
UNCERTAINTY AND IMPERFECT INFORMATION IN MARKETS*

BENJAMIN E. HERMALIN[†]

CONTENTS

1	INTRODUCTION	1
1.1	What's Not in this Chapter	2
2	EXOGENOUS ASYMMETRIES OF INFORMATION AT TIME OF CONTRACT- ING: INFORMED BUYER	7
2.1	Tariff Construction via Mechanism Design	9
2.2	The Monopoly Provision of Insurance	25
3	TWO-SIDED ASYMMETRIC INFORMATION	29
3.1	Trade Subject to Ex Ante Individual Rationality	31
3.2	Trade Subject to Interim Individual Rationality	38
3.3	The Money-on-the-Table Problem	45
4	EXOGENOUS ASYMMETRIES OF INFORMATION AT TIME OF CONTRACT- ING: INFORMED SELLER	48
4.1	The Lemons Model	49
4.2	Signaling	55
4.3	Experience Goods and Seller Reputation	65
5	ENDOGENOUS ASYMMETRIES OF INFORMATION AT TIME OF CONTRACT- ING	72
5.1	Cooperative Seller Actions	73
5.2	Semi-Cooperative Seller Actions	79
5.3	Cooperative Buyer Actions	82
5.4	Semi-Cooperative Buyer Actions	84
5.5	Selfish Buyer Actions	89
6	FINAL THOUGHTS	95
	REFERENCES	109

*The author gratefully acknowledges the financial support of the Thomas and Alison Schneider Distinguished Professorship in Finance.

[†]University of California • Berkeley, CA, USA • email: hermalin@berkeley.edu.

1 INTRODUCTION

Almost every market transaction is plagued by one party, at least, being ignorant of pertinent information. In some instances, both parties are ignorant of the same thing: neither farmer nor commercial granary know what the weather will be over the coming season when agreeing to a futures contract. In many instances, however, while one party is ignorant, the other knows the relevant information (or at least has better information): you know your willingness to pay for items in a shop, but the shopkeeper does not; you know the reliability of the used car you are selling, a potential buyer does not; and so on.

Although instances of common ignorance are not without interest—they, for example, are critical to understanding insurance and financial markets—such markets can often be analyzed using “textbook” methods once the situation is reformulated in terms of state-contingent commodities (see, *e.g.*, Section 2.2.1 *infra*).¹ What are arguably of greater interest are settings in which ignorance (or, equivalently, knowledge) is asymmetric between the parties: one party has better information about payoff-relevant factors than the other. Those settings are the focus of this chapter.

Asymmetry of information can arise in many ways and at various points in a bilateral relationship. As some examples:

1. one party can simply be endowed with better information than the other prior to any transaction (*e.g.*, a used-car seller has experienced her car’s reliability);
2. a party takes a payoff-relevant action, unobservable to her counter party, prior to any transaction (*e.g.*, a seller knows the quality of materials used in the manufacture of the product she sells);
3. after entering into a relationship, a party acquires better information (*e.g.*, a contractor learns how easy a job will be once on it);
4. after entering into a relationship, a party takes a payoff-relevant action, unobservable to her counter party (*e.g.*, an insured knows what precautions he takes to avoid a loss).

{item:HI-agency}

{item:HA-agency}

Asymmetry of information can also be relevant in *multilateral* settings (*e.g.*, an auction with many privately informed bidders). A complete treatment of all settings in which asymmetric information matters would entail a sizable volume in itself. Hence, of necessity, this chapter will be more tightly focused: for the most part, attention is limited to bilateral settings in which asymmetries of information exist prior to any transactions. The chapter will, thus, have little to say about scenarios #3 and #4 (*i.e.*, situations of hidden-information agency

¹Chapter 19 of Mas-Colell et al. (1995) provides a good introduction to state-contingent commodities and markets for their trade. See also Huang and Litzenberger (1988).

and moral hazard, respectively).² The next subsection briefly discusses some other areas of the literature that are being ignored, along with citations for readers interested in those areas.

Beyond that next subsection, the rest of the chapter is divided as follows: Section 2 considers situations in which the contract proposer is the ignorant party. Her problem is to design a *mechanism* that induces her counter party, who is exogenously endowed with his information, to reveal that information in such a way that maximizes the contract proposer's expected payoff. The quintessential example of this is uninformed seller designing a profit-maximizing price discrimination scheme. As will become clear, to induce the informed party to reveal his information, the uninformed contract proposer will need leave him some surplus (an information rent). Because she cannot, therefore, capture all the surplus, the contract proposer will not have appropriate incentives to maximize surplus (welfare). Hence, in contrast to settings of symmetric information, inefficiencies will tend to arise in equilibrium. Section 3 considers the situation when each side of the transaction is endowed with his or her own payoff-relevant information. Unlike the preceding section, the focus will be on whether and how a social planner could design a contract to achieve efficiency.³ Section 4 assumes it is the contract proposer who is endowed with the payoff-relevant information. The quintessential example of this a seller who knows the quality of the good she seeks to sell. Because she cannot commit fully to not deceive her counter party, inefficiencies arise. Section 5 turns to the problems that arise if the asymmetry of information arises *endogenously*, because one party's prior-to-trade actions provides him or her payoff-relevant information.

1.1 WHAT'S NOT IN THIS CHAPTER

1.1.1 COMPETITION AMONG SELLERS

Among the topics not being covered are models in which multiple sellers compete by offering contracts to buyers.

If price competition within an oligopoly is sufficient fierce, then there is little scope for price discrimination. It is, for instance, readily shown that if the standard Bertrand equilibrium would hold if sellers were limited to linear pricing, then it continues to hold even if their strategy spaces encompass complicated tariffs. On the other hand, if competition is less fierce—as, say, is true of Hotelling competition—then equilibria can exist in which sellers engage in

²Hidden-information agency is similar, in terms of methods, to the mechanism-design analysis of Section 2.1 *infra*. This is especially true if the agent is free to quit after learning the payoff-relevant information because, then, the contract will have to satisfy individual-rationality constraints similar to those that arise in that section. Classic articles on moral hazard—also known as *hidden-action agency*—are Holmstrom (1979), Shavell (1979), and Grossman and Hart (1983). Good textbook treatments can be found in Chapter 4 of Bolton and Dewatripont (2005) and Chapters 4 and 5 of Laffont and Martimort (2002). For a web-based resource see Caillaud and Hermalin (2000).

³If, in contrast, one side had the ability to design the contract, subject only to the other's acceptance of it, then the problem would be little different than the analysis of Section 2.

price discrimination. The topic of price discrimination in oligopolistic settings is, however, a chapter in itself, as the excellent surveys by Armstrong (2006b) and Stole (2007) attest. The interested reader would do well to start his or her study there.

Insurance markets with multiple insurers are particularly complex markets to analyze. Under the most natural formulation of competition and using standard notions of equilibrium, such markets can even fail to have equilibria (the well-known Rothschild and Stiglitz, 1976, non-existence result). Although modifying the models or the equilibrium concept can avoid non-existence (see, *e.g.*, Wilson, 1977, or Hellwig, 1987), such modifications are not always appealing.

Many of those modifications can, in a sense, be seen as expanding the space of potential contracts that insurers can offer. Implicitly or explicitly, insurers are allowed to make offers that are contingent on the offers of other insurers. Unfortunately, the economic modeling of competition in contracts (offers) is not well developed and the results highly sensitive to the extent to which competitors can make their offers contingent on those of others. As an illustration, extend the basic Bertrand model of price competition as follows. Suppose each firm could make the offer:

1. If my rivals all make the same offer as me (*i.e.*, an offer with these points #1 and #2), then I am offering to sell my product at price equal to the monopoly price; but
2. if any rival makes a different offer, then I am offering to sell my product at price equal to marginal cost.

Clearly, a best response to all rivals' offering the above is to do so yourself: if you don't, then you can make at most zero profit, because your rivals would, then, be pricing at marginal cost; but if you make that offer, then you get a fraction of the monopoly profit, which is a positive amount. To be sure, one can legitimately object to such an equilibrium as being unrealistic (at least with a large number of sellers), possibly illegal under relevant antitrust statutes, or dependent on an implausible level of commitment by sellers to their offers.⁴ But as a matter of game theory *per se*, this is a perfectly legitimate equilibrium. Given the unsettled nature of modeling competition in contracts and the almost anything-goes results of many of such models—to say nothing of the limited space afforded this chapter—I have chosen not to consider competition among sellers. The reader interested in competing contract offers is well advised to begin with Katz (1991, 2006).

1.1.2 EXTERNALITIES ACROSS BUYERS

Although many of the models analyzed below are readily extended to allow for multiple buyers, such extensions are predicated on there being no direct or indirect externalities across buyers.

⁴On the other hand, it is claimed that retailers' guaranteed-low-price policies are essentially offers of this nature. See, for instance, Edlin (1997) and cites therein.

In particular, an extension to multiple buyers can be complicated if the seller's production technology does not exhibit constant marginal costs.⁵ Related complications arise if there are many potential buyers for a single item (*e.g.*, the seller has a unique piece of art to sell). A common way of selling such an item is an auction. The literature on auctions is vast and arguably goes beyond a survey chapter itself—to say nothing of being a portion of a survey chapter; consequently, this chapter does not consider auctions. For the interested reader, there are a number of excellent books on the topic (consider, *e.g.*, Krishna, 2002; Klemperer, 2004; and Milgrom, 2004).

With multiple buyers there is also the possibility that the actions of one buyer can convey information about other buyers. By exploiting such correlation, the seller can extract more of the buyers' surplus. This has, for example, been considered in a pair of articles by Crémer and McLean (1985, 1988).⁶ That topic is not covered in this chapter.

A growing area of research has been on so-called *two-sided markets*, such as telecommunications, payment cards, and singles bars.⁷ In such markets, one actor—the platform—provides a service that facilitates the interaction of two other actors (or classes of actors). A payment-card network, for instance, facilitates the exchanges of consumers and merchants. In its pricing, a platform often needs to contend with two externalities that exist between the two sides: a transaction externality and a membership externality. The former is the benefit a user on one side (*e.g.*, a merchant) derives when a user on the other (*e.g.*, a consumer) chooses to transact with it. The latter is the benefit a user on one side (*e.g.*, a man in a stereotypical singles bar) derives from having more options on the other side (*e.g.*, more women at the bar). In designing the tariffs it offers, the platform needs to be mindful of these externalities.⁸ Although an interesting topic, the limited space afforded here precludes further discussion.

1.1.3 INCOMPLETE CONTRACTS AND CONTRACT RENEGOTIATION

As noted, the focus here is when buyer and seller (the contractual parties) possess different information. Starting with Grossman and Hart (1986), a literature has arisen that studies the consequences of asymmetric information not between the contractual parties, but between those parties and some third party (*e.g.*, a judge) that enforces the contract between the contractual parties. In particular, the contractual parties know payoff-relevant information that the third

⁵See Crémer and Riordan (1987) for an extension when the seller's cost function is not linear.

⁶There has related work in hidden-information agency settings with multiple agents. See, for instance, Demougin and Garvie (1991) and McAfee and Reny (1992).

⁷For recent surveys of the two-sided markets literature see Rochet and Tirole (2006) or Rysman (2009).

⁸For instance, a platform, which can utilize two-part tariffs and which confronts a transaction externality, will wish to set the prices of the transactions to sum to less than the cost of facilitating the transaction (see, *e.g.*, Hermalin and Katz, 2004, for details). For an analysis of membership externalities, see Armstrong (2006a).

party does not (in the literature, such information is described as “observable, but unverifiable”). Although this kind of informational asymmetry has proved to be of great importance, it is not covered here. For a partial survey of the literature, see Hermalin et al. (2007), especially Section 4.

When payoff-relevant information is observable to the contractual parties, but unverifiable (unobservable to a third-party contract enforcer), the resulting contracts are often described as *incomplete*. As Grossman and Hart (1986) observed, one way the contractual parties will respond to contractual incompleteness is to renegotiate their contracts should the contractual parties learn the observable-but-unverifiable information after writing their initial contract, but before its full execution. There is a large literature on contract renegotiation (see Hermalin et al., 2007, for a partial survey). Among the debates in that literature is whether the contractual parties can commit to a renegotiation mechanism or must continue to renegotiate as long as there is “money on the table” (for a brief introduction to the “money-on-the-table” problem see §3.3 *infra*). This touches more generally on the problem of the parties committing to a contract that will prove *ex post* inefficient: if the parties come to understand that there is money to be had, it seems natural to imagine that they will seek to pick it up by, if necessary, renegotiating their contract. The problem is that the anticipation of such a lack of commitment can have adverse effects *ex ante*: it can be in the parties’s interest to commit to leaving some money on the table, at least off the equilibrium path, in order to provide themselves appropriate incentives *ex ante*.⁹

Dealing with the money-on-the-table problem has proved difficult because of the general difficulty of satisfactorily modeling bargaining given asymmetric information. For further discussion see Section 3.3. It is that difficulty—in addition to the overall length constraint on the chapter—that has led me to omit a more detailed discussion of the topic.

1.1.4 HARD INFORMATION

This chapter is essentially limited to what is known as *soft* information: although the informed party can make claims about what s/he knows, those claims are cheap talk insofar as her/his counter party cannot verify such claims. That is, the informed party can lie. The literature has also considered *hard* information: if the informed party chooses to reveal her/his information, then the counter party can verify it. As an example, a seller’s information about the reliability of her car is soft, but her information about its repair history is hard to the extent she can provide receipts from her mechanic for work done.¹⁰ Although hard information cannot be misstated, it can be concealed: unless the informed

⁹Such issues arise, for instance, in the bilateral investment literature (*e.g.*, Demski and Sappington, 1991; Nöldeke and Schmidt, 1995, 1998; and Edlin and Hermalin, 2000). Another important example, having to do with hidden-action agency, is Fudenberg and Tirole (1990).

¹⁰There is, of course, the possibility that the receipts are fake or she has conspired with the mechanic to commit fraud, but such possibilities are typically ruled out.

party reveals it, the uninformed party still does not know what it is (*e.g.*, a used-car seller can hide receipts from her mechanic).

In some instances, the *existence* of hard information is the informed party's private information (*e.g.*, the seller knows what receipts she has) and in others it is commonly known (*e.g.*, a college graduate can produce a transcript). When existence is commonly known, it may be difficult for the informed party to conceal her/his hard information due to the unraveling argument of Grossman (1981): the uninformed party's expectation of the information conditional on concealment is necessarily less than the true value for some informed players; hence, those players, who are presumably concealing, would, in fact, do better to reveal. Extending this logic, only someone with the worst information would be willing to conceal.¹¹ When the existence of information is uncertainty, then concealment may occur in equilibrium.¹²

1.1.5 SIGNAL JAMMING

In some situations, the parties may initially be symmetrically informed, but actions of one may hamper the ability of the other to learn new information. A sizable literature of such *signal-jamming* models exists.¹³ For instance, suppose one party wishes to infer the value of a payoff-relevant parameter that is known to have been drawn from a specific distribution. A signal, s , that is informative of that parameter will be generated, but one party may only see $s + x$, where x is an action taken by the other party to distort the signal. For example, as in Holmstrom (1999), a manager's efforts affect a firm's revenues, which are a signal of his ability (the payoff-relevant parameter). In some situations, such signal-jamming efforts can be welfare enhancing (as suggested by Fama, 1980, and as evidenced in some models in Holmstrom, 1999); in others they can be welfare reducing (as in Fudenberg and Tirole, 1986, Stein, 1989, or other models in Holmstrom, 1999).

¹¹As a simple model, suppose the informed party's (his) hard information is $\theta \in [0, 1]$ and the uninformed party's (her) prior belief is it's distributed uniformly on that interval. Suppose her best-response action to her posterior belief is $a = \mathbb{E}\{\theta\}$ and that is worth $v(a)$ to the informed party, $v(\cdot)$ strictly increasing. Suppose there were an equilibrium in which the informed party conceals if $\theta \leq \bar{\theta}$ and he reveals otherwise. Because the uninformed party's best response would be $\bar{\theta}/2$ if information was not revealed, the informed party would do better to reveal than to conceal if $\theta > \bar{\theta}/2$. This can be consistent with the purported equilibrium only if $\bar{\theta} = 0$; that is, if all reveal.

¹²To extend the example of the previous footnote: suppose that informed party acquires the hard information with probability $1/2$. It is readily seen that an equilibrium exists in which those who acquire the information reveal if $\theta > 1/3$ and otherwise conceal. The uninformed party's estimate of θ given no information is revealed is $1/3$.

¹³The term "signal jamming" appears to be due to Fudenberg and Tirole (1986). There are some earlier examples of the phenomenon in the literature, though (see, *e.g.*, Holmstrom, 1999—the original version of which appear in 1982).

2 EXOGENOUS ASYMMETRIES OF INFORMATION AT TIME OF CONTRACTING: INFORMED BUYER

{sect:InformedBuyer}

To start, consider a situation in which one party proposes a contract to another. The offer is take-it-or-leave-it (TIOLI): the offer's recipient can accept, in which case the contract is binding on both parties; or he can reject, in which case there is no further negotiation and the parties receive their default (no-trade) payoffs. In many market settings, the contract offeror is a seller and the recipient a buyer. For the sake of brevity, let's employ the names seller and buyer, with the understanding that much of what follows is more general (*e.g.*, the contract proposer could be an employer and the recipient an employee).

A prominent example is the basic monopoly model, the one taught in introductory economics. To be sure, as typically taught, the model presumes many buyers, but the number of buyers is (essentially) irrelevant for the conclusions reached. In that model, the seller offers the buyer the opportunity to buy the quantity he wishes, x , in exchange for paying the seller $T(x)$, where the function $T : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is stipulated in the offer. For example, in the most basic monopoly model, the function T is a *linear tariff*: if the buyer announces (buys) x , he pays the seller px , where the rate (price) p is quoted in units of currency per unit of the good. Of course, as one learns in more advanced economics courses, the tariff needn't be linear. Depending on the various informational and other constraints she faces, the seller can derive greater profit by engaging in *price discrimination* via a *nonlinear* tariff.

The design of the tariff is, in part, a function of the informational asymmetry that exists between seller and buyer. In particular, the buyer is assumed to have been endowed with payoff-relevant information—his *type*, β . The buyer's type is typically his private information. It is, however, common knowledge that his type is drawn from a known type space, B , according to a specific distribution. In some instances, such as simple monopoly pricing, this may be all that the seller knows. In other instances, such as third-degree price discrimination, the seller is somewhat better informed, having observed some signal that is informative about the buyer's type.¹⁴

As an example, the buyer may wish to purchase, at most, one unit of the seller's product. The buyer's payoff if he does is $\beta - p$. If he doesn't (there's no trade), his payoff is zero. Suppose that his type space is $B = [0, 1]$ and β is drawn by "nature" from that space according to a known distribution $1 - D$ (the reason for writing it in this unusual manner will become clear shortly). Assume this is all the seller knows. The seller's payoff is p if she sells a unit (for convenience, normalize her costs to zero) and it is 0 otherwise. Suppose the seller offers the contract "one unit for p ." The buyer does better buying than not buying if $\beta - p \geq 0$. The probability that a seller who quotes a price of p makes a sale is, therefore, $\Pr\{\beta \geq p\}$; that is, it equals the survival function of the buyer's type evaluated at p . The expected quantity sold is $D(p)$. Observe

¹⁴This chapter omits a discussion of third-degree price discrimination—the interested reader is directed to Varian (1989) or Tirole (1988, §3.2).

this survival function is equivalent to what, in introductory economics, would be called a demand function.¹⁵ The seller will choose p to maximize her expected payoff, $pD(p)$. For instance, if the buyer's type is distributed uniformly, then the seller chooses p to solve

$$\max_p p(1 - p).$$

The solution is readily seen to be $p^* = 1/2$. In short, the seller would make the buyer the TIOLI offer “one unit at price $1/2$.”¹⁶

Although somewhat trivial, this example reveals a basic point about efficiency when contracting occurs in the shadow of asymmetric information. Welfare is β if trade occurs, but zero if it doesn't. Because $\beta \geq 0 = \text{cost}$, welfare is maximized if trade always occurs. The seller, however, sets a price of $1/2$: trade thus occurs with probability $1/2$. This is the standard result that linear pricing by a monopolist tends to yield too little trade. Observe this inefficiency is the result of (i) asymmetry of information and (ii) the bargaining game assumed: if the seller knew β —information were symmetric—then she would offer a type- β buyer “one unit at price β .” Because $\beta - \beta \geq 0$, the buyer would accept. Trade would always occur and maximum welfare, thus, achieved. Alternatively, if the buyer were to make a TIOLI offer to the seller, then welfare would also be maximized: the buyer would offer “one unit at price equal cost (*i.e.*, 0).” Because $0 \geq 0$, the seller would accept. Trade would be certain to occur and welfare thus maximized.

These observations generalize: consider a contract $\langle x, t \rangle$, where $x \in \mathcal{X}$ is an allocation and $t \in \mathbb{R}$ a monetary transfer. Let the seller's payoff be $t - c(x)$ and the buyer's $u(x, \beta) - t$, where $c : \mathcal{X} \rightarrow \mathbb{R}$ and $u : \mathcal{X} \times B \rightarrow \mathbb{R}$. Let x_0 be the “no trade” allocation; that is, if no agreement is reached, then the allocation is x_0 . Assume $x_0 \in \mathcal{X}$: among the feasible contracts is one that replicates no trade.

{prop:Efficiency1}

Proposition 1. *Consider the model set forth above. One player's (e.g., the buyer's) payoff is type dependent, but the other's (e.g., the seller's) is not. The player with the type-dependent payoff knows his type. Assume that for all types, there is a welfare-maximizing allocation.¹⁷ Then if one party gets to offer a contract on a take-it-or-leave-it basis and s/he knows the type when doing so, then welfare will be a maximum for all values of type in equilibrium.*

Proof: Because the offeror could always offer the no-trade allocation and zero transfer, there is no loss of generality in assuming the offeror's contract is

¹⁵It can be shown that any well-behaved demand function is proportional to a survival function, where “well-behaved” means finite demand at zero price, zero demand at an infinite price, and demand is everywhere non-increasing in price. The proportionality factor is just demand at zero price.

¹⁶If the bargaining is other than seller makes a TIOLI offer, then the outcome would be different. In particular, a key assumption is that the seller can commit to never make the buyer a subsequent offer should he reject the seller's initial offer. The literature on the Coase conjecture (Coase, 1972) explores what happens if the seller is unable to so commit. See, for example, Gul et al. (1986)). See also the discussion in Section 3.3 *infra*.

¹⁷That is, using the notation introduced above, assume the program $\max_{x \in \mathcal{X}} u(x, \beta) - c(x)$ has a solution for all $\beta \in B$.

accepted in equilibrium. If the seller makes the offer, she does best to offer

$$\operatorname{argmax}_{x \in \mathcal{X}, t \in \mathbb{R}} t - c(x)$$

subject to

$$u(x, \beta) - t \geq u(x_0, \beta). \quad (2.1) \quad \{\text{eq:BuyerIR-simp}\}$$

For any x , if the buyer's participation (offer-acceptance) constraint (2.1) did not bind, then the seller could raise t slightly, keeping x fixed, and do better. Hence, in equilibrium, (2.1) must bind. Hence, the seller's choice of x must solve

$$\max_{x \in \mathcal{X}} u(x, \beta) - c(x) - u(x_0, \beta). \quad (2.2) \quad \{\text{eq:Welfare-simp2}\}$$

Because (2.2) is welfare less a constant, the result follows.

The analysis when the buyer makes the offer is similar and, thus, omitted for the sake of brevity. ■

What, in fact, is *not* general is the conclusion that giving the bargaining power to the uninformed player (*i.e.*, the seller) leads to inefficiency. To see this, return to the simple example, maintaining all assumptions except, now, let $B = [1, 2]$. Because the seller can guarantee herself a profit of 1 by quoting a price of 1, it would be irrational of her to quote either a price less than 1 or greater than 2 (the latter because it would mean no sale). Hence, the seller's pricing problem can be written as

$$\max_{p \in [1, 2]} p(2 - p).$$

The solution is $p^* = 1$. At that price, all types of buyer will buy—welfare is maximized.

The astute reader will, at this point, object that efficiency has been shown only under the assumption that the seller is restricted to linear pricing. What, one may ask, if the seller were unconstrained in her choice of contract? To answer this, we need to derive the profit-maximizing tariff for the seller, the topic of the next subsection. Anticipating that analysis, however, it will prove—*for this example*—that linear pricing is the profit-maximizing tariff and, thus, the seller's ignorance of the buyer's type need not always lead to inefficiency.

2.1 TARIFF CONSTRUCTION VIA MECHANISM DESIGN

{sect:Tariff_Via_Mechanisms}

Now consider a mechanism-design approach to tariff construction. Payoffs are as above. The buyer's type, β , is fixed exogenously and is his private information. The seller knows the distribution from which β was drawn. Denote that distribution by F and its support by B . The structure of the game is common knowledge.

2.1.1 MECHANISM DESIGN: A CRASH COURSE

There are numerous texts that cover mechanism design.¹⁸ Consequently, the discussion here will be brief.

Call an allocation-transfer pair, (x, t) , a *contractual outcome*. The set of all such outcomes is $\mathcal{X} \times \mathbb{R}$. Let $\Delta(\mathcal{X} \times \mathbb{R})$ denote the set of all possible probability distributions over outcomes.

Definition. A mechanism is a game form, $\langle \mathcal{M}, \mathcal{N}, \sigma \rangle$, to be played by the parties. The set $\mathcal{M}$ is the informed player's (e.g., the buyer's) strategy set, the set $\mathcal{N}$ the uninformed player's (e.g., the seller's) strategy set, and σ maps any pair of strategies, (m, n) , to a probability distribution over contractual outcomes; that is, $\sigma : \mathcal{M} \times \mathcal{N} \rightarrow \Delta(\mathcal{X} \times \mathbb{R})$.

Assume the choice of mechanism is the uninformed player's (e.g., the seller's). Observe that any conceivable contract can be viewed as a mechanism. For example, linear pricing is the mechanism in which $\mathcal{M} = \mathcal{X} \subseteq \mathbb{R}_+$, $\mathcal{N} = \emptyset$, and $\sigma : m \mapsto (m, pm)$.¹⁹

Let $\mathbb{E}_{\sigma(m,n)}\{\cdot\}$ denote expectation with respect to the random vector (x, t) when it is distributed $\sigma(m, n)$. Define

$$U(\sigma(m, n), \beta) = \mathbb{E}_{\sigma(m,n)}\{u(x, \beta) - t\};$$

that is, $U(\sigma(m, n), \beta)$ is the informed player's expected payoff if he is type β , he plays m , and the uninformed player plays n .

A mechanism is *direct* if $\mathcal{M} = B$; that is, if the informed player's action is an announcement of type. It is a *direct-revelation mechanism* if, in equilibrium, the informed player announces his type truthfully. For truth-telling to be an equilibrium strategy, it must be a best response to the informed player's type and his expectation of the uninformed player's action n .²⁰

$$U(\sigma(\beta, n), \beta) \geq U(\sigma(\beta', n), \beta) \quad \forall \beta' \in B. \quad (2.3) \quad \{\text{eq:BR-GenDRMechanism}\}$$

The uninformed player seeks to design a mechanism that will, in equilibrium, maximize her expected payoff. Designing a mechanism entails choosing $\mathcal{M}$, $\mathcal{N}$, and σ . The classes of spaces and outcome functions are incomprehensibly large. How can the optimal mechanism be found within them? Fortunately, a simple, yet subtle, result—the *Revelation Principle*—allows us to limit attention to direct-revelation mechanisms.

¹⁸ A partial list is Gibbons (1992), Salanié (1997), Laffont and Martimort (2002), and Bolton and Dewatripont (2005). See also Chapters 13 and 23 of Mas-Colell et al. (1995).

¹⁹ If one wished to get technical, one could write that σ maps m to the distribution that assigns all weight to the outcome (m, pm) .

²⁰ Here, the equilibrium concept is Bayesian Nash. In other situations, different equilibrium concepts, such as solution in dominant strategies or perfect Bayesian equilibrium, are relevant. The discussion here, including the Revelation Principle, extends to other equilibrium concepts.

`\fn:BayesianNash`

{prop:RevelationPrinciple}

Proposition 2 (The Revelation Principle).²¹ *For any general mechanism $\langle \mathcal{M}, \mathcal{N}, \sigma \rangle$ and associated Bayesian Nash equilibrium, there exists a direct-revelation mechanism such that the associated truthful Bayesian Nash equilibrium generates the same distribution over outcomes in equilibrium as the general mechanism.*

Proof: A Bayesian Nash equilibrium of the game $\langle \mathcal{M}, \mathcal{N}, \sigma \rangle$ is a pair of strategies $(m(\cdot), n)$, $m(\cdot)$ a mapping from the type space to $\mathcal{M}$.²² Consider the direct mechanism: $\hat{\sigma}(\cdot) = \sigma(m(\cdot), n)$. The claim is that $\hat{\sigma}(\cdot)$ induces truth-telling (is a *direct-revelation* mechanism). To see this, suppose not. Then there must exist a type β that does better to lie—announce some $\beta' \neq \beta$; that is, formally, there must exist β and $\beta' \neq \beta$ such that

$$U(\hat{\sigma}(\beta'), \beta) > U(\hat{\sigma}(\beta), \beta).$$

Using the definition of $\hat{\sigma}(\cdot)$, this means, however, that

$$U(\sigma(m(\beta'), n), \beta) > U(\sigma(m(\beta), n), \beta).$$

But if that expression is true, then the informed player prefers to play $m(\beta')$ instead of $m(\beta)$ in the *original* mechanism. This contradicts the assumption that $m(\cdot)$ is an equilibrium best response to n in the original game. It follows, *reductio ad absurdum*, that $\hat{\sigma}$ induces truth-telling.

Moreover, because $\hat{\sigma}(\beta) = \sigma(m(\beta), n)$, the same distribution over outcomes is implemented in equilibrium. ■

An intuitive way to grasp the Revelation Principle is to imagine that, before he plays some general mechanism, the informed player could delegate his play to some trustworthy third party. Suppose the third party knew the agent's equilibrium strategy—the mapping $m : B \rightarrow \mathcal{M}$ —so the informed player need only reveal his type to the third party with the understanding that the third party should choose the appropriate action, $m(\beta)$. But, because this third party can be “incorporated” into the design of the mechanism, there is no loss of generality in restricting attention to direct-revelation mechanisms.

²¹The Revelation Principle is often attributed to Myerson (1979), although Gibbard (1973) and Green and Laffont (1977) could be identified as earlier derivations. Suffice it to say that the Revelation Principle has been independently derived a number of times and was a well-known result before it received its name. Proposition 2 states the Revelation Principle for a situation with one informed player. Extending the Revelation Principle to many informed players is straightforward. Further, as observed in footnote 20, the Revelation Principle holds under different solution concepts.

²²Observe that the informed player's strategy can be conditioned on β , which he knows, while the uninformed player's cannot be (since she is ignorant of β).

2.1.2 THE STANDARD ASSUMPTIONS

So far, the seller's cost has assumed to be independent of the buyer's type. This can be relaxed—let $c(x, \beta)$ denote her cost from now on.²³ Welfare is now

$$w(x, \beta) = u(x, \beta) - c(x, \beta).$$

To derive the seller's profit-maximizing mechanism, certain assumptions are necessary to make the problem tractable. The following are standard.

Assumption 1 (Spence-Mirrlees Condition). *There exist complete orders $\succ_\beta$ and $\succ_x$ on B and $\mathcal{X}$, respectively, such that the implication*

{ass:Spence-MirrleesGenl}

$$u(x, \beta') - t \geq u(x', \beta') - t' \implies u(x, \beta) - t > u(x', \beta) - t' \quad (2.4)$$

{eq:Spence-Mirrlees}

is valid whenever $\beta \succ_\beta \beta'$ and $x \succ_x x'$.

Assumption 2 (Trade is Desirable). *If β is not the infimum of B under the complete order $\succ_\beta$, then there exists an $x \in \mathcal{X}$ such that $w(x, \beta) > w(x_0, \beta)$.*

{ass:TradeDesirable}

Assumption 3 (Too Much of a Good Thing). *Either (i) the set of possible allocations is bounded or (ii) for all $\beta \in B$ there exists an $\bar{x}(\beta) \in \mathcal{X}$ such that $x \succ_x \bar{x}(\beta)$ implies $w(x, \beta) < w(\bar{x}(\beta), \beta)$.*

{ass:TooMuch}

Assumption 4 (No Countervailing Incentives). *There exists a constant u_R such that, for all $\beta \in B$, $u(x_0, \beta) = u_R$.*

{ass:NoCounterVail}

Assumption 5 (Minimum Element). *The no-trade allocation x_0 is the minimum element of $\mathcal{X}$ under the complete order $\succ_x$.*

{ass:MinElement}

The requirement of complete orders on the type space, B , and allocation space, $\mathcal{X}$, means little further loss of generality from treating each as a subset of the real line, $\mathbb{R}$. Henceforth, assume $B \subseteq \mathbb{R}$ and $\mathcal{X} \subseteq \mathbb{R}$. In general, expanding the analysis to allow for a (meaningful) multi-dimensional allocation space is difficult because of the issues involved in capturing how the buyer's willingness to make tradeoffs among the dimensions (including payment) varies with his type. The reader interested in multi-dimensional allocation spaces should consult Rochet and Choné (1998).²⁴

Given that $B \subseteq \mathbb{R}$ and $\mathcal{X} \subseteq \mathbb{R}$, it is meaningful to say β is a higher type than β' if $\beta > \beta'$. Similarly, the allocation x is greater than allocation x' if $x > x'$.

The Spence-Mirrlees condition (Assumption 1) says that if a lower type prefers the outcome (x, t) to (x', t') , $x > x'$, then a higher type will strictly prefer (x, t) to (x', t') . The Spence-Mirrlees condition has other interpretations. For instance, suppose $x > x'$ and $\beta > \beta'$. By choosing t and t' so that

$$u(x, \beta') - u(x', \beta') = t - t',$$

²³With this change in assumptions, Proposition 1 no longer holds if the buyer makes a TIOLI offer to the seller unless the seller knows the buyer's type.

²⁴Armstrong and Rochet (1999) and Basov (2010) are also useful references for those interested in multi-dimensional screening problems.

it follows, from (2.4), that

$$u(x, \beta) - u(x', \beta) > u(x, \beta') - u(x', \beta'); \quad (2.5) \quad \{\text{eq:IncreaseDif}\}$$

in other words, the Spence-Mirrlees condition implies $u(\cdot, \cdot)$ exhibits increasing differences. As is well known, increasing indifferences is equivalent to the condition that the marginal utility of allocation increase with type. If

$$u(x, \beta') - t \geq u(x', \beta') - t',$$

then

$$u(x, \beta') - u(x', \beta') \geq t - t';$$

hence, (2.5) implies (2.4). These arguments establish:

$\{\text{prop:SM-equivalence}\}$

Proposition 3. *The following are equivalent:*

- (i) *The Spence-Mirrlees condition holds.*
- (ii) *The function $u(\cdot, \cdot)$ exhibits strictly increasing differences.*
- (iii) *The marginal utility of allocation is greater for a higher type than a lower type.²⁵*

In light of Proposition 3, the following is immediate.

$\{\text{corr:CrossPartial-SM}\}$

Corollary 1. *Suppose $\mathcal{X}$ and B are intervals in $\mathbb{R}$. If the cross-partial derivative of $u(\cdot, \cdot)$ exists everywhere on $\mathcal{X} \times B$, then the Spence-Mirrlees condition holds if and only if*

$$\frac{\partial^2 u(x, \beta)}{\partial \beta \partial x} > 0. \quad (2.6) \quad \{\text{eq:SM-cross-partial}\}$$

Corollary 1 explains why the Spence-Mirrlees condition is sometimes referred to as the *cross-partial condition*.

Assumptions 2 and 3 ensure that trade is desirable with all types except, possibly, the lowest and trade never involves allocating an infinite amount. Assumption 4 says that all types of buyer enjoy the same utility, u_R , if there isn't trade. The value u_R is called the *reservation utility*. In many contexts, a common reservation utility is a reasonable assumption. For instance, if x denotes the amount of a good the buyer obtains, with $x_0 = 0$, then there is no obvious reason why different types would enjoy different levels of utility when they don't purchase. On the other hand, in other situations of contractual screening (*e.g.*,

²⁵If we think of a buyer's indifference curve in allocation-transfer (x - t space), its marginal rate of substitution is minus the marginal utility of allocation divided by the marginal utility of paying an additional dollar. The latter is -1 ; hence the marginal rate of substitution (MRS) equals the marginal utility of allocation. Given Spence-Mirrlees, a higher type has a greater MRS than a lower type. In other words, a higher type's indifference through a given point is steeper than a lower type's. Hence, an indifference curve for a higher type can cross that of a lower type at most once. This discussion explains why the Spence-Mirrlees condition is often referred to as a *single-crossing condition*.

an uninformed principal seeks to hire an informed agent and higher agent types have better alternatives *vis-à-vis* working for the principal than lower agent types), this assumption is less innocuous. Relaxing Assumption 4 requires a more extensive analysis than fits within this chapter. The reader interested in such models—in particular, so-called models of *countervailing incentives* in which $u(x_0, \cdot)$ is increasing in type—should consult Lewis and Sappington (1989) and Maggi and Rodriguez-Clare (1995), among other articles.

Assumption 5 reflects that the buyer is acquiring something more than what he would have absent trade. It is a natural assumption in contractual screening settings more generally (*e.g.*, no trade could constitute zero hours worked when an uninformed principal seeks to hire an informed agent).

2.1.3 CHARACTERIZING MECHANISMS

The Revelation Principle (Proposition 2) implies no loss of generality in restricting attention to direct-revelation mechanisms: the buyer's announcement of his type, β , maps to $(x(\beta), t(\beta))$ and, in equilibrium, the buyer announces truthfully. Because $(x(\beta), t(\beta))$ could equal $(x_0, 0)$, the no-trade allocation and “transfer,” there is no loss of generality in assuming that all types participate in equilibrium. These equilibrium conditions can be written as

$$u(x(\beta), \beta) - t(\beta) \geq u_R \text{ for all } \beta \in B \quad (\text{IR})$$

and

$$u(x(\beta), \beta) - t(\beta) \geq u(x(\beta'), \beta) - t(\beta') \text{ for all } \beta, \beta' \in B. \quad (\text{IC})$$

Condition (IR) is the requirement that all types participate. It is known in the literature as a participation or individual rationality (hence, IR) constraint. Condition (IC) is the requirement of truth telling in equilibrium. It is known in the literature as a truth-telling or incentive compatibility (hence, IC) constraint.

A well-known “trick” in mechanism design is to work with equilibrium utilities rather than directly with transfers. Observe a type- β buyer's equilibrium utility is

$$v(\beta) = u(x(\beta), \beta) - t(\beta). \quad (2.7) \quad \{\text{eq:Def-v-beta}\}$$

By adding and subtracting $u(x(\beta'), \beta')$ from the righthand side of (IC) and using (2.7), the IC constraint can be rewritten as

$$v(\beta) \geq v(\beta') + u(x(\beta'), \beta) - u(x(\beta'), \beta') \text{ for all } \beta, \beta' \in B. \quad (\text{IC}')$$

Consider two types, β and β' , $\beta > \beta'$. The relevant IC constraints must hold between these two (*i.e.*, β cannot do better to announce he's β' and *vice versa*):

$$v(\beta) \geq v(\beta') + u(x(\beta'), \beta) - u(x(\beta'), \beta') \text{ and} \quad (2.8) \quad \{\text{eq:RP1-beta}\}$$

$$v(\beta') \geq v(\beta) + u(x(\beta), \beta') - u(x(\beta), \beta). \quad (2.9) \quad \{\text{eq:RP1-beta-prime}\}$$

(This line of argumentation is known as a *revealed-preference argument*.) Expressions (2.8) and (2.9) together imply

$$u(x(\beta), \beta) - u(x(\beta), \beta') \geq v(\beta) - v(\beta') \geq u(x(\beta'), \beta) - u(x(\beta'), \beta'). \quad (2.10) \quad \{\text{eq:Pinch1}\}$$

Ignoring the middle term for the moment and recalling that the Spence-Mirrlees condition implies $u(\cdot, \cdot)$ exhibits increasing differences (Proposition 3), we have

{lemma:IncAll-necessary}

Lemma 1. *A necessary condition for a mechanism to be a direct-revelation mechanism (i.e., to induce truthful announcements in equilibrium) is that the allocation be nondecreasing with type (i.e., $x(\beta) \geq x(\beta')$ if $\beta > \beta'$).*

Recall that if the seller knew the buyer's type, her problem would be to choose x and t to maximize

$$t - c(x, \beta) \text{ subject to } u(x, \beta) - t \geq u_R.$$

As argued in the proof of Proposition 1, the constraint must bind. So if there were no asymmetry of information, the buyer would enjoy no rent (in equilibrium he would get just his reservation utility). With asymmetric information, that conclusion need no longer hold—some buyer types may earn a rent (have equilibrium utilities greater than their reservation utility). Because such a rent is due to the asymmetry of information, it is called an *information rent*. The following lemma is key to determining which types earn a rent and which don't.²⁶

{lemma:InfoRent}

Lemma 2. *Consider a direct-revelation mechanism such that $\beta' \in B$ is induced to buy; that is, such that $x(\beta') > x_0$. Then any higher-type buyer, $\beta \in B$, will also be induced to buy under this mechanism (i.e., $x(\beta) > x_0$) and will enjoy a greater equilibrium utility (i.e., $v(\beta) > v(\beta')$).*

An immediate corollary given the IR constraint is:

{corr:InfoRent}

Corollary 2. *Consider a direct-revelation mechanism such that $\beta' \in B$ is induced to buy. Then any higher-type buyer, $\beta \in B$, must earn an information rent in equilibrium under this mechanism.*

In other words, all types that trade, except possibly the lowest type who trades, capture an information rent in equilibrium.

What about the lowest type to trade? If the seller has designed the mechanism to maximize her expected payoff, then that buyer earns no rent:

{lemma:NoRentatBottom}

Lemma 3. *Consider a direct-revelation mechanism that maximizes the seller's expected payoff and assume a positive measure of types purchase in equilibrium. If, under this mechanism, $\tilde{\beta}$ is the lowest buyer type to buy (formally, $x(\beta) > x_0$, for all $\beta > \tilde{\beta}$; $x(\tilde{\beta}) \geq x_0$; and $x(\beta') = x_0$ for all $\beta' < \tilde{\beta}$), then the type- $\tilde{\beta}$ buyer captures no rent (i.e., $v(\tilde{\beta}) = u_R$).*

Further characterization of mechanisms is facilitated by—but not really dependent on—deciding whether the type space is discrete or continuous.

²⁶The proof of this lemma can be found in the Appendix. As a rule, proofs not given in the text can be found in the Appendix.

A Discrete Type Space. The type space is discrete if B is denumerable; that is, it can be put in a one-to-one mapping with an index set (set of consecutive integers), $\mathbb{Z}_B$. Let the type to which $n \in \mathbb{Z}_B$ maps be denoted β_n . The mapping is increasing. There is, thus, no type between β_n and β_{n+1} for any n and $n+1 \in \mathbb{Z}_B$. Call β_n and β_{n+1} *adjacent types*.

Because the index set is arbitrary, we can renormalize it in whatever way is convenient. In particular, adopt, for the moment, the convention that β_1 is the lowest type to purchase in equilibrium (*i.e.*, $x(\beta_1) > x_0$ and $x(\beta) = x_0$ for all $\beta < \beta_1$). Because the seller offers the mechanism on a TIOLI basis, we can restrict attention to mechanisms that maximize the seller's expected payoff; hence, from Lemma 3, $v(\beta_1) = u_R$.

For all $n \in \mathbb{Z}_B$ (except the minimum n should it exist), define the *rent function* as

$$R_n(x) = u(x, \beta_n) - u(x, \beta_{n-1}).$$

To understand this name, observe, from (2.8), that

$$v(\beta_n) \geq v(\beta_{n-1}) + R_n(x(\beta_{n-1})). \quad (2.11) \quad \{\text{eq:DownAdjIC}\}$$

Expression (2.11) is sometimes called the *downward adjacent IC constraint* (see, *e.g.*, Caillaud and Hermalin, 1993). Observe

$$v(\beta_2) \geq v(\beta_1) + R_2(x(\beta_1)) = u_R + R_2(x(\beta_1)),$$

$$v(\beta_3) \geq v(\beta_2) + R_3(x(\beta_2)) \geq u_R + \sum_{j=2}^3 R_j(x(\beta_{j-1})),$$

and, so on; hence,

$$v(\beta_n) \geq u_R + \sum_{j=2}^n R_j(x(\beta_{j-1})). \quad (2.12) \quad \{\text{eq:SumAdjIC}\}$$

One can, thus, describe $R_n(x(\beta_{n-1}))$ as the contribution to the type- β_n buyer's information rent necessary to keep him from mimicking his adjacent downward neighbor. (Admittedly, this would be clearer if (2.12) were an equality; fortunately, as shown below, it is under the seller's optimal mechanism.)

Recall $v(\beta) = u(x(\beta), \beta) - t(\beta)$. Hence, (2.12) implies

$$t(\beta_n) \leq u(x(\beta_n), \beta_n) - u_R - \sum_{j=2}^n R_j(x(\beta_{j-1})). \quad (2.13) \quad \{\text{eq:DiscreteEq-t-bound}\}$$

There are no countervailing incentives, so $R_j(x_0) \equiv 0$. By assumption $x(\beta_j) = x_0$ for $j < 1$. So, the inequality in (2.13) is not reversed if $\sum_{j \leq 1} R_j(x(\beta_{j-1}))$ is subtracted from the RHS of (2.13). This demonstrates the necessity of

$$t(\beta_n) \leq u(x(\beta_n), \beta_n) - u_R - \sum_{j \leq n} R_j(x(\beta_{j-1})). \quad (2.14) \quad \{\text{eq:DiscreteEq-t-bound-full}\}$$

The seller maximizes her revenue by having $t(\cdot)$ be as large as possible, which means, for any allocation profile $x(\cdot)$, she cannot do better than to have transfer function $t(\cdot)$ defined so that the inequality in (2.14) binds.

So far, only the necessary conditions for a direct-revelation mechanism have been characterized. Fortunately for purposes of calculating such mechanisms, these conditions are sufficient too:

Proposition 4. *Suppose the type space is discrete. A mechanism $\langle x(\cdot), t(\cdot) \rangle$ in which $x(\cdot)$ is nondecreasing and $t(\cdot)$ is given by*

$$t(\beta_n) = u(x(\beta_n), \beta_n) - u_R - \sum_{j \leq n} R_j(x(\beta_{j-1})). \quad (2.15) \quad \{\text{eq:t-discrete}\}$$

is a direct-revelation mechanism.

Proof: Consider an arbitrary type β_n . We wish to show he participates and won't lie about his type.

Increasing differences (Spence-Mirrlees) implies $R_n(x) \geq 0$ for all $x \geq x_0$. Hence, (2.15) implies

$$u(x(\beta_n), \beta_n) - t(\beta_n) = u_R + \sum_{j \leq n} R_j(x(\beta_{j-1})) \geq u_R;$$

that is, all types participate. Increasing differences also implies that $R_n(\cdot)$ is an increasing function.

Consider $n < m$. The goal is to show that a β_n -type buyer doesn't wish to pretend to be a β_m -type buyer. His utility were he to do so is

$$\begin{aligned} & u(x(\beta_m), \beta_n) - t(\beta_m) \\ &= u(x(\beta_m), \beta_n) - u(x(\beta_m), \beta_m) + u_R + \sum_{j \leq n} R_j(x(\beta_{j-1})) + \sum_{j=n+1}^m R_j(x(\beta_{j-1})) \\ &= u(x(\beta_m), \beta_n) - u(x(\beta_m), \beta_m) + v(\beta_n) + \sum_{j=n+1}^m R_j(x(\beta_{j-1})) \\ &= v(\beta_n) - \sum_{j=n+1}^m \left(R_j(x(\beta_m)) - R_j(x(\beta_{j-1})) \right) \leq v(\beta_n), \end{aligned} \quad (2.16) \quad \{\text{eq:discrete-TT1}\}$$

where the first two equalities follow from (2.15) and the third because

$$u(x(\beta_m), \beta_m) - u(x(\beta_m), \beta_n) = \sum_{j=n+1}^m R_j(x(\beta_m)).$$

The inequality in (2.16) follows because $x(\cdot)$ is nondecreasing and $R_j(\cdot)$ is an increasing function.

The proof for the case $n > m$ is in the Appendix. ■

A Continuous Type Space. The other common assumption is the type space is an interval $(\underline{\beta}, \bar{\beta}) \in \mathbb{R}$. Assume—primarily to avoid technical complications—that both limits are finite. To facilitate working with this space, take the buyer’s utility, $u(\cdot, \cdot)$, to be twice continuously differentiable in both arguments. Assume for any $x \in \mathcal{X}$ that $\partial u(x, \hat{\beta})/\partial \beta$ is bounded above for all $\hat{\beta} \in B$.

There is now no well-defined notion of adjacent types. However, an obvious calculus-based analogue to the rent function is the partial derivative

$$\frac{\partial u(x(\beta), \beta)}{\partial \beta}.$$

The corresponding analogue to the sum of rent functions (*e.g.*, as appears in (2.15) above) is the integral:

$$\int_{\underline{\beta}}^{\bar{\beta}} \frac{\partial u(x(z), z)}{\partial \beta} dz.$$

(Because z appears twice in that expression, it is readily seen that the integral is *not* a simple anti-derivative of the partial derivative.) Further reasoning by analogy suggests that the seller’s optimal direct-revelation mechanism is characterized by a nondecreasing allocation profile $x(\cdot)$ and transfer function

$$t(\beta) = u(x(\beta), \beta) - u_R - \int_{\underline{\beta}}^{\bar{\beta}} \frac{\partial u(x(z), z)}{\partial \beta} dz. \quad (2.17) \quad \{\text{eq:t-continuous}\}$$

This conclusion is, in fact, correct:

Proposition 5. *Suppose the type space is a continuous bounded interval in $\mathbb{R}$, that the buyer’s utility function is twice continuously differentiable in both arguments, and the partial derivative with respect to type is bounded above. A mechanism $\langle x(\cdot), t(\cdot) \rangle$ is a direct-revelation mechanism if and only if $x(\cdot)$ is nondecreasing and*

$\{\text{prop:SuffMechCont}\}$

$$t(\beta) = u(x(\beta), \beta) - \tau - \int_{\underline{\beta}}^{\bar{\beta}} \frac{\partial u(x(z), z)}{\partial \beta} dz, \quad (2.18) \quad \{\text{eq:t-continuous-gen1}\}$$

where $\tau \geq u_R$ is a constant.

Given a nondecreasing allocation profile, $x(\cdot)$, the seller can choose any $t(\cdot)$ satisfying (2.18) to implement it. She does better the greater is $t(\cdot)$, so it follows she wants τ as small as possible; that is, equal to u_R . Hence, as claimed, her expected-payoff-maximizing mechanism has a transfer function satisfying (2.17).

2.1.4 THE EQUILIBRIUM MECHANISM AND TARIFF

Consider, now, the question of the seller’s choice of mechanism to offer.

Two additional points before proceeding. First, it follows from the economic definition of cost that $c(x_0, \beta) \equiv 0 \ \forall \beta$. Second, for the case of a discrete type

space, “match” the continuous-space case by assuming the space is bounded above and below. Given the denumerability of the space, this is equivalent to assuming it has a finite number of elements, N . Using the ability to renormalize the index function, let the lowest type be β_1 .

Recall that $F : B \rightarrow [0, 1]$ is the distribution of buyer types. Given a mechanism $\langle x(\cdot), t(\cdot) \rangle$, the seller’s expected payoff is

$$\mathbb{E} \left\{ t(\beta) - c(x(\beta), \beta) \right\}.$$

In light of Proposition 4 or 5, as appropriate, there is no loss in limiting the seller to nondecreasing allocation profiles $x(\cdot)$ and transfer functions given by (2.15) or (2.17), as appropriate. Hence, her choice of mechanism is the program

$$\begin{aligned} \max_{x(\cdot)} \mathbb{E} \left\{ (u(x(\beta), \beta) - \int_{\{z \in B | z \leq \beta\}} R(z) dz - c(x(\beta), \beta)) \right\} - u_R \\ \equiv \max_{x(\cdot)} \mathbb{E} \left\{ (w(x(\beta), \beta) - \int_{\{z \in B | z \leq \beta\}} R(z) dz) \right\} - u_R \quad (2.19) \quad \{\text{eq: SellerProfit-mech-max}\} \end{aligned}$$

subject to $x(\cdot)$ ’s being nondecreasing; where $R(\beta)$ is the rent function (equal to $R_n(x(\beta_{n-1}))$ in the discrete case and equal to $\partial u(x(\beta), \beta) / \partial \beta$ in the continuous case) and $\int_{\{z \in B | z \leq \beta\}}$ is to be read as the appropriate summation notation in the discrete case.

Going forward let $f(\cdot)$ denote the density function implied by the distribution function $F(\cdot)$.²⁷ When B is a discrete space, it is often convenient to write f_n for $f(\beta_n)$ and *vice versa*. In the discrete case, there is no loss of generality in assuming that $f_n > 0$ for all n (to assume otherwise would be equivalent to assuming that type β_n simply didn’t exist—drop that type and reindex). The analogous assumption in the continuous case, $f(\beta) > 0$ for all $\beta \in (\underline{\beta}, \bar{\beta})$, is less general, but standard. Make both assumptions.

In the discrete case, observe

$$\begin{aligned} \sum_{n=1}^N \left(\sum_{k=2}^n R_k(x(\beta_{k-1})) \right) f_n &= \sum_{n=1}^N \left(\sum_{k=1}^{n-1} R_{k+1}(x(\beta_k)) \right) f_n \\ &= \sum_{n=1}^{N-1} \left(R_{n+1}(x(\beta_n)) \sum_{k=n+1}^N f_k \right) = \sum_{n=1}^{N-1} R_{n+1}(x(\beta_n)) (1 - F(\beta_n)) \end{aligned}$$

²⁷Observe, in the continuous-space case, this essentially implies the distribution function is differentiable (*i.e.*, a density function is defined for all $\beta \in B$). This assumption is readily relaxed; the added complexity, though, of doing so makes such a generalization too costly relative to its benefit to include in this chapter.

(note $\sum_{k=1}^0 \equiv 0$). Hence, in the discrete case, (2.19) becomes

$$\begin{aligned} \max_{x(\cdot)} \sum_{n=1}^{N-1} \left(w(x(\beta_n), \beta_n) f_n - R_{n+1}(x(\beta_n))(1 - F(\beta_n)) \right) + w(x(\beta_N), \beta_N) f_N \\ = \max_{x(\cdot)} \sum_{n=1}^N \left(w(x(\beta_n), \beta_n) - \frac{1 - F(\beta_n)}{f_n} R_{n+1}(x(\beta_n)) \right) f_n, \end{aligned} \quad (2.20) \quad \{\text{eq: SellerProfit-mech-max-disc}\}$$

where the fact that $1 - F(\beta_N) \equiv 0$ (the probability of drawing a type higher than the greatest is zero) permits the inclusion of the undefined term $R_{N+1}(x(\beta_N))$.

In the continuous case, observe, via integration by parts, that

$$\begin{aligned} \int_{\underline{\beta}}^{\bar{\beta}} \left(\int_{\underline{\beta}}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz \right) f(\beta) d\beta \\ = -(1 - F(\beta)) \int_{\underline{\beta}}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz \Big|_{\underline{\beta}}^{\bar{\beta}} + \int_{\underline{\beta}}^{\bar{\beta}} \frac{\partial u(x(\beta), \beta)}{\partial \beta} (1 - F(\beta)) d\beta \\ = \int_{\underline{\beta}}^{\bar{\beta}} \frac{\partial u(x(\beta), \beta)}{\partial \beta} (1 - F(\beta)) d\beta. \end{aligned}$$

Hence, in the continuous case, (2.19) becomes

$$\max_{x(\cdot)} \int_{\underline{\beta}}^{\bar{\beta}} \left(w(x(\beta), \beta) - \frac{1 - F(\beta)}{f(\beta)} \frac{\partial u(x(\beta), \beta)}{\partial \beta} \right) f(\beta) d\beta. \quad (2.21) \quad \{\text{eq: SellerProfit-mech-max-cont}\}$$

In both cases, underscoring the great similarity between them, the principal chooses an allocation profile to maximize

$$\mathbb{E} \left\{ w(x(\beta), \beta) - \frac{1 - F}{f} R \right\}$$

subject to $x(\cdot)$'s being nondecreasing. The expression inside the curly brackets—welfare generated by trade with a β -type buyer less a term reflecting the contribution the allocation $x(\beta)$ has on the information rents of types higher than β —is often referred to as *virtual surplus* (less often, virtual welfare). The information rent component reflects the inability of the seller to fully capture all the welfare generated by trade because she cannot escape “paying” an information rent to certain types. The ratio $(1 - F)/f$ is the multiplicative inverse of the hazard rate of the distribution of types.²⁸

Were it not for the constraint that the allocation profile be nondecreasing, the solution to the principal's mechanism-design problem could be found by

²⁸In statistics, the ratio $(1 - F)/f$ is known as the Mills ratio. That term is not widely used in economics, however.

point-wise optimization; that is, for each β , the optimal $x(\beta)$ would maximize the virtual surplus; that is, be the solution to

$$\max_{x \in \mathcal{X}} w(x, \beta) - \frac{1 - F(\beta)}{f(\beta)} R(x, \beta), \quad (2.22) \quad \{\text{eq:mech-pointwise-gen}\}$$

where $R(\cdot, \cdot)$ is the relevant rent function (*i.e.*, $R(x, \beta_n) = R_{n+1}(x)$ in the discrete case; and $R(x, \beta) = \partial u(x, \beta) / \partial \beta$ in the continuous case). Of course, if solving (2.22) for each β yields a profile that is nondecreasing, then we've solved the seller's problem—in such a case, the constraint on the allocation profile is simply not binding. In what follows, attention is limited to situations in which the nondecreasing-allocation condition is not binding.²⁹

Let $\Omega(x, \beta)$ denote virtual surplus (*i.e.*, the expression to be maximized in (2.22)). By the usual comparative statics, if $\Omega(\cdot, \cdot)$ exhibits increasing differences, then the x that solves (2.22) is necessarily nondecreasing in β and the constraint on the allocation profile can be ignored. This observation motivates the following series of assumptions.

Assumption 6. *The welfare function (*i.e.*, $w(x, \beta) = u(x, \beta) - c(x, \beta)$) exhibits increasing differences.* {\ass:welfare-incdif}

Observe Assumption 6 is implied by the Spence-Mirrlees condition (Proposition 3) if cost is invariant with respect to buyer type (an assumption true of most price-discrimination models).

Given Assumption 6, virtual surplus will exhibit increasing differences if, for $\beta > \beta'$ and $x > x'$,

$$\frac{1 - F(\beta)}{f(\beta)} (R(x, \beta) - R(x', \beta)) < \frac{1 - F(\beta')}{f(\beta')} (R(x, \beta') - R(x', \beta')), \quad (2.23) \quad \{\text{eq:part-virtsurp}\}$$

This observation motivates:

Assumption 7 (Monotone Hazard Rate Property). *The hazard rate associated with distribution of buyer types is nondecreasing in type.* {\ass:MHRP}

Assumption 8. *The rent function exhibits nonincreasing differences.* {\ass:rent-decdif}

As the discussion shows, one can conclude:

Lemma 4. *Given Assumptions 6–8, virtual surplus exhibits increasing differences.* {\lemma:noironing}

²⁹For many problems of interest in economics, the nondecreasing-allocation condition does not bind, which is why I'm limiting attention to that case here.

For the interested reader, let me briefly note that when the condition binds, the seller's optimal mechanism-design problem becomes an optimal-control problem. In particular, if $X(\cdot)$ is the schedule implied by solving (2.22) for each β and it is *not* nondecreasing, then the actual allocation schedule is a flattening—or, as it is typically described *ironing*—out of the hills and valleys in $X(\cdot)$ to achieve a non-decreasing allocation profile. Section 2.3.3.3 of Bolton and Dewatripont (2005) provides details on ironing.

As noted, Assumption 6 is, in many instances, implied by the Spence-Mirrlees condition. Because many common distributions (including the normal, uniform, and exponential) exhibit the monotone hazard rate property (MHRP), Assumption 7 is typically also seen as innocuous. Only Assumption 8 is difficult to justify as generally true because it does not correspond to any obvious *economic* principle (although it does not necessarily contradict any either). On the other hand, Assumptions 6–8 are merely sufficient conditions; it is possible that virtual surplus could exhibit increasing differences even if one or more of these assumptions fail.

Provided point-wise optimization is valid, the fact that there is zero probability of drawing a type greater than the highest type means

$$x(\beta_N) = \operatorname{argmax}_{x \in \mathcal{X}} w(x, \beta_N) \quad \text{and} \quad x(\bar{\beta}) = \operatorname{argmax}_{x \in \mathcal{X}} w(x, \bar{\beta})$$

for the discrete and continuous cases, respectively. This result is often described as *no distortion at the top*.

Proposition 6 (No distortion at the top). *Assume the seller can solve her mechanism-design problem via point-wise optimization (assume, e.g., Assumptions 6–8). Then the allocation provided the highest type is the welfare-maximizing allocation given that type.*

{prop:NoDistortTop}

On the other hand, if (i) $\beta \in B$ is not the highest type and (ii) all functions are differentiable, then

$$\frac{\partial \Omega(x, \beta)}{\partial x} = \frac{\partial w(x, \beta)}{\partial x} - \frac{1 - F(\beta)}{f(\beta)} \frac{\partial R(x, \beta)}{\partial x} < \frac{\partial w(x, \beta)}{\partial x}, \quad (2.24)$$

{eq:DistortBottom}

where the inequality follows because $\partial R(x, \beta)/\partial x > 0$ by Spence-Mirrlees. Consequently, if point-wise optimization is valid and $x(\beta)$ is an *interior* solution to (2.22), then $x(\beta)$ does *not* maximize welfare given that type. This is sometimes referred to as *distortion at the bottom*, which is slightly misleading insofar as this distortion affects *all* types other than the highest. Expression (2.24) also tells us the direction of the distortion: the type- β buyer will be allocated less than the welfare-maximizing amount given his type. To summarize:

Proposition 7. *Assume the seller can solve her mechanism-design problem via point-wise optimization (assume, e.g., Assumptions 6–8). Assume all functions are differentiable.³⁰ Consider a type other than the highest. If his allocation is an interior maximizer of virtual surplus, then his allocation is not welfare maximizing and is less than the welfare-maximizing amount.*

{prop:DistortBottom}

Intuitively, allocating more to a low-type buyer costs the seller insofar as it raises the information rent she must “pay” all higher types (equivalently, reduces what she can charge all higher types). Hence, she distorts the allocation to any type, but the highest, as she balances the gains from trade with that type against the forgone revenue from higher types.

³⁰It should be clear that analogous results can be derived when the functions aren’t differentiable. The value of such an extension is too small to justify its inclusion here.

2.1.5 EXAMPLES

{sect:PDexamples}

Linear Pricing. Recall the example, given at the beginning of this section, in which the buyer seeks at most one unit of the good and his benefit from the good is β , where $\beta \in [\underline{\beta}, \bar{\beta}]$, $\underline{\beta} \geq 0$ and the distribution of types is uniform. At its most general, an allocation, $x(\beta)$, is a probability of the good in question ending up in the buyer's hands. Hence, $\mathcal{X} = [0, 1]$, with $x_0 = 0$ (the buyer only gets the good if trade). Earlier, it was assumed that $c(x, \beta) \equiv 0$; let's generalize that slightly by assuming $c(x, \beta) \equiv xc \geq 0$, c a constant. Note $u(x, \beta) = x\beta$. The cross-partial derivative of that is 1; given $1 > 0$, the Spence-Mirrlees condition is met, as is Assumption 6. It is readily verified that Assumptions 2–5 are satisfied. The uniform distribution satisfies MHRP. Observe $R(x, \beta) = \partial u(x, \beta) / \partial \beta = x$, which trivially satisfies nonincreasing differences. Hence, the optimal mechanism can be found by point-wise maximization of virtual surplus:

$$\max_{x \in [0, 1]} x\beta - xc - \underbrace{(\bar{\beta} - \beta)x}_{\frac{1-F}{f}}.$$

It follows that $x(\beta) = 1$ if $2\beta \geq \bar{\beta} + c$ and $x(\beta) = 0$ otherwise. Observe if $\underline{\beta} \geq (\bar{\beta} + c)/2$, then the profit-maximizing mechanism entails trade with all types; this validates the claim made earlier that the seller's ignorance of the buyer's type need not necessarily result in inefficiency. Of course this problem is somewhat special insofar as maximizing virtual surplus always yields a corner solution and, hence, Proposition 7 does not apply. It was also claimed earlier that linear pricing is the profit-maximizing tariff. To confirm this, let $\tilde{\beta} = \max\{\underline{\beta}, (\bar{\beta} + c)/2\}$. Type $\tilde{\beta}$ is the lowest type to buy. Clearly, $u_R = 0$. From Lemma 3, $u_R = v(\tilde{\beta})$. The latter quantity is $\tilde{\beta} - t(\tilde{\beta})$. Hence, $t(\tilde{\beta}) = \tilde{\beta}$.³¹ Given $x(\beta) = x(\tilde{\beta}) = 1$ for all $\beta > \tilde{\beta}$, the IC constraint implies $t(\beta) = t(\tilde{\beta})$. So the tariff is $T(x) = \tilde{\beta}$ if $x > 0$ and $T(x) = 0$ if $x = 0$; that is, it is linear pricing with a price $p = \tilde{\beta}$.

Second-degree Price Discrimination via Quality Distortion.³² Many goods can be obtained in one of many versions or classes. This is true, for example, of software (business vs. home editions), travel (first vs. economy class), theater (orchestra vs. balcony seating), etc. In such cases, the seller is using quality differences to discriminate across buyers. To have a sense of this, suppose there are two types of buyer, β_1 and β_2 , $\beta_2 > \beta_1 > 1$. Let not receiving

³¹Of course, we can derive $t(\cdot)$ directly from (2.17):

$$t(\beta) = x(\beta)\beta - \int_{\underline{\beta}}^{\beta} x(z)dz = \begin{cases} \beta - \beta + \tilde{\beta} = \tilde{\beta}, & \text{if } \beta \geq \tilde{\beta} \\ 0 - 0 = 0, & \text{if } \beta < \tilde{\beta} \end{cases}.$$

³²In the usual labeling of price discrimination (*i.e.*, first, second, and third degree), discrimination in which the consumer must be induced to reveal his preferences is called second-degree discrimination. Such discrimination is, as seen, equivalent to tariff construction via mechanism design.

the good correspond to quality $x_0 \equiv 1$ and assume that $\mathcal{X} = [1, \infty)$. Let $u(x, \beta) = \beta \log(x)$ and $c(x) = x - 1$. Assumptions 1–5 are readily verified. The validity of pointwise optimization will be shown directly. Observe

$$x(\beta_1) = \operatorname{argmax}_{x \in \mathcal{X}} \beta_1 \log(x) - (x - 1) - \frac{1 - f_1}{f_1} \log(x) \quad \text{and} \quad (2.25) \quad \{\text{eq:qual-distort1}\}$$

$$x(\beta_2) = \operatorname{argmax}_{x \in \mathcal{X}} \beta_2 \log(x) - (x - 1). \quad (2.26) \quad \{\text{eq:qual-distort2}\}$$

Hence, $x(\beta_2) = \beta_2$ and

$$x(\beta_1) = \begin{cases} 1, & \text{if } f_1 \leq \frac{1}{\beta_1} \\ \frac{f_1(\beta_1 + 1) - 1}{f_1}, & \text{otherwise} \end{cases}.$$

The solution is nondecreasing, so pointwise optimization is valid. Note if low types are too small a proportion of the population, then the seller prefers to exclude them altogether (sometimes referred to as *shutting out the low types*). Hence, observing a seller offering a single version does not mean that she has eschewed second-degree price discrimination; she could be engaged in discrimination, but has chosen to shut out the low types. In some settings, the set $\mathcal{X}$ of possible qualities may be limited by technology (it could even be binary—an air traveler, *e.g.*, can be required to stay over Saturday night or not). Again the seller solves (2.25) and (2.26), but perhaps without being able to use calculus.

Second-degree Price Discrimination via Quantity Discounts. A familiar phenomenon is quantity discounts: the price ratio of a larger package to a smaller is less than their size ratio. To understand such pricing, suppose a buyer's benefit from x mobile-phone minutes is $\beta\gamma \log(x + 1)$, $\gamma > 0$ a known constant. Assume a constant marginal cost of supplying minutes, $c \in (0, 1)$. Assume β is distributed uniformly on $[0, 1]$. The various assumptions are readily verified for this example. Virtual surplus is

$$\beta\gamma \log(x + 1) - cx - \underbrace{(1 - \beta)}_{\frac{1-F}{f}} \underbrace{\gamma \log(x + 1)}_{\partial u / \partial \beta}.$$

The derivative of virtual surplus is

$$\frac{(2\beta - 1)\gamma - c(1 + x)}{1 + x}.$$

It immediately follows that

$$x(\beta) = \begin{cases} 0, & \text{if } \beta \leq \frac{\gamma + c}{2\gamma} \\ \frac{(2\beta - 1)\gamma - c}{c}, & \text{otherwise} \end{cases}.$$

Observe some low types are shut out. The transfer function is, from (2.17),

$$\begin{aligned} t(\beta) &= \beta\gamma \log(x(\beta) + 1) - \int_0^\beta \gamma \log(x(z) + 1) dz \\ &= \begin{cases} 0, & \text{if } \beta \leq \frac{\gamma+c}{2\gamma} \\ \frac{\gamma}{2} \left(2\beta - 1 + \log\left(\frac{\gamma(2\beta-1)}{c}\right) \right) - \frac{c}{2}, & \text{otherwise} \end{cases} \end{aligned}$$

Observe the transfer can be reexpressed as a traditional tariff:^{33,34}

$$T(x) = \frac{1}{2}(cx + \gamma \log(x + 1)). \quad (2.27) \quad \{\text{eq:Tariff-2ndQuantDisc}\}$$

It is readily verified that there are quantity discounts (*e.g.*, $2T(x) > T(2x)$).

2.2 THE MONOPOLY PROVISION OF INSURANCE

{sect:mono-insurance}

The analysis of the previous subsection covers a wide variety of trading situations and, via suitable redefinition of variables, an even wider array of contractual situations, but it does not cover all situations of interest. In particular, because the analysis exploits the buyer's quasi-linear utility, it is not suited to situations in which quasi-linear utility is a poor assumption. One such situation is an insurance market, where attitudes toward risk are of primary importance.

Suppose, now, the buyer's utility in a given state is $u(y_i)$, where y_i is his income in state i and $u(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is an increasing function (people prefer greater incomes to smaller incomes). A buyer's type is, now, his distribution over states. As an example, a buyer could have income y_H when healthy and y_I when injured, where $y_H > y_I$ because of lost wages, the need to pay for medical care, and so forth. His type, β , is his probability of injury.

Assume the buyer is risk averse: $u(\cdot)$ is strictly concave. Let $\mathbb{E}$ be the relevant expectation operator over income. Given a non-degenerate distribution

³³That a deterministic direct-revelation mechanism can always be reexpressed as a standard tariff is a general result known in the literature as the *taxation principle* (Rochet, 1985, attributes the name "taxation principle" to Roger Guesnerie). Formally if $x(B)$ is the set of allocations feasible under a direct-revelation mechanism and if it is possible to penalize the buyer sufficiently so that he would never choose an $x \notin x(B) \cup \{x_0\}$, then an equilibrium outcome under any deterministic direct-revelation mechanism defined by $\beta \mapsto (x(\beta), t(\beta))$, is also an equilibrium outcome of the game in which the seller allows the buyer to choose the allocation, x , in exchange for payment $T(x)$, where $T(\cdot)$ is defined by

$$T(x) = \begin{cases} 0, & \text{if } x = x_0 \\ t(\beta), & \text{if } \beta \in x^{-1}(x) \text{ (i.e., such that } x = x(\beta) \text{ for some } \beta \in B) \\ \text{buyer pays large penalty,} & \text{if } x \notin x(B) \cup \{x_0\} \end{cases}.$$

³⁴An alternative method of designing the profit-maximizing tariff to engage in price discrimination via quantity discounts is the so-called *demand profile* approach. This approach is particularly useful if the seller has access to actual data (*e.g.*, of the sort generated by grocery store checkout scanners). See Wilson (1993) for details.

over income, Jensen's inequality implies:

$$\mathbb{E}\{u(y)\} < u(\mathbb{E}\{y\}). \quad (2.28) \quad \{\text{eq:Jensen1}\}$$

Because $u(\cdot)$ is monotone, it is invertible. Define the *certainty equivalence* by

$$y_{\text{CE}} = u^{-1}(\mathbb{E}\{u(y)\}).$$

In words, if the buyer received y_{CE} for certain (*i.e.*, in all states), then his utility would be equal to his expected utility given the relevant distribution over income. Expression (2.28) implies $y_{\text{CE}} < \mathbb{E}\{y\}$.

2.2.1 THE BASICS OF INSURANCE

`\{sect:Basics_of_Insurance\}`

Consider a risk-neutral seller (*i.e.*, a party for whom the utility from a transfer of t is just t). As a benchmark, suppose the buyer's type were known to the seller prior to trade. The seller could offer to insure the buyer via a contract in which the buyer agreed to transfer his income in each state to the seller and the seller agreed to transfer a fixed $\hat{y}$ to the buyer in each state.³⁵ Hence, the buyer's utility would be $u(\hat{y})$ with certainty and the seller's expected utility would be $\mathbb{E}\{y\} - \hat{y}$ (normalize the seller's no-trade income to zero). Provided $\hat{y} \geq y_{\text{CE}}$, the buyer does at least as well trading as he would not trading. Provided $\mathbb{E}\{y\} \geq \hat{y}$, the seller does at least as well trading as she would not trading. Because the interval $(y_{\text{CE}}, \mathbb{E}\{y\})$ is nonempty, there exist values for $\hat{y}$ such that each side does at least weakly better to trade than not.

The last paragraph established that gains to trade exist. It didn't, though, establish that *full* insurance—a constant income for the risk-averse party across all states—is optimal. That full insurance is optimal is, however, readily shown.³⁶

`\{prop:Borch\}`

Proposition 8. *Let there be a discrete set of states. An insurance contract between risk-neutral and risk-averse parties that does not fully insure the risk-averse party is strictly Pareto dominated by one that does.*

Proof: Let there be N states, indexed by n . Let π_n denote the probability of state n occurring. Because otherwise state n does not really exist, we may take $\pi_n > 0$ for all n . Let y_n denote the risk-averse party's income in state n . An insurance contract, $\mathbf{Y} = \langle Y_1, \dots, Y_N \rangle$, yields the risk-averse party a payoff of Y_n in state n and the risk-neutral party a payoff of $y_n - Y_n$. Suppose $\mathbf{Y}$ were such that $Y_n \neq Y_m$ for some n and m . The buyer's expected utility is

$$\sum_{n=1}^N \pi_n u(Y_n) \equiv \bar{U}.$$

³⁵Obviously this contract can be replicated by one stated in the more familiar terms of premia and benefits. See footnote 38 *infra*.

³⁶The analysis here assumes a discrete set of states. The extension to a continuum of states is straightforward, but slightly more involved because of issues of measurability and the irrelevance of sets of states of measure zero. Because the economic intuition is the same, the extension is omitted for the sake of brevity.

If $\mathbf{Y}$ is *not* strictly Pareto dominated, then it must be a solution to the program

$$\max_{X_1, \dots, X_N} \sum_{n=1}^N \pi_n (y_n - X_n) \quad \text{subject to} \quad \sum_{n=1}^N \pi_n u(X_n) = \bar{U}. \quad (2.29) \quad \{\text{eq:Borch1}\}$$

Let λ denote the Lagrange multiplier on the constraint. The Lagrangian is

$$\sum_{n=1}^N \pi_n (y_n - X_n + \lambda u(X_n)) - \lambda \bar{U}.$$

The corresponding first-order condition is

$$-1 + \lambda u'(X_n) = 0 \quad \forall n. \quad (2.30) \quad \{\text{eq:Borch2}\}$$

Because $u'(\cdot)$ is strictly monotone, it follows that a necessary condition for a contract to solve (2.29) is $X_1 = \dots = X_N$. By assumption, $\mathbf{Y}$ does not satisfy this and is, thus, strictly dominated by the solution to (2.29). ■

Expression (2.30), the condition for a Pareto-optimal insurance contract, is sometimes referred to as the Borch (1968) sharing rule. In terms of economics, the Pareto-optimal contract must be one in which the ratio of the insurer's marginal utility of income to the insured's marginal utility of income is a constant across all states.³⁷ This makes sense: if the ratios varied, then income could be transferred across states in a way that made both parties better.

2.2.2 ADVERSE SELECTION

{sect:AdverseSelection1}

Now return to the situation in which there are multiple buyer types and the seller is ignorant of the buyer's type at the time of contracting.

Although the analysis could be done for more than two income states, limiting attention to two is with little loss because, to make the analysis tractable, a sufficiently strong order condition—essentially a variation of the Spence-Mirrlees condition—needs to be assumed. Given such an assumption, there is little further loss of generality in limiting attention to two states. Doing so also permits a graphical analysis. Call the two states healthy (H) and injured (I).

Assume that a buyer's injury probability, β , is the sole dimension of variation. Regardless of type, a buyer has income y_H if healthy and y_I if injured, $y_H > y_I$. A type- β buyer's expected income is $\bar{y}(\beta) = (1 - \beta)y_H + \beta y_I$. Let

$$y_{CE}(\beta) = u^{-1}((1 - \beta)u(y_H) + \beta u(y_I))$$

³⁷The rule generalizes to a mutual insurance arrangement between two risk-averse parties. Replace $y_n - X_n$ in (2.29) with $\psi(y_n - X_n)$, where $\psi(\cdot)$ is strictly increasing and strictly concave. Solving the new (2.29) would yield the following analogue of (2.30):

$$-\psi'(y_n - X_n) + \lambda u'(X_n) = 0 \quad \forall n \implies \frac{\psi'(y_n - X_n)}{u'(X_n)} = \lambda \quad \forall n.$$

denote his certainty equivalence. Note $y_{CE}(\cdot)$ is a decreasing function.

From Proposition 8, the seller, if she knew the buyer's type, would maximize her expected payoff by offering full insurance at the level $y_{CE}(\beta)$. When she doesn't know type, she cannot offer type-specific full insurance to multiple types: if $\beta' < \beta$, then, because $y_{CE}(\cdot)$ is a decreasing function, type β would do better to sign the type- β' contract. This potential for higher-risk insureds to pretend to be lower-risk insureds is known in the insurance literature as the problem of *adverse selection*. The seller's response to this adverse selection will be to offer a menu of insurance contracts, not all of which will provide full insurance.

The Revelation Principle (Proposition 2) still applies, so one can view the seller as designing a direct-revelation mechanism in which a buyer's announcement of type, β , maps to a contract $\langle Y_I(\beta), Y_H(\beta) \rangle$; that is, one in which the buyer transfers his income in state n , y_n , to the seller in exchange for $Y_n(\beta)$.³⁸

As earlier, a Spence-Mirrlees condition holds:

Lemma 5 (Spence-Mirrlees). *Suppose $\beta > \beta'$ and $Y_I > Y'_I$. Then if a β' -type buyer prefers $\langle Y_I, Y_H \rangle$ to $\langle Y'_I, Y'_H \rangle$, a β -type buyer strictly prefers $\langle Y_I, Y_H \rangle$.*

{lemma:SM-insurance}

Figure 1 illustrates Lemma 5. Note it also illustrates why the Spence-Mirrlees condition is known as a single-crossing condition.

Figure 1 also demonstrates that, in equilibrium, if the seller sells a lower-risk type (e.g., β') a policy (contract) with some amount of insurance, then the higher-risk type (e.g., β) could achieve strictly greater expected utility buying that policy than going without any insurance. For example, as shown, type β' is willing to buy **A** if offered and type β would enjoy strictly greater expected utility under **A** than if he went without any insurance. Consequently, it must be that if the seller serves lower-risk types in equilibrium, then higher-risk types earn an information rent, a result similar to Corollary 2. This establishes:

Proposition 9. *If a lower-risk type purchases an insurance contract in equilibrium, then all higher-risk types earn an information rent in equilibrium.*

{prop:Insurance-InfoRent}

Figure 2 illustrates the tradeoffs faced by the seller. Suppose there are only two types, β_h and β_ℓ —high and low risk types, respectively. Suppose the seller offered a full-insurance contract, denoted **F**, acceptable to the low-risk type. The high-risk would also purchase this contract and enjoy a considerable information rent. Suppose, instead, the seller moved along the low-risk type's indifference curve (the curve labeled $\mathcal{I}_{\beta_\ell}$) and offered a contract with less than full insurance (e.g., contract **D**). For a contract **D** near **F**, the efficiency loss from doing this would be second order given **F** is an optimum. But by doing this, the seller could offer the high-risk type a full-insurance contract that gave him a smaller information rent (e.g., contract **E**). Reducing the information

³⁸Observe we can readily reexpress this contract in the familiar form of a premium, t , the buyer pays in both states and a benefit, b , he receives in the insured-against state (e.g., sickness):

{fn:InsureContract}

$$t = y_H - Y_H \text{ and } b = y_H - Y_H + Y_I - y_I.$$

Working directly with income is simpler, though.

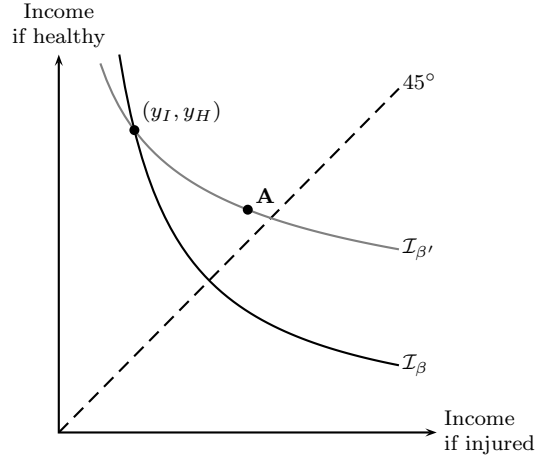


Figure 1: Illustration of Spence-Mirrlees Condition. An indifference curve, $\mathcal{I}_\beta$, for a high-risk type, β , and one, $\mathcal{I}_{\beta'}$, for a low-risk type, β' , through the uninsured point, (y_I, y_H) , are shown. Because the low-risk type prefers (weakly) the partial-insurance contract **A** to no insurance, the high-risk type strictly prefers the contract **A** to no insurance. fig:SM-insurance

rent represents a first-order gain for the seller. Because the efficiency loss is second order, while the rent-reduction benefit is first order, moving away from offering just **F** is in the seller's interest. The expected profit-maximizing choices of **D** and **E** depend on the relative proportions of high and low-risk types in the population. The greater the proportion of high-risk types, the closer **D** will be to the no-insurance point (y_I, y_H) . Reminiscent of Proposition 7 above, there is distortion at the bottom—to reduce the information rent of the high-risk type, the low-risk type is offered less than full insurance.

3 TWO-SIDED ASYMMETRIC INFORMATION

{sect:Two-Sided_Asymmetric_Info}

Now, consider a setting in which both buyer and seller have private information relevant to trade.

A key issue in the study of trading mechanisms is the point at which the buyer and seller become committed to play the mechanism. Figure 3 illustrates the three possible points. The first is *before* they learn their types. This corresponds to a situation in which there is symmetric uncertainty about how much the good or goods will cost to supply and the benefit it or they will yield. Such a situation arises when yet-to-be-realized market conditions will affect cost and benefit. Whether or not a party wishes to participate depends on his or her *expectation* of his or her type. When the participation decision is made prior to learning type, the situation is one of *ex ante* individual rationality.

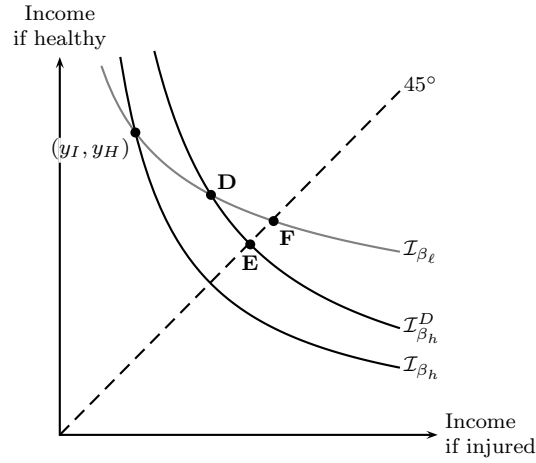


Figure 2: Illustration of Distortion of Low-Risk Type's Contract to Reduce Information Rent of High-Risk Type. Indifference curve $\mathcal{I}_{\beta_h}^D$ is the high-risk type's indifference curve through the contract **D** offered the low-risk type.
fig:insurance-solution

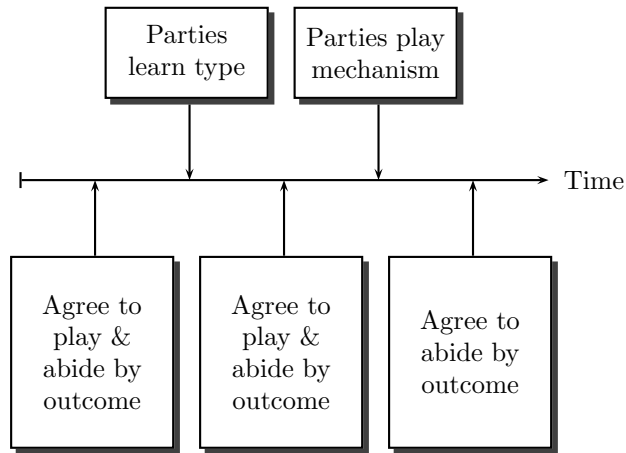


Figure 3: Possible Points of Commitment
fig:TradingTiming

The second point at which commitment could occur is *after* the parties learn their types, but before they play the mechanism. This is a situation in which either (i) contracting occurs after the parties learn their types or (ii) they learn their types after contracting, but can exit the relation without penalty before playing the mechanism. Interpretation (i) applies when a buyer and seller, who know their types, meet and seek to trade. Interpretation (ii) could be reasonable given certain aspects of real-life legal systems. For instance, courts may allow parties to walk away from contracts if they can prove that a change in circumstances makes the contract commercially impractical (*e.g.*, cost proves to be higher than expected) or purpose has been frustrated (*e.g.*, the buyer's benefit has significantly fallen).³⁹ The participation constraint at this point is known as interim individual rationality.

Finally, it might be possible for the parties to refuse to accept the results of the mechanism: the corresponding constraint that they accept the results is known as *ex post* individual rationality. In the literature, it is typically assumed that a mechanism (contract) will be enforced by some third party, which means the mechanism is *ex post* individually rational because neither party is willing to suffer the punishment the third party would impose for non-compliance. In reality, there are limits to enforcement. In such cases, those limits would be relevant to the design of the mechanism.

Another issue is whether the mechanism is *balanced*: transfers among the parties to the mechanism or contract always sum to zero. If a mechanism is *unbalanced*, then either an outside source must be providing funds if the sum of transfers is positive or the parties must be committed to ridding themselves of excess funds if the sum is negative. For most buyer-seller relations, it is unreasonable to imagine there is a third party willing to subsidize their trading. Similarly, it is unreasonable to imagine that the parties will actually rid themselves of excess funds ("burn money"). Should the mechanism call for them to rid themselves of excess funds, the parties will presumably renegotiate their agreement and divide the excess funds among themselves. In essence, one can think of burning money as violating a collective *ex post* rationality constraint.

That noted, one can, however, imagine a few trading relations in which unbalanced mechanisms could be reasonable. If buyer and seller are different divisions of the same firm, then headquarters could be willing to subsidize trade or vacuum up excess funds if need be.

3.1 TRADE SUBJECT TO EX ANTE INDIVIDUAL RATIONALITY

To begin, assume that the relevant participation constraint is *ex ante* individual rationality; that is, each party's *expected* payoff must exceed his or her no-trade payoff. In line with the previous analysis, normalize the no-trade payoff to zero for each party.

³⁹ An illustration of frustration of purpose is the case *Krell v. Henry*, in which the benefit of renting an apartment to watch a coronation parade was vastly reduced by the postponement of the coronation (Hermalin et al., 2007, p. 95).

3.1.1 SINGLE-UNIT EXCHANGE

Suppose the parties will exchange at most a single unit. At the time of contracting, both buyer and seller know that the buyer's value for the unit, b , will be drawn from $[0, \infty)$ according to the distribution F . Similarly, they both know that the seller's cost (equivalently, her value for the unit), c , will be drawn from $[0, \infty)$ according to the distribution G . Assume b and c are independent random variables. Once drawn, the buyer's value is his private information and likewise the seller's cost is her private information.

As before, a mechanism is an allocation rule $x : \mathbb{R}_+^2 \rightarrow [0, 1]$ and transfer rule. Limit attention to balanced mechanisms, so the amount the buyer pays is necessarily the seller's payment. Let $p : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ denote the seller's payment as a function of the buyer and seller's announcements.

Welfare is maximized provided trade occurs if and only if $b \geq c$. A mechanism is efficient therefore if

$$x(b, c) = \begin{cases} 0, & \text{if } b < c \\ 1, & \text{if } b \geq c \end{cases}.$$

A bit of intuition facilitates the mechanism-design problem. To wit, if the seller has been induced to reveal her cost, c , truthfully, then we only need the buyer, if exchange is to be efficient, to announce—having heard the seller's announced c —whether he wants the good, provided the mechanism is such that he wants it if and only if $b \geq c$. Buyer rationality implies he will do so if his payment to the seller is exactly c greater if he says he wants the good than if he says he doesn't. The issue then, at least with respect to truthful revelation, is to induce the seller to announce c truthfully.

Notice that the mechanism is such that one party's announcement follows—and can be conditioned on—the announcement of another party. Such mechanisms are known as *sequential mechanisms* for obvious reasons.⁴⁰

The buyer would behave efficiently if his payment were simply c if he buys. Linear pricing would not, however, induce truthful revelation from the seller: were the payment just c , then the seller would announce the $\hat{c}$ that solved

$$\max_{\hat{c}} \hat{c}(1 - F(\hat{c})) + cF(\hat{c}) \quad \text{equivalently} \quad \max_{\hat{c}} (\hat{c} - c)(1 - F(\hat{c})). \quad (3.1) \quad \{\text{eq:LP-trademech}\}$$

Lemma 6. *Unless her cost exceeds the buyer's maximum value, the seller would not announce her price truthfully if her payment were just her announced cost.*

Proof: By assumption, there exists a $c' > c$ such that $1 - F(c') > 0$. Clearly, either expression in (3.1) is strictly greater if $\hat{c} = c'$ than if $\hat{c} = c$. ■

Because, in a sequential mechanism, the buyer chooses whether to buy or not given the seller's announcement, we can write the payment as $p(x, c)$. Efficiency,

⁴⁰Cr  mer and Riordan (1985) was among the first use of sequential mechanisms (sometimes called *sequential-announcement mechanisms*) and consequently such mechanisms are sometimes referred to as Cr  mer-Riordan mechanisms.

as well as incentive compatibility for the buyer, requires that

$$b - p(1, c) \geq -p(0, c)$$

for all $b \geq c$ and

$$b - p(1, c) \leq -p(0, c)$$

for all $b < c$. This requirement can be met if and only if

$$c - p(1, c) = -p(0, c),$$

from which it follows that $p(\cdot, \cdot)$ must be of the form

$$p(x, c) = xc + T(c),$$

where $T(c)$ depends only on the seller's announcement and is independent of the buyer's purchase decision. If the seller claims her cost is $\hat{c}$ when it is truly c , then her expected profit (utility) is

$$\hat{c}(1 - F(\hat{c})) + cF(\hat{c}) + T(\hat{c}), \text{ alternatively } (\hat{c} - c)(1 - F(\hat{c})) + T(\hat{c}). \quad (3.2) \quad \{\text{eq:NoInvestment_Fill-in-the-price}\}$$

The first expression in (3.2) applies if c is the seller's value from consuming the item herself if no sale occurs, while the second applies if she only expends c if trade occurs (*e.g.*, in the latter case, c is the cost of manufacture). But since the same $\hat{c}$ will maximize either—see (3.1) above—the analysis applies to either alternative. (There is no “magic” here: the equivalence simply follows from the concept of opportunity cost.)

For convenience, consider the second alternative. Take any two types c and c' , $c > c'$. By the Revelation Principle, attention can be limited to mechanisms that induce the seller's truthful revelation. A standard revealed-preference argument implies

$$T(c) \geq (c' - c)(1 - F(c')) + T(c') \text{ and } T(c') \geq (c - c')(1 - F(c)) + T(c).$$

Combining these expressions yields

$$1 - F(c') \geq -\frac{T(c) - T(c')}{c - c'} \geq 1 - F(c).$$

This suggests—via integration—the following:⁴¹

$$T(c) = T(0) - \int_0^c (1 - F(z)) dz. \quad (3.3) \quad \{\text{eq:Fill-in-the-price_mechNI}\}$$

⁴¹The mechanism derived here was originally derived in Hermalin and Katz (1993), where it was referred to as a *fill-in-the-price mechanism*.

Remark 1. Recognize that the argument to this point does *not* establish that it is sufficient to limit attention to mechanisms in which $T(\cdot)$ is given by (3.3). Consequently, if this mechanism failed to maximize welfare, then one couldn't conclude that no mechanism would. Fortunately, as will be seen shortly, this mechanism does maximize welfare.

Remark 2. The expression $\int_0^c (1 - F(z))dz$ has an economic interpretation. Recall the survival function, $1 - F(\cdot)$, is equivalent to a demand curve. The area beneath a demand curve between two prices p' and p is the amount by which consumer benefit is reduced if the price is raised from p' to p . Recall the effective price faced by the buyer is the seller's announced cost $\hat{c}$ (the $T(\hat{c})$ component is sunk from the buyer's perspective). If the seller announces $\hat{c} > c$, the latter being her true cost, then she reduces consumer benefit by $\int_c^{\hat{c}} (1 - F(z))dz$. But this is $T(c) - T(\hat{c})$; that is, she reduces her transfer by exactly the amount she reduces the consumer's benefit. This causes her to internalize the externality that her announcement has on the buyer.

The mechanism is incentive compatible for the seller if the seller solves

$$\max_{\hat{c}} (\hat{c} - c)(1 - F(\hat{c})) + T(\hat{c}) \quad (3.4) \quad \{\text{eq: SellerProfit_FIP_NI}\}$$

by choosing $\hat{c} = c$ for all c . To see that is her choice, suppose it were not. Then there would exist a c and $\hat{c}$ such that

$$(\hat{c} - c)(1 - F(\hat{c})) + T(\hat{c}) > T(c).$$

Suppose $\hat{c} > c$, then this last expression implies

$$1 - F(\hat{c}) > \frac{1}{\hat{c} - c} \int_c^{\hat{c}} (1 - F(z))dz = 1 - F(\tilde{c}) \quad (3.5) \quad \{\text{eq:FIP_NIa1}\}$$

where $\tilde{c} \in (c, \hat{c})$. (Such a $\tilde{c}$ exists by the mean-value theorem.) Survival functions are non-increasing, hence (3.5) is impossible. *Reductio ad absurdum*, the seller wouldn't prefer to announce $\hat{c}$ rather than c . A similar argument reveals that there is no $\hat{c} < c$ that the seller would rather announce than c .

To summarize to this point: the sequential mechanism in which the seller announces her cost, c , and the buyer then decides to buy or not, with the buyer's payment being $p(x, c) = xc + T(c)$, with $T(\cdot)$ given by (3.3), induces the seller to announce her cost truthfully in equilibrium and induces the buyer to buy if and only if his valuation, b , exceeds the seller's cost. The one remaining issue is can $T(0)$ be set so as to satisfy *ex ante* IR for both buyer and seller? The seller and buyer's respective expected utilities under this mechanism are

$$U_S = \int_0^\infty T(c)dG(c) = T(0) - \int_0^\infty \left(\int_0^c (1 - F(b))db \right) dG(c)$$

and

$$\begin{aligned} U_B &= \int_0^\infty \left(-T(c) + \int_c^\infty (b - c) dF(b) \right) dG(c) \\ &= -U_S + \int_0^\infty \int_c^\infty (b - c) dF(b) dG(c). \end{aligned} \quad (3.6) \quad \{\text{eq:BuyerUtility_FIP_NI}\}$$

Set

$$T(0) = \int_0^\infty \left(\int_0^c (1 - F(b)) db \right) dG(c), \quad (3.7) \quad \{\text{eq:T0-FIP-NI}\}$$

making U_S zero; IR for the seller is satisfied. Plugging that into (3.6) reveals $U_B \geq 0$, so the buyer's IR is also satisfied. To conclude:

Proposition 10. *Assume a buyer and seller wish to exchange a single unit of a good efficiently. Assume, too, that the parties can enter into a contract prior to learning their valuation and cost. Then a mechanism exists that achieves first-best efficiency and is ex ante individually rational for both parties. An example of such a mechanism is one in which the seller announces her cost, c , and the buyer subsequently decides to purchase ($x = 1$) or not ($x = 0$), with the buyer's payment to the seller equal to $xc + T(c)$, where $T(\cdot)$ is defined by expressions (3.3) and (3.7).*

$\{\text{prop:FIP_NI-efficiency}\}$

3.1.2 MULTI-UNIT EXCHANGE

Now consider the exchange of multiple units of a good. Assume the buyer's utility is $U(x, \beta) - t$ and the seller's $t - C(x, \gamma)$, where β and γ are the buyer and seller's types, respectively; and, as before, x and t are units of the good and a monetary transfer, respectively. Assume that U and C are twice continuously differentiable in both arguments. Assume β and γ are independent random variables distributed, respectively, $F : [\beta_L, \beta_H] \rightarrow [0, 1]$ and $G : [\gamma_L, \gamma_H] \rightarrow [0, 1]$.

Welfare is

$$W(x, \beta, \gamma) = U(x, \beta) - C(x, \gamma).$$

Assume, for all β and γ , that $U(\cdot, \beta)$ is a concave function (buyer's marginal benefit is non-increasing) and $C(\cdot, \gamma)$ a convex function (non-increasing returns to scale), with one function at least strictly so. Hence, $W(\cdot, \beta, \gamma)$ is strictly concave for all β and γ . There is, thus, a unique welfare maximizing amount of trade, $x^*(\beta, \gamma)$, for all β and γ . To insure interior maxima assume:

- $\partial U(0, \beta) / \partial x > 0$ (marginal benefit at 0 is positive) for all $\beta > \beta_L$;
- $\partial C(0, \gamma) / \partial x = 0$ (marginal cost at 0 is 0) for all $\gamma < \gamma_H$;
- For all β and γ , there exists a finite $\bar{x}(\beta, \gamma)$ such that $x > \bar{x}(\beta, \gamma)$ implies $\partial W(x, \beta, \gamma) / \partial x < 0$ (infinite trade is never desirable).

As in the previous subsection, the goal is an efficient sequential mechanism. To “mix things up,” let the buyer be the one to announce first. The objective is

for the seller to respond to the buyer's announcement, $\hat{\beta}$, by truthfully revealing her type, γ ; that is, the solution to

$$\max_{\gamma} p(\hat{\beta}, \gamma) - C(x(\hat{\beta}, \gamma), \gamma) \quad (3.8) \quad \{\text{eq:MU-FIP-SellerObj}\}$$

needs to be $\hat{\gamma} = \gamma$ for all γ and $\hat{\beta}$. This can be achieved by defining

$$x(\beta, \gamma) = x^*(\beta, \gamma) \quad \text{and} \quad p(\beta, \gamma) = U(x^*(\beta, \gamma), \beta) - \tau(\beta) \quad (3.9) \quad \{\text{eq:MU-FIP-mech}\}$$

(the role of $\tau : [\beta_L, \beta_H] \rightarrow \mathbb{R}$ will be made clear below). To verify (3.9) induces truth-telling by the seller, observe (3.8) is equivalent to

$$\max_{x \in x^*(\hat{\beta}, [\gamma_L, \gamma_H])} U(x, \hat{\beta}) - C(x, \gamma), \quad (3.10) \quad \{\text{eq:MU-FIP-SellerObj2}\}$$

where $x^*(\hat{\beta}, [\gamma_L, \gamma_H])$ is the *image* of $x^*(\hat{\beta}, \cdot)$.⁴² Because $x^*(\hat{\beta}, \gamma)$ is the unconstrained maximizer of (3.10) and it's in $x^*(\hat{\beta}, [\gamma_L, \gamma_H])$, a best response for the seller is to announce her type truthfully.

The additional function, $\tau(\cdot)$, is necessary to induce truth-telling by the buyer. The buyer's expected utility in equilibrium is

$$\tau(\hat{\beta}) + \int_{\gamma_L}^{\gamma_H} \left(U(x^*(\hat{\beta}, \gamma), \beta) - U(x^*(\hat{\beta}, \gamma), \hat{\beta}) \right) dG(\gamma),$$

if his type is β but he announces $\hat{\beta}$. His utility is $\tau(\beta)$ if he tells the truth. Consider $\beta > \beta'$. By revealed preference:

$$\begin{aligned} \tau(\beta) &\geq \tau(\beta') + \int_{\gamma_L}^{\gamma_H} \left(U(x^*(\beta', \gamma), \beta) - U(x^*(\beta', \gamma), \beta') \right) dG(\gamma) \quad \text{and} \\ \tau(\beta') &\geq \tau(\beta) + \int_{\gamma_L}^{\gamma_H} \left(U(x^*(\beta, \gamma), \beta') - U(x^*(\beta, \gamma), \beta) \right) dG(\gamma). \end{aligned}$$

Rearranging,

$$\begin{aligned} \int_{\gamma_L}^{\gamma_H} \frac{U(x^*(\beta, \gamma), \beta') - U(x^*(\beta, \gamma), \beta)}{\beta - \beta'} dG(\gamma) &\geq \frac{\tau(\beta) - \tau(\beta')}{\beta - \beta'} \\ &\geq \int_{\gamma_L}^{\gamma_H} \frac{U(x^*(\beta', \gamma), \beta) - U(x^*(\beta', \gamma), \beta')}{\beta - \beta'} dG(\gamma) \end{aligned}$$

Via the implicit function theorem, it is readily shown that x^* is continuous in each of its arguments. One can, therefore, take the limit of the outer expressions as $\beta' \rightarrow \beta$. This yields⁴³

$$\tau'(\beta) = \int_{\gamma_L}^{\gamma_H} \frac{\partial U(x^*(\beta, \gamma), \beta)}{\partial \beta} dG(\gamma).$$

⁴²Recall that the *image* of a function is the set of all values the function can take.

⁴³Because the question is the sufficiency of the mechanism, there is no loss in assuming $\tau(\cdot)$ differentiable provided an efficient mechanism is derived.

Consequently,

$$\tau(\beta) = \tau(0) + \int_{\gamma_L}^{\gamma_H} \int_{\beta_L}^{\beta} \frac{\partial U(x^*(z, \gamma), z)}{\partial \beta} dz dG(\gamma). \quad (3.11) \quad \{\text{eq:tau-FIP-MU1}\}$$

We want to simplify (3.11). To that end, observe the envelope theorem entails

$$\frac{d}{d\beta} W(x^*(\beta, \gamma), \beta, \gamma) \equiv \frac{\partial U(x^*(\beta, \gamma), \beta)}{\partial \beta}.$$

Consequently, (3.11) can be rewritten as

$$\begin{aligned} \tau(\beta) &= \tau(0) + \int_{\gamma_L}^{\gamma_H} \left(W(x^*(\beta, \gamma), \beta, \gamma) - W(x^*(\beta_L, \gamma), \beta_L, \gamma) \right) dG(\gamma) \\ &= \xi(0) + \int_{\gamma_L}^{\gamma_H} W(x^*(\beta, \gamma), \beta, \gamma) dG(\gamma). \end{aligned} \quad (3.12) \quad \{\text{eq:tau-FIP-MU2}\}$$

It remains to be verified that this mechanism induces truth-telling. Consider $\hat{\beta} \neq \beta$, the latter being the buyer's true type. Is

$$\tau(\beta) \geq \tau(\hat{\beta}) + \int_{\gamma_L}^{\gamma_H} \left(U(x^*(\hat{\beta}, \gamma), \beta) - U(x^*(\hat{\beta}, \gamma), \hat{\beta}) \right) dG(\gamma)?$$

Suppose it weren't. Substituting for $\tau(\cdot)$ yields:

$$\begin{aligned} \int_{\gamma_L}^{\gamma_H} W(x^*(\beta, \gamma), \beta, \gamma) dG(\gamma) &< \int_{\gamma_L}^{\gamma_H} \left(U(x^*(\hat{\beta}, \gamma), \beta) - C(x^*(\hat{\beta}, \gamma), \gamma) \right) dG(\gamma) \\ &= \int_{\gamma_L}^{\gamma_H} W(x^*(\hat{\beta}, \gamma), \beta, \gamma) dG(\gamma) \end{aligned}$$

But this is a contradiction because $x^*(\beta, \gamma)$ maximizes $W(x, \beta, \gamma)$. *Reductio ad absurdum*, the supposition is false and the mechanism induces truth-telling.

The intuition for why the mechanism works can be seen from (3.9): the buyer's expected payment is

$$\int_{\gamma_L}^{\gamma_H} p(\hat{\beta}, \gamma) dG(\gamma) = \int_{\gamma_L}^{\gamma_H} C(x^*(\hat{\beta}, \gamma), \gamma) dG(\gamma) - \xi(0);$$

that is, he pays the expected cost of providing what would be the welfare-maximizing quantity were his type $\hat{\beta}$. Consequently, he maximizes actual expected welfare if he tells the truth, but fails to do so if he lies. Effectively, the buyer is made to face the social planner's optimization problem and, so, made to maximize welfare.

Finally, it needs to be verified that a $\tau(0)$ exists such that both buyer and seller are willing to participate. To that end, let

$$\tau(0) = \int_{\gamma_L}^{\gamma_H} \left(W(x^*(\beta_L, \gamma), \beta_L, \gamma) - \int_{\beta_L}^{\beta_H} W(x^*(\beta, \gamma), \beta, \gamma) dF(\beta) \right) dG(\gamma). \quad (3.13) \quad \{\text{eq:FIP-tau0}\}$$

It is readily seen that the seller's expected utility is positive—there are, in expectation, positive gains to trade and these expected gains to trade equal the sum of the buyer's and seller's expected utilities. To summarize:

Proposition 11. *Assume buyer and seller wish to exchange the welfare-maximizing quantity of some good, where the welfare-maximizing quantity depends on their types. Assume, too, that the parties can enter into a contract prior to learning their types. Then a mechanism exists that achieves first-best efficiency and is ex ante individually rational for both parties. An example of such a mechanism is given by expressions (3.9), (3.12), and (3.13).*

3.2 TRADE SUBJECT TO INTERIM INDIVIDUAL RATIONALITY

Suppose, now, a mechanism must satisfy interim individual rationality. Recall, this means that each party must wish to play the mechanism *knowing* his or her type. The no-trade payoffs continue to be zero for each party.

Limit attention to single-unit exchange. The result will prove to be that *no* balanced mechanism yields welfare-maximizing exchange for all buyer and seller types, a finding due originally to Myerson and Satterthwaite (1983).

At the time of contracting, both buyer and seller know the buyer's value for the unit, b , was drawn from $[0, \bar{b}]$ according to the distribution F . Only the buyer knows what value was drawn. Similarly, they both know the seller's cost (equivalently, value for the unit), c , was drawn from $[\underline{c}, \bar{c}]$ according to the distribution G . Only the seller knows what cost was drawn. Assume b and c are independent random variables. Assume both distributions F and G are differentiable. Let their derivatives (density functions) be denoted by f and g , respectively. Assume full support, so that $f(b) > 0$ for all $b \in [0, \bar{b}]$ and $g(c) > 0$ for all $c \in [\underline{c}, \bar{c}]$. Because trade is never efficient if $\underline{c} \geq \bar{b}$, assume $\bar{b} > \underline{c}$. Given c is a cost, $c \geq 0$; hence, $\bar{c} > 0$.

As before, a mechanism is an allocation rule $x : \mathbb{R}_+^2 \rightarrow [0, 1]$ and a transfer rule $p : \mathbb{R}_+^2 \rightarrow \mathbb{R}$.

Efficiency requires trade occur if and only if $b \geq c$. A mechanism is efficient, therefore, if and only if it satisfies

$$x(b, c) = \begin{cases} 0, & \text{if } b < c \\ 1, & \text{if } b \geq c \end{cases}.$$

To determine whether an efficient balanced mechanism exists, one either needs to derive a mechanism that works (the strategy employed in the previous subsection) or characterize the entire set of mechanisms and show no element of this set works (the strategy to be pursued here). From the Revelation Principle, attention can be limited to direct-revelation mechanisms. The first step is to characterize the set of balanced direct-revelation mechanisms; that is, mechanisms in which truth-telling by one party is a best response to truth-telling by the other and *vice versa*.

Define

$$\xi_B(b) = \int_{\underline{c}}^{\bar{c}} x(b, c)g(c)dc \quad \text{and} \quad \xi_S(c) = \int_0^{\bar{b}} x(b, c)f(b)db.$$

The quantity $\xi_i(\theta)$ is the probability of trade given party i 's announced type, θ , and truth-telling by the other party. Similarly, define

$$\rho_B(b) = \int_{\underline{c}}^{\bar{c}} p(b, c)g(c)dc \quad \text{and} \quad \rho_S(c) = \int_0^{\bar{b}} p(b, c)f(b)db.$$

The quantity $\rho_i(\theta)$ is the expected payment given party i 's announced type and truth-telling by the other party.

Let $u(b)$ and $\pi(c)$ be, respectively, the buyer and seller's expected utilities if they announce their types truthfully in equilibrium. Hence,

$$u(b) = b\xi_B(b) - \rho_B(b) \quad \text{and} \quad (3.14) \quad \{\text{eq:MS-defU}\}$$

$$\pi(c) = \rho_S(c) - c\xi_S(c). \quad (3.15) \quad \{\text{eq:MS-defpi}\}$$

A mechanism induces truth-telling and satisfies interim IR if and only if, for all $b, b' \in [0, \bar{b}]$ and all $c, c' \in [0, \bar{c}]$, we have

$$u(b) \geq b\xi_B(b') - \rho_B(b'), \quad (\text{IC}_B)$$

$$\pi(c) \geq \rho_S(c') - c\xi_S(c'), \quad (\text{IC}_S)$$

$$u(b) \geq 0, \quad \text{and} \quad (\text{IR}_B)$$

$$\pi(c) \geq 0. \quad (\text{IR}_S)$$

By revealed preference:

$$u(b) \geq u(b') + \xi_B(b')(b - b') \quad \text{and}$$

$$u(b') \geq u(b) + \xi_B(b)(b' - b).$$

Combining these expressions yields:

$$\xi_B(b)(b - b') \geq u(b) - u(b') \geq \xi_B(b')(b - b'). \quad (3.16) \quad \{\text{eq:MS-pinch-buyer}\}$$

A similar revealed-preference argument yields:

$$\xi_S(c')(c - c') \geq \pi(c') - \pi(c) \geq \xi_S(c)(c - c'). \quad (3.17) \quad \{\text{eq:MS-pinch-seller}\}$$

Suppose $b > b'$. It follows from (3.16) that $\xi_B(b) \geq \xi_B(b')$ —a necessary condition for the mechanism to be incentive compatible is that the probability of trade be non-decreasing in the buyer's type. Similarly, suppose $c > c'$. It follows from (3.17) that a necessary condition for the mechanism to be incentive compatible is that the probability of trade be non-increasing in the seller's cost.

The functions $u(\cdot)$ and $\pi(\cdot)$ are absolutely continuous.⁴⁴ It follows that they are differentiable almost everywhere. Where their derivatives exist, taking limits reveals that

$$u'(b) = \xi_B(b) \text{ and } \pi'(c) = -\xi_S(c).$$

Finally, because $u(\cdot)$ and $\pi(\cdot)$ are absolutely continuous, they can be expressed as the indefinite integral of their derivative (Yeh, 2006, Theorem 13.17, p. 283):

$$u(b) = u(0) + \int_0^b \xi_B(z) dz \text{ and} \quad (3.18) \quad \{\text{eq:MS-U-indefint}\}$$

$$\pi(c) = \pi(\bar{c}) + \int_c^{\bar{c}} \xi_S(z) dz. \quad (3.19) \quad \{\text{eq:MS-pi-indefint}\}$$

Because probabilities are non-negative, (3.18) implies that $u(b) \geq 0$ if $u(0) \geq 0$. Likewise, (3.19) implies that $\pi(c) \geq 0$ if $\pi(\bar{c}) \geq 0$. This analysis yields:

$\{\text{lemma:MS-necessity}\}$

Lemma 7. *If a mechanism is incentive compatible, then (i) the buyer's perceived probability of trade given his value (i.e., $\xi_B(\cdot)$) is non-decreasing in his value; (ii) the seller's perceived probability of trade given her cost (i.e., $\xi_S(\cdot)$) is non-increasing in her cost; (iii) the buyer's equilibrium expected utility conditional on his value is given by (3.18); and (iv) the seller's equilibrium expected utility conditional on her cost is given by (3.19). Moreover, necessary and sufficient conditions for the mechanism to be interim individually rational are that the buyer with the lowest value wish to participate and the seller with the highest cost wish to participate (i.e., $u(0) \geq 0$ and $\pi(\bar{c}) \geq 0$).*

What about sufficiency with respect to incentive compatibility? Suppose a mechanism satisfies conditions (i)–(iv) of Lemma 7. Suppose the mechanism were not IC for the buyer, then there would exist b and b' such that

$$u(b) < u(b') + \xi_B(b')(b - b'). \quad (3.20) \quad \{\text{eq:MS-Buyer_suff0}\}$$

Substituting for $u(\cdot)$ and canceling like terms, this last expression implies

$$\int_{b'}^b \xi_B(z) dz < \xi_B(b')(b - b') \quad (3.21) \quad \{\text{eq:MS-Buyer_suff1}\}$$

if $b > b'$ or

$$\int_b^{b'} \xi_B(z) dz > \xi_B(b')(b' - b) \quad (3.22) \quad \{\text{eq:MS-Buyer_suff2}\}$$

if $b < b'$. By the intermediate value theorem, (3.21) and (3.22) imply

$$\xi_B(\tilde{b})(b - b') < \xi_B(b')(b - b') \text{ and} \quad (3.23) \quad \{\text{eq:MS-Buyer_suff3}\}$$

$$\xi_B(\tilde{b})(b' - b) > \xi_B(b')(b' - b), \quad (3.24) \quad \{\text{eq:MS-Buyer_suff4}\}$$

⁴⁴Proof: Because ξ is a probability, it is less than one. Hence, for all b, b', c , and c' , it follows from (3.16) and (3.17), respectively, that $|u(b) - u(b')| < |b - b'|$ and $|\pi(c) - \pi(c')| < |c - c'|$. The functions $u(\cdot)$ and $\pi(\cdot)$ thus satisfy the Lipschitz condition. The result follows because all Lipschitzian functions are absolutely continuous (see, e.g., van Tiel, 1984, p. 5).

respectively, where $\tilde{b} \in (\min\{b', b\}, \max\{b', b\})$. But expression (3.23) implies $\xi_B(\tilde{b}) < \xi_B(b')$, which contradicts the fact that $\xi_B(\cdot)$ is non-decreasing (recall (3.23) is the relevant expression if $b' < b$ and, thus, $b' < \tilde{b}$). Similarly, (3.24) leads to a contradiction. Given that (3.20) leads to a contradiction, it follows that it cannot hold and, thus, that the mechanism cannot fail to be IC. This establishes that conditions (i) and (iii) of Lemma 7 are sufficient for a mechanism to induce truth-telling by the buyer. A similar analysis holds for the seller. To summarize:

{lemma:MS-sufficiency}

Lemma 8. *If a mechanism is such that (i) the buyer's perceived probability of trade given his value (i.e., $\xi_B(\cdot)$) is non-decreasing in his value; (ii) the seller's perceived probability of trade given her cost (i.e., $\xi_S(\cdot)$) is non-increasing in her cost; (iii) the buyer's equilibrium expected utility conditional on his value is given by (3.18); and (iv) the seller's equilibrium expected utility conditional on her cost is given by (3.19), then the mechanism is a direct-revelation mechanism.*

The following can now be shown:

{prop:Myerson-Satterthwaite}

Proposition 12 (Myerson-Satterthwaite). *No balanced interim individually rational mechanism exists that achieves the first best in a setting in which a single unit is potentially to be exchanged and the buyer's value and seller's cost are continuously distributed with full support over overlapping intervals.*

Given that an efficient mechanism doesn't exist, the next question is what is the *second-best* mechanism? To answer, it is helpful to restate the interim IR constraint: Lemma 8 implies a necessary condition for interim IR is that $u(0) + \pi(\bar{c}) \geq 0$. This condition is also *sufficient*. Why? Well if one component, say $u(0)$, were negative, the other must be positive; moreover, $-u(0)$ in surplus can be shifted from the seller to the buyer to obtain new constants of integration:

$$\tilde{u}(0) = 0 \quad \text{and} \quad \tilde{\pi}(\bar{c}) = \pi(\bar{c}) - u(0) \geq 0.$$

(Recall the constants of integration are arbitrary.) Next, the mechanism is balanced, so

$$\int_0^{\bar{b}} \rho_B(b) f(b) db = \int_{\underline{c}}^{\bar{c}} \rho_S(c) g(c) dc.$$

Expressions (3.14), (3.15), (3.18), and (3.19) imply

$$\rho_B(b) = b\xi_B(b) - u(0) - \int_0^b \xi_B(z) dz \quad \text{and} \quad \rho_S(c) = c\xi_S(c) + \pi(\bar{c}) + \int_{\underline{c}}^{\bar{c}} \xi_S(z) dz.$$

Hence,

$$\begin{aligned} & -u(0) + \int_0^{\bar{b}} \left(b\xi_B(b) - \int_0^b \xi_B(z) dz \right) f(b) db \\ & = \pi(\bar{c}) + \int_{\underline{c}}^{\bar{c}} \left(c\xi_S(c) + \int_{\underline{c}}^{\bar{c}} \xi_S(z) dz \right) g(c) dc. \end{aligned} \quad (3.25) \quad \{\text{eq:MS-derive-IR1}\}$$

From (3.25), a mechanism satisfies interim IR if and only if

$$\int_0^{\bar{b}} \left(b \xi_B(b) - \int_0^b \xi_B(z) dz \right) f(b) db - \int_{\underline{c}}^{\bar{c}} \left(c \xi_S(c) + \int_c^{\bar{c}} \xi_S(z) dz \right) g(c) dc \geq 0.$$

From the definitions of $\xi_B(\cdot)$ and $\xi_S(\cdot)$, the LHS of this last expression equals

$$\begin{aligned} & \int_0^{\bar{b}} \left(b \int_{\underline{c}}^{\bar{c}} x(b, c) g(c) dc - \int_0^b \int_{\underline{c}}^{\bar{c}} x(z, c) g(c) dc dz \right) f(b) db \\ & - \int_{\underline{c}}^{\bar{c}} \left(c \int_0^{\bar{b}} x(b, c) f(b) db + \int_c^{\bar{c}} \int_0^{\bar{b}} x(b, z) f(b) db dz \right) g(c) dc. \end{aligned}$$

In turn, this equals

$$\begin{aligned} & \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} (b - c) x(b, c) f(b) g(c) db dc + \int_0^{\bar{b}} \int_0^b \int_{\underline{c}}^{\bar{c}} x(z, c) g(c) dc dz \times (-f(b) db) \\ & - \int_{\underline{c}}^{\bar{c}} \int_c^{\bar{c}} \int_0^{\bar{b}} x(b, z) f(b) db dz \times g(c) dc \\ & = \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} (b - c) x(b, c) f(b) g(c) db dc - \int_0^{\bar{b}} (1 - F(b)) \int_{\underline{c}}^{\bar{c}} x(b, c) g(c) dc db \\ & - \int_{\underline{c}}^{\bar{c}} G(c) \int_0^{\bar{b}} x(b, c) f(b) db dc \\ & = \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} \left(\left(b - \frac{1 - F(b)}{f(b)} \right) - \left(c + \frac{G(c)}{g(c)} \right) \right) x(b, c) f(b) g(c) dc db, \quad (3.26) \quad \{\text{eq:MS-derive-IR2}\} \end{aligned}$$

where the first equality follows via integration by parts. To summarize the analysis to this point:

Lemma 9. *A necessary and sufficient condition for a direct-revelation mechanism to be interim individually rational is that expression (3.26) be non-negative.*

\{MS-suff-interimIR\}

The second-best problem can now be stated:

$$\max_{x(\cdot, \cdot)} \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} (b - c) x(b, c) f(b) g(c) dc db$$

subject to the constraints that (3.26) be non-negative, $\int_{\underline{c}}^{\bar{c}} x(b, c) g(c) dc$ be non-decreasing in b , and $\int_0^{\bar{b}} x(b, c) f(b) db$ be non-increasing in c . In light of Proposition 12, the constraint that expression (3.26) be non-negative is binding. Let

$\lambda > 0$ be the Lagrange multiplier on that constraint. As is often done in such problems, proceed by ignoring the monotonicity constraints on $\xi_B(\cdot)$ and $\xi_S(\cdot)$ and hope that the solution ends up satisfying them. The Lagrangean is

$$\begin{aligned} & \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} (b - c) x(b, c) f(b) g(c) dc db \\ & + \lambda \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} \left(\left(b - \frac{1 - F(b)}{f(b)} \right) - \left(c + \frac{G(c)}{g(c)} \right) \right) x(b, c) f(b) g(c) dc db \\ & \propto \int_0^{\bar{b}} \int_{\underline{c}}^{\bar{c}} \left(\left(b - \frac{\lambda}{1 + \lambda} \frac{1 - F(b)}{f(b)} \right) - \left(c + \frac{\lambda}{1 + \lambda} \frac{G(c)}{g(c)} \right) \right) x(b, c) f(b) g(c) dc db, \end{aligned}$$

where $\propto$ denotes “proportional to”; that is, the last line is the Lagrangean up to a positive multiplicative constant. Since such a constant is irrelevant to the solution, the optimal $x(\cdot, \cdot)$ maximizes that last line. The obvious solution is to set $x(b, c) = 1$ when

$$b - \frac{\lambda}{1 + \lambda} \frac{1 - F(b)}{f(b)} \geq c + \frac{\lambda}{1 + \lambda} \frac{G(c)}{g(c)} \quad (3.27) \quad \{\text{eq:MS-sell-condition}\}$$

and to set $x(b, c) = 0$ otherwise.

Does the solution given by condition (3.27) yield $\xi_B(\cdot)$ and $\xi_S(\cdot)$ that satisfy the monotonicity constraints?

Lemma 10. *A sufficient condition for the monotonicity conditions on $\xi_B(\cdot)$ and $\xi_S(\cdot)$ to be satisfied is that $x(\cdot, c)$ be a non-decreasing function for all c and that $x(b, \cdot)$ be a non-increasing function for all b .*

$\{\text{lemma:MS-x-monotone-suff}\}$

In light of Lemma 10, it remains to check if (3.27) yields an $x(\cdot, \cdot)$ that is non-decreasing in its first argument and non-increasing in its second. This will hold if the LHS of (3.27) is non-decreasing in b and the RHS is non-decreasing in c .

$\{\text{lemma:MonoVirtuals}\}$

Lemma 11. *If*

$$b - \frac{1 - F(b)}{f(b)} \quad \text{and} \quad c + \frac{G(c)}{g(c)} \quad (3.28)$$

are non-increasing in b and c , respectively, then the allocation rule satisfying (3.27) satisfies the monotonicity conditions on $\xi_B(\cdot)$ and $\xi_S(\cdot)$.

$\{\text{corr:MonoVirtuals}\}$

Corollary 3. *If $F(\cdot)$ satisfies the monotone hazard rate property (MHRP) and $G(\cdot)$ the monotone reverse hazard rate property (the latter property being that the reverse hazard rate be non-increasing),⁴⁵ then the allocation rule satisfying (3.27) satisfies the monotonicity conditions on $\xi_B(\cdot)$ and $\xi_S(\cdot)$.*

⁴⁵The reverse hazard rate is the density function divided by the distribution function.

As shown in the proofs given in the appendix, the key to the analysis are the expressions

$$V_B(b, \sigma) = b - \sigma \frac{1 - F(b)}{f(b)} \quad \text{and} \quad V_C(c, \sigma) = c + \sigma \frac{G(c)}{g(c)}.$$

These are similar to the virtual surplus function encountered in the previous section (consider, *e.g.*, expression (2.22) and surrounding discussion). We can view them, respectively, as the *virtual benefit function* and *virtual cost function*. As with the virtual surplus function, these functions differ from the true benefit (b) and true cost (c) because of the information rents the parties get: the cost of inducing truthful revelation is that high-benefit buyers and low-cost sellers must be left some amount of information rent. The need to satisfy *interim* IR prevents the mechanism designer from recapturing these rents, in expectation, via *ex ante* non-contingent transfers. The consequence is distortion in the allocation of the good.

Summarizing the analysis to this point:

Proposition 13. *Consider a setting in which a single unit is potentially to be exchanged, the buyer's value and seller's cost are continuously distributed with full support over overlapping intervals, and the mechanism must satisfy interim individual rationality. Assume, given $\sigma = 1$, the virtual benefit function is increasing in the buyer's valuation and the virtual cost function is increasing in the seller's cost. Then there exists a second-best direct-revelation mechanism. Moreover, there exists a $\sigma \in (0, 1)$ such that this second-best mechanism utilizes an allocation rule such that there is exchange if virtual benefit exceeds virtual cost and no exchange otherwise (i.e., there is exchange if $V_B(b, \sigma) \geq V_C(c, \sigma)$ and no exchange otherwise).*

{prop:MS-2ndBest}

As an example, suppose that b is distributed uniformly on $[0, \bar{b}]$ and c is distributed uniformly on $[0, \bar{c}]$ (note, here, $\underline{c} = 0$). Assume $\bar{c} \geq \bar{b}$. The uniform satisfies both MHRP and the monotone reverse hazard rate property. Consequently,

$$V_B(b, \sigma) = b - \sigma(\bar{b} - b) \quad \text{and} \quad V_C(c, \sigma) = c + \sigma c.$$

The allocation rule, given σ , is

$$x(b, c) = \begin{cases} 0, & \text{if } (1 + \sigma)b - \sigma\bar{b} < (1 + \sigma)c \\ 1, & \text{if } (1 + \sigma)b - \sigma\bar{b} \geq (1 + \sigma)c \end{cases}.$$

Interim IR binds; hence, (3.26) entails

$$\begin{aligned} 0 &= \int_0^{\bar{b}} \int_0^{\bar{c}} (V_B(b, 1) - V_C(c, 1)) x(b, c) (\bar{b}\bar{c})^{-1} dc db \\ &= \int_0^{\bar{b}} \int_0^{\bar{c}} (2b - \bar{b} - 2c) x(b, c) (\bar{b}\bar{c})^{-1} dc db \quad (3.29) \end{aligned} \quad \{\text{eq:MS-2ndBest-uniform1}\}$$

Note that $x(b, c) = 0$ if

$$b \leq \frac{\sigma}{1 + \sigma} \bar{b} \equiv \zeta.$$

It is also 0 if

$$c > b - \frac{\sigma}{1 + \sigma} \bar{b} = b - \zeta.$$

It is otherwise equal to 1. Knowing this, rewrite (3.29) as

$$0 = \int_{\zeta}^{\bar{b}} \int_0^{b-\zeta} (2b - \bar{b} - 2c)(\bar{b}\bar{c})^{-1} dc db = \frac{(4\zeta - \bar{b})(\bar{b} - \zeta)^2}{6\bar{b}\bar{c}} = \frac{\bar{b}^2(1 - 3\sigma)}{6\bar{c}(1 + \sigma)^3}.$$

It follows that $\sigma = 1/3$. Observe

$$V_B\left(b, \frac{1}{3}\right) - V_C\left(c, \frac{1}{3}\right) = \frac{4}{3}b - \frac{4}{3}c - \frac{1}{3}\bar{b}.$$

Hence, in this example, the second-best mechanism employs the allocation rule:

$$x(b, c) = \begin{cases} 0, & \text{if } b < c + \bar{b}/4 \\ 1, & \text{if } b \geq c + \bar{b}/4 \end{cases}.$$

In words, exchange occurs if and only if the buyer's valuation exceeds the seller's cost by at least one-quarter of the buyer's maximum valuation. Figure 4 illustrates for the case in which $\bar{b} = \bar{c}$ ("should" in the figure means if the first best is to be achieved).⁴⁶

3.3 THE MONEY-ON-THE-TABLE PROBLEM

{sect:CoaseConjecture}

From Proposition 13, the best outcome when subject to interim IR is exchange whenever $V_B(b, \sigma) \geq V_C(c, \sigma)$. Unless $\sigma = 0$, there exist values of b and c such that $b > c$, but $V_B(b, \sigma) < V_C(c, \sigma)$. Given that $\sigma = 0$ is impossible (Myerson-Satterthwaite), the following arises with positive probability: exchange should occur ($b > c$), but is blocked by the second-best mechanism. Moreover, if buyer and seller hear each other's announcement, then they know that an efficient exchange could have occurred but didn't. Assuming the possibility of exchanging the good has not vanished, the parties are confronted with a situation in which abiding by the mechanism means "leaving money on the table"; that is, there is surplus to be realized—and presumably split between the parties—if only they can go back and trade.

As a rule, it is difficult to see parties walking away and leaving money on the table. Presumably, after hearing that $b > c$, but exchange is not to occur, one party would approach the other and propose exchange at some price between b and c (e.g., $p = (b + c)/2$). This is fine if the parties are naïve; that is, somehow

⁴⁶For more on how much the second-best mechanism loses *vis-à-vis* the first-best solution see Williams (1987). See also Larsen (2012) for an empirical analysis that suggests that, at least in some contexts, real-life mechanisms are not too inefficient *vis-à-vis* the first-best ideal.

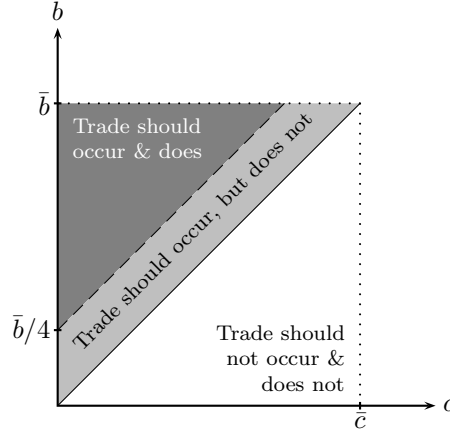


Figure 4: Welfare Loss Due to Interim Individual Rationality Constraint^{fig:MS-welfareloss}

they failed to anticipate the possibility of such future bargaining and, thus, announced their types truthfully as intended under the mechanism. We should not, though, expect such naïveté. Rather, each would anticipate that money wouldn't be left on the table. But, then, each would have incentive to lie about his or her type.

To see that lying would occur—formally, that truth-telling is no longer an equilibrium—suppose the buyer will tell the truth and consider whether the seller also wishes to tell the truth. To make the problem concrete, suppose that, if $b > c$ but no exchange occurs under the mechanism, then the price splits the difference; that is, $p = (b + c)/2$. The seller's optimization problem is

$$\begin{aligned} \max_{\hat{c}} \rho_S(\hat{c}) - c \left(1 - F(V_B^{-1}(V_C(\hat{c}, \sigma), \sigma)) \right) \\ + \frac{b^E + \hat{c} - 2c}{2} \left(F(V_B^{-1}(V_C(\hat{c}, \sigma), \sigma)) - F(\hat{c}) \right), \end{aligned}$$

where V_B^{-1} is defined by $V_B(V_B^{-1}(x, \sigma), \sigma) \equiv x$, b^E is the expected value of b given $b \in (\hat{c}, V_B^{-1}(V_C(\hat{c}, \sigma), \sigma))$, and the first line of that last expression would be the seller's payoff if there were no play beyond the end of the mechanism.⁴⁷ The derivative of the first line with respect to $\hat{c}$ is zero evaluated at $\hat{c} = c$ (the

⁴⁷Note that V_B^{-1} exists if we assume $V_B(\cdot, \sigma)$ is increasing. Stronger versions of the assumptions in Lemma 11 or Corollary 3 would be sufficient for this property to hold.

mechanism induces truth-telling). The derivative of the second line is

$$F(V_B^{-1}(V_C(\hat{c}, \sigma), \sigma)) - F(\hat{c}) \\ + \frac{b^E + \hat{c} - 2c}{2} \left(f(V_B^{-1}(V_C(\hat{c}, \sigma), \sigma)) \frac{\partial V_B^{-1}}{\partial V_C} \frac{\partial V_C}{\partial c} - f(\hat{c}) \right).$$

Generically, there is no reason to expect that expression to be zero at $\hat{c} = c$; in other words, the potential for later bargaining will cause the seller to deviate from truth-telling.

What should be made of this? The answer is that one should be suspicious of the mechanism of Proposition 13 unless there is good reason to believe the parties are committed to playing the mechanism as given (*i.e.*, not negotiating if exchange does not occur, but $b > c$). One good reason would be if there is literally one point in time at which the good can be exchanged (it is, *e.g.*, highly perishable and negotiations to pick up the money left on the table would conclude only after the good has perished). A second is that the one or both parties wish to develop a reputation for never negotiating after the play of the mechanism:⁴⁸ because the mechanism is second best, it could be in the interest of one or both parties to develop a reputation for fully committing to the mechanism to avoid a third-best outcome. In general, these reasons are likely to be the exception, not the rule;⁴⁹ hence, another solution is necessary.

Unhappily, the current state of economic theory is weak with respect to the modeling of bargaining between two asymmetrically informed parties. One reason is that such games are highly sensitive to the extensive form that is assumed. For example, a game in which one player repeatedly makes offers and the other merely says yes or no can have quite a different outcome than a game in which the players alternate making offers. Nevertheless, some such bargaining game is played and one should expect that it ends with trade if $b > c$ and no trade otherwise.

As an example of such a game, suppose that the seller makes repeated offers to the buyer until the buyer accepts or the seller concludes $c > b$.⁵⁰ Each round of bargaining takes a unit of time and both parties discount at a rate $\delta \in (0, 1)$ per unit of time. Assume that $\bar{b} = 1$ and F is the uniform distribution. Let $p_t(c)$ be the price offered by a seller of cost c at time t . The objective is an equilibrium in which $p_1(c) \neq p_1(c')$ for $c \neq c'$; that is, after the seller's first

⁴⁸Such reputations could be realistic in some settings. For instance, some divorce lawyers, referred to as “bombers,” have developed a reputation for making take-it-or-leave-it offers in pre-trial settlement talks; that is, they are committed to never negotiate beyond putting their initial demands on the table. (Lest the reader worry that the author obtained this knowledge through painful personal experience, he did not—he learned this by talking to divorce lawyers in a social setting.)

⁴⁹Larsen (2012) can be seen as evidence in favor of this claim. He shows the parties don't commit to a mechanism (an auction) and post-mechanism bargaining occurs.

⁵⁰The following discussion also serves as a demonstration of the Coase (1972) conjecture. For a more general approach, see Gul et al. (1986).

offer, the buyer knows, in equilibrium, the seller's cost. That is, the game from that point on is a game of one-sided asymmetric information.

Consider the following renormalization. Define $\omega = \max\{b - c, 0\}$. Given c , one can think of ω as distributed uniformly on $[0, 1 - c]$ because the mass at 0 is irrelevant (no exchange should occur there). Let $m = p - c$ denote the seller's markup. The goal will be an equilibrium in which the parties play linear strategies: a buyer who hasn't yet bought buys if $\omega \geq \alpha m$ and, if ω is the highest renormalized type who hasn't purchased prior to the current period, the seller sets her markup (price less cost) $m = \gamma\omega$.

Proposition 14. *Assume the buyer's valuation is uniformly distributed. Then there exists a subgame-perfect equilibrium in which the parties play linear strategies.*

{prop:Coase}

As shown in the proof of Proposition 14, $\gamma < 1$; hence, $m_t = (1 - c)\gamma^t$ tends to 0 as $t \rightarrow \infty$. In words, the markup approaches zero, which means the seller's price approaches her cost. Consequently, eventually all buyers for whom $b > c$ will buy. The equilibrium of this bargaining game is, thus, such that no money gets left on the table.

Although money is not left on the table, this does not mean that the first best is achieved. In fact, it isn't because, with positive probability, welfare-enhancing exchange (*i.e.*, when $b > c$) is delayed. Indeed, this was known: Proposition 12 (Myerson-Satterthwaite) rules out the first best. Arguments about money left on the table don't change that conclusion: if there were some bargaining game that achieved the first best—achieved efficient trade immediately and at no bargaining cost—then, by the Revelation Principle, that game would be replicable by a direct-revelation mechanism, but Proposition 12 establishes that is impossible. In other words, the true importance of Proposition 12 is it implies that, for any procedure in which exchange is guaranteed to happen (eventually) if $b > c$, there must be costs that cause the buyer and seller's collective welfare to be less than the first-best level, at least with positive probability.

4 EXOGENOUS ASYMMETRIES OF INFORMATION AT TIME OF CONTRACTING: INFORMED SELLER

{sect:InformedSeller}

Now consider the situation of a contract proposer endowed with payoff-relevant information that her counter party does not know. In keeping with Section 2's naming convention, call the contract proposer the seller and the counter party the buyer (although, again, the techniques and insights developed here have wider applicability). As in that earlier section, the contract offer is made on a take-it-or-leave-it (TIOLI) basis.

As before, call the informed party's information her type. Denote it by γ and assume it is an element of Γ . Let $G : \Gamma \rightarrow [0, 1]$ be the distribution function by which “nature” chooses γ . Although γ is the informed party's (here, seller's) private information, Γ and $G(\cdot)$ are common knowledge.

4.1 THE LEMONS MODEL

The classic informed-seller model is Akerlof's (1970) "lemons model": the seller's type (*i.e.*, γ) describes the quality of the good or service she is offering. For example, as in Akerlof's original paper, the seller knows the quality of the used car she is offering, whereas the buyer does not.⁵¹ It is further assumed that the buyer cannot simply observe quality prior to trade: he only learns quality once he has taken possession or received the service. Because of this feature, these models are often called *experience-good models*.

Although quality could be multidimensional, it is standard to assume either that is unidimensional or buyer and seller have a common complete preference order over the elements of Γ , the seller's type space.⁵² Consequently, there is no further loss in treating Γ as a subset of $\mathbb{R}$. To avoid pathological cases, assume $\Gamma \subseteq (\underline{\gamma}, \bar{\gamma})$, both limits finite. As a convention, $\gamma > \gamma'$ is understood to mean that type γ is higher quality than γ' . Assume the buyer and seller's payoffs are, respectively,

$$U_B = x(u(\gamma) - p) \text{ and } U_S = xp + (1 - x)\gamma,$$

where $x \in \{0, 1\}$ denotes the amount of trade, p is payment in case of trade, and $u : \Gamma \rightarrow \mathbb{R}$ is an increasing function.⁵³ Trade would never be efficient if $u(\gamma) < \gamma$ for all $\gamma \in \Gamma$; hence, assume there exists a $\Gamma^* \subseteq \Gamma$, $G(\Gamma^*) > 0$, such that $u(\gamma) > \gamma$ if $\gamma \in \Gamma^*$.

In the basic lemons model, the buyer plays a pure-strategy response: his strategy is an $x : \mathbb{R} \rightarrow \{0, 1\}$ —a mapping from prices into purchase decisions. Observe that, for $p > p'$ such that $x(p) = x(p') = 1$, there can no equilibrium in which the seller quotes price p' : in any equilibrium in which trade may occur, there is a single price, $\hat{p}$, offered by seller types who wish to trade.

The rationality of the seller is assumed to be common knowledge. The buyer can thus infer from a price quote of p that $\gamma \not> p$. If there is an equilibrium in which the buyer accepts $\hat{p}$, it follows the buyer must believe that the seller is playing the strategy⁵⁴

$$p(\gamma) = \begin{cases} \hat{p}, & \text{if } \gamma \leq \hat{p} \\ \infty, & \text{if } \gamma > \hat{p} \end{cases}.$$

⁵¹A low-quality car is colloquially known as a "lemon," hence the title of Akerlof's original article.

⁵²The analyzes of multidimensional signaling in Quinzii and Rochet (1985) and Engers (1987) are two notable exceptions.

⁵³Note there is no gain in generality from assuming

$$U_S = xp + (1 - x)v(\gamma),$$

$v : \Gamma \rightarrow \mathbb{R}$ increasing, because one could simply renormalize the type space and the buyer's utility: $\tilde{\Gamma} = v(\Gamma)$ and $\tilde{u}(\tilde{\gamma}) \mapsto u(v^{-1}(\tilde{\gamma}))$. Given the notion of opportunity cost, an equivalent analysis is possible assuming

$$U_S = x(p - c(\gamma)),$$

where $c : \Gamma \rightarrow \mathbb{R}$, an increasing function, is the cost of supplying a product of quality γ .

⁵⁴A price of ∞ is equivalent to not making an offer.

Buying, $x(\hat{p}) = 1$, is a best response if and only if

$$\mathbb{E}\{u(\gamma)|\gamma \leq \hat{p}\} \geq \hat{p}; \quad (4.1) \quad \{\text{eq:Lemons1}\}$$

that is, if and only if the buyer's expected utility exceeds $\hat{p}$ conditional on knowing quality is no greater than $\hat{p}$. To conclude: there is an equilibrium with trade if and only if condition (4.1) has a solution.

As an example, suppose that $u(\gamma) = m\gamma + b$, $\Gamma = [0, 1]$, and $G(\cdot)$ is the uniform distribution. It is readily shown that

$$\mathbb{E}\{u(\gamma)|\gamma \leq \hat{p}\} = \begin{cases} b + \frac{m\hat{p}}{2}, & \text{if } \hat{p} \leq 1 \\ b + \frac{m}{2}, & \text{if } \hat{p} > 1 \end{cases}.$$

For instance, suppose that $b = 0$ and $m = 3/2$. Condition (4.1) is satisfied only for $\hat{p} = 0$; hence, the probability of trade is zero. This is, however, inefficient because $u(\gamma) > \gamma$ for all $\gamma > 0$ (i.e., $\Gamma^* = (0, 1]$). For these parameters, the *lemons problem* is severe: asymmetric information destroys the market.

Consider different parameters: $b = 1/8$ and $m = 3/2$. Condition (4.1) is now satisfied for all $\hat{p} \leq 1/2$. There is, thus, an equilibrium in which types $\gamma \leq 1/2$ offer to sell at price $1/2$, other types make no offer, and the buyer purchases if offered the product at price $1/2$.⁵⁵ Trade occurs 50% of the time rather than never, so this is a more efficient outcome than under the initial set of parameters; but it is still inefficient *vis-à-vis* the symmetric-information ideal.

The analysis to this point is premised on the buyer's playing pure strategies only. Are there mixed-strategy equilibria? To explore this, suppose $\Gamma = \{\gamma_L, \gamma_H\}$, $\gamma_H > \gamma_L$, and let $g = G(\gamma_L) \in (0, 1)$. Define $u_i = u(\gamma_i)$. If $u_i \leq \gamma_i$ both i , there are no gains to trade. If $u_L > \gamma_L$, but $u_H \leq \gamma_H$, there is an efficient pure-strategy equilibrium in which $\hat{p} = u_L$. Hence, the only case of interest is $u_H > \gamma_H$. Moreover, to truly be interesting, it must be that

$$gu_L + (1 - g)u_H < \gamma_H, \quad (4.2) \quad \{\text{eq:LemonsInteresting}\}$$

as otherwise there is a pure-strategy equilibrium in which both seller types sell.

Suppose that there is an equilibrium in which there is trade with positive probability at *two* prices p_h and p_ℓ , $p_h > p_\ell$. For p_ℓ to be offered, it must be that $x(p_h) < x(p_\ell)$, where, now, $x : \mathbb{R} \rightarrow [0, 1]$. Let $U_j = \mathbb{E}\{u(\gamma)|p = p_j\}$; that is, U_j is the buyer's expectation of his utility upon being offered the product at price p_j , where the expectation is calculated given the seller's equilibrium strategy. Because $0 < x(p_h) < 1$, the buyer mixes over accepting or rejecting in response to p_h ; hence, $U_h - p_h = 0$.

⁵⁵Formally, a perfect Bayesian equilibrium is seller and buyer, respectively, play strategies

$$p(\gamma) = \begin{cases} 1/2, & \text{if } \gamma \leq 1/2 \\ \infty, & \text{if } \gamma > 1/2 \end{cases} \quad \text{and} \quad x(p) = \begin{cases} 1, & \text{if } p = 1/2 \\ 0, & \text{if } p \neq 1/2 \end{cases};$$

and the buyer believes types $\gamma \leq 1/2$ offer at price $1/2$, types $\gamma > 1/2$ make no offer, and any other price offer was made by type 0.

What if the γ_H -type seller offers p_h and the γ_L -type offers p_ℓ in equilibrium (*i.e.*, the seller does not mix)? Buyer rationality implies $p_\ell \leq u_L$ and his mixing implies $p_h = u_H$. Suppose p_ℓ in fact equals u_L and suppose $x(p_\ell) = 1$. Given earlier assumptions and condition (4.2), for this to be an equilibrium, it is necessary that

$$u_L \geq x(u_H)u_H + (1 - x(u_H))\gamma_L \quad (4.3) \quad \{\text{eq:LemonsMix1}\}$$

(*i.e.*, the γ_L -type seller must prefer to offer the product at $p = u_L$ than at $p = u_H$). Expression (4.3) holds if and only if $u_L > \gamma_L$. Assuming that condition, (4.3) can be solved for $x(u_H)$:

$$x(u_H) \leq \frac{u_L - \gamma_L}{u_H - \gamma_L}. \quad (4.4) \quad \{\text{eq:LemonsMix2}\}$$

Efficiency, here, means maximizing the probability of trade; hence, the most efficiency equilibrium is the one in which (4.4) holds as equality. To summarize:

Proposition 15. *For the two-type model considered here, there is an equilibrium in which the low-quality seller (γ_L) offers the product at price u_L , the high-quality seller (γ_H) offers it at price u_H , the buyer accepts a price of u_L or less with certainty, mixes over accepting a price of u_H with probability equal to the RHS of expression (4.4), and rejects all other offers. The buyer believes an offer at price u_H comes from the high-quality seller and he believes any other offer is from the low-quality seller.*

$\{\text{prop:LemonsMix1}\}$

As an example, suppose $\gamma_L = 1$, $\gamma_H = 3$, $u(\gamma) = \gamma + 1$, and $g = 3/5$. Limiting both buyer and seller to pure strategies, the equilibrium is one with a price of 2 and expected welfare of $36/15$. In contrast, under the Proposition 15 equilibrium, the prices are 2 and 4, the buyer accepts an offer of 4 with probability $1/3$, and expected welfare is $38/15$. In other words, the lemons problem proves less severe than might originally have been thought once account is taken of the possibility that the buyer can mix.

4.1.1 CREDENCE GOODS

A phenomenon related to experience goods is the following: a buyer knows he has a problem (*e.g.*, with his car, of a medical nature, etc.), but not its cause. He seeks treatment from the seller (an expert, such as a mechanic or physician), who can diagnosis the cause and administer a treatment. The buyer knows if the the problem has been fixed or not, but not whether the diagnosis and treatment regime were correct. This is relevant insofar as long as the seller corrects the problem, she can claim a more severe diagnosis (*e.g.*, engine needs replacing) even when the truth is less severe (*e.g.*, just the carburetor needs replacing) and provide the more expensive treatment (*e.g.*, replace the engine rather than just the carburetor). Models that explore such situations are known as *credence-good* models.⁵⁶

⁵⁶Darby and Karni (1973) are often credited with introducing the notion of credence goods into the literature. For a relatively recent treatment, as well as citations to earlier literature, see Fong (2005).

One issue in such models is whether the buyer has recourse if promised one treatment but receives another. For instance, suppose, based on his mechanic's diagnosis, the buyer agrees to a new engine (rather than just a new carburetor). When the buyer picks his car up, it runs well (*i.e.*, he knows the problem has been fixed). Question: does he know if the engine inside his car is new (as the mechanic claims) or is it his old engine with just a new carburetor (the mechanic has cheated him)? If the answer is he knows (and can obtain recourse in case of fraud), then we have one kind of credence-good model. If he doesn't, then we have the other. Call the two scenarios, verifiable and unverifiable treatment, respectively. Here, attention is limited to the verifiable-treatment scenario. For an analysis of the unverifiable-treatment scenario see Fong (2005).

As a basic model, suppose the buyer's gross benefit if his problem is corrected is $\bar{v}$. Without loss, we can and will normalize $\bar{v} = 0$. If his problem goes *uncorrected* and it is of type (severity) σ , $\sigma \in \{L, H\}$ (low and high severity, respectively), then his gross benefit is $-\ell_\sigma$; that is, he loses ℓ_σ if his problem is not fixed. Assume

$$0 \leq \ell_L \leq \ell_H,$$

with at least one inequality strict.⁵⁷ The seller's cost of treatment is c_σ , where $c_H > c_L$. Let $q = \Pr\{\sigma = L\}$. Assume $0 < q < 1$. This probability is common knowledge, but only the seller can determine actual severity.

The extent of possible fraud by the seller is limited given the assumption of verifiable treatment. Specifically, assume that the treatment for the more severe problem cures the less severe problem (*e.g.*, replacing the entire engine "fixes" a broken carburetor), but the opposite is not true (*e.g.*, changing a broken carburetor is not a fix when the entire engine must be replaced).

If the buyer had to pay for treatment prior to receiving a diagnosis, the most he would pay is

$$\bar{p} = q\ell_L + (1 - q)\ell_H.$$

If $\bar{p} \geq c_H$ and $\ell_\sigma \geq c_\sigma$ for both σ , then efficiency will result: the seller offers to fix any problem for $\bar{p}$ and the buyer accepts the offer. The seller has the appropriate incentive to employ the correct treatment and it is efficient for problems of both severity levels to be treated. Issues arise if $\bar{p} < c_H$ or if $\ell_\sigma < c_\sigma$ for one σ .

Cost to correct severe problem exceeds average benefit of repair.

Suppose that $\bar{p} < c_H$, but $\ell_\sigma \geq c_\sigma$ for both σ . It is, thus, efficient to fix both problems, but the seller is unwilling to fix the severe problem if she is paid only $\bar{p}$.⁵⁸ Observe, $\ell_L < c_H$ given $\bar{p} < c_H$. If

$$\ell_H - c_H \leq \ell_L - c_L \tag{4.5} \quad \{\text{eq:Credence-NetCost}\}$$

⁵⁷The case $\ell_L = 0$ corresponds to a situation in which the problem quickly clears up without treatment. The case $\ell_L = \ell_H$ corresponds to one in which the problem (*e.g.*, car doesn't run) has the same effect on the buyer's wellbeing regardless of cause.

⁵⁸The assumption is the seller can quit if fixing a problem at a quoted price would cause her a loss; that is, an *ex post* IR constraint is in effect for the seller.

(i.e., the social benefit of fixing the more severe problem does not exceed that of fixing the less severe problem), then efficiency can be attained using two prices: $p_\sigma = \ell_\sigma$, $\sigma \in \{L, H\}$. In other words, the seller offers to fix a problem of severity σ for price p_σ . Because, given (4.5), her *margin* is greater on fixing the low-severity problem than on fixing the high-severity one and she cannot fraudulently use the low-severity treatment for the high-severity problem, the seller will behave honestly.⁵⁹

If (4.5) does not hold, then the seller would have an incentive to behave fraudulently were $p_\sigma = \ell_\sigma$ for both σ . Anticipating such fraud, the buyer infers he will pay p_H regardless of his true problem. This exceeds $\bar{p}$, so the buyer would not do business with the seller. Nonetheless, there remains a two-price solution that achieves efficiency: to wit, let m_σ be the seller's margin on fixing problem σ given her quoted prices (i.e., $m_\sigma = p_\sigma - c_\sigma$). As we've seen, seller honesty requires that $m_L \geq m_H$. In choosing her prices, the seller seeks to solve:

$$\max_{m_L, m_H} qm_L + (1 - q)m_H \quad (4.6) \quad \{\text{eq:Cred-Max1}\}$$

subject to

$$m_L \geq m_H, \quad (\text{Seller IC})$$

$$q\left(\ell_L - \underbrace{(m_L + c_L)}_{p_L}\right) + (1 - q)\left(\ell_H - \underbrace{(m_H + c_H)}_{p_H}\right) \geq 0, \quad (\text{Buyer IR})$$

and

$$m_H \geq 0 \quad (\text{Seller IR})$$

(given the seller's IC condition, the additional condition that $m_L \geq 0$ is superfluous and, so, can be ignored). Let S denote expected surplus from trade:

$$S = q(\ell_L - c_L) + (1 - q)(\ell_H - c_H).$$

⁵⁹Observe this conclusion relies on the verifiability of treatment. Were treatment unverifiable and $p_H > p_L$, then the seller would always have an incentive to claim $\sigma = H$. As Fong (2005) shows, the solution in this case would be similar to Proposition 15: the buyer agrees to treatment with certainty if the diagnosis is $\sigma = L$ and pays p_L , but agrees to treatment with probability

$$\frac{p_L - c_L}{p_H - c_L}$$

and pays p_H if the diagnosis is $\sigma = H$. The seller who faces a buyer with a severe ($\sigma = H$) problem never lies and a seller who faces a buyer with a minor problem ($\sigma = L$) states her diagnosis honestly with probability

$$\frac{p_H - \ell_L}{\ell_H - \ell_L}. \quad (\star)$$

It is readily verified this an equilibrium for a given (p_L, p_H) pair of announced prices. It can be shown that the price pair (ℓ_L, ℓ_H) maximizes the seller's expected profit. Given $(\star)$ this means that the seller is always honest in equilibrium (but the buyer still mixes because he is indifferent about obtaining treatment or not). Because of the mixing, the equilibrium is inefficient. See Fong for further details.

It is readily seen that the solutions to the seller's linear-programming problem is any pair (m_L, m_H) , $m_L \geq m_H$ and $m_H \geq 0$ on the line

$$S = qm_L + (1 - q)m_H. \quad (4.7)$$

In other words, there is an efficient outcome if the seller commits to her prices prior to learning the buyer's actual problem and the buyer commits to pay those prices prior to learning the seller's diagnosis. Reflecting her bargaining power, the seller captures all surplus in expectation.

Unlike the earlier analysis of the market for lemons, efficiency is achieved here. This is because, here, prices are set *before* the seller learns her type. As such, the comparison between the two analyses reflects a general point: inefficiencies due to asymmetric information are more pronounced if the asymmetry information predates the parties ability to write contracts than if it arises after such contracting.⁶⁰

In many situations, the buyer is allowed to walk away after receiving a diagnosis. In such situations, the program (4.6) would also need to satisfy *ex post* IR constraints for the buyer:

$$\ell_L - (m_L + c_L) \geq 0 \text{ and} \quad (\text{IR-L})$$

$$\ell_H - (m_H + c_H) \geq 0. \quad (\text{IR-H})$$

Because (4.5) does not hold, the *ex post* constraint (IR-L) is binding. To see this, observe the smallest value of m_L that satisfies the program (4.6) as originally given is

$$m_L = m_H = S.$$

Substituting that into the LHS of (IR-L) yields

$$(1 - q) \left((\ell_L - c_L) - (\ell_H - c_H) \right), \quad (4.8)$$

which is negative given (4.5) does not hold. Given (IR-L) is binding, it follows that

$$m_L = \ell_L - c_L; \text{ equivalently that } p_L = \ell_L.$$

It will also be that $m_H = m_L$, hence

$$p_H = \ell_L + (c_H - c_L) < \ell_H,$$

where the inequality follows because (4.5) does not hold. Observe, now, that although fully efficiency is still achieved, the seller no longer captures all the surplus generated: the need to provide the seller strong enough incentives to make an honest diagnosis conflicts with the need to keep the buyer from walking away.

⁶⁰See Hermalin and Katz (1993) and Hermalin et al. (2007, §§2-3) for further development of this point.

If

$$\ell_H - c_H < 0, \quad (4.9) \quad \{\text{eq:Cred-NetCost2}\}$$

then efficiency will also attain: either $\ell_L < c_L$, in which case the market should not exist and won't; or $\ell_L \geq c_L$ and the market should exist for fixing the less-severe problem only and it will exist (the seller promises to fix any problem she diagnoses as L at price ℓ_L).

The remaining case when $\bar{p} < c_H$ is

$$\ell_H - c_H \geq 0 > \ell_L - c_L : \quad (4.10) \quad \{\text{eq:Cred-NetCost3}\}$$

efficiency dictates that only the more-severe problem be treated. The previous logic continues to apply: for the seller to have an incentive to make an honest diagnosis, $m_L \geq m_H$. Here, given no market for fixing less-severe problems, $m_L = 0$; hence, efficiency is achievable only if $p_H = c_H$. In other words, in this case, the seller announces she won't fix L problems, but will fix H problems at cost, c_H . In this case, the incentive problem is so severe that it results in all surplus going to the buyer.

Cost to correct severe problem less than average benefit of repair.

Now suppose that $\bar{p} \geq c_H$. Suppose $\ell_\sigma < c_\sigma$ for one σ , so it is inefficient to fix all problems. If (4.9) holds, then an efficient equilibrium exists: rather than quote a price of $\bar{p}$ and offer to fix all problems, the seller does better to quote a price of $p_L = \ell_L$ and offer to fix only L problems. To see this, observe

$$\bar{p} - (qc_L + (1-q)c_H) = q(\ell_L - c_L) + (1-q)(\ell_H - c_H) < q(\ell_L - c_L), \quad (4.11)$$

where the first term is expected profit from offering to fix all problems and the last is expected profit from offering to fix L problems only.

Finally, suppose $\bar{p} \geq c_H$, but $\ell_L < c_L$. It is now impossible to provide the seller incentives to diagnosis honestly (absent setting $p_H = c_H$, which the seller wouldn't do). In this scenario, the seller offers to fix all problems and fixes all problems even though it is inefficient to fix the L problems.

4.2 SIGNALING

The welfare loss due to the lemons problem is essentially borne by high-quality (high- γ) sellers. This is immediate in the two-type case leading up to Proposition 15 on page 51: the low-quality seller is always able to sell her product—a rational buyer is willing to pay up to $u(\gamma_L)$ regardless of his beliefs about the seller's quality since, for any beliefs, $\mathbb{E}\{u(\gamma)\} \geq u(\gamma_L)$ —the low-quality seller is sure to receive at least what she would have under symmetric information. It is the high-quality seller who is at risk of not realizing the profit, $u(\gamma_H) - \gamma_H$, she would have were the setting one of symmetric information.

This insight suggests that were there a way for a high-quality seller to prove who she was, she would do so provided it were not too expensive. In some instances, such proof is direct—the seller, for example, employs a reliable rating

agency to certify publicly that her product is high quality. In other instances, the proof is more indirect—the high-quality seller undertakes an action that is worthwhile for her if it convinces the buyer she’s high quality, but would not be worthwhile for a low-quality seller even if it misled the buyer into believing she was high quality. An example of such indirect proof would be a seller who offers a warranty with her product: if repair costs are high enough, then it could be too costly for a low-quality seller to offer a warranty, even if it misled the buyer, but not so costly for a high-quality seller. Hence, a buyer would accurately infer that a seller offering a warranty is high quality. In the language of information economics, such indirect proof is known as a *signal*.

4.2.1 A BRIEF REVIEW OF SIGNALING

Spence (1973) was the first analysis of the use of signals or *signaling*. In that original paper, the seller was an employee and quality was her ability. The buyer—potential employer—could not directly observe a would-be employee’s ability, but he could observe the employee’s educational attainment. The key assumption of the model was that the marginal cost of obtaining an additional level of attainment (*e.g.*, year in school) was always less for high-ability employees than for low-ability employees. Hence, by obtaining enough education, a high-ability employee could signal her ability to the potential employer.

Since Spence’s seminal work, signaling models have been utilized to explore a wide variety of economic and other social phenomena. The topic is, thus, of great importance. On the other hand, due to its importance, there are now many excellent texts that cover signaling (see, *e.g.*, Fudenberg and Tirole, 1991; Gibbons, 1992; and Mas-Colell et al., 1995). Consequently, despite its importance, the treatment here will be brief.

A signaling game is one between an informed party, who plays first, and an uninformed party, who responds. In keeping with this chapter’s nomenclature, call the former party the seller and the latter the buyer, although the analysis applies more generally. The seller has utility $V(a, x, \gamma)$, where $a \in \mathcal{A}$ is the seller’s action, $x \in \mathcal{X}$ is the buyer’s response, and $\gamma \in \Gamma$ is the seller’s type. In the typical buyer-seller relation, $\mathcal{X} = \{0, 1\}$ —the buyer rejects or accepts the seller’s offer, respectively. In that relation, $\mathcal{A}$ might be a two-dimensional space consisting of a price and signaling action (*e.g.*, a warranty). As before, γ could be a measure of quality. The buyer’s utility is $U(a, x, \gamma)$. A pure strategy for the seller is a mapping $a : \Gamma \rightarrow \mathcal{A}$. A pure strategy for the buyer is a mapping $x : \mathcal{A} \rightarrow \mathcal{X}$. Mixed strategies are the usual extension of pure strategies. An outcome is *separating* if $\gamma \neq \gamma'$ implies $a(\gamma) \neq a(\gamma')$ for any $\gamma, \gamma' \in \Gamma$; that is, an outcome is separating if different seller types choose different actions. An outcome is *pooling* if $a(\gamma) = a(\gamma')$ for all $\gamma, \gamma' \in \Gamma$; that is, it is pooling if all types choose the same action. Various hybrid outcomes are also possible (*e.g.*, one type plays a with certainty, while another mixes between a and a').

Although not normally characterized as a signaling game, the treatment of the lemons problem in the previous subsection nonetheless fits the structure of a signaling game. In particular, the equilibrium of Proposition 15 is a separating

equilibrium: the two types offer different prices (*i.e.*, $a(\gamma_L) = p_\ell = u_L$ and $a(\gamma_H) = p_h = u_H$). For the same game, if (4.2) were instead reversed, then a pooling equilibrium would exist in which both types offered

$$\hat{p} = gu_L + (1 - g)u_H$$

(*i.e.*, $a(\gamma_L) = a(\gamma_H) = \hat{p}$). Returning to the assumption that (4.2) holds, one can construct a hybrid (partial-separating) equilibrium in which the high-quality type offers p_h only, but the low-quality type mixes between p_ℓ and p_h .

As just indicated, signaling games often admit multiple solutions (equilibria). For example, under the assumptions underlying Proposition 15, another equilibrium is

$$p(\gamma) = \begin{cases} u_L, & \text{if } \gamma = \gamma_L \\ \infty, & \text{if } \gamma = \gamma_H \end{cases}; \quad x(p) = \begin{cases} 1, & \text{if } p \leq u_L \\ 0, & \text{if } p > u_L \end{cases};$$

and the buyer believes all offers come from the low-quality type. This is because the standard solution concept for such games, perfect Bayesian equilibrium (PBE), does not tie down the uninformed player's beliefs in response to an out-of-equilibrium move by the informed player.⁶¹ In particular, it is possible to construct equilibria that are supported by what the uninformed player “threatens” to believe in response to out-of-equilibrium play. To reduce the set of equilibria—in particular, to eliminate equilibria supported by “unreasonable” out-of-equilibrium beliefs—various equilibrium refinements can be employed (see, *e.g.*, Fudenberg and Tirole, 1991, especially Chapter 11, for an introduction to refinements). A prominent refinement is the “Intuitive Criterion” of Cho and Kreps (1987).

To help illustrate the issue further, as well as provide a basic understanding of the Intuitive Criterion, consider the signaling game in Figure 5, which is based on Cho and Kreps's famous beer and quiche game. The game starts at the chance node in the middle, where Nature determines whether the seller is low quality or high quality. The probability she selects low quality is g . The seller moves next (hollow decision node), deciding whether to offer a warranty, but charge a high price; or to offer no warranty, but charge a low price. The buyer moves last (filled-in decision node), deciding whether to buy or not. The game, then, ends with payoffs as indicated; the first number in each pair is the seller's payoff and the second the buyer's. The seller knows Nature's move; that is, she knows the quality of what she's selling. The buyer does not—he only observes what he is offered. That the buyer is ignorant of the seller's type is indicated by the dashed lines connecting his decision nodes, which indicate they are in the same information set for the buyer.

The game has two equilibria: (1) a low-quality seller offers no warranty and a low price, a high-quality seller a warranty and a high price, and the buyer

⁶¹In contrast, the uninformed's beliefs following a move or action that can occur in equilibrium must be consistent with Bayes Law given the distribution of types and the informed player's equilibrium strategy.

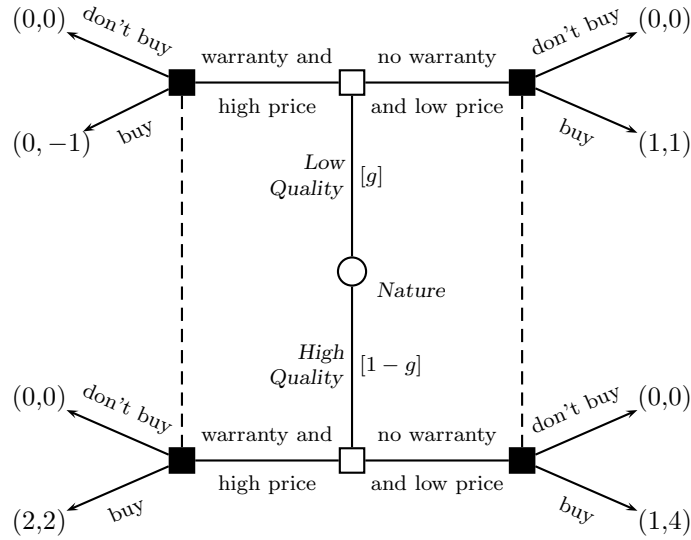


Figure 5: A simple signaling game. fig:BeerQuiche

accepts both offers; and (2) both seller types offer no warranty and the low price, the buyer accepts that offer, but will reject the warranty-and-high-price offer. Equilibrium (1) is separating and equilibrium (2) is pooling. The pooling equilibrium is supported by the buyer's believing that the out-of-equilibrium play of an offer of a warranty and high price indicates he is facing the *low*-quality seller with a probability, $\hat{g}$, in excess of $2/3$. Given such a belief, his rejecting that offer is rational (*i.e.*, $-1\hat{g} + 2(1 - \hat{g}) < 0$).

The second equilibrium is a valid PBE, but nonetheless strikes most observers as unreasonable. No matter what response offering a warranty and high price triggered in the buyer, a low-quality seller is better off playing her equilibrium strategy of no warranty and low price; that is, a low-quality seller has no incentive whatsoever to deviate. The same is not true of a high-quality seller: if the buyer bought in response to the deviation of a warranty and high price, then a high-quality seller is better off—she gets 2 instead of 1. Given this, it seems unreasonable to postulate $\hat{g} > 2/3$; indeed, $\hat{g} = 0$ seems most sensible. As Cho and Kreps observe, one can even imagine the seller helping the buyer to make such a forward induction: if the parties expected the second PBE to be played, a high-quality seller could make the following speech to accompany a deviating offer of a warranty and high price.

Dear buyer, I know you were expecting an offer of no warranty and a low price. Instead, I'm offering you a warranty, but a higher price. What should you make of this innovation? If I were a low-quality seller—which I am assuredly not—I would have absolutely no reason to make this offer: even if you accepted, I would be worse off than

I would be in the equilibrium you thought we were playing. On the other hand, if you accept my offer, then, because I'm truly high quality, I am better off. Hence, because only a high-quality seller could have a motive to make this offer, it is only reasonable of you to believe that I am high quality.

This reasoning can be formalized: assume Γ is finite with N elements. Let Δ^N be the N -dimensional unit simplex (*i.e.*, the subset of vectors in $\mathbb{R}_+^N$ whose elements sum to 1). Let $\mathcal{X}(a)$ be the set of responses available to the uninformed player (*e.g.*, buyer) if the informed player (*e.g.*, seller) takes action $a \in \mathcal{A}$ (it could be—as in Figure 5 and many games of interest—that $\mathcal{X}(a) = \mathcal{X}$ for all a ; the formulation $\mathcal{X}(a)$ simply permits a more general analysis). As a second generalization, it may be that an informed player's action space depends on her type; let $\Gamma(a)$ denote the types that have $a \in \mathcal{A}$ in their action space. Let $\mu : \Gamma \rightarrow \Delta^N$ denote beliefs over types; that is, μ is a density function over Γ . Define

$$\text{br}(\mu, a) = \operatorname{argmax}_{x \in \mathcal{X}(a)} \sum_{\gamma \in \Gamma} U(a, x, \gamma) \mu(\gamma).$$

In words, $\text{br}(\mu, a)$ are the uninformed player's best responses, given belief μ , to action a by the informed player. If $\hat{\Gamma} \subseteq \Gamma$, define

$$\text{BR}(\hat{\Gamma}, a) = \bigcup_{\{\mu \mid \mu(\hat{\Gamma})=1\}} \text{br}(\mu, a);$$

that is, $\text{BR}(\hat{\Gamma}, a)$ are actions for the uninformed player that are best responses to the informed player's action a for some beliefs provided those beliefs assign all weight to types in the subset $\hat{\Gamma}$. For example, in the Figure 5 game, if $\hat{\Gamma} = \{\gamma_H\}$ — γ_H again denoting the high-quality type—then

$$\text{BR}(\hat{\Gamma}, \text{warranty \& high price}) = \{\text{buy}\},$$

because if the buyer assigns no weight to the low-quality type, his best response to that offer is to buy. In contrast, if $\hat{\Gamma} = \Gamma$, then

$$\text{BR}(\hat{\Gamma}, \text{warranty \& high price}) = \{\text{buy, don't buy}\},$$

because there are beliefs (*e.g.*, $\mu(\gamma_H) \leq 1/3$) such that not buying is a best response and there are beliefs (*e.g.*, $\mu(\gamma_H) \geq 1/3$) such that buying is a best response. Finally, let $V^*(\gamma)$ denote the payoff to a type- γ informed player in the equilibrium under analysis. The Intuitive Criterion can now be formally given:

Definition (Intuitive Criterion). *Consider an equilibrium. For each out-of-equilibrium action of the informed player, a , let*

$$\Gamma^0 = \left\{ \gamma \in \Gamma(a) \mid V^*(\gamma) > \max_{x \in \text{BR}(\Gamma(a), a)} V(a, x, \gamma) \right\}.$$

If there is a type $\gamma' \in \Gamma(a) \setminus \Gamma^0$ such that

$$V^*(\gamma') < \min_{x \in \text{BR}(\Gamma(a) \setminus \Gamma^0, a)} V(a, x, \gamma'), \quad (4.12) \quad \{\text{eq:IntuitiveCriterion}\}$$

then the equilibrium fails the Intuitive Criterion.

The set Γ^0 are types who do better in the equilibrium in question than they could reasonably hope to do by deviating to a . For example, in the pooling PBE of Figure 5, if a is warranty and high price, then $\Gamma^0 = \{\gamma_L\}$ (i.e., the set containing only the low-quality type). Types in Γ^0 have no incentive to play the deviation a . In contrast, if there is a type γ' not in Γ^0 who has the ability to play a (i.e., $\gamma' \in \Gamma(a) \setminus \Gamma^0$) and that type would do better than its equilibrium payoff playing a , even if that triggered the worst response, from its perspective, from the uninformed player given he believes no type in Γ^0 played a (i.e., condition (4.12) holds), then the equilibrium is unreasonable (not “intuitive”) and should be rejected.

The pooling equilibrium of Figure 5 fails the Intuitive Criterion: there is a type, namely the high-quality seller, in $\Gamma(a) \setminus \Gamma^0$ whose equilibrium payoff, 1, is less than the worst she would receive if the buyer, recognizing the low-quality seller would not offer a warranty and high price, plays his best response to the offer of a warranty and high price (which is to buy, yielding the high-quality seller a payoff of 2). The separating equilibrium trivially passes the Intuitive Criterion because there is no out-of-equilibrium action (offer).

4.2.2 APPLICATION OF SIGNALING TO TRADING RELATIONS

Consider, as an initial model, a seller who can take a public action $s \in \mathcal{S}$ prior to trade. For instance, the seller could be a website designer and s is a measure of the quality of her own website. Let her utility be $xp - c(s, \gamma)$, where p is the payment she may receive from the buyer, $\gamma \in \Gamma \subset \mathbb{R}$ is her type, and $c : \mathcal{S} \times \Gamma \rightarrow \mathbb{R}_+$. Assume the cost function induces a *common* complete ordering, $\succ$, over $\mathcal{S}$, where $s \succ s'$ if and only if $c(s, \gamma) > c(s', \gamma)$ for all $\gamma \in \Gamma$ (e.g., $\succ$ denotes “better design” and a better-designed site costs all types of website designers more than a worse-designed site). To avoid issues of countervailing incentives, assume there is a minimum element, $s_0 \in \mathcal{S}$, such that $c(s_0, \gamma) < c(s, \gamma)$ for all γ and $s \in \mathcal{S} \setminus \{s_0\}$ and, moreover,

$$c(s_0, \gamma) = c(s_0, \gamma') \quad (4.13) \quad \{\text{eq:NoCounterVail-Signal}\}$$

for all $\gamma, \gamma' \in \Gamma$. There is no loss of generality in normalizing the common value $c(s_0, \gamma)$ to zero and that normalization is made henceforth.

The buyer’s utility if he buys is $\gamma - p$ and 0 if he does not. Observe that s does not enter his utility function directly; hence, this is a model of *wasteful signaling* insofar as (first-best) efficiency dictates $s \equiv s_0$.

A critical assumption is that the seller’s utility satisfies the Spence-Mirrlees condition. To wit, if $s \succ s'$ and $\gamma > \gamma'$, then

$$c(s, \gamma) - c(s', \gamma) < c(s, \gamma') - c(s', \gamma'); \quad (4.14) \quad \{\text{eq:Sig1-SM}\}$$

that is, a higher-quality type of seller finds it less costly to raise the value of the signal than does a lower-quality type. Letting $s' = s_0$, conditions (4.13) and (4.14) together imply that if $\gamma > \gamma'$ and $s \neq s_0$, then

$$c(s, \gamma) < c(s, \gamma'); \quad (4.15) \quad \{\text{eq:Sig1-SMx}\}$$

that is, a higher type has a lower cost of signaling than does a lower type.

The timing is that the seller chooses s , which the buyer observes, and then she makes a TIOLI price offer.

Assume there is a lowest type, $\underline{\gamma} \geq 0$. Observe that type can guarantee herself a payoff of $\underline{\gamma}$: following any play by the seller, the worst belief possible—from the seller’s perspective—is that the seller’s quality is $\underline{\gamma}$. So if $\underline{\gamma}$ -type seller chooses s_0 and price $\underline{\gamma}$, the buyer has no reason to reject. It follows that in a *separating* equilibrium, the $\underline{\gamma}$ -type seller’s payoff is $\underline{\gamma}$: because the seller’s type is revealed in equilibrium, the buyer will rationally never pay more than $\underline{\gamma}$. It further follows that such a seller must play s_0 in a separating equilibrium. This is a general result: the worst-type seller gets the same payoff in a separating equilibrium as she would have under symmetric information. This underscores that it is not the worst type that is harmed by asymmetric information.

Suppose there are two types, $\Gamma = \{\underline{\gamma}, \bar{\gamma}\}$, and two signals, $\mathcal{S} = \{s_0, s_1\}$. Let $G : \Gamma \rightarrow [0, 1]$ again denote the distribution “nature” uses when determining the seller’s type. Let $g = G(\underline{\gamma})$. Define

$$\gamma_P = g\underline{\gamma} + (1 - g)\bar{\gamma};$$

that is, γ_P is the expected value of γ given the prior distribution. A pooling PBE is $a(\gamma) = (s_0, \gamma_P)$ for both γ (*i.e.*, each type chooses signal s_0 and prices at γ_P); the buyer’s belief is $\mu = (g, 1 - g)$ (*i.e.*, the prior) regardless of the seller’s action; and the buyer accepts all $p \leq \gamma_P$, but rejects all $p > \gamma_P$. Does this PBE satisfy the Intuitive Criterion? Consider the out-of-equilibrium action $\bar{a} = (s_1, \bar{\gamma})$. If

$$\gamma_P > \bar{\gamma} - c(s_1, \bar{\gamma}), \quad (4.16) \quad \{\text{eq:SimpSigIC1}\}$$

then (4.15) implies $\Gamma^0 = \Gamma$, so $\Gamma(\bar{a}) \setminus \Gamma^0 = \emptyset$, which means the Intuitive Criterion is satisfied. If (4.16) is an equality, then (4.15) implies $\Gamma^0 = \{\underline{\gamma}\}$. But if (4.16) is an equality, then (4.12) doesn’t hold, so the Intuitive Criterion is again satisfied. To summarize to this point:

Lemma 12. *For the two-type-two-signal game of the previous paragraph, the pooling equilibrium described above satisfies the Intuitive Criterion if*

$$\bar{\gamma} - c(s_1, \bar{\gamma}) \leq \gamma_P. \quad (4.17) \quad \{\text{eq:SimpSig-GenIC}\}$$

What if (4.17) does not hold? Observe, given the Spence-Mirrlees condition (4.14), there must therefore be a $\tilde{p} \in (\gamma_P, \bar{\gamma})$ such that

$$\tilde{p} - c(s_1, \underline{\gamma}) < \gamma_P < \tilde{p} - c(s_1, \bar{\gamma}). \quad (4.18) \quad \{\text{eq:SimpSigIC2}\}$$

Consider the out-of-equilibrium action $\tilde{a} = (s_1, \tilde{p})$. Expression (4.18) entails $\Gamma^0 = \{\underline{\gamma}\}$ and $\Gamma(\tilde{a}) \setminus \Gamma^0 = \{\bar{\gamma}\}$. Given that $\tilde{p} < \bar{\gamma}$, $\text{BR}(\Gamma(\tilde{a}) \setminus \Gamma^0) = \{1\}$ (i.e., the only possible best response for the buyer is to accept). It follows that (4.12) holds; that is, the pooling PBE *fails* the Intuitive Criterion when (4.17) doesn't hold. To summarize:

Lemma 13. *For the two-type-two-signal game described above, the pooling equilibrium described there satisfies the Intuitive Criterion if and only if condition (4.17) holds.*

Next consider separating PBE. From the discussion above, $a(\underline{\gamma}) = (s_0, \underline{\gamma})$. Clearly if the equilibrium is separating, the high-quality seller must choose signal s_1 . Let her action be $a(\bar{\gamma}) = (s_1, p)$. A necessary condition for this to constitute an equilibrium is that neither type can wish to mimic the other:

$$\underline{\gamma} \geq p - c(s_1, \underline{\gamma}) \quad \text{and} \quad (4.19) \quad \{\text{eq:SimpSig-LowTypeIC}\}$$

$$p - c(s_1, \bar{\gamma}) \geq \underline{\gamma} \quad (4.20) \quad \{\text{eq:SimpSig-HighTypeIC}\}$$

Combining these conditions, a necessary condition is that

$$p \in [\underline{\gamma} + c(s_1, \bar{\gamma}), \underline{\gamma} + c(s_1, \underline{\gamma})] \equiv \mathcal{P}_S.$$

Given that the buyer will never accept a $p > \bar{\gamma}$, it follows from (4.20) that a necessary condition for a separating PBE to exist is that

$$\bar{\gamma} - \underline{\gamma} \geq c(s_1, \bar{\gamma}). \quad (4.21) \quad \{\text{eq:SimpSig-NecSep}\}$$

Because (4.17) can be rewritten as

$$(\bar{\gamma} - \underline{\gamma})g \leq c(s_1, \bar{\gamma}),$$

it follows that if (4.21) doesn't hold, then a pooling PBE satisfying the Intuitive Criterion exists. Define

$$\mathcal{P}^* = \{p | p \leq \bar{\gamma}\} \cap \mathcal{P}_S.$$

If (4.21) holds, then $\mathcal{P}^*$ is non-empty; that is, there exists at least one p that satisfies (4.19) and (4.20), which is acceptable to the buyer. The following is a separating PBE: a low-quality seller plays $(s_0, \underline{\gamma})$ and a high-quality seller plays $(s_1, \bar{p})$, $\bar{p} \in \mathcal{P}^*$; the buyer believes $s = s_0$ or $p > \bar{p}$ means the seller is low quality and that $s = s_1$ and $p \leq \bar{p}$ means the seller is high quality; and the buyer accepts $p \leq \underline{\gamma}$ if $s = s_0$, he accepts $p \leq \bar{p}$ if $s = s_1$, and he otherwise rejects the seller's offer.

Because a separating equilibrium can be constructed for any $\bar{p} \in \mathcal{P}^*$, it follows that there is a continuum of such equilibria if $\mathcal{P}^*$ contains more than a single element. Only one, however, satisfies the Intuitive Criterion. Define $p^* = \max \mathcal{P}^*$. By construction

$$p^* = \min \{\bar{\gamma}, \underline{\gamma} + c(s_1, \underline{\gamma})\}. \quad (4.22) \quad \{\text{eq:SimpSig-pstar}\}$$

Consider some $\bar{p} \in \mathcal{P}^*$, $\bar{p} < p^*$, and the separating PBE in which $\bar{p}$ is the high-quality seller's price offer. Observe there exists a $\tilde{p} \in (\bar{p}, p^*)$. Consider the out-of-equilibrium action $\tilde{a} = (s_1, \tilde{p})$. Given (4.22),

$$\underline{\gamma} > \tilde{p} - c(s_1, \underline{\gamma}).$$

It follows that $\Gamma^0 = \{\underline{\gamma}\}$. Because $\tilde{p} < \bar{\gamma}$, $\text{BR}(\Gamma(\tilde{a}) \setminus \Gamma^0) = \{1\}$. Because $\tilde{p} > \bar{p}$, it follows that (4.12) holds; that is, the PBE fails the Intuitive Criterion. To summarize:

Proposition 16. *Consider the two-type-two-signal game described above. If the difference in quality, $\bar{\gamma} - \underline{\gamma}$, does not exceed $c(s_1, \bar{\gamma})/g$, then a pooling equilibrium in which both types choose the lower-cost signal, s_0 , and price at average quality, γ_P , exists and satisfies the Intuitive Criterion. If the difference in quality is not less than $c(s_1, \bar{\gamma})$, then separating equilibria exist; the unique separating equilibrium to satisfy the Intuitive Criterion is the one in which the low-quality seller chooses the lower-cost signal and prices at $\underline{\gamma}$ and the high-quality seller chooses the higher-cost signal, s_1 , and prices at p^* as given by (4.22).*

As an extension of this model, expand the space of possible signals to $[s_0, \infty)$. The set Γ remains $\{\underline{\gamma}, \bar{\gamma}\}$. Assume $c(\cdot, \gamma)$ is a continuous function for both γ s, with $\lim_{s \rightarrow \infty} c(s, \bar{\gamma}) = \infty$. As will be seen, a consequence of these changes is that there is only one PBE that satisfies the Intuitive Criterion.

Lemma 14. *In a PBE, the low-quality seller plays, with positive probability, at most one signal that is not the lowest signal, s_0 , and only if the high-quality seller also plays that signal with positive probability.*

{lemma:AtMost2Signals}

Lemma 15. *A PBE in which the low-quality seller sends a signal other than the lowest signal does not satisfy the Intuitive Criterion.*

{lemma:NoLowTypeHighSignal}

The logic used in the proof of Lemma 15 (see Appendix) can be extended to prove:

Lemma 16. *A pooling equilibrium in which both seller types send signal s_0 does not satisfy the Intuitive Criterion.*

Lemma 15 and the last lemma together establish that the only PBE that could satisfy the Intuitive Criterion are separating equilibria. Among the separating PBE, only one satisfies the Intuitive Criterion:

Proposition 17. *Consider the signaling game described above in which there are two types and a continuum of signals with a continuous and unbounded cost-of-signaling function. Then the only PBE of that game that survives the Intuitive Criterion is the least-cost separating equilibrium.⁶² (i) the low-quality seller chooses the lowest signal, s_0 , and charges a price equal to her quality, $\underline{\gamma}$; (ii) the high-quality seller chooses the signal s^* , the unique solution to*

{prop:Riley}

$$\bar{\gamma} - \underline{\gamma} = c(s^*, \underline{\gamma}), \quad (4.23) \quad \{\text{eq:Def-sstar-sig}\}$$

⁶²This separating equilibrium is also known as the Riley equilibrium from Riley (1979).

and charges a price equal to her quality, $\bar{\gamma}$; (iii) the buyer believes any signal less than s^* is sent by a low-quality seller and any signal s^* or greater is sent by a high-quality seller; and (iv) the buyer accepts an offer if and only if it yields him a non-negative payoff given his beliefs.

To take stock: Proposition 17 establishes that when the seller can incur a continuum of signaling costs, the only “reasonable” equilibrium is a separating one in which the low-quality seller admits who she is (doesn’t signal) and sets a price equal to the buyer’s value for a low-quality product; and in which the high-quality seller signals just enough to be convincing—the minimum signal that the low-quality seller would be unwilling to mimic—and sets a price equal to the buyer’s value for a high-quality product. Because the high-quality seller chooses the smallest effective signal, the equilibrium is known as *least-cost separating*.

This logic can be extended to a setting in which there are N quality levels, $\{\gamma_0, \dots, \gamma_{N-1}\} = \Gamma$. Assume the same payoffs as before and maintain the Spence-Mirrlees condition. Then, in a least-cost separating equilibrium, the γ_n -type seller plays (s_n^*, γ_n) , where $s_0^* = s_0$ and s_n^* , $n > 0$, is defined recursively as the solution to

$$c(s_n^*, \gamma_{n-1}) = \gamma_n - \gamma_{n-1} + c(s_{n-1}^*, \gamma_{n-1}).$$

A Continuum of Types. To conclude this subsection, suppose that $\Gamma = [\underline{\gamma}, \bar{\gamma}] \subset \mathbb{R}_+$ (i.e., Γ is an interval). Maintain the same payoffs, but assume now that $c(\cdot, \cdot)$ is twice differentiable in each argument. The Spence-Mirrlees condition can now be given by

$$\frac{\partial^2 c(s, \gamma)}{\partial \gamma \partial s} < 0; \quad (4.24)$$

that is, the marginal cost of the signal is falling in the seller’s type. The goal is to derive a separating equilibrium in which a γ -type seller plays $(s(\gamma), \gamma)$, where $s(\cdot)$ is a differentiable function.

Because all $s \in s([\underline{\gamma}, \bar{\gamma}])$ are potentially played in equilibrium, a necessary condition is

$$\gamma \in \operatorname{argmax}_{\gamma' \in \Gamma} \gamma' - c(s(\gamma'), \gamma) \quad (4.25) \quad \{\text{eq:ContSig-Argmax}\}$$

for all γ . All functions are differentiable, so consider replacing (4.25) with the first-order condition

$$0 = 1 - \frac{\partial c(s(\gamma), \gamma)}{\partial s} s'(\gamma) \quad (4.26) \quad \{\text{eq:ContSig-DE}\}$$

for all γ . Because $\partial c / \partial s > 0$, (4.26) implies $s(\cdot)$ is strictly increasing—higher types signal more than lower types. In addition, recall the lowest-quality type doesn’t signal at all in a separating PBE:

$$s(\underline{\gamma}) = s_0. \quad (4.27) \quad \{\text{eq:ContSig-Initial}\}$$

It follows that a function $s^*(\cdot)$ that solves the differential equation (4.26) given initial condition (4.27) is part of a separating PBE provided (4.26), with $s(\gamma) =$

$s^*(\gamma)$, is a *sufficient* condition for a maximum. That it follows from the Spence-Mirrlees condition:

$$1 - \frac{\partial c(s^*(\gamma'), \gamma)}{\partial s} \underbrace{\left(\frac{\partial c(s^*(\gamma'), \gamma')}{\partial s} \right)^{-1}}_{=s^{*'}(\gamma')} \quad (4.28)$$

is negative for all $\gamma' > \gamma$ and positive for all $\gamma' < \gamma$.

As an example, suppose $s_0 = 0$ and $c(s, \gamma) = s^2/\gamma$. Expression (4.26) implies

$$\gamma = 2s(\gamma)s'(\gamma).$$

The class of solutions to this differential equation is $s(\gamma) = \gamma^2/2 + k$, k a constant. Expression (4.27) implies $k = -\underline{\gamma}^2/2$; that is,

$$s^*(\gamma) = \frac{\gamma^2 - \underline{\gamma}^2}{2}.$$

For more on signaling games with a continuum of types, see Mailath (1987).

4.3 EXPERIENCE GOODS AND SELLER REPUTATION

Buyers often frequent the same seller repeatedly. If the quality of the seller's good is constant, then buyers will learn its quality over time through experience. Beyond issues of getting buyers to try the product in the first place—which resemble the issues of the one-shot analysis considered so far—such settings are straightforward and, thus, of little independent interest. A more interesting setting is one in which the quality of the seller's good varies over time; as might occur if there are variations across batches (*e.g.*, random fluctuations in chemical processes) or the product is agricultural (*e.g.*, variation across vintages of a given vineyard's wine).⁶³ In such settings, a seller may be able to develop a reputation for truthfully revealing quality.

Let $\gamma_t \in \Gamma$ denote the quality of the seller's good in period t . The set of possible qualities, Γ , is time invariant, but realized quality can vary period to period. To keep the analysis straightforward, assume each period's quality is independently drawn from the same distribution, $G : \Gamma \rightarrow [0, 1]$. Assume the unit cost of producing a quality- γ good is $c(\gamma)$.⁶⁴

⁶³Quality is here determined exogenously. This is what distinguishes the analysis in this section from that of models of seller reputation with *endogenous* quality, such as those of Klein and Leffler (1981) and Shapiro (1982); see Section 5.1 *infra*.

⁶⁴A natural assumption, given that quality is exogenous, is $c(\gamma)$ is a constant. On the other hand, recalling that what is relevant is opportunity cost, it is also plausible that $c(\cdot)$ is an increasing function (*e.g.*, rather than selling its grapes as wine under its own label, a vineyard can sell the grapes to another winery or it sell the wine under a generic label, with the return from such activity being $c(\gamma)$).

Assume the per-period payoffs of the buyer and seller, respectively, to be

$$U_B = x(\gamma - p) \text{ and } U_S = x(p - c(\gamma)),$$

where $x \in \{0, 1\}$ again indicates trade.⁶⁵

4.3.1 A BASIC MODEL

{sect: SellerReputation1BasicModel}

Suppose that $\Gamma = \{\gamma_L, \gamma_H\}$, $\gamma_L < \gamma_H$, and define $g = G(\gamma_L)$. Let $c(\gamma) = \hat{c}$ for both γ . Assume, critically, that at least some trade is desirable: $\gamma_H > \hat{c}$. For convenience, consider pure-strategy equilibria only. In a one-shot game, there are two possibilities: if

$$\gamma_P = g\gamma_L + (1 - g)\gamma_H \geq \hat{c}, \quad (4.29) \quad \{\text{eq:Def-gammaP}\}$$

then an equilibrium exists in which both types of seller set $p = \gamma_P$ and the buyer purchases; or, if the inequality in (4.29) does not hold, then the seller sets $p > \gamma_P$ and the buyer does not purchase in equilibrium (there is no market).

Now suppose that the game is repeated infinitely. Let $\delta \in (0, 1)$ be the seller's discount factor. Consider the following strategy for the buyer to play in any given period:

- if the seller has never lied, believe the seller's announcement of her type and buy if the price she quotes does not exceed her announced type; but
- if the seller has ever lied, disregard the seller's announcement of her type and buy if the price she quotes does not exceed γ_P .

Clearly if the seller has never lied and her type is γ_H , her best response to the buyer's strategy is to announce her type as γ_H and set $p = \gamma_H$. What if her type is γ_L ? If she tells the truth, then the expected present discounted value (PDV) of her payoffs is

$$\begin{aligned} & (\gamma_L - \hat{c})^+ + \sum_{t=1}^{\infty} \delta^t (g(\gamma_L - \hat{c})^+ + (1 - g)(\gamma_H - \hat{c})) \\ &= (\gamma_L - \hat{c})^+ + \frac{\delta}{1 - \delta} (g(\gamma_L - \hat{c})^+ + (1 - g)(\gamma_H - \hat{c})), \end{aligned} \quad (4.30) \quad \{\text{eq:TPayoff-SimpleRepModel}\}$$

where $(z)^+ = z$ if $z \geq 0$ and equals 0 if $z < 0$. If she lies, then the expected PDV of her payoffs is

$$\gamma_H - \hat{c} + \sum_{t=1}^{\infty} \delta^t (\gamma_P - \hat{c})^+ = \gamma_H - \hat{c} + \frac{\delta}{1 - \delta} (\gamma_P - \hat{c})^+. \quad (4.31) \quad \{\text{eq:LPayoff-SimpleRepModel}\}$$

⁶⁵Given the assumption of quasi-linear utility, there is no further loss of generality in letting γ denote the buyer's utility from a good of quality γ (recall footnote 53 *supra*). Given the opportunity-cost definition of cost, the seller's utility could equivalently be written

$$U_S = xp + (1 - x)c(\gamma).$$

Suppose $\gamma_P \geq \hat{c}$ (*i.e.*, trade will incur in the equilibrium of the one-shot game). From (4.29), if $\gamma_L \geq \hat{c}$, then (4.30) is less than (4.31): a low-quality seller will lie. Hence, there is no equilibrium in which the seller is truthful. Note, the condition under which this holds, $\gamma_L \geq \hat{c}$, means the one-shot equilibrium is efficient—trade should always occur and does. There is no efficiency loss from being unable to support truth-telling in equilibrium.

Continue to suppose $\gamma_P \geq \hat{c}$, but now assume $\gamma_L < \hat{c}$: trade will always occur in the one-shot equilibrium, but it is inefficient with a low-quality seller. Straightforward algebra reveals that (4.30) is at least (4.31) provided:

$$\delta \geq \frac{\gamma_H - \hat{c}}{(\gamma_H - \hat{c}) - g(\gamma_L - \hat{c})}. \quad (4.32) \quad \{\text{eq:delta-SimpleRepModel1}\}$$

In other words, if the seller is sufficiently patient (has a discount factor greater than the RHS of (4.32)), then her best response to the buyer's strategy is to always tell the truth (here, seek to sell if and only if she has high quality to offer). Given she is telling the truth, the buyer's strategy is clearly a best response for him—this is an equilibrium. Observe that the RHS of (4.32) is decreasing in g —the greater the likelihood of low quality, the easier it is to sustain an efficient equilibrium.⁶⁶

Finally, suppose $\gamma_P < \hat{c}$; that is, trade never occurs in the equilibrium of the one-shot game, which is inefficient. Necessarily, $\gamma_L < \hat{c}$. Straightforward algebra reveals that (4.30) is at least (4.31) provided:

$$\delta \geq \frac{1}{2 - g}. \quad (4.33) \quad \{\text{eq:delta-SimpleRepModel2}\}$$

For the same reasons just given, if (4.33) holds, then an equilibrium exists in which the seller tells the truth (*i.e.*, seeks to sell if and only if she has high quality to offer). Observe that the RHS of (4.33) is decreasing in g —the less likely low quality is, the easier it is to sustain an efficient equilibrium.⁶⁷

As is the rule with repeated games, the feasibility of sustaining a desired outcome (here, efficient trade) is easier the worse the punishment for deviating. Here, a deviating seller is effectively punishing herself (regardless of equilibria, the buyer's expected surplus is always zero). Put slightly differently, the issue is whether a current seller sufficiently internalizes the externality that her lying imposes on her future selves who will have a high-quality product. The punishment that deters dishonesty is losing the future gains from efficient trade. Hence, when the one-shot equilibrium is already efficient, there is no scope for sustaining truthful revelation. When the one-shot equilibrium is inefficient (*i.e.*, $\gamma_L < \hat{c}$), efficiency gains arise from the use of valuable information. It follows that the more valuable is the information, the more feasible an efficient outcome is (*i.e.*, the lower the cutoff δ below which efficiency is not sustainable). From

⁶⁶Given that $\gamma_P \geq \hat{c}$, it must be that $g \leq (\gamma_H - \hat{c})/(\gamma_H - \gamma_L)$.

⁶⁷Given that $\gamma_P < \hat{c}$, it follows $g > (\gamma_H - \hat{c})/(\gamma_H - \gamma_L)$.

the analysis of (4.32) and (4.33), efficient trade in a repeated context (given $\gamma_L < \hat{c}$) is most feasible (the δ cutoff is smallest) when

$$g = \frac{\gamma_H - \hat{c}}{\gamma_H - \gamma_L}. \quad (4.34)$$

Straightforward algebra shows that when g takes that value, $\gamma_P = \hat{c}$; that is, absent information revelation, the seller is indifferent between selling and not selling. This reflects a general result: information is most valuable when, absent the information, the decision maker is indifferent between alternative decisions (see Proposition 19 *infra*). To summarize:

Proposition 18. *A seller's type (quality) varies independently period to period. If trade would be efficient in a one-shot game despite the buyer's ignorance of quality, then there is no equilibrium of a repeated game with truthful revelation of type. If trade would be inefficient with positive probability in a one-shot game, then an equilibrium with truthful revelation exists in a repeated game if the seller is sufficiently patient (has a high enough discount factor, δ). It is easier to sustain such an equilibrium (i.e., δ can be lower) the more valuable is information about type.*

{prop: SellerRep1}

4.3.2 AN ASIDE: THE VALUE OF INFORMATION

It was noted that information is most valuable when the decision maker would otherwise be indifferent between different actions in the absence of that information. This subsection establishes that insight more formally.

A decision maker can make one of two decisions: d_0 or d_1 . Given there are only two possible decisions, there is no loss of generality in assuming two possible states, s_0 and s_1 .⁶⁸ Assume that decision d_i is best for the decision maker in state s_i . Formally, let her payoff if she makes decision d_i in state s_j be $U(d_i|s_j)$. The assumption d_i is best given s_i means $U(d_i|s_i) > U(d_i|s_j)$ for any i and j , $i \neq j$. Define

$$L(d_j|s_i) = U(d_i|s_i) - U(d_j|s_i)$$

as the loss from choosing d_j in state s_i . Let q_i denote the probability that the true state is s_i .

Absent information as to the state, the decision maker decides d^* , where

$$d^* = \operatorname{argmax}_{d \in \{d_0, d_1\}} q_0 U(d|s_0) + q_1 U(d|s_1).$$

The value of information, V , is the difference between the decision maker's expected payoff if she makes her decision with information about the state and her expected payoff if she makes it in ignorance:

$$V = (q_0 U(d_0|s_0) + q_1 U(d_1|s_1)) - (q_0 U(d^*|s_0) + q_1 U(d^*|s_1))$$

⁶⁸If there are more than two states, we can view the state space as partitioned into two: the set of states in which one decision is optimal and the set of states in which the other decision is optimal. Relabel these two sets (events) as the relevant states.

Without loss of generality, suppose $d^* = d_i$. Algebra reveals

$$V = (1 - q_i)L(d_i|s_{-i});$$

hence, the value of information is decreasing in the probability of state in which d_i is the correct action and increasing in the loss from choosing d_i in state s_{-i} . The probability q_i cannot be too low *ceteris paribus*, because otherwise d^* would no longer be d_i . Similarly, $L(d_i|s_j)$ cannot be too great *ceteris paribus*, because otherwise $d^* \neq d_i$. These insights yield

Proposition 19. *Consider a binary decision problem. Holding constant all other parameters, the value of information is greatest when the probabilities of the two states are such that the decision maker would be indifferent between her two alternatives in the absence of information. Similarly, holding constant all other parameters, the value of information is greatest when the losses from making the wrong decision are such that the decision maker would be indifferent between her two alternatives in the absence of information.*

{prop:Value-of-Info}

4.3.3 SELLER REPUTATION WITH SIGNALING

Return to the model of Section 4.3.1. Suppose, now, though the seller can signal: specifically, at cost $C(s, \gamma)$, she can send signal s . Assume $C : [0, \infty) \times \{\gamma_L, \gamma_H\}$ satisfies the Spence-Mirrlees condition:

$$\frac{\partial C(s, \gamma_H)}{\partial s} < \frac{\partial C(s, \gamma_L)}{\partial s}.$$

For example, s could be the number of positive reviews or tests the seller can produce.⁶⁹ Consistent with the usual notion of cost, $C(0, \gamma) = 0$ for both γ and $\partial C(s, \gamma)/\partial s > 0$ both γ . Assume for all $k \in \mathbb{R}$ an $s \in \mathbb{R}$ exists such that $k = C(s, \gamma_L)$.

Define s^* as the value of s that solves

$$\max\{0, \gamma_L - \hat{c}\} = \gamma_H - \hat{c} - C(s, \gamma_L). \quad (4.35) \quad \{\text{eq:LeastCostSignalSellerRep}\}$$

It is readily shown that the assumptions given imply that s^* exists and is unique. The Spence-Mirrlees condition and (4.35) ensure that

$$\gamma_H - \hat{c} - C(s^*, \gamma_H) > \max\{0, \gamma_L - \hat{c}\};$$

hence, a high-type seller would prefer to send signal s^* than to shutdown or be seen as a low-quality seller. It is readily seen that the one-shot game satisfies the conditions of Proposition 17. Consequently, the only reasonable PBE of the

⁶⁹Consider γ to be the *average* quality of the product in a given period and assume all parties are risk neutral. Suppose reviews or tests are *hard information*: the seller can suppress those she does not like, but if she reports a review or test it must be accurate. A seller with higher average quality (*i.e.*, a greater γ) will find it easier *ceteris paribus* to generate s favorable reports or reviews than will a seller lower average quality.

one-shot game is the least-cost separating equilibrium in which a low-quality seller either does not sell (if $\gamma_L < \hat{c}$) or sells without signaling (if $\gamma_L \geq \hat{c}$), while a high-quality seller sells at price γ_H and transmits signal s^* .

Consider, now, a infinitely repeated version of this game.⁷⁰ Again $\delta \in (0, 1)$ denotes the seller's discount factor. The goal is to determine conditions such that an equilibrium exists in which the seller announces her type truthfully *without* signaling. To that end, consider the following strategy for the buyer:

- if the seller has never lied, believe the seller's announcement of her type and buy if the price she quotes does not exceed her announced type; but
- if the seller has ever lied, disregard the seller's announcement of her type. Believe her type that period is γ_H if she sends a signal s^* or greater and believe her type is γ_L otherwise. Buy if the price she quotes does not exceed the quality her signal indicates.

In other words, if the seller ever lies, then the game reverts to infinite repetition of the one-shot game with play defined by the sole PBE that satisfies the Intuitive Criterion of Cho and Kreps (1987).

As in Section 4.3.1, a high-quality seller has no incentive to lie about her type. The issue is whether a low-quality type would lie. Her payoff from telling the truth is

$$\begin{aligned} & (\gamma_L - \hat{c})^+ + \sum_{t=1}^{\infty} \delta^t (g(\gamma_L - \hat{c})^+ + (1 - g)(\gamma_H - \hat{c})) \\ &= (\gamma_L - \hat{c})^+ + \frac{\delta}{1 - \delta} (g(\gamma_L - \hat{c})^+ + (1 - g)(\gamma_H - \hat{c})). \end{aligned} \quad (4.36) \quad \{\text{eq:TSellerRepSignal}\}$$

Her payoff from lying is

$$\begin{aligned} & \gamma_H - \hat{c} + \sum_{t=1}^{\infty} \delta^t (g(\gamma_L - \hat{c})^+ + (1 - g)(\gamma_H - \hat{c} - C(s^*, \gamma_H))) \\ &= \gamma_H - \hat{c} + \frac{\delta}{1 - \delta} (g(\gamma_L - \hat{c})^+ + (1 - g)(\gamma_H - \hat{c} - C(s^*, \gamma_H))). \end{aligned} \quad (4.37) \quad \{\text{eq:LSellerRepSignal}\}$$

Truth-telling dominates lying (*i.e.*, (4.37) does not exceed (4.36)) if

$$\frac{\delta}{1 - \delta} (1 - g)C(s^*, \gamma_H) \geq \gamma_H - \hat{c} - (\gamma_L - \hat{c})^+; \quad (4.38) \quad \{\text{eq:TTconditionSellerRepSignal}\}$$

that is, should the expected discounted cost of future signaling (if her reputation for truth-telling be lost) exceed the one-period gain from fooling the buyer, then the seller's best response is to tell the truth every period. Given she will tell the truth, the buyer is also playing a best response and we have an equilibrium.

⁷⁰Mester (1992) considers a finitely repeated game of signaling in which the signaler's type is correlated across periods (in contrast to here, where it is independently drawn each period).

The fraction $\frac{\delta}{1-\delta}$ is increasing in δ . It follows, therefore, from (4.38) that the greater the likelihood the seller is high type or the greater the cost of signaling, the lower is the cutoff value of δ for which truth-telling can be sustained. To summarize:

Proposition 20. *A seller's type (quality) varies independently period to period. Suppose the seller has access to a costly signal and the cost of signaling function satisfies the Spence-Mirrlees condition. An equilibrium of the repeated game exists in which there is no signaling in equilibrium, but rather the seller simply announces her type truthfully each period if the seller is sufficiently patient (has a high enough discount factor, δ). It is easier to sustain such an equilibrium (i.e., δ can be lower) the more likely the seller is to be the high type or the greater the high type's cost of signaling.*

{prop: SellerRepSignal}

In Proposition 18, a truth-telling equilibrium did not exist if trade was efficient in the one-shot game. Here, trade is efficient in the one-shot game insofar as it occurs if and only if $\gamma \geq \hat{c}$. But the one-shot game here is never wholly efficient because the high-type seller engages in costly signaling. If $\gamma_L \geq \hat{c}$, then this signaling is completely wasteful from a welfare perspective. If $\gamma_L < \hat{c}$, then signaling enhances efficiency insofar as it ensures trading efficiency, but the first-best is nevertheless not achieved.

In this model, as opposed to the one of Section 4.3.1, what deters a low-quality seller from lying is the knowledge that her future selves will have to pay signaling costs in the future. The effectiveness of this deterrent is increasing in both the amount of those costs and the likelihood that they will have to be paid. This is why, in contrast to the result in Proposition 18, the ease of sustaining a truth-telling equilibrium is monotonically increasing in the probability of the high quality.

A final question is what if (4.38) does not hold: does this imply the equilibrium is infinite repetition of the equilibrium of the one-shot game (i.e., the seller sends signal s^* when she is high quality)? To answer this, we need to ask if there could be an equilibrium of the repeated game with only some signaling. Specifically, we seek an equilibrium of the following sort: when high type, the seller announces her type as γ_H and sends signal $\hat{s} \in (0, s^*)$; when low type, she announces her types as γ_L and sends no signal. The buyer believes the seller's announcement if accompanied by the appropriate signal, provided he's never been lied to. If he's been lied to, then the game reverts to infinite repetition of the one-shot equilibrium. As before, the issue is whether a low-type seller's payoff from truth telling,

$$(\gamma_L - \hat{c})^+ + \frac{\delta}{1-\delta} \left(g(\gamma_L - \hat{c})^+ + (1-g)(\gamma_H - \hat{c} - C(\hat{s}, \gamma_H)) \right), \quad (4.39) \quad \{\text{eq:TSellerRepPartSig}\}$$

exceeds her payoff from lying,

$$\gamma_H - \hat{c} - C(\hat{s}, \gamma_L) + \frac{\delta}{1-\delta} \left(g(\gamma_L - \hat{c})^+ + (1-g)(\gamma_H - \hat{c} - C(s^*, \gamma_H)) \right). \quad (4.40) \quad \{\text{eq:FSellerRepPartSig}\}$$

Expression (4.39) exceeds (4.40) if

$$\begin{aligned} \frac{\delta}{1-\delta}(1-g)(C(s^*, \gamma_H) - C(\hat{s}, \gamma_H)) &\geq (\gamma_H - \hat{c}) - (\gamma_L - \hat{c})^+ - C(\hat{s}, \gamma_L) \\ &= C(s^*, \gamma_L) - C(\hat{s}, \gamma_L), \end{aligned} \quad (4.41) \quad \{\text{eq:TFPartSig}\}$$

where the equality follows because s^* is the least-cost level of signaling in the equilibrium of the one-shot game. Observe that (4.41) reduces to (4.38) if $\hat{s} = 0$. Given Spence-Mirrlees, if (4.41) holds for $\hat{s}$ it must hold for all $\hat{s}' \in (\hat{s}, s^*]$. Consequently, if, as conjectured, (4.38) does not hold, then (4.41) fails to hold for all $\hat{s} \in (0, s^*)$. To conclude:

Proposition 21. *Given the signaling game of this section, the equilibrium of the infinitely repeated game is either one with honest announcements of type and no signaling or it is simply infinite repetition of the equilibrium of the one-shot game.*⁷¹

5 ENDOGENOUS ASYMMETRIES OF INFORMATION AT TIME OF CONTRACTING

{sect:Endo_Asymmetry}

To this point, the parties have been endowed with private information. In many situations, their private information arises because of actions they take; that is, it is *endogenous*. For instance, a seller's decisions about materials, production methods, and the like could determine the quality of her product. As a second example, a buyer could make investments in complementary assets that affect his utility from purchasing the seller's product (*e.g.*, a buyer's utility from buying a new DVD player could depend on the number of DVDs previously acquired).

Suppose that private information is the consequence of actions. These actions—typically investments—fall into two broad categories: selfish and cooperative (to use the terminology of Che and Hausch, 1999). Selfish actions directly affect the actor's payoff from trade. Cooperative actions directly affect the payoffs of the actor's trading partner. For instance, a seller's investment in the quality of her product is cooperative, while a buyer's acquisition of complementary assets is selfish. Given the focus here on situations in which the seller possesses all the bargaining power (makes TIOLI offers), only three possibilities are of interest:⁷² cooperative action by the seller, cooperative action by the buyer, and selfish action by the buyer. Selfish actions by the seller are not of interest because the seller's (contract proposer's) private information about her own payoffs generally do not create distortions.

⁷¹This result is, in part, due to the signal's not being directly productive. In a game in which signaling is partially productive, equilibria with honest announcements and limited signaling are possible. See Hermalin (2007) for an example.

⁷²Recall, in this chapter, that “seller” and “buyer” are shorthand for contract proposer and contract recipient, respectively.

5.1 COOPERATIVE SELLER ACTIONS

{sect:endo-quality}

Suppose the seller takes an action, $q \in \mathcal{Q}$, that affects the buyer's payoff should trade occur. In a buyer-seller relationship, a natural interpretation is that q affects or is the quality of the seller's product.

As a basic model: the buyer wants at most one unit, $\mathcal{Q} = [\underline{q}, \infty)$, $\underline{q} > 0$, and the payoffs of seller and buyer are

$$U_S = x(p - c(q)) \text{ and } U_B = x(q - p),$$

respectively; where $x \in \{0, 1\}$ indicates whether trade occurs, p is price, and $c : \mathcal{Q} \rightarrow \mathbb{R}_+$. Critically, assume the seller's cost is increasing in the quality of the good she produces: $q > q' \Rightarrow c(q) > c(q')$. Assume there exist $q \in \mathcal{Q}$ such that $q > c(q)$ (i.e., trade is efficient for some quality levels), but that there exists a finite $\bar{q}$ such that $q < c(q)$ for all $q > \bar{q}$ (i.e., too much quality is inefficient to produce).

If the buyer could observe the seller's choice of q , then the seller would do best to choose p and q to solve

$$\max_{p, q} p - c(q)$$

subject to

$$q - p \geq 0.$$

It was earlier established that the buyer's participation constraint binds (see Proposition 1), so substituting the constraint the problem is

$$\max_q q - c(q). \quad (5.1) \quad \{\text{eq:WelfareMaxQuality}\}$$

In other words, were the buyer able to observe quality, the seller would have an incentive to choose a quality that maximizes welfare: efficiency would attain.

If the buyer cannot observe quality, then one of two possibilities arises: if $\underline{q} \geq c(\underline{q})$, the seller offers a product of that quality and sets a price of $p = \underline{q}$; or, if $\underline{q} < c(\underline{q})$, then there is no market. To understand these conclusions, observe that if $\hat{p}$ is the maximum price the buyer will accept, then, anticipating a sale at $\hat{p}$, the seller's choice of quality is

$$\max_q \hat{p} - c(q).$$

The sole solution is $q = \underline{q}$. Hence, the highest quality the buyer can expect is $\underline{q}$. If trade is to occur, the largest price he will accept is $p = \underline{q}$. Knowing this, the seller either shuts down if $\underline{q} < c(\underline{q})$ (she cannot make a profit at that price); or she charges $p = \underline{q}$ and provides the lowest possible quality. Unless $\underline{q}$ is a solution to (5.1), the outcome is inefficient. To summarize:

{prop:BadQuality}

Proposition 22. *Suppose the quality of an experience good is endogenous, with higher quality costing the seller more. Then the equilibrium of the one-shot game is inefficient unless minimum quality is welfare maximizing.*

5.1.1 SELLER REPUTATION MODELS

A better outcome than the Proposition 22 outcome can attain in an infinitely repeated game if the seller is sufficiently patient.⁷³ To wit, suppose q^* is a solution to (5.1). To make the problem of interest, assume $\underline{q}$ is not a solution. Define

$$\underline{\pi} = \max \{0, \underline{q} - c(\underline{q})\};$$

that is, $\underline{\pi}$ is the seller's profit in the equilibrium of the one-shot game. Observe that repetition of the one-shot game means the seller is free to change her quality each period if she wishes.

Observe that if the buyer believes quality will be $\underline{q}$, then the seller's best response is either shutdown ($\underline{q} < c(\underline{q})$) or offer quality $\underline{q}$ at price $\underline{q}$. Hence, even in the infinitely repeated game, there is a subgame-perfect equilibrium in which the seller's per-period payoff is $\underline{\pi}$.

Consider the following strategy for the buyer:

- if the seller has never provided quality less than q^* , then believe she is offering quality q^* this period; and buy if and only if the price she charges does not exceed q^* ;
- but if the seller has ever provided quality less than q^* , then believe she is offering quality $\underline{q}$ if she offers to sell; and buy if and only if the price she charges does not exceed $\underline{q}$.

If the buyer is expecting quality q^* , then the seller's best deviation from offering quality q^* remains $\underline{q}$. Hence, offering quality q^* and charging price q^* is the seller's best response to the buyer's strategy if

$$\sum_{t=0}^{\infty} \delta^t (q^* - c(q^*)) \geq q^* - c(\underline{q}) + \sum_{t=1}^{\infty} \delta^t \underline{\pi},$$

where $\delta \in (0, 1)$ is again the seller's discount factor. Straightforward algebra reveals that condition holds provided

$$\delta \geq \frac{c(q^*) - c(\underline{q})}{q^* - \underline{\pi} - c(\underline{q})}. \quad (5.2) \quad \{\text{eq:KL-delta}\}$$

If the seller's discount factor satisfies condition (5.2), then there is an equilibrium in which welfare-maximizing quality is provided. That and readily done comparative statics yield:

Proposition 23. *Consider an infinitely repeated game in which the seller chooses quality each period. If the seller is sufficiently patient (as defined by condition*

⁷³This analysis is along the lines of that in Klein and Leffler (1981). A difference is that Klein and Leffler allow for multiple sellers to enter the market and compete. This enriches the analysis, but does not matter for the points being made here.

(5.2)), then an equilibrium exists in which the welfare-maximizing level of quality is provided each period. Such an equilibrium is supported for a larger set of discount factors the smaller is the cost difference between producing welfare-maximizing quality and producing minimal quality *ceteris paribus*.

The last part of the proposition follows because what tempts the seller to cheat on quality is the cost savings from producing minimal quality versus higher quality: the smaller the temptation, the easier it is to sustain an equilibrium with honest provision of high quality.

If $\underline{q} \geq c(\underline{q})$, then (5.2) becomes

$$\delta \geq \frac{c(q^*) - c(\underline{q})}{q^* - \underline{q}}. \quad (5.3) \quad \{\text{eq:KL-delta2}\}$$

If $c(\cdot)$ is a strictly convex function, then the RHS of expression (5.3) is increasing in q^* .⁷⁴ Hence, if the seller's discount factor is less than the RHS of (5.3), then there may exist a $\hat{q} \in (\underline{q}, q^*)$ such that

$$\frac{c(\hat{q}) - c(\underline{q})}{\hat{q} - \underline{q}} = \delta. \quad (5.4) \quad \{\text{eq:KL-delta3}\}$$

It follows, from now familiar logic, that if (5.4) holds, then there is an equilibrium of the infinitely repeated game in which the seller supplies quality $\hat{q}$. Given the assumed convexity of $c(\cdot)$ and the optimality of q^* , it must be that supplying $\hat{q}$ is welfare superior to supplying minimum quality. To summarize:

Proposition 24. *Assume an infinitely repeated game in which the seller chooses quality each period. Assume the cost-of-quality function, $c(\cdot)$, is strictly convex and the production of minimum quality, $\underline{q}$, is welfare superior to shutting down (i.e., $\underline{q} \geq c(\underline{q})$). Then there is an equilibrium of the repeated game in which quality greater than minimal quality is supplied provided the seller is sufficiently patient; that is, provided her discount factor δ satisfies⁷⁵*

$$\delta > \lim_{q \downarrow \underline{q}} \frac{c(q) - c(\underline{q})}{q - \underline{q}}. \quad (5.5) \quad \{\text{eq:KL-delta4}\}$$

⁷⁴Proof: Let $q' > q > \underline{q}$. Convexity entails

$$c(q) < \frac{q' - q}{q' - \underline{q}} c(\underline{q}) + \frac{q - \underline{q}}{q' - \underline{q}} c(q')$$

Subtracting $c(\underline{q})$ from both sides yields

$$c(q) - c(\underline{q}) < \frac{q - \underline{q}}{q' - \underline{q}} (c(q') - c(\underline{q})).$$

The claim follows.

⁷⁵Because $c(\cdot)$ is convex, the limit in (5.5) (the right derivative) exists. See van Tiel (1984, Theorem 1.6).

5.1.2 ONE-TIME QUALITY CHOICE WITH CONSUMER LEARNING

Suppose that the seller chooses quality, q , once and for all at time 0.⁷⁶ The cost of producing a unit of quality q remains $c(q)$. Assume that the seller will potentially trade with the buyer in T periods (where T could be infinity). Once the buyer has experienced the good, he knows for sure what its quality will be in future periods. To keep the analysis straightforward, assume trading minimal quality yields non-negative surplus (*i.e.*, assume $\underline{q} \geq c(\underline{q})$).

To construct an equilibrium, suppose that the seller plays a pure strategy with respect to her choice of quality. In equilibrium, the buyer must correctly anticipate this.⁷⁷ Let q^e be the quality he anticipates the seller will choose. Hence, if the buyer has not yet purchased, he will buy if the seller offers the good at price q^e or less and won't purchase otherwise.⁷⁸ Once the buyer has purchased, he knows quality for sure and the seller can, therefore, charge him price equal to actual quality, q . Because $q^e \geq \underline{q} \geq c(\underline{q})$, the seller can have no reason to avoid a sale in the initial period, 0. Hence, if her actual choice is q , her discounted profit is

$$\Pi = q^e - c(q) + \sum_{t=1}^{T-1} \delta^t (q - c(q)) = q^e + \frac{1}{1-\delta} ((\delta - \delta^T)q - (1 - \delta^T)c(q)).$$

Maximizing Π is equivalent to maximizing

$$\max_q \frac{\delta - \delta^T}{1 - \delta^T} q - c(q). \quad (5.6) \quad \{\text{eq:One-TimeMax}\}$$

Because $\delta < 1$, $\frac{\delta - \delta^T}{1 - \delta^T} < 1$. Hence, the solution to (5.6) cannot exceed the minimum level of quality that maximizes welfare. Under standard assumptions— $c(\cdot)$ is strictly convex and everywhere differentiable—the solution to (5.6)—call it q^{SB} —is strictly less than the, then, unique welfare-maximizing quality, q^* . Observe that if $T = 1$ —this is just the one-period game—then $q^{\text{SB}} = \underline{q}$, as is to be expected. In equilibrium, the buyer must correctly anticipate the seller's quality choice; that is, $q^e = q^{\text{SB}}$. To summarize:

Proposition 25. *Suppose buyer and seller potentially trade in T periods ($1 \leq T \leq \infty$), but the seller sets quality for all time prior to the initial period of trade. Suppose trade is welfare superior to no trade even at minimal quality (*i.e.*, $\underline{q} \geq c(\underline{q})$). Then there is an equilibrium in which the seller sets the same price every period, which equals the quality of the good, and there is trade in*

`{prop:OneTimeQuality}`

⁷⁶The analysis in this section is similar in spirit to that in Section 2 of Shapiro (1982).

⁷⁷The notion that the buyer—consumers more generally—can engage in such game-theoretic reasoning separates the analysis here from some of the analysis in Shapiro (1982) and from Shapiro (1983) (as well as some earlier literature), in which the buyer's estimation of quality follows a more exogenously given path. The analysis below leading up to Proposition 26 illustrates how the analysis changes with less rational consumers.

⁷⁸Given his beliefs, the buyer anticipates no gain from experimenting.

every period. The quality of the good solves the program (5.6) and is never greater than any welfare-maximizing quality. It is strictly less for all T if $c(\cdot)$ is strictly convex and everywhere differentiable.

Continuing to assume $c(\cdot)$ differentiable and strictly convex, the following is readily shown:

Corollary 4. *Maintain the assumptions of Proposition 25. Assume, in addition, that $c(\cdot)$ is everywhere differentiable and strictly convex. Let q^{SB} be the solution to the program (5.6); that is, equilibrium quality and price. Then*

{corr:OneTimeQuality}

(i) $q^{\text{SB}} = \underline{q}$ (minimal quality) if $\frac{\delta - \delta^T}{1 - \delta^T} \leq c'(\underline{q})$;

(ii) $q^{\text{SB}} > \underline{q}$ if $\frac{\delta - \delta^T}{1 - \delta^T} > c'(\underline{q})$; and

(iii) q^{SB} is nondecreasing in the number of periods, T , and strictly increasing at T if $\frac{\delta - \delta^T}{1 - \delta^T} > c'(\underline{q})$.

The last result follows because $\frac{\delta - \delta^T}{1 - \delta^T}$ is increasing in T . Because the limit of that ratio is strictly less than 1 as $T \rightarrow \infty$, maximum welfare will not attain even if there are infinite periods of trade.

A prediction of Corollary 4 is that consumers will expect higher quality from a product they anticipate a manufacturer selling for a long time than from a product they anticipate will be sold for a short time.

Proposition 25 might strike one as odd insofar as there is no deception on the equilibrium path—the buyer always “knows” the quality he will receive and the seller provides him that quality—yet somehow welfare is not maximized. The reason it is not maximized is that the seller cannot *commit* not to cheat the buyer. In particular, the seller’s discounted profit, Π , is the sum of discounted total surplus plus a payment that is independent of any choice she makes (q^e) less a cost that does depend on what she does (the first $c(q)$ term in the expression for Π). If she chose $q = q^*$, then she would maximize the sum of the discounted surplus. Since that is a maximum, lowering q slightly from that level would represent a second-order loss. But she would enjoy a first-order gain by lowering her initial period cost. It follows, therefore, that she cannot be expected to choose the welfare-maximizing level of quality.

In some of the original literature in this area (*e.g.*, Shapiro, 1982, 1983), the buyer was less sophisticated than modeled above. As an example, suppose that the buyer’s estimate of quality after consumption at time $t - 1$, given a prior estimate of q_{t-1}^e and actual quality q , is

$$q_t^e = \lambda q + (1 - \lambda)q_{t-1}^e, \quad (5.7) \quad \{\text{eq:BuyerLearnRecursion}\}$$

where $\lambda \in [0, 1]$. A $\lambda = 0$ represents no learning and a $\lambda = 1$ represents immediate learning. Assuming the buyer has bought in all periods $0, \dots, t - 1$, solving the recursive expression (5.7) reveals his estimate of quality at the beginning of period t is

$$q_t^e = (1 - (1 - \lambda)^t)q + (1 - \lambda)^t q_0^e, \quad (5.8)$$

where q_0^e is his estimate prior to any exchange.

For the sake of brevity, limit attention to $c(\cdot)$ strictly convex and everywhere differentiable.

Assume the seller knows q_0^e . Because even selling minimal quality is weakly profitable, the seller will wish to sell in every period and her price in each period will be q_t^e . The seller's choice of q solves:

$$\begin{aligned} \max_q \sum_{t=0}^{T-1} \delta^t (q_t^e - c(q)) &\equiv \max_q q_0^e \frac{1 - \delta^T (1 - \lambda)^T}{1 - \delta + \delta \lambda} \\ &+ q \frac{\delta \lambda + \delta^T (1 - \delta)(1 - \lambda)^T - \delta^T (1 - \delta(1 - \lambda))}{(1 - \delta)(1 - \delta(1 - \lambda))} \\ &- c(q) \frac{1 - \delta^T}{(1 - \delta)(1 - \delta(1 - \lambda))}. \end{aligned}$$

That program is equivalent to

$$\max_q q \underbrace{\frac{\delta \lambda + \delta^T (1 - \delta)(1 - \lambda)^T - \delta^T (1 - \delta(1 - \lambda))}{1 - \delta^T}}_{R(\delta, \lambda, T)} - c(q). \quad (5.9) \quad \{\text{eq:BuyerLearnsMax}\}$$

It is readily shown that $\partial R(\delta, \lambda, T)/\partial T > 0$ and $\lim_{T \rightarrow \infty} R(\delta, \lambda, T) = \delta \lambda < 1$. Consequently, (5.9) implies the seller will choose a quality level less than the welfare-maximizing quantity.

Let $\hat{q}(\lambda)$ denote the solution to (5.9). Observe that if $\lambda = 1$ (learning is immediate), then

$$R(\delta, 1, T) = \frac{\delta - \delta^T}{1 - \delta^T}.$$

The program (5.9) thus reduces to the program (5.6) if $\lambda = 1$; hence, $\hat{q}(1) = q^{\text{SB}}$. Because $R(\delta, \lambda, T) < \delta \lambda$, it follows that as $\lambda \downarrow 0$ (the buyer ceases to learn), $R(\delta, \lambda, T) \rightarrow 0$, which implies $\hat{q}(0) = \underline{q}$: if the buyer never learns, then the seller will cheat him by providing minimal quality. Finally, observe

$$\frac{\partial R(\delta, \lambda, T)}{\partial \lambda} \propto \frac{1 - \delta^T}{1 - \delta} - T \delta^{T-1} (1 - \lambda)^{T-1} \geq \frac{1 - \delta^T}{1 - \delta} - T \delta^{T-1}.$$

The rightmost term is positive for $T = 1$. For $T \geq 2$, the rightmost term is the difference between the slope of the chord between (δ, δ^T) and $(1, 1^T)$ and the derivative of δ^T at δ . Because the power function is a convex function for powers greater than 1, the slope of the chord must exceed the derivative; that is, the rightmost term is positive. It follows, therefore, that $\partial R/\partial \lambda > 0$, which in turn entails that the quicker the buyer is at learning true quality, the greater will be the quality the seller provides (assuming quality is not a corner solution, $\underline{q}$). In other words, $\hat{q}(\lambda)$ is nondecreasing in λ and strictly increasing at any point at which $\hat{q}(\lambda) > \underline{q}$. To summarize:

Proposition 26. *Consider the model of gradual buyer learning just articulated. If learning is immediate (i.e., $\lambda = 1$), then the equilibrium is identical to the one of Proposition 25 with respect to the seller's choice of quality (i.e., she will choose q^{SB}). If the buyer never learns, the seller will provide minimal quality (i.e., $\underline{q}$). Finally, the faster the buyer's rate of learning (i.e., λ), the weakly greater will be the seller's choice of quality.*

{prop:Shapiro-like}

From expression (5.9), it follows that the seller's choice of quality is independent of the buyer's initial expectation, q_0^e . In particular, even if the buyer correctly anticipates the seller's quality initially in equilibrium, $\hat{q}(\lambda) < q^{\text{SB}}$ if $\lambda < 1$. In other words, the possibility that the buyer will not immediately detect that the seller has provided quality other than what he expected puts a downward pressure on quality. In essence, the ability to fool the buyer (at least in a limited fashion for a limited time), further erodes the seller's ability to commit to high quality.

Observe for any λ (including $\lambda = 1$), the seller would do better if she could commit to the welfare-maximizing level of quality, q^* . To the extent that warranties or similar measures provide such commitment, the seller would have an incentive to offer them.

5.2 SEMI-COOPERATIVE SELLER ACTIONS

{sect:Semi-Coop_Seller}

The selfish-cooperative dichotomy set forth above can, in some contexts, be too stark. Consider a scenario where the seller is a homebuilder. Again, she chooses the quality of the house—perhaps through the quality of the materials she uses. Assume, critically, however, that she (i) incurs the costs *prior* to sale and (ii) has, as an alternative to sale, living in the house herself. That is, the seller's action is cooperative if sale occurs, but selfish if it does not.⁷⁹ Sale is, however, always welfare superior to no sale.

To study such a scenario, let the payoffs of seller and buyer be, respectively,

$$U_S = x(t - q) + (1 - x)b_S(q) - q \text{ and } U_B = x(b_B(q) - t),$$

where $x \in \{0, 1\}$ indicates the amount of trade, $t \in \mathbb{R}$ is payment in the event of sale, and $b_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ is party i 's benefit (possibly expected) from possession of a good of quality q . Observe, as a normalization, the cost of supplying quality q is, now, just q .

Some assumptions on the benefit functions:

- The functions $b_i(\cdot)$ are twice continuously differentiable, strictly increasing, and strictly concave functions (i.e., there is a positive, but diminishing, marginal benefit to increased quality).
- For all $q > 0$, $b_B(q) > b_S(q)$ (i.e., trade is strictly welfare superior, at least if the seller has chosen positive quality).

⁷⁹The analysis in this section draws heavily from Hermalin (forthcoming).

- Zero quality is not privately optimal for the seller if she is certain to retain possession: $b'_S(0) > 1$.
- Infinite quality is not optimal: there exists a $\bar{q} < \infty$ such that $b'_B(q) < 1$ if $q > \bar{q}$.

In light of these assumptions, a unique welfare-maximizing quality, q^* , exists (*i.e.*, q^* maximizes $b_B(q) - q$). If trade were impossible (*i.e.*, autarky held), then the seller would choose quality to maximize

$$b_S(q) - q. \quad (5.10)$$

The earlier given assumptions ensure that program has a unique interior solution: call it $\hat{q}$. To eliminate a case of minor interest, assume⁸⁰

$$b_B(0) < b_S(\hat{q}) - \hat{q}; \quad (5.11) \quad \{\text{eq:Autarky_beats_ZeroInvestment}\}$$

that is, maximum welfare under autarky exceeds welfare given trade but zero quality.

As is common in settings such as these (see, *e.g.*, Gul, 2001, for a discussion), no pure-strategy equilibrium exists:

`{prop:NoPS-HermalinHoldup}`

Proposition 27. *No pure-strategy equilibrium exists.*

Proof: Suppose, to the contrary, the seller played a pure-strategy of q . Suppose $q > 0$. In equilibrium, the buyer must correctly anticipate the seller's action. Hence, he is willing to pay up to $b_B(q)$ for the good. As the seller has the bargaining power, that is the price she will set. However, knowing she can receive a price of $b_B(q)$ regardless of the quality she actually chooses, the seller would do better to deviate to 0 quality ($b_B(q) - q < b_B(q)$ if $q > 0$). Suppose $q = 0$. The seller can then obtain only $b_B(0)$ for the good. But given (5.11), the seller would do better to invest $\hat{q}$ and keep the good for herself. The result follows *reductio ad absurdum*. ■

Because welfare maximization requires the seller to choose q^* with certainty and trade to occur with certainty, Proposition 27 implies that the first-best outcome is unattainable in equilibrium.

What about the second best? As demonstrated by Proposition 27, there is a tradeoff between trading efficiently and providing the seller investment incentives. The second-best welfare-maximization program can be written as

$$\max_{x,q,t} x b_B(q) + (1-x) b_S(q) - q \quad (5.12) \quad \{\text{eq:SB-HermalinHoldup}\}$$

subject to

$$q \in \operatorname{argmax}_q x t + (1-x) b_S(q) - q, \quad (5.13) \quad \{\text{eq:SellerIC-HermalinHoldup}\}$$

$$b_B(q) \geq t, \text{ and} \quad (5.14) \quad \{\text{eq:BuyerIR-HermalinHoldup}\}$$

$$x t + (1-x) b_S(q) - q \geq b_S(\hat{q}) - \hat{q}. \quad (5.15) \quad \{\text{eq:SellerIR-HermalinHoldup}\}$$

⁸⁰See Hermalin (forthcoming) for an analysis with this case included.

Constraint (5.13) is the requirement that the choice of q be incentive compatible for the seller. Constraints (5.14) and (5.15) are, respectively, the buyer and seller's participation (IR) constraints.

The previously given assumptions imply that the program in (5.13) is globally concave in q with an interior solution. Hence, that constraint can be replaced with the corresponding first-order condition:

$$(1 - x)b'_S(q) - 1 = 0.$$

This, in turn, defines the probability of trade as

$$x = 1 - \frac{1}{b'_S(q)}. \quad (5.16) \quad \{\text{eq:x-HermalinHoldup}\}$$

Note that $x = 0$ if $q = \hat{q}$. Because marginal benefit is decreasing in quality, it follows that $q \in [0, \hat{q}]$. Substituting that back into (5.12) makes the program:

$$\max_{q \in [0, \hat{q}]} \left(1 - \frac{1}{b'_S(q)}\right) b_B(q) + \frac{1}{b'_S(q)} b_S(q) - q. \quad (5.17) \quad \{\text{eq:Exp6-HH}\}$$

Because the domain is compact and the function to be maximized continuous, the program must have at least one solution. Let $\mathcal{M}$ equal the maximized value of (5.17). Let $\mathcal{Q}_{\mathcal{M}}$ denote the set of maximizers of (5.17). A second-best level of quality is, therefore, $q^{\text{SB}} \in \mathcal{Q}_{\mathcal{M}}$.

This analysis has ignored the participation constraints, expressions (5.14) and (5.15). There is no loss in having done so: given q^{SB} is played, these will hold for a range of transfers, including $t = b_B(q^{\text{SB}})$.

The next proposition finds that the second-best solution is supportable as an equilibrium:

\{\text{prop:HH-SellerOffer}\}

Proposition 28. *A perfect Bayesian equilibrium exists in which the second best is achieved.⁸¹ Specifically, it is an equilibrium for the seller to choose a second-best quality (an q^{SB}) with certainty and offer the good to the buyer at price $b_B(q^{\text{SB}})$. The buyer plays the mixed strategy in which he accepts the seller's offer with probability x ,*

$$x = 1 - \frac{1}{b'_S(q^{\text{SB}})}.$$

The buyer believes a price less than $b_B(q^{\text{SB}})$ means the seller has chosen quality 0; a price of $b_B(q^{\text{SB}})$ means the seller has chosen quality q^{SB} ; and a price greater than $b_B(q^{\text{SB}})$ means the seller has chosen a quality no greater than q^{SB} .

Because of the buyer's playing of a mixed strategy, the seller may retain ownership and that possibility gives her an incentive to provide quality. Hence,

⁸¹As Hermalin (forthcoming) notes, this equilibrium is not unique. The outcome (level of quality and probability of trade) of this PBE can, however, be part of an essentially unique equilibrium if the contract space is expanded to allow the seller to make TIOLI offers of mechanisms. See Hermalin for details.

it is possible to have an equilibrium in which quality is provided. The problem—and the reason the first best cannot be attained—is that the final allocation may prove to be inefficient: the good may remain in the seller's hands. This reflects the fundamental tradeoff in this situation: the seller's incentives to provide quality are greatest when she is certain to retain ownership, but her retention of ownership is inefficient; conversely, if trade is certain, she has no incentive to provide quality. The second-best solution balances these competing tensions.

5.3 COOPERATIVE BUYER ACTIONS

Suppose that prior to trade, the buyer can make an investment, $I \in \mathcal{I} \subseteq \mathbb{R}_+$, that affects the seller's cost. Assume $0 \in \mathcal{I}$ (i.e., the buyer can choose not to invest). One interpretation is that by investing I , the buyer facilitates delivery by the seller. Another interpretation is that the investment lowers the seller's cost of customizing her product for the buyer. Assume that the parties may trade $x \in [0, \bar{x}]$, $\bar{x} \leq \infty$ units.⁸² Let the payoffs be

$$U_S = t - c(x, I) \text{ and } U_B = b(x) - t - I,$$

respectively, for seller and buyer, where $t \in \mathbb{R}$ is a transfer between them (possibly contingent on x), $b : \mathbb{R}_+ \rightarrow \mathbb{R}$ is an increasing function, and $c : \mathbb{R}_+ \times \mathcal{I} \rightarrow \mathbb{R}_+$ is increasing in its first argument and decreasing in its second. Assume, for all $I > 0$, a positive and finite quantity solves

$$\max_{x \in [0, \bar{x}]} b(x) - c(x, I).$$

Denote the solution to that equation by $x^*(I)$.

The welfare-maximizing level of investment solves

$$\max_{I \in \mathcal{I}} b(x^*(I)) - c(x^*(I), I) - I. \quad (5.18)$$

Assume the program has a unique, finite, and positive solution. Denote it as I^* .

Suppose the seller believes the buyer has invested I^e . Her best response is to offer the buyer $x^*(I^e)$ units for a total payment of $b(x^*(I^e))$. Anticipating this, the buyer understands his choice of investment as the program

$$\max_{I \in \mathcal{I}} b(x^*(I^e)) - \underbrace{b(x^*(I^e))}_t - I \equiv \max_{I \in \mathcal{I}} -I. \quad (5.19)$$

The obvious solution is $I = 0$. Because the seller appropriates the benefit of his investment, the buyer has no incentive to invest. In other words, a *holdup*

⁸²The analysis here also encompasses the case in which there is, at most, one unit to trade. In this case $\bar{x} = 1$ and x denotes the probability of trade. The benefit function would be $b(x) = xv$, v a constant, and the cost function $xC(I)$.

problem arises.⁸³

Could a better outcome be obtained by allowing the buyer to play mixed strategies? The answer is no: suppose the buyer played a mixed strategy over his investments. The seller can, at no cost, induce the buyer to reveal his investments by offering the mechanism $\langle x^*(I), t(I) = b(x^*(I)) \rangle$. Since, regardless of what he announces, the buyer always gets surplus zero, he has no incentive to lie and can be presumed to tell the truth in equilibrium.⁸⁴ But, as just seen, if he gains no surplus, he has no incentive to invest.

To conclude:

{prop:NoInvestBuyerCoop}

Proposition 29. *If the buyer's actions (e.g., investments) are purely cooperative, then there is no equilibrium in which he takes an action other than the one that is least costly to him (e.g., he makes no investment) when the seller has all the bargaining power (makes TIOLI offers).*

5.3.1 BUYER INVESTMENT IN AN INFINITELY REPEATED GAME

As has become evident, infinitely repeated play often yields a better outcome than the one-shot game. Unlike earlier analyses in this chapter, here both buyer and seller must be induced to cooperate: the buyer must invest and the seller cannot gouge him.⁸⁵

Let δ be the common discount factor. Define

$$U_S^0 = b(x^*(0)) - c(x^*(0), 0).$$

The quantity U_S^0 is the seller's payoff should the parties revert to repetition of the equilibrium of the one-shot game. Given he has no bargaining power, the buyer's payoff under such reversion is 0.

Suppose the timing within each period is the buyer chooses his investment for that period, which the seller cannot observe. The seller makes an offer $\langle x, t \rangle$. The buyer accepts or rejects it. If he accepts, then there is trade. Because of this trade, the seller will know her costs, from which she learns what the buyer's investment was. Let the parties wish to support an outcome in which the buyer invests $\hat{I} > 0$ each period, the seller offers $\langle x^*(\hat{I}), \hat{t} \rangle$ each period, and the buyer accepts that offer. If either player deviates from that, then each

⁸³A holdup problem arises when the party making an investment cannot fully capture its benefit because some portion of that benefit is being captured by another party. Williamson (1976) is often credited as introducing the holdup problem into the literature, although in essence the holdup problem is another version of the familiar under-provision of positive externalities.

⁸⁴Although the fact that he obtains zero surplus might suggest consideration of allowing the buyer to mix over accepting the seller's offer, that cannot represent an equilibrium: the seller can trivially induce truth telling and acceptance by offering $\langle x^*(I), b(x^*(I)) - \varepsilon \rangle$, $\varepsilon > 0$ but arbitrarily small.

⁸⁵In its logic, the analysis here is similar to the analysis of followers paying a leader tribute in Hermalin (2007), among other literatures (including the literature on relational contracting; see MacLeod, 2007, for a survey).

expects future play to revert to infinite repetition of the one-shot equilibrium and plays accordingly.⁸⁶

Given the threat of reversion, buyer rationality dictates he accepts an offer $\langle x, t \rangle$ —even if not $\langle x^*(\hat{I}), \hat{t} \rangle$ —provided

$$b(x) - t \geq 0.$$

As noted, the buyer must be given incentives to invest and the seller incentives not to overcharge. These conditions are, respectively, equivalent to⁸⁷

$$b(x^*(\hat{I})) - \hat{t} - \hat{I} \geq (1 - \delta)(b(x^*(\hat{I})) - \hat{t}) \quad \text{and} \quad (5.20) \quad \{\text{eq:BuyerInvestRep}\}$$

$$\hat{t} - c(x^*(\hat{I}), \hat{I}) \geq (1 - \delta)(b(x^*(\hat{I})) - c(x^*(\hat{I}), \hat{I})) + \delta U_S^0. \quad (5.21) \quad \{\text{eq:SellerNoGougeRep}\}$$

Taking the limit as $\delta \rightarrow 1$, (5.20) and (5.21) imply

$$b(x^*(\hat{I})) - c(x^*(\hat{I}), \hat{I}) - \hat{I} \geq U_S^0. \quad (5.22) \quad \{\text{eq:BuyerInvestCoopBigDelta}\}$$

By definition (5.22) holds—indeed is a strict inequality—if $\hat{I} = I^*$. By continuity, it follows that if the parties are sufficiently patient (δ is large enough), then an equilibrium of the repeated game exists in which the buyer invests. Observe, critically, that such an equilibrium requires that the seller not capture all the surplus—she must leave the buyer with some.⁸⁸

5.4 SEMI-COOPERATIVE BUYER ACTIONS

As noted in Section 5.2, the selfish-cooperative dichotomy is sometimes too stark. Suppose a buyer can produce the relevant good himself or buy it from the seller. In either scenario, assume an investment by the buyer reduces production costs (*e.g.*, the buyer needs a service performed—which he can do himself or have provided by the seller—and the cost of the service is reduced for either by the buyer’s preparatory investment). As in Section 5.2, sale is welfare superior to no sale (*i.e.*, the seller is the more efficient provider).

To study this situation, let the payoffs of seller and buyer be, respectively,

$$U_S = x(t - c_S(I)) \quad \text{and} \quad U_B = b - xt - (1 - x)c_B(I) - I,$$

⁸⁶To be complete, certain subtle issues can arise when parties deviate in terms of offers, but not actions. See, for example, Halac (2012).

⁸⁷Observe

$$\sum_{\tau=0}^{\infty} \delta^{\tau} w \geq y + \sum_{\tau=1}^{\infty} \delta^{\tau} z \iff \frac{1}{1-\delta} w \geq y + \frac{\delta}{1-\delta} z.$$

The claimed equivalence follows immediately.

⁸⁸Expression (5.20) is equivalent to

$$\hat{I} \leq \delta(b(x^*(\hat{I})) - \hat{t});$$

the cost of investing today must not exceed the present discounted value of consumer surplus tomorrow.

where $x \in [0, 1]$ denotes the probability of trade, $t \in \mathbb{R}$ is payment in the event of trade, b is the inherent benefit the buyer obtains from the good or service, and $c_i : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is party i 's cost of providing the good or service as a function of the buyer's preparatory investment, I . As the parameter b has no bearing on the analysis to follow, there is no loss in setting it to 0.⁸⁹

Some assumptions on the cost functions:

- The functions $c_i(\cdot)$ are twice continuously differentiable, strictly decreasing, and strictly convex functions (*i.e.*, there is a positive, but diminishing, marginal *benefit* to greater preparatory investment).
- For all $I > 0$, $c_B(I) > c_S(I)$ (*i.e.*, trade is strictly welfare superior to no trade, at least if the buyer has invested).
- Zero investment is not privately optimal for the buyer if no trade is certain: $c'_B(0) < -1$.
- Infinite investment is not optimal: there exists an $\bar{I} < \infty$ such that $c'_S(I) > -1$ if $I > \bar{I}$.

These assumptions imply a unique interior welfare-maximizing investment level, I^* , exists (*i.e.*, I^* maximizes $-c_S(I) - I$). If there were autarky, the buyer would invest to maximize

$$-c_B(I) - I. \quad (5.23)$$

The assumptions just given ensure that program has a unique interior solution: call it $\hat{I}$. To eliminate a case of minor interest, assume⁹⁰

$$-c_S(0) < -c_B(\hat{I}) - \hat{I}; \quad (5.24) \quad \{\text{eq:BuyerInvestsAutarky}\}$$

that is, maximum welfare under autarky exceeds welfare given trade but zero investment.

For the same reasons as in Section 5.2, no pure-strategy equilibrium exists:

Proposition 30. *No pure-strategy equilibrium exists. Moreover, there is no equilibrium in which the buyer invests a given amount as a pure strategy.*

$\{\text{prop:NoPS-positiveI}\}$

Because welfare maximization entails the buyer's investment of I^* and certain trade, this proposition implies that the first-best outcome is unattainable in equilibrium. It also implies that, in equilibrium, the buyer must mix over different investment levels.

The analysis is facilitated by working with the buyer's production cost, C , if no trade occurs: Define $C = c_B(I)$. Because $c_B(\cdot)$ is strictly monotone, it is invertible. Define $\iota(\cdot)$ as the inverse; that is, $\iota(c_B(I)) \equiv I$ for all I . Earlier given assumptions entail that $\iota(\cdot)$ is strictly decreasing, strictly concave, and

⁸⁹The model set forth in this section yields the same results as would the model of Section 5.2 if the buyer possessed all the bargaining power in that model.

⁹⁰Again, see Hermalin (forthcoming) for an analysis with this case included.

twice continuously differentiable. The analysis proceeds by acting as if the buyer chooses his cost should trade not occur, C .

Recall the seller does not observe the buyer's investment; that is, C is the buyer's private information. It can, thus, be considered to be the buyer's type. Let $\mathcal{C} \equiv [\underline{C}, \bar{C}]$ denote his type space. As will become evident,

$$\mathcal{C} \subseteq [c_B(\hat{I}), c_B(0)];$$

that is, the buyer will never invest more than the autarky level, $\hat{I}$. To economize on notation, let $\hat{C} = c_B(\hat{I})$ and $C^0 = c_B(0)$.

Assuming the buyer has played a mixed strategy in terms of his investment, the seller's problem of what contract to offer is a mechanism-design problem, akin to those considered in Section 2.1. A mechanism is, here, a pair $\langle x(\cdot), \tau(\cdot) \rangle$, where $x : \mathcal{C} \rightarrow [0, 1]$ is a probability of trade and $\tau : \mathcal{C} \rightarrow \mathbb{R}$ is the seller's payment. (By the Revelation Principle—Proposition 2—there is no loss of generality in restricting attention to direct-revelation mechanisms.)

Let

$$U(C) = -(1 - x(C))C - \tau(C) \quad (5.25) \quad \{\text{eq:BuyerUtility_inCost}\}$$

denote the buyer's utility if he truthfully announces his type. Note (i) at the time the mechanism is being played, the buyer's investment is sunk; and (ii) consistent with the approach in Section 2.1, what the buyer pays the seller is contingent on his announcement only (*i.e.*, it does not depend on whether trade actually occurs).

As preliminaries to studying such mechanisms, we have:

$\{\text{lemma:MonoX-HH}\}$

Lemma 17. *Given an incentive-compatible mechanism, $x(\cdot)$ is non-decreasing and $U(\cdot)$ is a convex function.*

We can now characterize incentive-compatible mechanisms:

$\{\text{prop:MechChar-HH}\}$

Proposition 31. *Necessary conditions for a mechanism to be incentive compatible (induce truth-telling by the buyer) are (i) that the probability of trade, $x(\cdot)$, be non-decreasing in the buyer's cost and (ii) that the buyer's utility as a function of his type be given by*

$$U(C) = \underline{U} - \int_{\underline{C}}^C (1 - x(z)) dz, \quad (5.26) \quad \{\text{eq:SolvedDiffEq-HH}\}$$

where $\underline{U}$ is a constant.

Moreover, any mechanism in which $x(\cdot)$ is non-decreasing and expression (5.26) holds is incentive compatible (*i.e.*, conditions (i) and (ii) are also sufficient).

Anticipating the mechanism the seller will offer him, the buyer is willing to invest $\iota(C)$ if and only if it maximizes $U(C) - \iota(C)$. It follows:

$\{\text{prop:HH-EqProbTrade}\}$

Proposition 32. *If $\iota(C) > 0$ is a level of investment chosen by the buyer with positive probability in equilibrium, then the subsequent probability of trade given that investment is $1 + \iota'(C)$.*

A corollary is

Corollary 5. *There is no equilibrium in which the buyer invests more than his autarky level of investment, $\hat{I}$.*

We seek to construct an equilibrium. To that end, suppose the buyer mixes over $\mathcal{C}$ according to the distribution function $F : \mathcal{C} \rightarrow [0, 1]$. To facilitate the analysis take F to be differentiable on $(\underline{C}, \bar{C})$, with derivate (density) $f(C)$. The possibility that there is a “mass” at $\bar{C}$ is permitted: defining $\Sigma(\cdot)$ as the corresponding survival function, it may be that $\Sigma(\bar{C}) > 0$.⁹¹ Take $f(C) > 0$ for all $C \in [\underline{C}, \bar{C})$, $\hat{C} \leq \underline{C} < \bar{C} \leq C^0$. Of course, these properties will need to be verified.

The first-order condition for maximizing $U(C) - \iota(C)$ (expression (A.23) in the appendix) holds for any $C < C^0$ that the buyer chooses with positive probability (*i.e.*, such that $f(C) > 0$). Moreover, because the buyer could invest $\hat{I}$ and refuse to trade, his equilibrium utility cannot be less than $-\hat{C} - \iota(\hat{C})$. Hence,

$$U(C) - \iota(C) = \underline{U} - \iota(\underline{C}) \geq -\hat{C} - \iota(\hat{C})$$

for all $C \in [\underline{C}, \bar{C})$, where use has been made of both (5.26) and Proposition 32. Because the seller makes a TIOLI offer, this constraint is binding.

The seller choose $x(\cdot)$ and $\underline{U}$ to maximize her expected profit:

$$\left(\tau(\bar{C}) - x(\bar{C})c_S(\iota(\bar{C})) \right) \Sigma(\bar{C}) + \int_{\underline{C}}^{\bar{C}} \left(\tau(C) - x(C)c_S(\iota(C)) \right) f(C) dC. \quad (5.27)$$

Using (5.25), (5.26), and defining $\sigma(C) = -f(C)$ (hence, $\sigma(C) = \Sigma'(C)$), this last expression can be rewritten as

$$\begin{aligned} & -\underline{U} - \left(x(\bar{C})c_S(\iota(\bar{C})) + (1 - x(\bar{C}))\bar{C} - \int_{\underline{C}}^{\bar{C}} (1 - x(z)) dz \right) \Sigma(\bar{C}) \\ & + \int_{\underline{C}}^{\bar{C}} \left(x(C)c_S(\iota(C)) + (1 - x(C))C - \int_{\underline{C}}^C (1 - x(z)) dz \right) \sigma(C) dC. \end{aligned} \quad (5.28)$$

Integration by parts permits rewriting that expression as

$$\begin{aligned} & -\underline{U} - \left(x(\bar{C})c_S(\iota(\bar{C})) + (1 - x(\bar{C}))\bar{C} \right) \Sigma(\bar{C}) \\ & + \int_{\underline{C}}^{\bar{C}} \left(x(C)c_S(\iota(C)) + (1 - x(C))C + (1 - x(C)) \frac{\Sigma(C)}{\sigma(C)} \right) \sigma(C) dC. \end{aligned} \quad (5.29) \quad \{\text{eq:BuyerEqUtility-HH}\}$$

⁹¹Here,

$$\Sigma(C) \equiv 1 - \int_{\underline{C}}^C f(z) dz.$$

From Proposition 32, if the buyer plays $C < \bar{C} \leq C^0$ with positive probability, then $x(C)$ must equal $1 + \iota'(C)$. Differentiating, pointwise, the seller's expected profit (*i.e.*, expression (5.29)) with respect to $x(C)$ reveals that consistency with Proposition 32 and seller optimization is met if and only if

$$\frac{\Sigma(C)}{\sigma(C)} = c_S(\iota(C)) - C \quad (5.30) \quad \{\text{eq:Hazard-HH}\}$$

for $C < \bar{C} \leq C^0$, because, then, the seller is indifferent as to her choice of $x(\cdot)$ and might as well choose $x(\cdot)$ to be consistent with the buyer's mixing (*i.e.*, such that $x(\cdot) = 1 + \iota'(\cdot)$).

Using (5.30), expression (5.29) can be rewritten as

$$\begin{aligned} -\underline{U} - \left(x(\bar{C})c_S(\iota(\bar{C})) + (1 - x(\bar{C}))\bar{C} \right) \Sigma(\bar{C}) + \int_{\underline{C}}^{\bar{C}} (C\sigma(C) + \Sigma(C))dC \\ = -\underline{U} - \underline{C} + \left(\bar{C} - c_s(\iota(\bar{C})) \right) x(\bar{C})\Sigma(\bar{C}) \\ = -\underline{C} - \iota(\underline{C}) - \left(-\hat{C} - \iota(\hat{C}) \right) + \left(\bar{C} - c_s(\iota(\bar{C})) \right) x(\bar{C})\Sigma(\bar{C}), \quad (5.31) \quad \{\text{eq:magic-HH}\} \end{aligned}$$

where the last equality follows because the buyer's participation constraint binds. Because $\hat{C}$ maximizes $-C - \iota(C)$, (5.31) cannot exceed

$$\left(\bar{C} - c_s(\iota(\bar{C})) \right) x(\bar{C})\Sigma(\bar{C}).$$

At the same time, the seller could deviate from offering the mechanism by simply offering to sell at price $\bar{C}$, which would net her expected profit

$$\left(\bar{C} - c_s(\iota(\bar{C})) \right) \Sigma(\bar{C}).$$

From (5.31), it follows she will do so unless $\underline{C} = \hat{C}$ and, if $\Sigma(\bar{C}) > 0$, $x(\bar{C}) = 1$. The following can now be established:

Proposition 33. *There exists a subgame-perfect equilibrium in which the seller makes the buyer a TIOLI offer in which the buyer plays a mixed strategy whereby he chooses $C \in [\hat{C}, C^0]$ according to the distribution function*

$\{\text{prop:Unseen-Cost}\}$

$$F(C) = 1 - \exp \left(\int_{\hat{C}}^C \frac{1}{c_S(\iota(z)) - z} dz \right) \quad (5.32) \quad \{\text{eq:mixF-HH}\}$$

and seller offers the mechanism $\langle x(\cdot), \tau(\cdot) \rangle$ such that

$$x(C) = \begin{cases} 1 + \iota'(C), & \text{if } C < C^0 \\ 1, & \text{if } C = C^0 \end{cases}$$

and

$$\tau(C) = \hat{C} + \iota(\hat{C}) - \iota(C) - (1 - x(C))C. \quad (5.33) \quad \{\text{eq:tau-HH}\}$$

Because the buyer is playing a non-degenerate mixed strategy, the equilibrium of Proposition 33 is not even second best insofar as the buyer is, with positive probability, making investments that are welfare inferior to the second-best investment level. If, contrary to the maintained assumption, he had the bargaining power, then there is an equilibrium in which he invests at the second-best level with certainty (the equilibrium would be similar to the one in Proposition 28). In other words, the holdup problem that arises when the non-investing party has the bargaining power further exacerbates a situation already made imperfect by asymmetric information. To reduce the degree to which he is held up, the buyer must mix—otherwise he would be vulnerable to being held up completely—and this undermines his incentive to choose the right level of investment.⁹²

5.5 SELFISH BUYER ACTIONS

We now turn to wholly selfish actions by the buyer. As a somewhat general framework suppose that the timing of the game between buyer and seller is the following:

- Buyer sinks an investment $I \in \mathbb{R}_+$. This affects his benefit, $b \in \mathbb{R}_+$, should he obtain a unit of some good, asset, or service from the seller. The buyer is assumed to want at most one unit.
- The seller observes a signal, s , that may contain information about b .
- The seller makes a TIOLI offer to sell one unit at price p .
- After observing b and s , the buyer decides whether to buy or not.

The payoffs to buyer and seller are, respectively,

$$U_B = x(b - p) - I \text{ and } U_S = xp,$$

where $x \in \{0, 1\}$ is the amount of trade. Note, for convenience, the seller's cost has been normalized to zero.

5.5.1 NO ASYMMETRIC INFORMATION

Suppose that the seller's signal, s , is just b ; that is, there is no asymmetry of information. The seller will obviously set $p = s$ (equivalently, $p = b$) in equilibrium. The buyer will buy in equilibrium. Hence, his equilibrium payoff is $-I$. He maximizes this by choosing $I = 0$. In this case, holdup destroys all investment incentives:

Proposition 34. *If the seller can observe the buyer's benefit, then the buyer will invest nothing in equilibrium.*

⁹²Note, critically, this distortion does not mean he “underinvests”: his average investment can actually be higher when he doesn't have the bargaining power than when he does. See Hermalin (forthcoming) for details.

To the extent positive investment is welfare superior to no investment (*i.e.*, if $0 \notin \operatorname{argmax}_I \mathbb{E}\{b|I\} - I$), this outcome is undesirable.

5.5.2 DETERMINISTIC RETURN

Suppose there is a function $B : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that an investment of I returns a benefit $b = B(I)$. Assume, primarily to make the problem interesting and straightforward, that $B(\cdot)$ is twice differentiable, strictly increasing (benefit increases with investment), and strictly concave (diminishing returns to investment). In addition, assume there is a finite $I^* > 0$ such that $B'(I^*) = 1$. Observe that I^* is the unique welfare-maximizing level of investment.

Unlike the previous subsection assume the seller does not observe b . Nor does she observe I (which, here, would be equivalent to observing b). In fact, let her signal, s , be pure noise (or, equivalently, assume she observes nothing).

For reasons similar to those explored above (see, *e.g.*, Proposition 30), there is no equilibrium in which the buyer invests a positive amount as a pure strategy:

{prop:HK09-1}

Proposition 35. *There is no equilibrium in which the buyer invests a positive amount as a pure strategy.*

Proof: Suppose not. Then, in equilibrium, the seller would set $p = B(I)$, where I is the buyer's pure-strategy investment level. The buyer's equilibrium payoff would, thus, be $-I < 0$. Given the buyer can secure a payoff of 0 by not investing at all, it follows this is not an equilibrium. The result follows *reductio ad absurdum*. ■

As Gul (2001) observed, the equilibrium in a game such as this depends critically on whether $B(0) = 0$ or $B(0) > 0$.⁹³ In the former case, trade is worthless unless the buyer invests; in that latter, it has value even there is no investment.

{lemma:NotGul}

Lemma 18. *Suppose $B(0) = 0$ (trade is worthless absent buyer investment). Then there is an equilibrium in which $I = 0$ and the seller charges some price p , $p > B(I^*)$.*

Note, critically, that what sustains the Lemma 18 equilibrium is the willingness of the seller to charge an exorbitant price. That willingness goes away if $B(0) > 0$ (trade is valuable even absent investment). Knowing that the buyer will always accept $B(0)$ regardless of how much he has invested, the seller can ensure herself a positive profit by offering $p = B(0)$.

{lemma:Gul01-1}

Lemma 19. *Suppose trade is valuable even absent investment (*i.e.*, $B(0) > 0$). Then, in equilibrium, neither the buyer nor the seller plays a pure strategy.*

To analyze the mixed strategies played in equilibrium, it is easier to treat the buyer as choosing a benefit, b , and, thus, making investment $B^{-1}(b)$. For

⁹³Technically, whether $B(0)$ equals or exceeds the seller's cost of production.

convenience, denote $B^{-1}(\cdot)$ by $\iota(\cdot)$. The seller's strategy is a distribution function, G , over prices and the buyer's strategy is a distribution function, F , over benefits. Let p_ℓ and p_h be the lowest and highest prices the seller might play (technically, $p_\ell = \sup \{p | G(p) = 0\}$ and $p_h = \inf \{p | G(p) = 1\}$). Let b_ℓ and b_h be the lowest and highest benefits the buyer might play (they have analogous technical definitions). Some initial observations:

- By construction, $b_\ell \geq B(0)$ (no negative investment).
- From Lemma 19, $p_h > p_\ell$ and $b_h > b_\ell$. Let $\mathcal{B}^+$ denote the set of $b > B(0)$ played by the buyer with positive probability. Note that set is nonempty by Lemma 19.
- The seller can guarantee herself a profit of b_ℓ by playing $p = b_\ell$. Hence, $p_\ell \geq b_\ell$.
- The buyer's utility, u_B , from playing b_ℓ is $-\iota(b_\ell)$:

$$u_B = \max\{0, b_\ell - p\} - \iota(b_\ell) = -\iota(b_\ell),$$

where the second equality follows from the previous bullet point.

- Hence, $b_\ell = B(0)$.

Note that the last bullet point means the buyer's equilibrium expected payoff is zero.

{lemma:pell}

Lemma 20. *In equilibrium, $p_\ell = B(0)$.*

An immediate corollary is that the seller's expected profit in equilibrium is $B(0)$.

{lemma:Gu101-11}

Lemma 21. *In equilibrium, $b_h = B(I^*)$ if $p_h \leq B(I^*)$.*

Consider the following strategies for the seller and buyer, respectively:

$$G(p) = G(B(0)) + \int_{B(0)}^p g(z)dz \text{ and } F(b) = 1 - \int_b^{b_h} f(z)dz. \quad (5.34)$$

Because the buyer's expected utility given any b he plays is zero, we have

$$bG(B(0)) + \int_{B(0)}^b (b-p)g(p)dp - \iota(b) \equiv 0. \quad (5.35)$$

Since that is an identity, differentiating implies

$$G(b) - \iota'(b) = 0.$$

Hence, the strategy for the seller is

$$G(p) = \iota'(p). \quad (5.36)$$

Because $\iota(\cdot)$ is convex and $\iota'(B(I^*)) = 1$, this strategy requires $p_h = B(I^*)$. Similarly, the seller's expected profit given any p she plays is $B(0)$; hence,

$$p(1 - F(p)) \equiv B(0).$$

It follows that

$$F(p) = 1 - \frac{B(0)}{p};$$

in other words, the strategy for the buyer is

$$F(b) = \begin{cases} 1 - \frac{B(0)}{b}, & \text{if } b < b_h \\ 1, & \text{if } b = b_h \end{cases}. \quad (5.37) \quad \{\text{eq:Gul01-eq2}\}$$

Expression (5.37) does not directly pin down b_h . But given $p_h = B(I^*)$, it follows from Lemma 21 that $b_h = B(I^*)$. To summarize:⁹⁴

Proposition 36. *An equilibrium is the seller mixes over price in the interval $[B(0), B(I^*)]$ according to the distribution*

$$G(p) = \iota'(B(0)) + \int_{B(0)}^p \iota''(z) dz$$

and the buyer mixes over benefit in the interval $[B(0), B(I^)]$ according to the distribution*

$$F(b) = \begin{cases} 1 - \frac{B(0)}{b}, & \text{if } b < B(I^*) \\ 1, & \text{if } b = B(I^*) \end{cases}. \quad (5.38) \quad \{\text{eq:Gul01-eqF}\}$$

Proof: As the analysis in the text shows, the parties are indifferent over all actions in $[B(0), B(I^*)]$. As established in the proof of Lemma 21 any $b > B(I^*)$ is dominated for the buyer if $p_h = B(I^*)$. That proof also established that any $p > b_h$ is dominated for the seller. ■

Observe that as $B(0) \downarrow 0$, the buyer's strategy, expression (5.38), converges to his *not* investing with probability 1. This suggests a link between this proposition and Lemma 18. Indeed, when $B(0) = 0$, there is no equilibrium in which the buyer invests with positive probability:⁹⁵

Proposition 37. *If $B(0) = 0$, then the buyer's expected level of investment is zero in equilibrium.*

{prop:xx1}

5.5.3 STOCHASTIC RETURN

Suppose now that the buyer's benefit b is stochastic, with a distribution that depends on his investment, I . Denote the survival function for this conditional

⁹⁴This proposition is essentially Proposition 1 of Gul (2001).

⁹⁵This is essentially Proposition 4 of Hermalin and Katz (2009).

distribution by $D(\cdot|I)$; that is, the probability of the buyer's benefit equaling or exceeding b is $D(b|I)$. The use of the letter "D" is not accidental. Let the set of possible b be $[0, \bar{b}]$, where $0 < \bar{b} \leq \infty$.

Assume D is twice differentiable in each argument. Using integration by parts, it follows, for a given $\underline{b}$, that

$$\mathbb{E}\{\max\{b - \underline{b}, 0\}\} = \int_{\underline{b}}^{\bar{b}} (b - \underline{b}) \left(-\frac{\partial D(b|I)}{\partial b} \right) db = \int_{\underline{b}}^{\bar{b}} D(b|I) db.$$

Further assume: the buyer gets no benefit with certainty absent investment (*i.e.*, $D(0|0) = 0$) and there is a finite positive level of investment that maximizes welfare (*i.e.*, there is an $I^* \in (0, \infty)$ that maximizes $\int_0^{\bar{b}} D(b|I) db - I$).

Two possibilities will be considered here about what the seller knows prior to trade: (i) the seller observes the buyer's investment, but not his benefit; or (ii) the seller observes neither the buyer's investment nor his his benefit. In terms of earlier introduced notation, case (i) corresponds to $s = I$ and case (ii) corresponds to s being pure noise.⁹⁶

In case (ii), a degenerate equilibrium exists along the lines of the Lemma 18 equilibrium in which the buyer does not invest because he anticipates the seller will offer an exorbitant price and because she believes the buyer has not invested, the seller may as well charge such an exorbitant price given $D(0|0) = 0$.

Our interest here, though, is non-degenerate equilibria. To see that there are functions that support non-degenerate equilibria, suppose $\bar{b} = \infty$ and

$$D(b|I) = \exp\left(-\frac{\alpha b}{\sqrt{I}}\right), \quad (5.39) \quad \{\text{eq:ExampleD}\}$$

where α is constant. In case (i), the seller chooses p to maximize

$$pD(p|I) = p \exp\left(-\frac{\alpha p}{\sqrt{I}}\right),$$

the solution to which is $p(I) = \sqrt{I}/\alpha$. Anticipating this response, the buyer chooses I to maximize

$$\int_{p(I)}^{\infty} D(b|I) db - I = \frac{\sqrt{I}}{\alpha e} - I,$$

where $e = \exp(1)$, the base of the natural logarithm. The solution to the maximization problem is

$$\hat{I} = \frac{1}{4\alpha^2 e^2}. \quad (5.40) \quad \{\text{eq:HK09-I1}\}$$

⁹⁶Hermalin and Katz (2009) consider the case in which s is an imperfect, but informative signal of b ; that is, there is some joint distribution of s and b given I . See Hermalin and Katz for details.

In case (ii), if the seller anticipates the buyer has chosen $\tilde{I}$, then she will play $p(\tilde{I})$. To have an equilibrium, $\tilde{I}$ must in fact be a best response for the buyer to $p(\tilde{I})$. It will be if

$$\tilde{I} \in \operatorname{argmax}_I \int_{p(\tilde{I})}^{\infty} D(b|I)db - I;$$

equivalently if $\tilde{I}$ is a solution to the first-order condition

$$\begin{aligned} 0 &= -1 + \frac{p(\tilde{I}) \exp(-p(\tilde{I})\alpha/\sqrt{\tilde{I}})}{2I} + \frac{\exp(-p(\tilde{I})\alpha/\sqrt{\tilde{I}})}{2\alpha\sqrt{\tilde{I}}} \\ &= -1 + \frac{\sqrt{\tilde{I}} \exp\left(-\sqrt{\tilde{I}/I}\right)}{2I\alpha} + \frac{\exp\left(-\sqrt{\tilde{I}/I}\right)}{2\alpha\sqrt{\tilde{I}}}. \end{aligned}$$

The solution is

$$\tilde{I} = \frac{1}{\alpha^2 e^2}. \quad (5.41) \quad \{\text{eq:HK09-I2}\}$$

Comparing the two equilibria for this example—expressions (5.40) and (5.41)—we see that, when the seller can observe the buyer's investment (case (i)), the buyer invests less in equilibrium than when the seller cannot observe it (case (ii)). At first blush this might seem a general phenomenon: being able to observe the buyer's investment allows the seller to better holdup the buyer, correspondingly reducing the buyer's incentives to invest. This logic, although tempting, proves to be incomplete. As Hermalin and Katz (2009) show, the result depends on the properties of the demand function, D : in particular, the buyer invests less when his investment is observable if the buyer's demand becomes less price elastic the greater is his investment.⁹⁷

The welfare consequences of observable investment are also less than clear cut. Although Hermalin and Katz provide conditions under which welfare is greater with unobservable investment than with observable investment, that cannot be seen as a general result. What is a general result is that the buyer actually does better when his investment is observable than when it is not:

Proposition 38. *The buyer's equilibrium expected utility is at least as great when the seller can observe his investment than when she cannot.*

Proof: Define

$$U(p, I) = \int_p^{\infty} D(b|I)db.$$

Suppose the buyer's investment is observable. Then the buyer can “pick” the price he faces in the sense that he knows an investment of I will result in a price of $p(I)$. By revealed preference:

$$U(p(\hat{I}), \hat{I}) - \hat{I} \geq U(p(\tilde{I}), \tilde{I}) - \tilde{I}, \quad (5.42) \quad \{\text{eq:HK09-RP}\}$$

⁹⁷For this example, elasticity is $p\alpha/\sqrt{I}$, which is clearly decreasing in I .

where $\hat{I}$ and $\tilde{I}$ are as defined above. But the RHS of (5.42) is the buyer's expected utility in equilibrium when investment is unobservable. ■

Intuitively, when investment is observable, the buyer becomes the Stackelberg leader of the game. This result illustrates a very general point: consider, as in case (ii), a game in which the second mover does not observe the first mover's action. If that game has an equilibrium in which the first mover plays a pure strategy, then the first mover would necessarily be better off in equilibrium in the variant of that game in which the second mover can observe the first mover's action.⁹⁸

6 FINAL THOUGHTS

As the introduction sought to make clear, the economics of asymmetric information is a vast topic. A complete survey, even if that were a feasible task, would necessarily yield a massive tome. This chapter has, of necessity, had to consider a narrow subset of literature: attention has been limited to buyer-seller relations and, then, largely to those involving a single buyer and a single seller, in which the asymmetry of information between them arises prior to their establishing a trading relationship.⁹⁹

Much of the literature surveyed in this chapter, especially in Sections 2–4, although novel when I began my career, has become part of the established literature. More recent literature has, as hinted at in Section 5, involved asymmetries that arise endogenously.¹⁰⁰ This topic remains an area of active research.

With respect to ongoing and future research, a key issue is the modeling of bargaining. This chapter has focused almost exclusively on take-it-or-leave-it (TIOLI) bargaining. Assuming that bargaining vastly simplifies matters. In some contexts, however, it is unrealistic. Hence, despite the complications of non-TIOLI bargaining, allowing for such bargaining is an important area of exploration.

Related to the issue of bargaining is the “money on the table” problem (recall, *e.g.*, Section 3.3 *supra*): many mechanisms require the parties to commit to honor *ex post* inefficient outcomes (possibly only off the equilibrium path). But if it is common knowledge that an outcome is inefficient and there is no impediment to renegotiation, then it is unreasonable to expect the parties will not take action to remedy the situation. Anticipation of such action would, in many cases, undermine the original mechanism, creating some doubt as to whether such mechanisms are good predictors of actual behavior. This, thus,

⁹⁸This is true even if in the unobservable-action version the second mover plays a mixed strategy in equilibrium: think of U in expression (5.42) as expected payoff given the mixed strategy $p(\cdot)$.

⁹⁹As observed previously, however, many of the techniques considered are more broadly applicable.

¹⁰⁰In addition to some of the articles cited above in this regard, the reader is also directed to González (2004) and Lau (2008).

represents an important area of research.¹⁰¹

Beyond the suggestions for future work contained in this section and also made earlier, I hope the reader has been motivated to ask his or her own questions about alternative assumptions or possible extensions. At the very least, I hope this chapter has provided the reader a good foundation from which to his or her own research.

APPENDIX: PROOFS

Proof of Lemma 2: The IR constraint for β' and the Spence-Mirrlees condition imply

$$u(x(\beta'), \beta) - t(\beta') > u(x_0, \beta) - 0 = u_R;$$

hence, β must do better to buy than not buy. The Spence-Mirrlees condition also implies strictly increasing differences (Proposition 3):

$$u(x(\beta'), \beta) - u(x(\beta'), \beta') > u(x_0, \beta) - u(x_0, \beta') = 0.$$

That $v(\beta) > v(\beta')$ then follows from (2.10). ■

Proof of Lemma 3: Suppose not. Let $\langle x(\cdot), t(\cdot) \rangle$ be the seller's expected-payoff-maximizing mechanism. Consider a new mechanism $\langle x(\cdot), \tilde{t}(\cdot) \rangle$, where

$$\tilde{t}(\beta) = t(\beta) + v(\tilde{\beta}) - u_R$$

for all $\beta \geq \tilde{\beta}$ and $\tilde{t}(\beta) = t(\beta) = 0$ for $\beta < \tilde{\beta}$. Using (IC), it is readily verified that that if type β participates, he does best to purchase the same $x(\beta)$ as he would have under the original mechanism. By design, $\tilde{\beta}$ is still willing to purchase. From Lemma 2,

$$u(x(\beta), \beta) - t(\beta) > v(\tilde{\beta}) \implies u(x(\beta), \beta) - \tilde{t}(\beta) > u_R.$$

So all types $\beta > \tilde{\beta}$ who participated under the original mechanism participate under the new mechanism. (Clearly those types that “participated” by purchasing nothing continue to participate in the same way.) Given the buyer's true type, the seller incurs, in equilibrium, the same cost as she would have under the old (*i.e.*, $c(x(\beta), \beta)$ remains unchanged), but her transfer is strictly greater under the new; this contradicts the supposition that the original mechanism maximized her expected payoff. The result follows *reductio ad absurdum*. ■

Proof of Proposition 4 (completion): Consider $n > m$. A chain of reason-

¹⁰¹Beaudry and Poitevin, in a series of articles in the 1990s (see, *e.g.*, Beaudry and Poitevin, 1993, 1995), made some progress on the topic, but much remains unsettled.

ing similar to that in the main text provides:

$$\begin{aligned}
& u(x(\beta_m), \beta_n) - t(\beta_m) \\
&= u(x(\beta_m), \beta_n) - u(x(\beta_m), \beta_m) + u_R + \sum_{j \leq n} R_j(x(\beta_{j-1})) - \sum_{j=m+1}^n R_j(x(\beta_{j-1})) \\
&= u(x(\beta_m), \beta_n) - u(x(\beta_m), \beta_m) + v(\beta_n) - \sum_{j=m+1}^n R_j(x(\beta_{j-1})) \\
&= v(\beta_n) - \sum_{j=m+1}^n \left(R_j(x(\beta_{j-1})) - R_j(x(\beta_m)) \right) \leq v(\beta_n)
\end{aligned}$$

(note, in this case,

$$u(x(\beta_m), \beta_n) - u(x(\beta_m), \beta_m) = \sum_{j=m+1}^n R_j(x(\beta_m))$$

and $x(\beta_{j-1}) \geq x(\beta_m)$ for $j \geq m+1$). ■

Proof of Proposition 5: Lemma 1 established the necessity of $x(\cdot)$ being nondecreasing.

Using the fundamental theorem of calculus, (2.10) can be rewritten as

$$\int_{\beta'}^{\beta} \frac{\partial u(x(\beta), z)}{\partial \beta} dz \geq v(\beta) - v(\beta') \geq \int_{\beta'}^{\beta} \frac{\partial u(x(\beta'), z)}{\partial \beta} dz. \quad (\text{A.1}) \quad \{\text{eq:DoubleIneq}\}$$

The function $v(\cdot)$ is absolutely continuous.¹⁰² An absolutely continuous function has a derivative almost everywhere (see, *e.g.*, Royden, 1968, p. 109). Dividing

¹⁰²**Proof:** The function $\partial u(x(\bar{\beta}), \cdot) / \partial \beta$ is continuous and bounded, it thus has a finite maximum, M , and a maximizer β_M . By Spence-Mirrlees,

$$M = \frac{\partial u(x(\bar{\beta}), \beta_M)}{\partial \beta} \geq \frac{\partial u(x(\bar{\beta}), \beta')}{\partial \beta} > \frac{\partial u(x(\beta), \beta')}{\partial \beta},$$

for any β and $\beta' \in B$. Pick any $\varepsilon > 0$ and let $\delta = \varepsilon / M$. Consider any finite non-overlapping collection of intervals $\{(\beta'_i, \beta_i)\}_{i=1}^I$ such that

$$\sum_{i=1}^I (\beta_i - \beta'_i) < \delta.$$

The conclusion follows if

$$\sum_{i=1}^I |v(\beta_i) - v(\beta'_i)| < \varepsilon$$

(see, *e.g.*, Royden, 1968, p. 108). Because $v(\cdot)$ is nondecreasing (Lemma 2), the absolute value can be ignored (*i.e.*, $|v(\beta_i) - v(\beta'_i)|$ can be replaced with just $v(\beta_i) - v(\beta'_i)$). From (A.1) and the intermediate value theorem, there is some $\hat{\beta}_i \in (\beta'_i, \beta_i)$ such that

$$M(\beta_i - \beta'_i) \geq \frac{\partial u(x(\beta_i), \hat{\beta}_i)}{\partial \beta} (\beta_i - \beta'_i) = \int_{\beta'_i}^{\beta_i} \frac{\partial u(x(\beta_i), z)}{\partial \beta} dz \geq v(\beta_i) - v(\beta'_i).$$

(A.1) through by $\beta - \beta'$ (recall this is positive) and taking the limit as $\beta' \rightarrow \beta$, it follows the derivative of $v(\cdot)$ is given by

$$\frac{dv(\beta)}{d\beta} = \frac{\partial u(x(\beta), \beta)}{\partial \beta}$$

almost everywhere. An absolutely continuous function is the integral of its derivative (see, *e.g.*, Yeh, 2006, p. 283), hence

$$v(\beta) = v(\underline{\beta}) + \int_{\underline{\beta}}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz. \quad (\text{A.2}) \quad \{\text{eq:GenMech-v}\}$$

Because $v(\beta) = u(x(\beta), \beta) - t(\beta)$, (A.2) implies

$$t(\beta) = u(x(\beta), \beta) - v(\underline{\beta}) - \int_{\underline{\beta}}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz. \quad (\text{A.3}) \quad \{\text{eq:GenMech-t}\}$$

Setting $\tau = v(\underline{\beta})$, the necessity of (2.18) follows.

The logic for establishing sufficiency is similar to that in the proof of Proposition 4; hence, I will be brief. The proof of participation is the same. To establish incentive compatibility, one needs to show that a β -type buyer won't pretend to be a β' -type buyer. His utility were he to do so is

$$\begin{aligned} & u(x(\beta'), \beta) - t(\beta') \\ &= u(x(\beta'), \beta) - u(x(\beta'), \beta') + \tau + \int_{\underline{\beta}}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz - \int_{\beta'}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz \\ &= u(x(\beta'), \beta) - u(x(\beta'), \beta') + v(\beta) - \int_{\beta'}^{\beta} \frac{\partial u(x(z), z)}{\partial \beta} dz \\ &= v(\beta) - \int_{\beta'}^{\beta} \left(\frac{\partial u(x(z), z)}{\partial \beta} - \frac{\partial u(x(\beta'), z)}{\partial \beta} \right) dz \leq v(\beta), \quad (\text{A.4}) \quad \{\text{eq:ContCase-suff1}\} \end{aligned}$$

where the first two equalities follow from (2.18) and the last from the fundamental theorem of calculus. That the integral in the last line is positive follows from Spence-Mirrlees because $x(\cdot)$ is nondecreasing: the integrand is positive if $\beta > \beta'$ and integration is the positive direction; it is negative if $\beta < \beta'$ and integration is in the negative direction. $\blacksquare$

Summing over i yields

$$\varepsilon = M\delta > M \sum_{i=1}^I (\beta_i - \beta'_i) \geq \sum_{i=1}^I (v(\beta_i) - v(\beta'_i)).$$

Proof of Lemma 5: The result is immediate if $Y_H \geq Y'_H$, so suppose $Y_H < Y'_H$. By assumption,

$$\beta' u(Y_I) + (1 - \beta') u(Y_H) \geq \beta' u(Y'_I) + (1 - \beta') u(Y'_H).$$

Hence,

$$\beta' \left((u(Y_I) - u(Y'_I)) + (u(Y'_H) - u(Y_H)) \right) \geq u(Y'_H) - u(Y_H).$$

Because $\beta > \beta'$, it follows that

$$\beta \left((u(Y_I) - u(Y'_I)) + (u(Y'_H) - u(Y_H)) \right) > u(Y'_H) - u(Y_H);$$

and, thus, that $\beta u(Y_I) + (1 - \beta) u(Y_H) > \beta u(Y'_I) + (1 - \beta) u(Y'_H)$. ■

Proof of Proposition 12: Suppose such a mechanism existed. Because it is efficient $x(b, c) = 1$ if $b \geq c$ and $= 0$ otherwise. Hence, from (3.18),

$$\begin{aligned} u(b) &= u(0) + \int_0^b \underbrace{\int_{\underline{c}}^{\bar{c}} x(z, c) g(c) dc}_{=\xi_B(z)} dz = u(0) + \int_0^b \int_{\underline{c}}^z g(c) dc dz \\ &= u(0) + \int_{\underline{c}}^b G(z) dz, \quad (\text{A.5}) \quad \{\text{eq:MS-EffU}\} \end{aligned}$$

where the second equality follows because $x(b, c) = 1$ for $c \leq b$ and $= 0$ for $c > b$. A similar analysis reveals

$$\pi(c) = \pi(\bar{c}) + \int_c^{\bar{b}} (1 - F(z)) dz. \quad (\text{A.6}) \quad \{\text{eq:MS-Effpi}\}$$

Realized surplus is $b - c$ if $b \geq c$ and 0 otherwise. Hence, *ex ante* expected surplus is

$$\begin{aligned} S^* &= \int_0^{\bar{b}} \left(\int_{\underline{c}}^b (b - c) g(c) dc \right) f(b) db \\ &= \int_0^{\bar{b}} \underbrace{\left((b - b) G(b) - (b - 0) G(0) + \int_{\underline{c}}^b G(c) dc \right)}_{\text{integration by parts}} f(b) db \\ &= \int_0^{\bar{b}} \int_{\underline{c}}^b G(c) dc f(b) db \quad (\text{A.7}) \quad \{\text{eq:MS-ex_ante_surp}\} \end{aligned}$$

If the mechanism is balanced, then $S^* = \mathbb{E}_b \{u(b)\} + \mathbb{E}_c \{\pi(c)\}$. But, from (A.5),

$$\mathbb{E}_b \{u(b)\} = \int_0^{\bar{b}} \left(u(0) + \int_{\underline{c}}^b G(z) dz \right) f(b) db = u(0) + S^*, \quad (\text{A.8})$$

where the last equality follows from (A.7). The mechanism can, thus, be balanced only if $u(0) = 0$ and $\mathbb{E}_c\{\pi(c)\} = 0$; but, from (A.6), $\mathbb{E}_c\{\pi(c)\} > 0$. *Reductio ad absurdum*, no balanced, efficient, and interim IR mechanism can exist. ■

Proof of Lemma 10: Consider $b > b'$. Then, for all c , $x(b, c) \geq x(b', c)$. Hence,

$$0 \leq \int_{\underline{c}}^{\bar{c}} (x(b, c) - x(b', c))g(c)dc = \xi_B(b) - \xi_B(b').$$

So $\xi_B(\cdot)$ is non-decreasing as desired. The proof for $\xi_S(\cdot)$ is similar. ■

Proof of Lemma 11 and Corollary 3: Consider the functions:

$$V_B(b, \sigma) = b - \sigma \frac{1 - F(b)}{f(b)} \quad \text{and} \quad V_C(c, \sigma) = c + \sigma \frac{G(c)}{g(c)}. \quad (\text{A.9}) \quad \{\text{eq:MS-c-funcs}\}$$

The premise of the lemma is that $V_B(\cdot, 1)$ and $V_C(\cdot, 1)$ are non-decreasing functions. Observe, given previous discussion, that one needs to show

$$V_B\left(\cdot, \frac{\lambda}{1 + \lambda}\right) \quad \text{and} \quad V_C\left(\cdot, \frac{\lambda}{1 + \lambda}\right)$$

are non-decreasing functions. Note $\lambda/(1 + \lambda) \in (0, 1)$. We have

$$\frac{\partial^2 V_B(b, \sigma)}{\partial \sigma \partial b} = -\frac{d}{db} \left(\frac{1 - F(b)}{f(b)} \right) \quad \text{and} \quad \frac{\partial^2 V_C(c, \sigma)}{\partial \sigma \partial c} = \frac{d}{dc} \left(\frac{G(c)}{g(c)} \right). \quad (\text{A.10}) \quad \{\text{eq:MS-c-funcs-dif}\}$$

If the expressions in (A.10) are non-negative, then

$$0 < 1 = \frac{\partial V_B(b, 0)}{\partial b} \leq \frac{\partial V_B(b, \sigma)}{\partial b}$$

(and similarly for V_C) for all $\sigma \geq 0$. So the result would follow. If the expressions in (A.10) are negative, then

$$0 \leq \frac{\partial V_B(b, 1)}{\partial b} \leq \frac{\partial V_B(b, \sigma)}{\partial b}$$

(and similarly for V_C) for all $\sigma \leq 1$ (where the first inequality follows by assumption). This proves the lemma.

The corollary follows because, given the assumptions of the corollary, $V_B(\cdot, 1)$ and $V_C(\cdot, 1)$ are non-decreasing functions. ■

Proof of Proposition 14: Suppose that $p_1(c) \neq p_1(c')$ for $c \neq c'$ (this is verified below). Assume the buyer's belief upon seeing $p_1(c)$ is the seller's type is c with probability 1. If $p_1 \neq p_1(c)$ for any c , assume the buyer believes that $c = \underline{c}$. Suppose the buyer is playing a linear strategy. The objective is to show that a linear strategy is the seller's best response regardless of c . Consider an

arbitrary period t . Assume all normalized types greater than ω_t have purchased by t . The expected present value of the seller's profit discounted back to t is

$$(1-c)\Pi_t = (\omega_t - \alpha m_t)m_t + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t}(\alpha m_{\tau-1} - \alpha m_{\tau})m_{\tau},$$

where the scaling $(1-c)$ reflects that ω is distributed uniformly on $[0, 1-c]$. Maximizing Π_t with respect to $m_t, m_{t+1}, \dots$ yields the first-order conditions:

$$\begin{aligned} \omega_t - 2\alpha m_t + \delta \alpha m_{t+1} &= 0 \\ &\vdots \\ \delta^{\tau}(\alpha m_{t+\tau-1} - 2\alpha m_{t+\tau} + \delta \alpha m_{t+\tau+1}) &= 0 \\ &\vdots \end{aligned}$$

Solving this system of difference equations yields:

$$m_{t+\tau} = \frac{\omega}{\alpha} \left(\frac{1 - \sqrt{1-\delta}}{\delta} \right)^{\tau+1} = \left(\frac{1 - \sqrt{1-\delta}}{\alpha \delta} \right) \times \underbrace{\alpha m_{t+\tau-1}}_{=\omega_{t+\tau-1}} \quad (\text{A.11})$$

if $\tau \geq 1$ and

$$m_t = \left(\frac{1 - \sqrt{1-\delta}}{\alpha \delta} \right) \omega_t. \quad (\text{A.12})$$

So a linear rule is, indeed, a best response for the seller regardless of her cost. The linear rule has

$$\gamma = \frac{1 - \sqrt{1-\delta}}{\alpha \delta}.$$

Next, verify that the buyer wishes to follow a linear rule in response to the seller's linear rule: let b be the indifferent type at time t , then we have

$$b - p_t = \delta(b - p_{t+1}) \iff \omega - m_t = \delta(\omega - \gamma \omega),$$

where the equivalence follows by adding and subtracting c on the LHS and recognizing that $p_{t+1} = m_{t+1} + c = \gamma \omega + c$. Solving:

$$\omega = \frac{1}{1 - \delta + \delta \gamma} m_t.$$

So we also have a linear rule for the buyer with

$$\alpha = \frac{1}{1 - \delta + \delta \gamma}.$$

We can solve our linear rules:

$$\gamma = \frac{-(1-\delta) + \sqrt{1-\delta}}{\delta} \quad \text{and} \quad \alpha = \frac{1}{\sqrt{1-\delta}}.$$

Note, critically, that neither γ nor α depend on c . Finally, observe that

$$p_1(c) = m_1(c) + c = \gamma(1 - c) + c = \gamma + (1 - \gamma)c,$$

so $p_1(c) \neq p_1(c')$ if $c \neq c'$. ■

Proof of Lemma 14: First, suppose, contrary to the lemma, that there at least two distinct pairs the $\underline{\gamma}$ type plays that do not include s_0 : (s, p) and (s', p') . The order is arbitrary, hence one is free to assume $s_0 < s < s'$. Given the $\underline{\gamma}$ type plays the two pairs with positive probability (mixes):

$$xp - c(s, \underline{\gamma}) = x'p' - c(s', \underline{\gamma}),$$

where x and x' are the equilibrium probabilities the buyer accepts the seller's offer. Spence-Mirrlees (*i.e.*, expression (4.14)) implies:

$$xp - c(s, \bar{\gamma}) < x'p' - c(s', \bar{\gamma}).$$

Consequently, the $\bar{\gamma}$ type never offers (s, p) in equilibrium. The buyer must, therefore, believe the seller is low quality upon seeing (s, p) . Because $s > s_0$, seller rationality dictates that $x > 0$; so buyer rationality dictates $p \leq \underline{\gamma}$. But the $\underline{\gamma}$ type can guarantee herself a payoff $\underline{\gamma}$ by playing $(s_0, \underline{\gamma})$. The chain

$$\underline{\gamma} \geq x\underline{\gamma} - c(s_0, \underline{\gamma}) \geq xp - c(s_0, \underline{\gamma}) > xp - c(s, \underline{\gamma})$$

indicates that (s, p) is strictly dominated for the low-quality type, which contradicts that she plays it with positive probability. The first part of the lemma follows *reductio ad absurdum*.

We just saw that an action (s, p) , $s > s_0$, not played by the high-quality type is strictly dominated for the low-quality type, which proves the lemma's second part. ■

Proof of Lemma 15: Given the previous lemma, there is at most one such signal, $s > s_0$, that the low-quality seller could send. Let π be the expected *payment* sellers who play that signal get. Because $s > s_0$, seller rationality dictates that $\pi > 0$, which entails buyer acceptance with strictly positive probability. It must be that $\pi < \bar{\gamma}$. To see this, note the buyer never accepts a price greater than $\bar{\gamma}$; hence, $\pi \not\geq \bar{\gamma}$. If $\pi = \bar{\gamma}$, then the buyer always accepts when faced with $(s, \bar{\gamma})$. Hence, the only price that would be rational for the seller to offer given signal s is $p = \bar{\gamma}$. But, by supposition, the $\underline{\gamma}$ plays the signal s with positive probability; hence, the buyer's Bayesian beliefs dictate that $\mathbb{E}\{\gamma|s\} < \bar{\gamma}$ and so the buyer should not accept $p = \bar{\gamma}$. *Reductio ad absurdum*, $\pi < \bar{\gamma}$ follows.

Define $\hat{s}$ by

$$\bar{\gamma} - c(\hat{s}, \underline{\gamma}) = \pi - c(s, \underline{\gamma});$$

that is, a low-quality seller would be indifferent between playing signal $\hat{s}$ in exchange for receiving $\bar{\gamma}$ for sure and playing s in exchange for receiving expected

payment π . That such a $\hat{s}$ exists follows because $c(\cdot, \underline{\gamma})$ is unbounded and continuous. Because $\pi < \bar{\gamma}$, it follows that $\hat{s} > s$. Spence-Mirrlees (4.14) entails

$$\bar{\gamma} - c(\hat{s}, \bar{\gamma}) > \pi - c(s, \bar{\gamma}).$$

These last two expression imply that there exists $\delta > 0$ such that

$$\bar{\gamma} - \delta - c(\hat{s}, \underline{\gamma}) < \pi - c(s, \underline{\gamma}) \quad \text{and} \quad (\text{A.13}) \quad \{\text{eq:LowTypeNoMotive}\}$$

$$\bar{\gamma} - \delta - c(\hat{s}, \bar{\gamma}) > \pi - c(s, \bar{\gamma}). \quad (\text{A.14}) \quad \{\text{eq:HighTypeMotive}\}$$

Consider the deviation $(\hat{s}, \bar{\gamma} - \delta)$:¹⁰³ expressions (A.13) and (A.14) imply that $\Gamma^0 = \{\underline{\gamma}\}$. But as

$$\Gamma((\hat{s}, \bar{\gamma} - \delta)) \setminus \Gamma^0 = \{\bar{\gamma}\},$$

$$\text{BR}(\Gamma((\hat{s}, \bar{\gamma} - \delta)) \setminus \Gamma^0, (\hat{s}, \bar{\gamma} - \delta)) = \{1\},$$

which, given (A.14), means (4.12) holds. As claimed, the equilibrium fails the Intuitive Criterion. $\blacksquare$

Proof of Proposition 17: It needs to be shown (a) this is an equilibrium; (b) it satisfies the Intuitive criterion; and (c), in light of the earlier lemmas, that no other *separating* PBE satisfies the Intuitive criterion.

With respect to (a): the buyer's beliefs and strategy are obviously consistent with a PBE. Given that, a seller's expected payment, π , is at most $\underline{\gamma}$ if $s < s^*$ and at most $\bar{\gamma}$ if $s \geq s^*$. A low-quality seller will not wish to deviate:

$$\underline{\gamma} > \underline{\gamma} - c(s, \underline{\gamma}) \quad (\text{A.15}) \quad \{\text{eq:RileyLH-IC}\}$$

for all $s > s_0$ and, by (4.23),

$$\underline{\gamma} = \bar{\gamma} - c(s^*, \underline{\gamma}) \geq \bar{\gamma} - c(s, \underline{\gamma})$$

for all $s \geq s^*$. Nor will a high-quality seller:

$$\underline{\gamma} - c(s, \bar{\gamma}) \leq \underline{\gamma} < \bar{\gamma} - c(s^*, \bar{\gamma}) \quad (\text{A.16}) \quad \{\text{eq:RileyHL-IC}\}$$

for all $s < s^*$, where the second inequality follows from (4.14) (*i.e.*, Spence-Mirrlees) and (4.23), and

$$\bar{\gamma} - c(s, \bar{\gamma}) \leq \bar{\gamma} - c(s^*, \bar{\gamma}) \quad (\text{A.17})$$

for all $s \geq s^*$.

With respect to (b), $\Gamma^0 = \Gamma$ if $p \notin (\underline{\gamma}, \bar{\gamma}]$ or $s \geq s^*$, in which case the Intuitive Criterion holds trivially. Consider $s \in (s_0, s^*)$ and $p \in (\underline{\gamma}, \bar{\gamma}]$. If $\underline{\gamma} > p - c(s, \underline{\gamma})$, then, given $\underline{\gamma} = \bar{\gamma} - c(s^*, \underline{\gamma})$, Spence-Mirrlees implies

$$\bar{\gamma} - c(s^*, \bar{\gamma}) > p - c(s, \bar{\gamma});$$

¹⁰³This must be a deviation—out-of-equilibrium play—because otherwise (A.14) implies the high-quality seller would never play s .

in other words, $\underline{\gamma} \in \Gamma^0$ implies $\bar{\gamma} \in \Gamma^0$. Hence, $\Gamma(a) \setminus \Gamma^0$ can equal $\emptyset$, $\{\underline{\gamma}\}$, or Γ . If the first, the Intuitive Criterion holds trivially. If the latter two,

$$\min_{x \in \text{BR}(\Gamma(a) \setminus \Gamma^0, a)} V(a, x, \gamma) \leq \underline{\gamma} - c(s, \gamma) < \underline{\gamma};$$

hence, (4.12) could not hold given (A.15) and (A.16); that is, the Intuitive Criterion is satisfied.

Turning to (c): in a separating PBE, the $\underline{\gamma}$ type plays $(s_0, \underline{\gamma})$ and the buyer accepts all $p \leq \underline{\gamma}$. Let $(\bar{s}, \bar{p})$ be an action of the $\bar{\gamma}$ type that she plays with positive probability in a separating PBE.

Claim. *In a separating PBE that satisfies the Intuitive Criterion, $\bar{p} = \bar{\gamma}$ and the buyer must accept the high-quality seller's price offer with probability one.*

Proof of Claim: Suppose not. The parties' rationality rules out an equilibrium in which $\bar{p} > \bar{\gamma}$. Let π be the expected payment the seller receives if she plays $(\bar{s}, \bar{p})$. By supposition, $\pi < \bar{\gamma}$. There exists a $\delta > 0$ such that

$$\bar{\gamma} - c(\bar{s} + \delta, \underline{\gamma}) = \pi - c(\bar{s}, \underline{\gamma}). \quad (\text{A.18}) \quad \{\text{eq:NoMixIC-claima}\}$$

By Spence-Mirrlees,

$$\bar{\gamma} - c(\bar{s} + \delta, \bar{\gamma}) > \pi - c(\bar{s}, \bar{\gamma}).$$

Hence, there exists an $\varepsilon > 0$ such that

$$\bar{\gamma} - \varepsilon - c(\bar{s} + \delta, \bar{\gamma}) > \pi - c(\bar{s}, \bar{\gamma}). \quad (\text{A.19}) \quad \{\text{eq:NoMixIC-claimb}\}$$

It follows from (A.18) and the definition of equilibrium that

$$\underline{\gamma} \geq \pi - c(\bar{s}, \underline{\gamma}) > \bar{\gamma} - \varepsilon - c(\bar{s} + \delta, \underline{\gamma}). \quad (\text{A.20}) \quad \{\text{eq:NoMixIC-claimc}\}$$

Expressions (A.19) and (A.20) reveal that the deviation $(\bar{s} + \delta, \bar{\gamma} - \varepsilon)$ is such that the PBE fails the Intuitive Criterion. $\square$

In light of the claim, there is no PBE satisfying the Intuitive Criterion such that $\bar{s} < s^*$. Consider $\bar{s} > s^*$. Define

$$\varepsilon = \frac{1}{2}(c(\bar{s}, \bar{\gamma}) - c(s^*, \bar{\gamma})).$$

Note $\varepsilon > 0$. In light of the claim, definition of ε , and condition (4.23):

$$\underline{\gamma} > \bar{\gamma} - \varepsilon - c(s^*, \underline{\gamma}) \quad \text{and} \quad \bar{\gamma} - c(\bar{s}, \bar{\gamma}) < \bar{\gamma} - \varepsilon - c(s^*, \bar{\gamma}),$$

where the terms on the left of the inequalities are the relevant payoffs under the PBE. Considering the deviation $(s^*, \bar{\gamma} - \varepsilon)$, it is readily seen that these last two inequalities mean the PBE fails the Intuitive Criterion. $\blacksquare$

Proof of Proposition 28: Given his beliefs, the buyer is indifferent between accepting and rejecting an offer at price $b_B(q^{\text{SB}})$. He is, thus, willing to mix.

Given the probabilities with which the buyer mixes, it is a best response for the seller to choose quality q^{SB} if she intends to offer the good at price $b_B(q^{\text{SB}})$. All that remains is to verify that the seller does not wish to deviate with respect to investment *and* price given the buyer's beliefs. Given that $b_B(\cdot)$ is increasing, the buyer, given his beliefs, will reject any price greater than $b_B(q^{\text{SB}})$ or in the interval $(b_B(0), b_B(q^{\text{SB}}))$. Because the seller's expected payoff is $\mathcal{M}$ on the purported equilibrium path, she cannot do strictly better under autarky; that is, she has no incentive to induce the buyer to reject her offer. Finally, setting a price of $b_B(0)$ cannot be a profitable deviation for the seller given

$$b_B(0) - q < \mathcal{M}$$

for all q . ■

Proof of Proposition 30: Given the seller makes a TIOLI offer, if the buyer invests a given amount, I , as a pure-strategy, the seller's best response is to offer trade at price

$$t = c_B(I) \geq c_S(I).$$

But Proposition 29 rules out a pure-strategy equilibrium in which the buyer invests a positive amount and trade is certain to occur. There is thus no equilibrium in which the buyer plays $I > 0$ as a pure strategy. If the buyer doesn't invest, then trade would occur. But from (5.24), the buyer would do better to deviate, invest $\hat{I}$, and not trade. ■

Proof of Lemma 17: A revealed-preference argument yields:

$$U(C) \geq -(1 - x(C'))C - \tau(C') = U(C') - (1 - x(C'))(C - C') \quad \text{and} \quad (\text{A.21}) \quad \{\text{eq:RP-RnotRPrime}\}$$

$$U(C') \geq -(1 - x(C))C' - \tau(C) = U(C) - (1 - x(C))(C' - C).$$

Hence,

$$(1 - x(C))(C' - C) \geq U(C) - U(C') \geq (1 - x(C'))(C' - C). \quad (\text{A.22}) \quad \{\text{eq:HH-pinch}\}$$

Without loss take $C' > C$, then $x(C') \geq x(C)$ if (A.22) holds.

Pick C and $C' \in \mathcal{C}$ and a $\lambda \in (0, 1)$. Define $C_\lambda = \lambda C + (1 - \lambda)C'$. Revealed preference implies:

$$\lambda U(C) \geq \lambda U(C_\lambda) - \lambda(1 - x(C_\lambda))(C - C_\lambda) \quad \text{and}$$

$$(1 - \lambda)U(C') \geq (1 - \lambda)U(C_\lambda) - (1 - \lambda)(1 - x(C_\lambda))(C' - C_\lambda).$$

Add those two expressions:

$$\lambda U(C) + (1 - \lambda)U(C') \geq U(C_\lambda) - (1 - x(C_\lambda)) \underbrace{(\lambda C + (1 - \lambda)C' - C_\lambda)}_{=0}.$$

The result follows. ■

Proof of Proposition 31: The necessity of condition (i) follows immediately from Lemma 17. Convex functions are absolutely continuous (see, *e.g.*, van Tiel, 1984, p. 5). As noted earlier (see proof of Proposition 5), an absolutely continuous function has a derivative almost everywhere and, moreover, is the integral of its derivative. Hence, dividing (A.22) through by $C' - C$ and taking the limit as $C' \rightarrow C$ reveals that $U'(C) = -(1 - x(C))$ almost everywhere. Expression (5.26) follows.

To establish sufficiency, suppose the buyer's type is C and consider any $C' > C$. We wish to verify (A.21):

$$U(C) - U(C') = \int_C^{C'} (1 - x(z)) dz \geq \int_C^{C'} (1 - x(C')) dz = (1 - x(C'))(C' - C),$$

where the inequality follows because $x(\cdot)$ is non-decreasing. Expression (A.21) follows. The case $C' < C$ is provided similarly and, so, omitted for the sake of brevity. ■

Proof of Proposition 32: By supposition, the buyer chooses C with positive probability in equilibrium, hence $C \in \mathcal{C}$. Consequently, C must satisfy the first-order condition

$$0 = U'(C) - \iota'(C) = -(1 - x(C)) - \iota'(C). \quad (\text{A.23}) \quad \{\text{eq:HH-buyerC-FOC}\}$$

The result follows. ■

Proof of Proposition 33: Expression (5.32) solves the differential equation (5.30). By assumption, $c_s(\iota(C)) < C$ for all $C < C^0$. So, if $\bar{C} < C^0$, then the integral in (5.32) exceeds $-\infty$, implying $\Sigma(\bar{C}) > 0$. That, in turn, would entail $x(\bar{C}) = 1$, but that is inconsistent with Proposition 32 when $\bar{C} < C^0$. *Reductio ad absurdum*, $\bar{C} = C^0$. Expression (5.33) follows from Proposition 31 because $1 - x(C) = -\iota'(C)$. The remaining steps were established in the text that preceded the statement of the proposition. ■

Proof of Lemma 18: If the buyer has not invested, he will reject all offers at positive prices. Hence, the seller cannot expect to earn more than zero from any price offer. Given this, p is a best response for her. Expecting a price of p , the buyer would accept the offer only if he has invested $I > I^*$. In that case, his payoff would be

$$B(I) - I - p < B(I) - I - B(I^*) + I^* < 0,$$

where the second inequality follows because $I^* = \arg\max_z B(z) - z$. The buyer would not deviate to an $I > I^*$. Because he won't buy if $I \leq I^*$, he prefers not investing to investing an $I \in (0, I^*]$. So not investing and rejecting is a best response for the buyer. ■

Proof of Lemma 19: From Proposition 35, there is no equilibrium in which the buyer invests a positive amount as a pure strategy. Suppose, then, he

invests 0 as a pure strategy in equilibrium. The seller's unique best response is $p = B(0)$. But if the seller would play $p = B(0)$ as a pure strategy, the buyer's best response is I^* : because $B(I^*) - I^* > B(0) - 0$, $B(I^*) > B(0)$ (the buyer will buy if he invested) and his overall utility is positive (*i.e.*, $B(I^*) - I^* - B(0) > 0$). *Reductio ad absurdum*, there is no equilibrium in which the buyer plays a pure strategy.

Suppose the seller played a pure strategy of p . Because the seller can secure herself $B(0)$, the buyer must accept an offer of p with positive probability; that is, he must play an I such that $B(I) \geq p$ and $B(I) - I - p \geq 0$. Note that if the second condition holds, the first automatically holds; hence, the buyer must be playing investment levels that maximize

$$B(I) - I - p.$$

But there is a unique maximizer, I^* ; that is, the buyer can only play I^* as a best response to p . But this contradicts the first half of the proof, which established he doesn't play a pure strategy in equilibrium. *Reductio ad absurdum*, there is no equilibrium in which the seller plays a pure strategy. ■

Proof of Lemma 20: Suppose not. Observe then that buyer rationality rules out his playing any $b \in (B(0), p_\ell)$. Let $\underline{b} = \inf \mathcal{B}^+$. Observe $\underline{b}$ must strictly exceed p_ℓ :

$$0 \leq \mathbb{E}_G\{\max\{0, \underline{b} - p\}\} - \iota(\underline{b}) \leq \underline{b} - p_\ell - \iota(\underline{b}) < \underline{b} - p_\ell,$$

where the last inequality follows because $\iota(\underline{b}) > 0$. The seller's expected profit from playing p_ℓ is

$$(1 - F(\underline{b}))p_\ell < (1 - F(\underline{b}))\underline{b}.$$

But this means the seller would do better to deviate to $\underline{b}$, a contradiction. ■

Proof of Lemma 21: There is no equilibrium in which $p_h > b_h$, since otherwise the seller's profit is zero from p_h whereas she gets a profit of $B(0)$ by playing p_ℓ . So the buyer's payoff if he plays b_h is

$$u_B = b_h - \iota(b_h) - \mathbb{E}\{p\}. \quad (\text{A.24}) \quad \{\text{eq:Gu101-eq1}\}$$

If $b_h < B(I^*)$, the buyer would do better to deviate to $B(I^*)$. Hence, $b_h \geq B(I^*)$. Suppose $b_h > B(I^*)$. Given $p_h \leq B(I^*)$, then (A.24) remains a valid expression for the buyer's expected payoff with b_h replaced by any $b \in [B(I^*), b_h]$. But since such a b enhances the buyer's expected payoff, it follows the buyer would deviate if $b_h > B(I^*)$. ■

Proof of Proposition 37: Proposition 35 rules out an equilibrium in which the buyer invests a positive amount as a pure strategy.

Suppose, contrary to the proposition, that the buyer invests a positive amount in expectation. There must exist a $\hat{b} > 0$ such that $1 - F(\hat{b}) > 0$. It follows that the seller can guarantee herself an expected profit of

$$\hat{\pi} \equiv (1 - F(\hat{b}))\hat{b} > 0.$$

Hence, $p_\ell \geq \hat{\pi}$. Buyer rationality dictates that

$$b - \iota(b) - p_\ell \geq 0$$

for any $b > 0$ he is willing to play with positive probability. It follows that $\underline{b} = \inf \mathcal{B}^+ > p_\ell$. But then, as shown in the proof of Lemma 20, the seller could increase her profit by charging $\underline{b}$ rather than p_ℓ ; hence, p_ℓ isn't the infimum of prices the seller charges. *Reductio ad absurdum*, the buyer cannot invest a positive amount in expectation. ■

REFERENCES

- Akerlof, George**, “The Market for ‘Lemons’: Qualitative Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics*, 1970, *84* (3), 488–500.
- Armstrong, Mark**, “Competition in Two-Sided Markets,” *RAND Journal of Economics*, Autumn 2006, *37* (3), 668–691.
- , “Recent Developments in the Economics of Price Discrimination,” in Richard Blundell, Whitney K. Newey, and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications*, Vol. II, Cambridge, UK: Cambridge University Press, 2006.
- and **Jean-Charles Rochet**, “Multi-dimensional Screening: A User’s Guide,” *European Economic Review*, 1999, *43*, 959–979.
- Basov, Suren**, *Multidimensional Screening*, Berlin: Springer-Verlag, 2010.
- Beaudry, Paul and Michel Poitevin**, “Signalling and Renegotiation in Contractual Relationships,” *Econometrica*, 1993, *61*, 745–782.
- and —, “Competitive Screening in Financial Markets when Borrowers can Recontract,” *Review of Economic Studies*, 1995, *62*, 401–423.
- Bolton, Patrick and Mathias Dewatripont**, *Contract Theory*, Cambridge, MA: MIT Press, 2005.
- Borch, Karl H.**, *The Economics of Uncertainty*, Princeton, NJ: Princeton University Press, 1968.
- Caillaud, Bernard and Benjamin E. Hermalin**, “The Use of an Agent in a Signalling Model,” *Journal of Economic Theory*, June 1993, *60* (1), 83–113.
- and —, “Hidden Actions and Incentives,” <http://faculty.haas.berkeley.edu/hermalin/#WP>, 2000. Unpublished manuscript.
- Che, Yeon-Koo and Donald B. Hausch**, “Cooperative Investments and the Value of Contracting,” *American Economic Review*, March 1999, *89* (1), 125–147.
- Cho, In-Koo and David Kreps**, “Signaling Games and Stable Equilibria,” *Quarterly Journal of Economics*, 1987, *102*, 179–222.
- Coase, Ronald H.**, “Durability and Monopoly,” *Journal of Law and Economics*, 1972, *15*, 143–149.
- Crémer, Jacques and Michael H. Riordan**, “A Sequential Solution to the Public Goods Problem,” *Econometrica*, January 1985, *53* (1), 77–84.
- and —, “On Governing Multilateral Transactions with Bilateral Contracts,” *RAND Journal of Economics*, Autumn 1987, *18* (3), 436–451.

-
- and **Richard P. McLean**, “Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent,” *Econometrica*, March 1985, *53* (2), 345–361.
- and —, “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions,” *Econometrica*, November 1988, *56* (6), 1247–1257.
- Darby, Michael R. and Edi Karni**, “Free Competition and the Optimal Amount of Fraud,” *Journal of Law and Economics*, April 1973, *16* (1), 67–88.
- Demougin, Dominique M. and Devon A. Garvie**, “Contractual Design with Correlated Information under Limited Liability,” *RAND Journal of Economics*, Winter 1991, *22* (4), 477–489.
- Demski, Joel S. and David E.M. Sappington**, “Resolving Double Moral Hazard Problems with Buyout Agreements,” *RAND Journal of Economics*, 1991, *22* (2), 232–240.
- Edlin, Aaron S.**, “Do Guaranteed-Low-Price Policies Guarantee High Prices, and Can Antitrust Rise to the Challenge?,” *Harvard Law Review*, 1997, *111*, 528–575.
- and **Benjamin E. Hermalin**, “Contract Renegotiation and Options in Agency Problems,” *Journal of Law, Economics, and Organization*, 2000, *16* (2), 395–423.
- Engers, Maxim**, “Signalling with Many Signals,” *Econometrica*, May 1987, *55* (3), 663–674.
- Fama, Eugene F.**, “Agency Problems and the Theory of the Firm,” *Journal of Political Economy*, 1980, *88* (2), 288–307.
- Fong, Yuk-fai**, “When Do Experts Cheat and Whom Do They Target?,” *RAND Journal of Economics*, Spring 2005, *36* (1), 113–130.
- Fudenberg, Drew and Jean Tirole**, “A ‘Signal-Jamming’ Theory of Predation,” *RAND Journal of Economics*, Autumn 1986, *17* (3), 366–376.
- and —, “Moral Hazard and Renegotiation in Agency Contracts,” *Econometrica*, 1990, *58*, 1279–1320.
- and —, *Game Theory*, Cambridge, MA: MIT Press, 1991.
- Gibbard, Allan**, “Manipulation of Voting Schemes,” *Econometrica*, July 1973, *41* (4), 587–601.
- Gibbons, Robert S.**, *Game Theory for Applied Economists*, Princeton, NJ: Princeton University Press, 1992.

- González, Patrick**, “Investment and Screening Under Asymmetric Endogenous Information,” *RAND Journal of Economics*, Autumn 2004, 35 (3), 502–519.
- Green, Jerry and Jean-Jacques Laffont**, “Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods,” *Econometrica*, March 1977, 45 (2), 427–438.
- Grossman, Sanford J.**, “The Informational Role of Warranties and Private Disclosure about Product Quality,” *Journal of Law and Economics*, 1981, 21, 461–83.
- and **Oliver D. Hart**, “An Analysis of the Principal-Agent Problem,” *Econometrica*, 1983, 51, 7–46.
- and —, “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,” *Journal of Political Economy*, August 1986, 94 (4), 691–719.
- Gul, Faruk**, “Unobservable Investment and the Hold-Up Problem,” *Econometrica*, March 2001, 69 (2), 343–376.
- , **Hugo Sonnenschein**, and **Robert B. Wilson**, “Foundations of Dynamic Monopoly and the Coase Conjecture,” *Journal of Economic Theory*, 1986, 39, 155–190.
- Halac, Marina**, “Relational Contracts and the Value of Relationships,” *American Economic Review*, April 2012, 102 (2), 750–779.
- Hellwig, Martin**, “Some Recent Developments in the Theory of Competition in Markets With Adverse Selection,” *European Economic Review*, February–March 1987, 31 (1–2), 319–325.
- Hermalin, Benjamin E.**, “Leading for the Long Term,” *Journal of Economic Behavior and Organization*, 2007, 62, 1–19.
- , “Unobserved Investment, Endogenous Quality, and Trade,” *RAND Journal of Economics*, forthcoming.
- and **Michael L. Katz**, “Judicial Modification of Contracts between Sophisticated Parties: A More Complete View of Incomplete Contracts and Their Breach,” *Journal of Law, Economics, & Organization*, October 1993, 9 (2), 230–255.
- and —, “Sender or Receiver: Who Should Pay to Exchange an Electronic Message?,” *RAND Journal of Economics*, Autumn 2004, 35 (3), 423–448.
- and —, “Information and the Hold-up Problem,” *RAND Journal of Economics*, Autumn 2009, 40 (3), 405–423.

-
- , **Avery W. Katz**, and **Richard Craswell**, “Law & Economics of Contracts,” in A. Mitchell Polinsky and Steven Shavell, eds., *Handbook of Law and Economics*, Vol. 1 2007.
- Holmstrom, Bengt**, “Moral Hazard and Observability,” *Bell Journal of Economics*, Spring 1979, 10 (1), 74–91.
- , “Managerial Incentive Problems—A Dynamic Perspective,” *Review of Economic Studies*, January 1999, 66 (226), 169–182.
- Huang, Chi-fu** and **Robert H. Litzenberger**, *Foundations for Financial Economics*, Amsterdam: North-Holland, 1988.
- Katz, Michael L.**, “Game-Playing Agents: Unobservable Contracts as Precommitments,” *RAND Journal of Economics*, 1991, 22 (3), 307–328.
- , “Observable Contracts as Commitments: Interdependent Contracts and Moral Hazard,” *Journal of Economics & Management Strategy*, 2006, 15 (3), 685–706.
- Klein, Benjamin** and **Keith B. Leffler**, “The Role of Market Forces in Assuring Contractual Performance,” *Journal of Political Economy*, August 1981, 89 (4), 615–641.
- Klemperer, Paul**, *Auctions: Theory and Practice*, Princeton: Princeton University Press, 2004.
- Krishna, Vijay**, *Auction Theory*, San Diego: Academic Press, 2002.
- Laffont, Jean-Jacques** and **David Martimort**, *The Theory of Incentives: The Principal-Agent Model*, Princeton, NJ: Princeton University Press, 2002.
- Larsen, Bradley**, “The Efficiency of Dynamic, Post-Auction Bargaining: Evidence from Wholesale Used-Auto Auctions,” 2012. MIT Working Paper.
- Lau, Stephanie**, “Information and Bargaining in the Hold-up Problem,” *RAND Journal of Economics*, Spring 2008, 39 (1), 266–282.
- Lewis, Tracy R.** and **David E.M. Sappington**, “Countervailing Incentives in Agency Problems,” *Journal of Economic Theory*, December 1989, 49 (2), 294–313.
- MacLeod, W. Bentley**, “Reputations, Relationships, and Contract Enforcement,” *Journal of Economic Literature*, September 2007, 45 (3), 595–628.
- Maggi, Giovanni** and **Andres Rodriguez-Clare**, “On the Countervailing Incentives,” *Journal of Economic Theory*, June 1995, 66 (1), 238–263.
- Mailath, George J.**, “Incentive Compatibility in Signaling Games with a Continuum of Types,” *Econometrica*, November 1987, 55 (6), 1349–1365.

-
- Mas-Colell, Andreu, Michael Whinston, and Jerry Green**, *Microeconomic Theory*, Oxford, England: Oxford University Press, 1995.
- McAfee, R. Preston and Philip J. Reny**, “Correlated Information and Mechanism Design,” *Econometrica*, March 1992, 60 (2), 395–421.
- Mester, Loretta J.**, “Perpetual Signalling with Imperfectly Correlated Costs,” *RAND Journal of Economics*, Winter 1992, 23 (4), 548–563.
- Milgrom, Paul R.**, *Putting Auction Theory to Work*, Cambridge, UK: Cambridge University Press, 2004.
- Myerson, Roger B.**, “Incentive Compability and the Bargaining Problem,” *Econometrica*, 1979, 47, 61–73.
- and **Mark A. Satterthwaite**, “Efficient Mechanisms for Bilateral Trading,” *Journal of Economic Theory*, April 1983, 29 (2), 265–281.
- Nöldeke, Georg and Klaus M. Schmidt**, “Option Contracts and Renegotiation: A Solution to the Hold-up Problem,” *RAND Journal of Economics*, 1995, 26 (2), 163–179.
- and —, “Sequential Investments and Options to Own,” *RAND Journal of Economics*, 1998, 29 (4), 633–653.
- Quinzii, Martine and Jean-Charles Rochet**, “Multidimensional Signalling,” *Journal of Mathematical Economics*, 1985, 14, 261–284.
- Riley, John G.**, “Informational Equilibrium,” *Econometrica*, March 1979, 47 (2), 331–359.
- Rochet, Jean-Charles**, “The Taxation Principle and Multi-Time Hamilton-Jacobi Equations,” *Journal of Mathematical Economics*, 1985, 14, 113–128.
- and **Jean Tirole**, “Two-Sided Markets: A Progress Report,” *RAND Journal of Economics*, Autumn 2006, 37 (3), 645–667.
- and **Philippe Choné**, “Ironing, Sweeping, and Multidimensional Screening,” *Econometrica*, July 1998, 66 (4), 783–826.
- Rothschild, Michael and Joseph Stiglitz**, “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information,” *The Quarterly Journal of Economics*, 1976, 90, 629–650.
- Royden, Halsey L.**, *Real Analysis*, 2nd ed., New York: Macmillan Publishing, 1968.
- Rysman, Marc**, “The Economics of Two-Sided Markets,” *Journal of Economic Perspectives*, 2009, 23, 125–144.
- Salanié, Bernard**, *The Economics of Contracts: A Primer*, Cambridge, MA: MIT Press, 1997.

-
- Shapiro, Carl**, “Consumer Information, Product Quality, and Seller Reputation,” *Bell Journal of Economics*, Spring 1982, *13* (1), 20–35.
- , “Optimal Pricing of Experience Goods,” *Bell Journal of Economics*, Autumn 1983, *14* (2), 497–507.
- Shavell, Steven**, “Risk Sharing and Incentives in the Principal and Agent Relationship,” *Bell Journal of Economics*, Spring 1979, *10* (1), 55–73.
- Spence, A. Michael**, “Job Market Signaling,” *Quarterly Journal of Economics*, 1973, *87*, 355–374.
- Stein, Jeremy C.**, “Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior,” *Quarterly Journal of Economics*, November 1989, *104* (4), 655–669.
- Stole, Lars A.**, “Price Discrimination and Competition,” in Mark Armstrong and Robert H. Porter, eds., *Handbook of Industrial Organization*, Vol. 3, Amsterdam: North-Holland, 2007.
- Tirole, Jean**, *The Theory of Industrial Organization*, Cambridge, MA: MIT Press, 1988.
- van Tiel, Jan**, *Convex Analysis: An Introductory Text*, New York: John Wiley & Sons, 1984.
- Varian, Hal R.**, “Price Discrimination,” in Richard Schmalensee and Robert Willig, eds., *Handbook of Industrial Organization*, Vol. 1, Amsterdam: North-Holland, 1989.
- Williams, Steven R.**, “Efficient Performance in Two Agent Bargaining,” *Journal of Economic Theory*, 1987, *41* (1), 154–172.
- Williamson, Oliver E.**, “Franchise Bidding for Natural Monopolies—in General and with Respect to CATV,” *The Bell Journal of Economics*, Spring 1976, *7* (1), 73–104.
- Wilson, Charles**, “A Model of Insurance Markets with Incomplete Information,” *Journal of Economic Theory*, December 1977, *16* (2), 167–207.
- Wilson, Robert B.**, *Nonlinear Pricing*, Oxford: Oxford University Press, 1993.
- Yeh, James**, *Real Analysis: Theory of Measure and Integration*, 2nd ed., Singapore: World Scientific, 2006.