoffs are

below

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

which is below his equilibrium payoffs because $v - \tilde{c} < 0$ for all $\tilde{c} \in (v, \lambda(M_s)]$. A deviation with $\lambda(M_b) < v$ has three effects: the price decreases, the probability of trade decreases, and a price surcharge of $\nu\chi$ could be triggered. The first effect benefits the buyer while the second and third effect hurt him. In particular, payoffs of such a deviation are below

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

As the buyer's narrative must miss at least a mass of $v - \lambda(M_b)$ appreciated features, the probability that the seller observed one of these features is at least $(v - \lambda(M_b))/\nu$. Hence, $\text{Prob}(m_s \notin (M_b \cup \{\nu\})) \geq (v - \lambda(M_b))/\nu$ and such a deviation is unprofitable because $v - \tilde{c} > 0$ for all $\tilde{c} \in [\lambda(M_s), v)$. Therefore, any deviation in M_b is unprofitable. A deviation in m_b has no effect for the buyer. Therefore, any (joint) deviation in M_b and m_b is unprofitable. Consequently, his strategy is optimal for the buyer.

To sum up, the mechanism is individually rational, ex-post efficient and budget-balanced. Choosing χ , the seller's bargaining power, appropriately, the mechanism can implement any distribution of the surplus.

References

- Al-Najjar, N. I., Anderlini, L., and Felli, L. (2006). Undescribable Events. *The Review of Economic Studies*, 73(4):849–868.
- Ali, N. and Bénabou, R. (2020). Image Versus Information: Changing Societal Norms and Optimal Privacy. *American Economic Journal: Microeconomics*, forthcoming.
- Anderlini, L. and Felli, L. (1994). Incomplete Written Contracts: Undescribable States of Nature. *The Quarterly Journal of Economics*, 109(4):1085–1124.
- Ball, I. and Kattwinkel, D. (2019). Probabilistic Verification in Mechanism Design. *Working Paper*.
- Ben-Porath, E. and Lipman, B. L. (2012). Implementation with partial provability. *Journal of Economic Theory*, 147(5):1689–1724.
- Bénabou, R., Falk, A., and Tirole, J. (2019). Narratives, Imperatives and Moral Reasoning. *NBER Working Paper*, 24798.
- 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.
- Caragiannis, I., Elkind, E., Szegedy, M., and Yu, L. (2012). Mechanism Design: from Partial to Probabilistic Verification. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 266–283.
- Carroll, G. and Egorov, G. (2019). Strategic Communication With Minimal Verification. *Econometrica*, 87(6):1867–1892.
- Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, 11(1):17–33.
- Crémer, 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.
- Eliasz, K. and Spiegler, R. (2019). A Model of Competing Narratives. *Working Paper*.
- Fenn, P., Lowe, D., and Speck, C. (1997). Conflict and dispute in construction. *Construction Management and Economics*, 15(6):513–518.
- 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.
- 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. *The 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.
- Hagenbach, J., Koessler, F., and Perez-Richet, E. (2014). Certifiable pre-play communication: Full disclosure. *Econometrica*, 82(3):1093–1131.
- Hart, S., Kremer, I., and Perry, M. (2017). Evidence games: Truth and commitment. *American Economic Review*, 107(3):690–713.
- Holmström, B. (1979). Moral Hazard and Observability. *Bell Journal of Economics*, 10(1):74–91.
- Jakobsen, A. M. (2020). A Model of Complex Contracts. *American Economic Review*, 110(5):1243–1273.
- Kartik, N. and Tercieux, O. (2012). Implementation with evidence. *Theoretical Economics*, 7(2):323–355.
- Kim, S. K. (1995). Efficiency of an Information System in an Agency Model. *Econometrica*, 63(1):89–102.
- 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.
- Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press.
- Mathis, J. (2008). Full revelation of information in Sender-Receiver games of persuasion. *Journal of Economic Theory*, 143(1):571–584.
- 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. (1981). Optimal Auction Design. *Mathematics of Operations Research*, 6:58–73.
- 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.
- Schweighofer-Kodritsch, S. and Strausz, R. (2020). Principled mechanism design with evidence. *Working Paper*.
- 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.
- Shiller, R. J. (2017). Narrative Economics. *American Economic Review*, 107(4):967–1004.
- 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.