THE ENDOGENOUS VALUE OF INFORMATION

Thomas A. Weber

A DISSERTATION

in

Operations and Information Management

for the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2003

Supervisor of Dissertation

Graduate Group Chairperson

For my Parents.

Acknowledgments

I would like to express my deep gratitude to Professor Paul Kleindorfer, my advisor and teacher, for his guiding spirit and numerous fruitful discussions. His energy and comprehensive knowledge in economics and the management sciences were an invaluable resource in the endeavor of this doctoral thesis. I would also like to thank David Croson for his friendship and the long hours of enlightening exchanges. Chapter 3 of this thesis is based on our joint work. Krishnan Anand, Eric Clemons, Lorin Hitt, and Richard Kihlstrom as members of the dissertation committee were enormously helpful at various stages of this project, and I am very grateful for their support. I should also like to acknowledge my indebtedness to the great teachers at Wharton and the University of Pennsylvania, in particular David Cass, Richard Kihlstrom, Alexandre Kirillov, Steve Matthews, Rafael Rob, as well as Michael Steele, whose classes had a profound impact on my thinking.

My special thanks go to Ann & Wim for their sacrifice, endless love and understanding.

Thomas Alois Weber

Philadelphia, Pennsylvania
April, 2003

ABSTRACT

THE ENDOGENOUS VALUE OF INFORMATION

Thomas A. Weber

Paul R. Kleindorfer

This doctoral thesis examines the value of information in settings, where two or more agents interact. In such situations, contrary to one-person decision problems, a more informative signal is not necessarily more valuable, and it may be profit-maximizing for an information seller to deliberately garble or damage his signal before selling it to another agent. More generally, the value of information depends on the precise contractual arrangement under which the information is to be transferred and used. I examine the following applications: (i) value of information in portfolio decision problems; and (ii) the transfer of information to a wealth-constrained investor. In a multiagent setting I examine (iii) the value of shared information services; and (iv) the value of information and flexibility for screening a heterogeneous consumer base.

Contents

List of Tables

List of Figures

All human knowledge thus begins with intuitions, proceeds then to concepts, and ends with ideas.
— Immanuel Kant

Chapter 1

Introduction

1.1 Preliminaries

The design and management of information systems responds to the need of decision makers to efficiently gain access to critical information resources. The *decision value* of the information to be stored, accessed and managed partly justifies investments in management information systems.[1] In an uncertain environment with interacting agents, the value of information to any one of the decision makers generally depends on the precise nature of her strategic relationships to the other agents. A deeper inquiry into the issues connected with precisely determining this "endogenous" value of information (which may of course vary from agent to agent and may sometimes be negative) is the object of this dissertation. The information value is thereby called "endogenous," since it fundamentally depends on the actions of other agents and on the way the information it is procured.

In the Economics literature the value of information to a particular agent is often assessed using one of the two following methods, depending on the agent's risk posture (Raiffa and Schlaifer 1961, Marschak and Miyasawa 1968, Athey and Levin 2000): (i) for *risk-neutral* agents the information value is determined as difference between the optimal payoff with information

[1]There is a rich literature on assessing the value of information technology investments (Barua et al. 1991, Clemons 1991, Hitt and Brynjolfsson 1996, Benaroch and Kauffman 1999), cf. also Section 1.3. Instead of entering this discussion directly I am focusing here on the economic value of information for specific strategic decisions only. Issues such as the impact of IT on the reduction of coordination costs (Shin 1997) or on firm boundaries (Clemons et al. 1993, Dewan et al. 1998, Hitt 1999), the substitution of IT for other production factors (Dewan and Min 1997), the real-option value of IT projects (Kambil et al. 1993, Benaroch and Kauffman 1999), general productivity increases (Mukhopadhyay et al. 1997, Mukhopadhyay et al. 1997b) or changes in consumer surplus (Hitt and Brynjolfsson 1996) through IT are not explicitly considered here.

and the optimal "default" (or "no-information") payoff; (ii) for *risk-averse* agents, the cost of an informative signal is typically modeled as a "utility cost," which is then subtracted from the agent's full-information utility, *as if* the amount for the information was paid out of a separate budget. These two common approaches to computing the value of information neglect the "wealth effect" the payment itself has on the agent's budget set and thereby on constraining her feasible actions. They also often neglect the *strategic* utility or disutility the access to information may exhibit when agents interact.

In this doctoral thesis I investigate the *endogenous* value of information in multiagent situations, taking "wealth effects" fully into account. One purpose is to clarify what causes information to potentially *lose* value[2] when players interact and therefore trigger efficiency losses compared to its first-best use. For instance, Weber and Croson (2002) show that a revenue-maximizing risk-neutral seller of investment information may have an incentive to "damage" the information he sells, so as to render the buyer's investment in a risky asset (to which the information relates) less "aggressive" and thus to decrease the risk of default on an agreed ex-post payment for the information. Here the voluntary distortion of information – in spite of reducing the buyer's willingness to pay (WTP) – increases expected revenues (based on a limited-liability ex-post payment) for the information seller. More generally, if the payoff of agent A who controls a piece of valuable information depends on agent B's noncontractible actions, A might be able to influence B's actions to his advantage by distorting the information[3] and still charge B a (slightly diminished) rent for it.

The aforementioned "wealth effects," induced by the fact that the payment for information reduces the budget set and constrains the information buyer's feasible actions, are of considerable relevance at least in a portion of the wealth domain, irrespective of the assumed risk-aversion characteristics. For instance, if the payment for information occurs ex ante with respect to the realization of the uncertain event, then the agent's willingness to pay (WTP) for it can never exceed her current wealth. I demonstrate that these wealth effects *cannot* generally be eliminated by an "appropriate" choice of a class of utility functions possessing multiplicative or additive separability properties (such as those of constant absolute risk aversion as is often claimed).

As a consequence, the value of information needs to be carefully defined taking into account both the exact contractual terms governing the transfer of information from the seller

[2]It can naturally also gain value, for instance through the aggregation of information by several agents. I will also investigate such situations (e.g., in the context of shared information services), but in a somewhat less focused manner, as these effects are better accounted for in the current literature.

[3]Agent B is fully aware that this distortion takes place.

to the buyer *and* the utility variation due to wealth reduction. We define the monetary value or WTP for information in terms of Hicks' (1939) measure of welfare change, as the *compensating variation* that moves the agent's expected utility with information to the same level as without information (cf. Kihlstrom (1974) for the product-consumption treatment of information). This definition of the value of information to a potential acquirer of the information is very general, but depends on how exactly the agent is to pay for the information. Payment may occur *ex ante* or *ex post* with respect to the realization of the payoff-relevant underlying stochastic variable and may be contingent on the signal's realization or not. In addition, overall efficiency may be improved if the principal can make the price to be paid dependent on the agent's actions or a (possibly composite) signal correlated with the agent's actions (and/or the external event) and also dependent on the agent's payoff (or ex-post wealth). The resulting contracts generally aim at shifting the risk-averse agent's actions towards more risk-taking thereby increasing her (expected) willingness to pay for information.

If the use of the information is rival,[4] which means that by selling the information the principal loses the ability to use it himself, the seller has to be offered a payment that corresponds at least to his willingness to accept (WTA) or *equivalent variation* in utility for not observing the signal realization and passing it on to the agent. In the next chapter I derive an exact relationship between WTP and WTA and show that it is not necessary for one of these welfare measures to be larger than the other, so that rival information may or may not be traded depending on principal and agent's respective utility profiles.

Another purpose of this thesis is to demonstrate a new setting in which managerial flexibility and information can to some extent be substitutes. As an application I consider multiattribute product differentiation and I explicitly value the option to delay differentiation until perfect demand information arrives versus the value of imperfect demand information without the possibility to observe true demand before committing to a product portfolio. I show how much information is needed ex ante to make the firm indifferent between the flexibility of delaying the differentiation decision and observing a partially informative signal about demand.

1.2 Research Questions

The first part of the thesis focuses on a general definition of the value of information for a single agent in terms of her WTP and WTA.

QUESTION 1 *How can the value of information be generally defined in terms of an agent's*

[4] Examples are customized information or information that is protected by ownership rights, such as a patent.

willingness to pay and willingness to accept, and what is the precise relationship between the two welfare measures?

In the second part of this dissertation, I look at two-player situations, where an informed principal proposes to transfer information to an uninformed agent and examine how different sale contracts influence the revenues for the principal. If his revenues depend on an noncontractible action by the agent, which can be however indirectly influenced by modifying the information strategically, then there exists a potential incentive for the principal to garble information strategically, since at the margin the value of the ungarbled information is exceeded by the value of a favorable change of the agent's action. As pointed out before, such an action could be an overly aggressive investment behavior under an ex-post payment agreement for information about a risky investment opportunity, but more generally it can be anything that affects the principal's payoffs.

QUESTION 2 *Consider a principal-agent environment.*

(i) What are the incentives for the distortion of information in a principal-agent relationship where information exerts an externality?

(ii) What are the consequences for organizational design for an informed principal?

In the third part of the thesis, I focus attention on the endogenous value of information in multiplayer situations with both cooperative and strategic information sharing. Chapter 4 examines the value of shared information services. I then turn to situations where a firm would like to extract information from a consumer base with multiple hidden characteristics.

QUESTION 3 *Consider a multiagent environment.*

(i) How should a heterogeneous group of agents share cooperatively an investment for a common source of information?

(ii) What is the option value of being able to wait for information that helps a firm to screen a heterogeneous consumer base?

The *main objectives* of this thesis are (a) to provide a theoretical contribution that may influence the way managers think about the value of information and how it depends on the particular parameters of the problem; (b) to create a foundation for a justification of IT investments through an economic analysis of the endogenous information value; and (c) to examine a number of concrete applications in detail.

In an attempt to find answers to the above research questions I am using analytical tools of standard microeconomic theory (Mas-Colell et al. 1995), statistical decision theory (DeGroot 1970, Raiffa et al. 1995), information economics (Laffont 1989, Hirshleifer and Riley 92, Chambers and Quiggin 2000), and the theory of incentives (Salanié 1999, Laffont and Martimort 2002) to construct theoretical models addressing different features of the overall problem.

1.3 Literature Review

I will now review some of the literature pertaining to the main chapters of this doctoral thesis: the value of information (i) for a single decision maker, (ii) in two-agent (i.e., principal-agent) environments, and (iii) in multiagent environments. The reason for distinguishing between principal-agent and multiagent situations is that in multiagent environments the granularity of the analysis is necessarily higher due to an increased model complexity. In addition we will interpret a principal-agent setting in which the type of an agent is only imperfectly known to another agent as a multiagent situation, since the agent whose characteristic is not perfectly known can effectively be interpreted as a population of many agents, the distribution of which corresponds to the prior probability distribution in the characteristic. For instance I will consider screening methods (e.g., through the "versioning" of information) in that section.

Before discussing the economic value of information in these three regimes, I will briefly relate this dissertation topic to the literature on the value of information technology.

1.3.1 IT and the Value of Information

The value of IT has been discussed in three main streams of the research literature: (i) the classic "productivity paradox" contributions (Brynjolfsson and Hitt 1996, 2000); (ii) the options-value-of-IT literature (Benaroch and Kauffman 1999); (iii) the literature on optimal investments in shared information systems, in particular EDI systems (Clemons and Kleindorfer 1992, Riggins et al. 1994, Wang and Seidmann 1995). In addition, there are various contributions on the impact of IT on firm boundaries (Jensen and Meckling 1992, Clemons et al. 1993, Hitt 1999) or on the reduction of coordination cost (Shin 1997, Nault and Tyagi 2001).

A main conclusion of the first stream of literature is that there exists a link between IT investment and complementary investment. The motivation for the complementary investment can be explained via the information processing view of the firm (Galbraith 1974, Radner 1993, VanZandt 1997). Radner (1993) argues that in an efficient hierarchy the top decision node is always busy, so that information overload at the top necessitates the decentralization of decision rights. Jensen and Meckling (1992) show that as a result of the tradeoff between the ease of information transmission (favoring decentralization) and the need for costly monitoring (favoring centralization), there exists an optimal level of decentralization, moving the organization "to the middle" (Clemons et al. 1993).

The options-value-of-IT literature does address the issue that IT generates value by enabling a more flexible reaction to uncertainty. But most of this literature does not consider strategic interaction and treats the competitive environment in a *stochastic* manner in order to use the standard real-options valuation framework (e.g., Dixit and Pindyck 1994, Trigeorgis 1996).

As already pointed out in footnote 1 on page 1, I concentrate in this dissertation on the *economic value of information* as it pertains directly to managerial decision making, and not so much on the more encompassing assessment of the value of information technology. The strategic use of the information that actually passes through the information systems has not received very much attention in the standard information systems literature. However, there is a substantial literature on information systems and their endogenous value in the Economics literature. The term "information system" is here understood in the classical sense of Marschak and Miyasawa (1968) as a "set of potential messages to be received by the decision maker" (p. 137). The value of an information system derives from an expected payoff variation it can bring to an otherwise uninformed decision maker. An information system can be interpreted as the result of an "experiment" as it relates to a decision maker's payoff (Raiffa and Schlaifer 1961, DeGroot 1970). Such experiments can be processed and (sometimes) be generated by information technology (IT). In that context, the value of IT derives from the ability to cut down on the cost of generating experiments and/or from the capacity of improving on the decision value of the generated informative experiment outcomes (referred to as messages or "signals"). In addition, IT derives value from the aspects mentioned in footnote 1, which we do not explicitly consider in the theoretical developments in this dissertation. We limit ourselves to the intrinsic economic value of information where agents interact.

1.3.2 Environments with a Single Decision Maker

The origins of information theory reside in the work by Shannon (1948); a unified treatment of the main results of information and communications theory can be found in Gallager (1968). The systematic assessment of the value of an experiment has been pioneered by Bohnenblust et al. (1949). Based on this work Blackwell (1953) shows that two experiments can be ranked in terms of their informativeness if and only if one signal is a sufficient statistic of the other. DeGroot (1962) demonstrates that the signal ranking in terms of informativeness does indeed imply an equivalent ranking in economic terms when the signals are benchmarked over *all* possible decision problems.[5] The point is that in the sense of Bohnenblust et al., Blackwell, and DeGroot, for a one-person decision problem a distorted or garbled signal could not possibly be of higher value to the decision-maker, unlike in situations where agents interact strategically as we will see below. Kihlstrom (1984) unifies these mathematical approaches to the value of information in economic terms and from a Bayesian viewpoint.

In this dissertation I define the value of information as the Hicks' (1939) *compensating variation* that moves the a decision maker back to his maximum expected utility achievable without information. This welfare measure has first been used by Kihlstrom (1974) to compute the value of information in a consumption setting and by Treich (1997) in the context of a standard portfolio investment problem. As alluded to at the outset, the funds used to purchase information are in many models, describing for instance information acquisition (such as in Athey and Levin (2000)) segregated from the wealth used to generate utility, effectively assuming additive separability in utility from the "wealth to be invested" and "wealth earmarked for purchasing information." In these models the cash disbursed to acquire information does not constrain the feasible set of actions, nor does economizing on information or deferring its purchase settlement yield any investment benefit. The treatment here it inclusive in that it explicitly takes into account the wealth effects induced by the contraction of the budget set by the information expenditure.

1.3.3 Principal-Agent Environments and the Incentive to Degrade Information

When agents interact Blackwell's (1953) informativeness criterion does not necessarily coincide any longer with the economic value of information. For instance, Gal-Or (1988) shows in the situation of a two-stage Cournot duopoly that less precise information of one firm about its own cost can increase its incentives to increase production and therefore lead to higher expected profits than the other more informed firm, which – as a best response – reduces output in equilibrium. The key reason for the negative marginal value of information is that a lack of information provides a credible commitment mechanism that deters the informed party from producing at similar levels. More generally, the value of information can become negative if the information at the margin leads to a negative externality, for instance – as in

[5]For more specific situations, such as monotone decision problems with certain characteristics of the payoff function, there are stronger characterizations of the value of information available (Athey 2000), which allow for less restrictive criteria than statistical sufficiency.

Gal-Or (1988) – the reduction of the capability to commit or – in the case of a contractual transfer of information – moral hazard by the information buyer. A natural consequence is that in such settings value can be created to at least one party by distorting (or "garbling") the information.[6]

The elements of the research literature related to endogenous garbling of information fall into three major categories: (1) the concept of signal garbling and relative informativeness of multiple signals (Blackwell 1953, Marschak and Miyasawa 1968, Holmstrom 1979, Gjesdal 1982); (2) the concept of deliberately degrading the quality of a good to be sold, for the purpose of encouraging buyers to separate by type (Deneckere and McAfee 1996, Shapiro and Varian 1998); (3) diverse analyses of "perverse" situations in agency contexts, wherein more accurate information or more intensive monitoring induces less effort or worse outcomes for the principal (Cowen and Glazer 1996, Dubey and Haimanko 2000, Jacobides and Croson 2001).

1.3.4 Multiagent Environments

I examine multiagent environments from two different viewpoints. First, how should a heterogeneous group of agents share cooperatively an investment for a common source of information? Second, how does a firm facing a mass of agents with heterogeneous hidden characteristics build a multiattribute product portfolio to extract a maximum of profits and how does it value flexibility and bettern demand information in that context?

The pioneering contributions in *strategic* information sharing have been made by Novshek and Sonnenschein (1982) and Clarke (1983) for oligopolistic firms producing homogeneous goods (no incentive for information sharing), as well as Vives (1984) for a differentiated Cournot duopoly (incentive for information sharing exists if the goods are complements or weak substitutes). Raith (1996) unifies and generalizes these and most subsequent treatments. Kihlstrom and Vives (1989) look at information sharing where it may help firms to collude in a Cournot oligopoly. I also look at information sharing in a *cooperative* context, where sharing the cost of a fixed investment in a common information service implies increasing returns in the number of participants sharing the service. Nash (1953) founded the field of "cooperative bargaining theory," showing that based on a number of axioms a unique solution could be obtained that implements a Pareto-optimal allocation. A good overview is provided by Roth (1979).

Beginning with Hotelling's (1929) seminal paper on horizontal competition, numerous con-

[6] In the case of public (or "social") information, there may also an incentive for *both* players to garble the information, whenever the information destroys the market for mutual insurance (Hirshleifer 1971).

tributions have been made to product differentiation.[7] The corresponding literature can be divided into locational models in the tradition of Hotelling, where each firm is attributed an "address" in product space, and into so-called "non-address" models in the spirit of Chamberlin's (1933) monopolistic competition, where a representative consumer exhibits (probabilistic) preferences for different products.[8] An important distinction between the two groups of models is that in the latter group, *each* product is competing with each other, while in the former consumers are truly heterogeneous in their preferences, and some products may have no overlap, i.e., may never be in direct competition.[9] To find optimal second-degree differentiation strategies, I adopt the locational approach, which in my view better captures consumer heterogeneity and allows the explicit consideration of participation constraints that inevitably arise when dealing with a spatial distribution of endowed unobservable consumer charateristics. In fact, Lancaster (1966) first realized "[t]he good, per se, does not give utility to the consumer; it possesses characteristics, and these characteristics give rise to utility" (p. 65). In addition he pointed out that "[i]n general a good will possess more than one characteristic, and many characteristics will be shared by more than one good" (ibid.). We refer to these Lancasterian characteristics as *product attributes*, and naturally products contain a number of different such attributes, which – facing a heterogeneous consumer base of unknown types – allows a monopolist (the "principal") to *screen* the "agents." The product attributes can be used as instruments in the screening process. Using multiple instruments to screen consumers of a one-dimensional type has been earlier examined by Matthews and Moore (1987), whereas the inverse case of a single instrument (price) given consumers of multidimensional types has been considered among others by Laffont, Maskin, and Rochet (1987). This line of work on second-degree price discrimination dates back to Mussa and Rosen (1978), based on methods developed earlier by Mirrlees (1971) in the context of optimal income taxation, who treat the case for consumers of a single characteristic and single vertical-attribute products. Wilson (1993) and Armstrong (1996) provide generalizations for fully nonlinear pricing models in the multiproduct case. A multidimensional screening model generalizing these approaches has been advanced by Rochet and Choné (1998), based also on results in multidimensional signalling by Wilson (1985) as well as Quinzii and Rochet (1985). In contrast to Rochet and Choné (1998) I consider a *discrete* product structure (i.e., "natural" bunching) generated by

[7] The notion of product differentiation as such can be traced back at least to Launhardt (1885).

[8] For a good bibliography see Anderson et al. (1992), Beath and Katsoulacos (1991), as well as Tirole (1988).

[9] A notable exception in this dichotomy, is the model by Perloff and Salop (1985) that combines characteristics from both groups, driven by symmetry assumptions in the preferences of a representative consumer who is faced with localized products.

fixed versioning costs for each new product and deal with type-dependent participation constraints. This allows us to be less concerned about single-crossing properties than with the endogenous choice of the mode of product differentiation (vertical, horizontal, mixed). The particular setup of our model is inspired by Salop's (1979) circular city, which allows to avoid undesired boundary effects of solutions and thus to obtain cleaner results.

1.4 Outline

The outline of this dissertation is a follows. Chapter 2 presents some theoretical foundations on the value of information. In particular I focus on the value of information that is transferred from one agent to another and depending on the precise contractual agreement for the transfer the willingness to pay for information may vary dramatically, even though the underlying signal whose realization is to be observed remains unaltered. In Chapter 3, I examine the value of information given payment that can occur ex ante and/or ex post with respect to the realization of the for the information-user critical exogenous uncertain event. Chapter 4 deals with the value of information for heterogeneous agents when the investment in a common information source is to be shared. In Chapter 5 I turn attention to a problem of multidimensional screening of a consumer base that is heterogeneous with respect to both reservation prices and a horizontally differentiating criterion. In this context I determine for a multiproduct monopolist the value of the flexibility to obtain demand information before committing to a full differentiated product portfolio. Chapter 6 concludes this dissertation by revisiting the research questions outlined in Section 1.2 and provides some general directions for future research.

1.5 References

For a detailed list of references see the Bibliography at the end of this dissertation.

I have had my results for a long time, but I do not yet know how I am to arrive at them.
— Carl Friedrich Gauss

Chapter 2

The Value of Information: Foundations

In this chapter I first give some basic definitions underlying the general idea of the thesis and discuss some related applications. A fundamental concept for valuing information is the notion of "compensating variation," i.e., the monetary amount that makes the information user indifferent between having no information but full wealth and the information in question and wealth diminished by the price for the information. The compensating variation measures the WTP of an investor for information. The dual concept to the compensating variation of an information user is called the "equivalent variation," i.e., the minimal amount that a seller would accept in order to *not* use the object (e.g., rival information). I provide later in this chapter an *exact* relation between compensating and equivalent variations, essentially independent of the kind of good that is being transferred (other than that it is non readily substitutable by other goods purchased on a market).

2.1 Preliminaries

Let an agent's monetary payoff $\Pi(a, \tilde{x})$ depend on her action a and let an exogenous uncertain event be represented by the random variable $\tilde{x} \in \mathcal{X}$. The agent is endowed with an intial wealth $W \geq 0$ and her actions belong to a compact action set[1] $\mathcal{A}(W) \subset \mathbb{R}^n$ that is an image of W and is strictly contracting with decreasing wealth, i.e.,

$$W' < W \quad \Longrightarrow \quad \mathcal{A}' \subsetneq \mathcal{A}, \tag{2.1}$$

[1] We assume that the mapping $\mathcal{A} : \mathbb{R}_+ \to \mathbb{R}^n$ is upper hemicontinuous.

where $\mathcal{A}' = \mathcal{A}(W')$, $\mathcal{A} = \mathcal{A}(W)$, and n is some positive integer. For zero wealth we assume that the action set reduces to a singleton containing only the "zero-wealth" strategy a_0, which naturally is an element of *all* action sets. The agent's utility for money is given by the differentiably concave and strictly increasing function $U : \mathbb{R} \to \mathbb{R}$. We also assume that the payoff function $\Pi(\cdot, \cdot)$ is smooth with respect to both arguments as well as concave in the action a. It is therefore clear that the maximizer of the agent's expected utility maximization problem,

$$\hat{a}(W) \in \arg \max_{a \in \mathcal{A}(W)} E_x \left(\Pi(a, \tilde{x}) \right), \tag{2.2}$$

exists. As a special case, we can think of the action set as a linear budget set of the form

$$\mathcal{A}(W) = \{ a \in \mathbb{R}^n : a \leq W \} . \tag{2.3}$$

Information is conveyed via a *signal*, represented by the random variable $\tilde{s}$ with a measurable probability density defined on the compact support $\mathcal{S}$. The signal $\tilde{s}$ is thereby *correlated* with the exogenous random event $\tilde{x}$, from which its potential value for the decision maker arises.

2.2 The Value of Information

In the standard portfolio problem (Ingersoll 1987) the agent's payoff Π is determined by the portion $a \in \mathcal{A}(W) = [0, W]$ of her wealth W that she allocates to a risky asset of uncertain return $\tilde{x} = \tilde{r}$,[2]

$$\Pi(a, \tilde{r}) = W + a\tilde{r}.$$

In Chapter 3, we derive simple necessary and sufficient conditions under which any information seller prefers to contract for payment to occur *after* the signal is realized, as opposed to demanding prepayment. Fully secured payments may not be optimal, depending on the seller's risk preferences. We derive optimal pricing structures for this information, characterizing investment opportunity, signal quality and investor utility conditions under which the seller can receive more value from his information by deferring payment after the investment realization occurs. Further, we find that the seller of the information, in order to at least partially secure payment, may have an incentive to garble the signal, so as to exclude realizations that induce the investor to allocate a large share of her portfolio to risky assets.

Single-person decision problems. For single-person decision problems the value of information is strictly increasing in the informativeness of the signal, as is well known through the work of Bohnenblust et al. (1949), Blackwell (1953), and DeGroot(1962). It also depends

[2] Note that $\mathcal{A}(W)$ is strictly contracting as W decreases, in accordance with (2.1).

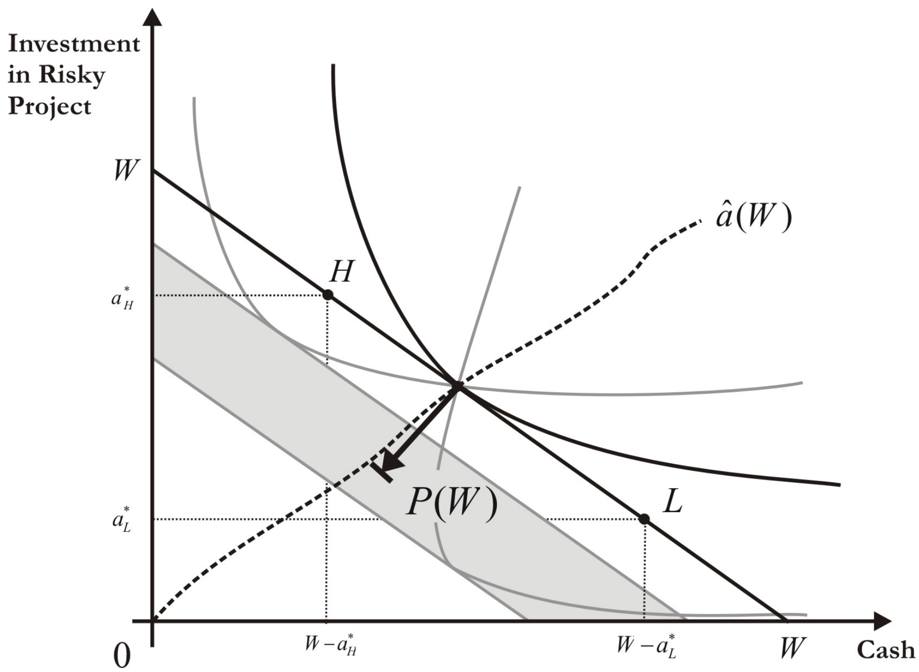

Figure 2.1: Willingness to Pay, P, for Information.

decisively on the "default" action (2.2) an agent would take without any information. For a simple informative signal with two realizations (L: "low" and H: "high"), the willingness to pay can be determined as *compensating variation* P in wealth that makes the agent indifferent between obtaining the signal or not, cf. Figure 2.1 (where a_s^* are the agent's optimal actions in a simple portfolio investment problem contingent on realizations $s \in \{L, H\}$).

The value of information can be rarely isolated from the precise way in which the agent obtains it. Thus, in this thesis I pay particular information to the *endogenous* value of information as it is realized for any particular transfer contract from a (trusted) information seller to the agent.

Value of transferred information. As pointed out above, the value of information depends on the endowed wealth of the economic agents involved in the transfer, their respective risk profiles, and the precise contingencies specified in the transfer contract. In addition, if information is rival this value may be different for buyer and seller so that no trade may actually occur. Below we provide some alternative definitions of the value of information to illustrate its dependence on the payment contract. Such a contract be in principle contingent on the signal realizations as we show below.[3]

[3]In following chapters of this thesis I will not discuss signal-contingencies in any detailed fashion. It is an

2.2.1 First-Best WTP

The agent's the "default" or "no-information" strategy $\hat{a}(W)$ is given by (2.2). Hence the certainty equivalent of the investment opportunity is determined by

$$\mathrm{CE}(W) = U^{-1}\left(E_x U(\Pi(\hat{a}(W), \tilde{x}))\right).$$

Thus we are able to compute the first-best WTP, or in other words the absolute maximum of the information value contained in the principal-agent "system":

$$P_{\mathrm{FB}} = E_s\left[\max_{a\in\mathcal{A}(W)} E_x\left[U(\Pi(a,\tilde{x}))|\,\tilde{s}\right]\right] - \mathrm{CE}(W).$$

This first-best price that the principal can charge a risk-neutral agent given an efficient mechanism is the benchmark against which we can compare other contract designs, of which I discuss some below.

2.2.2 Ex-Ante Noncontingent Payment

If full payment occurs *before* the realization of the signal is observed, then the agent's action set will be reduced corresponding to the amount she pays. Her maximum willingness to pay $P \in [0, W]$ is a monetary transfer to the principal that makes the agent indifferent between having access to the information or not, and

$$E_s\left[\max_{a\in\mathcal{A}(W-P)} E_x\left[U(\Pi(a,\tilde{x}))|\,\tilde{s}\right]\right] = E_x U(\Pi(\hat{a}(W), \tilde{x})) \tag{2.4}$$

is necessarily satisfied.[4] The so obtained ex-ante noncontingent WTP is strictly less than W, iff the zero-wealth singleton does not contain the optimal default strategy, i.e., iff $a_0 \neq \hat{a}$.

2.2.3 Ex-Ante Contingent Payment

If it is possible for both the agent and the principle to observe the signal realization before the agent selects her action, than the payment can be made in principle contingent on that realization and be contracted ex-ante.[5] For any $s \in \mathcal{S}$ the following relation therefore describes

interesting direction for future research with potential financial applications such as margin loans.

[4]Note that from a technical point of view (2.1) does not ensure that (2.4) is satisfied only for a single point in $[0, W]$, since the contraction of the action set might have no effect on the feasibility of the optimal action with information. Thus strictly speaking, one should think of P as a reservation-price interval. Of course P reduces to a singleton, if we use a linear budget set according to (2.3).

[5]Of course, this contract must be binding since the agent may have an incentive to renege once the signal realization has been observed. For instance the principal or a third party may take hold of the agent's wealth (surety bond) and then execute her actions after deducting the contingent payment promised to the principal

the agent's maximum contingent WTP:

$$\max_{a \in \mathcal{A}(W - P^c(s))} E_x \left[U(\Pi(a, \tilde{x})) \right| s] = E_x \left[U(\Pi(\hat{a}(W), \tilde{x})) | s \right]. \tag{2.5}$$

Then the expected compensation for the information seller is

$$P^c = E_s(P^c(\tilde{s})|\tilde{s}),$$

and this value may or may not be larger than P, depending on the agent's risk aversion and the specific problem at hand. For the portfolio investment problem discussed below it is possible to construct an example that produces both results as a function of the riskiness of the investment[6]

2.2.4 Ex-Post Noncontingent Payment

The principal may choose to relax the agent's budget constraint by allowing her to make the payment *after* she has taken her investment action. A contract to this effect is written ex ante, and in the case it is not contingent on the signal realization, the solution $\hat{P}$ to

$$E_s \left[\max_{a \in \mathcal{A}(W)} E_x \left[U \left(\left[\Pi(a, \tilde{x}) - \hat{P} \right]_+ \right) \middle| \tilde{s} \right] \right] = E_x U(\Pi(\hat{a}(W), \tilde{x})) \tag{2.6}$$

describes the agent's ex-post noncontingent "nominal" WTP. The WTP is called "nominal", since the principal can actually only collect at most all of the agent's ex-post wealth, so that he implicitly assumes the agent's default risk. On average the expected amount $\bar{P}$ the principal collects, and what I am referring to as the "ex-post noncontingent WTP," is given by

$$\bar{P} = E_s \left[\min \left\{ \hat{P}, \max_{a \in \mathcal{A}(W)} E_x \left[\Pi(a, \tilde{x}) \right| \tilde{s}] \right\} \right].$$

It is clear that $\hat{P} = \bar{P}$, iff there is no default risk, and further we have that $\bar{P}$ (which is greater than both P and P^c) can in principle exceed W, especially when payoffs are high enough and the signal is sufficiently informative.

2.2.5 Ex-Post Contingent Payment

If the ex-post payment is made contingent on the common observation of the signal, then the agent's WTP in the event of realization $s \in \mathcal{S}$ is determined by

$$\max_{a \in \mathcal{A}(W)} E_x \left[U(\Pi(a, \tilde{x}) - \bar{P}^c(s)) \right| s] = E_x \left[U(\Pi(\hat{a}(W), \tilde{x})) \right]. \tag{2.7}$$

[6] It turns out that for very risky investments any agent would prefer to pay a noncontingent price.

The expected compensation for the information seller is consequently

$$\bar{P}^c = E_s(\bar{P}^c(\tilde{s})|\tilde{s}).$$

We remark that there is no default risk for the information seller in this case and therefore no need to post a different "nominal" price as we had to in the non-contingent ex-post payment contract.

2.2.6 Incentives for the Garbling of Information

The default risk of the agent in the ex-post payment case may provide the principal with an incentive to garble certain realizations of the signal or otherwise distort the signal, such that the initial signal $\tilde{x}$ is a sufficient statistic for the modified signal $\tilde{y}$, but the converse does not hold. I examine the incentives of the principal to distort his signal and how this can be reconciled with Blackwell's theorem that predicts $\tilde{x}$ to be more valuable for "any" decision problem. Another incentive for the modification of a nonrival signal may be to offer a range of distortions in order to screen a heterogeneous base of agents and engage in second-degree price discrimination.

2.2.7 Other Forms of Bilateral Information Transfer

As previously mentioned, the principal can propose other nonlinear contracts contingent on the actions or on ex-post wealth, if either is observable and verifiable. This may increase overall surplus and distort the agent's actions towards higher risk-taking, the extreme case being when the principal compensates the agent by paying her certainty equivalent and then implementing the action on his own, which achieves a first-best outcome.

2.2.8 Multiagent Problems

The Value of Shared Information Services

In Chapter 4, I analyze the value of shared information services, both when they are operated by their members and when they are implemented by a monopoly provider. The value of information is defined as the compensating variation in price that makes a risk-averse agent indifferent between procuring an informative signal or not. I provide investment sharing rules that implement an individually rational Nash bargaining solution and compare this to the situation in which a nonscreening monopolist maximizes profits. We find that any efficient price schedule for information should take into account (i) the agent's confidence in the signal,

(ii) the agent's project risk, (iii) her risk aversion, and (iv) her wealth and the mean return of the project if they are small. Interestingly in a cooperative bargaining situation an agent's investment share may either increase or decrease when her risk aversion goes up, depending on if her demand for information decreases faster than her bargaining power relative to the other agents or vice versa. I further show that even for CARA utilities there are important wealth effects. Our results, including the definition of a "critical Nash network size," provide a benchmark for the value of information that is shared by a group of agents for use in their respective projects and not employed strategically against each other.

Screening and Delayed Differentiation

In Chapter 5, I investigate the multiattribute product versioning problem, the findings of which also apply to the maximization of profits from the sale of products to a heterogeneous consumer base. In this section I examine the value of the option to delay product differentiation until better information has arrived and compare it to the ex-ante quality of information needed to make the firm indifferent to having the flexibility to delay differentiation.

2.2.9 Willingness to Accept

I have pointed out earlier that if the use of the signal is rival, because for instance it contains customized information pertaining to some bilaterally shared idiosyncratic risk or it perishes directly after observation, then the seller of the information has a reservation price A given by his willingness to accept depending on his own utility of observing the signal himself. For a risk-neutral agent the ex-ante noncontingent WTA is determined by

$$E_s\left[\max_{a\in\mathcal{A}(W)} E_x\left[\Pi(a,\tilde{x})|\,\tilde{s}\right]\right] = \max_{a\in\mathcal{A}(W+A)} E_x\Pi(a,\tilde{x}) \tag{2.8}$$

and the extension of this definition to the risk-averse case is straightforward by introducing a strictly concave utility function for the principal. As for the WTP, the payment of A can be made contingent on the signal realizations and other contractual terms such as the timing of the payment. Definitions of $A^c, \bar{A}$, and $\bar{A}^c$ can therefore be given in complete analogy to $P^c, \bar{P}$, and $\bar{P}^c$ and the following discussion of WTP also directly applies, *mutatis mutandis*, to the WTA.

2.3 General Comment: Relation between WTP and WTA

I now present a general, simple and exact relationship between the compensating variation C and the equivalent variation E associated with an exogenous welfare change. Differences between willingness to pay (WTP) and willingness to accept (WTA) have been widely acknowledged in the empirical literature, where a persistent discrepancy in the two welfare measures is noted.[7] On the normative side, progress has been made in bounding the difference between WTA and WTP. The equivalent and compensating variation (denoted by E and C) are the welfare measures in standard demand theory that directly correspond to WTA and WTP. Willig (1976) noted that the difference between the two is likely to be small if the change in welfare is due to a price change of a market commodity. However, based on results by Randall and Stoll (1980), Hanemann (1991) shows that when the welfare change is induced by varying a nonmarket public good q, then differences between C and E can be arbitrarily large (infinite in the limit), depending on the degree of substitutability between q and the other ordinary market commodities. He restates the bounds on the difference obtained by Randall and Stoll in terms of elasticities, making plain the separate influence of substitution and income effect, each accounting for a portion of the deviation.

In this paper, I derive an explicit relation between C and E that holds for a large class of utility maximization problems. The idea is that, given a certain reference variation in the level of the nonmarket good, the induced equivalent variation $\hat{E}$ at an income level $\hat{y} = y - C$, reduced by the compensating variation, is equal to C. And this identity holds over the whole range of incomes, so that one welfare measure can be recovered from the other by the fundamental theorem of calculus. The obtained identity between compensating and equivalent variation allows bounding the difference, $E - C = \text{WTA} - \text{WTP}$, if limits on the changes of one welfare measure are available, for instance through direct computation or observation.

Suppose there are $n \geq 1$ conventional market goods $x_1, \ldots, x_n$ and one nonmarket good q. Let the consumer's preferences over the consumption of these goods be strictly convex, and represented by the increasing and strictly quasiconcave[8] utility function, $u : \mathbb{R}^{n+1} \to \mathbb{R}$. In addition, to simplify the ensuing analysis, assume that $u = u(\mathbf{x}, q)$ is continuously differentiable. Given a vector $\mathbf{p} \gg 0$ whose components represent prices for the respective market commodities, the consumer tries to find the optimal Hicksian commodity bundle $\mathbf{x}^*$ in some feasible convex set $\mathcal{X} \subset \mathbb{R}_+$, subject to her finite income $y \geq 0$. The classical utility

[7] See e.g., Kahneman et al. (1991).

[8] Strict quasiconcavity is required only with respect to the conventional market goods $\mathbf{x}$ to guarantee the existence of a unique solution to the utility maximization problem (2.9).

maximization problem is given by

$$\max_{\mathbf{x}\in\mathcal{X}} u(\mathbf{x}, q), \quad \text{subject to} \quad \mathbf{p}\cdot\mathbf{x} = y. \tag{2.9}$$

Because of the continuity of u there exists an optimal solution, $\mathbf{x}^*$, to the utility maximization problem (2.9), the components of which are described by the following Hicksian demand functions

$$x_i^* = h^i(\mathbf{p}, q, y), \quad i = 1, \ldots, n.$$

The resulting indirect utility is defined as $v(\mathbf{p}, q, y) = u(h(\mathbf{p}, q, y), q)$. Note that the indirect utility is strictly increasing in income, so that for any $(\mathbf{p}, q) \gg 0$,[9]

$$y^0 < y^1 \quad \Longrightarrow \quad v(\mathbf{p}, q, y^0) < v(\mathbf{p}, q, y^1). \tag{2.10}$$

To value a change in the provision of the nonmarket good q from q^0 to q^1 we define the compensating variation $C(\mathbf{p}, y)$ and equivalent variation $E(\mathbf{p}, y)$ by

$$v(\mathbf{p}, q^1, y - C(\mathbf{p}, y)) = v(\mathbf{p}, q^0, y), \tag{2.11}$$

$$v(\mathbf{p}, q^1, y) = v(\mathbf{p}, q^0, y + E(\mathbf{p}, y)). \tag{2.12}$$

Without loss of generality[10] we assume $q^1 > q^0$. The term $C(\mathbf{p}, y)$ is the income that a consumer would need to be *compensated* with in order to be indifferent between the higher level q^1 and the current level q^0. If the consumer is endowed with q^1, then $E(\mathbf{p}, y)$ is the income that yields *equivalent* utility to her as consuming at the lower level q^0. In other words, C corresponds to the WTP, and E to the WTA for the proposed welfare change.

PROPOSITION 2.1 *The willingness to pay, $C(\mathbf{p}, y) \in [0, y]$, and the willingness to accept, $E(\mathbf{p}, y) \in (0, \infty)$, for the welfare change from q^0 to q^1 exist, and are uniquely determined by (2.11) and (2.12) respectively.*

Proof. For any $(\mathbf{p}, q, y) \geq 0$ there is a unique $\mathbf{x}^* = h(\mathbf{p}, q, y)$ that solves the maximization problem (2.9). We obtain thus

$$v(\mathbf{p}, q^1, y - 0) \geq v(\mathbf{p}, q^0, y) \geq v(\mathbf{p}, q^0, y - y),$$

so that there exists a $C(\mathbf{p}, y) \in [0, y]$ that solves (2.11). It is unique, since the LHS of (2.11) is strictly monotone in C as can be seen by differentiating with respect to C and using the

[9] This result holds under weaker conditions on u (only continuity and local nonsatiation of the underlying preferences are required, cf. Mas-Colell et al. (1995), p. 56).

[10] The results presented here are untouched by this matter; simply the signs of C and E change.

envelope theorem, which yields using (2.10),

$$\partial_C v(\mathbf{p}, q, y - C) = -\partial_y v(\mathbf{p}, q, y - C) < 0.$$

The existence of $E(\mathbf{p}, y)$ follows from a similar analysis of (2.12). Note first,

$$\lim_{\lambda \to \infty} v(\mathbf{p}, q^0, y + \lambda) \geq v(\mathbf{p}, q^1, y) \geq v(\mathbf{p}, q^0, y + 0).$$

since $\left\{u(\mathbf{x}, q^1) : x \in \mathcal{X}\right\}$ is compact and therefore bounded. Hence, a solution to (2.12) exists and must be unique by virtue of the strict monotonicity of $v(\mathbf{p}, q^0, y)$ in y.

Let $C(\mathbf{p}, y) \in [0, y]$ be the unique solution of (2.11) and set $\hat{y} = y - C(\mathbf{p}, y)$. Then the willingness to accept, $\hat{E}$, at the reduced budget $\hat{y} \leq y$ can be determined using (2.12),

$$v(\mathbf{p}, q^1, \hat{y}) = v(\mathbf{p}, q^0, \hat{y} + \hat{E}).$$

Equivalently stated,

$$v(\mathbf{p}, q^1, y - C(\mathbf{p}, y)) = v(\mathbf{p}, q^0, y - C(\mathbf{p}, y) + \hat{E}),$$

so that (2.11), together with the injectivity of $v(\mathbf{p}, q^0, \cdot)$ (since $\partial_y v(\mathbf{p}, q^0, y) > 0$), implies $\hat{E} = C$. In other words $E(\mathbf{p}, y - C(\mathbf{p}, y)) = C(\mathbf{p}, y)$ for all $y \geq 0$. Differentiating the last identity with respect to y, we obtain

$$\partial_y E(\mathbf{p}, y - C(\mathbf{p}, y)) = \frac{\partial_y C(\mathbf{p}, y)}{1 - \partial_y C(\mathbf{p}, y)}. \tag{2.13}$$

Note, by differentiating (2.11) with respect to y and using the envelope theorem we find

$$0 < \partial_y v(\mathbf{p}, q^0, y) = (1 - \partial_y C(\mathbf{p}, y)) \partial_y v(\mathbf{p}, q^1, y - C(\mathbf{p}, y)),$$

so that $\partial_y C(\mathbf{p}, y) < 1$ for all $y \geq 0$. Hence, the relation $\hat{y} = y - C(\mathbf{p}, y)$ is strictly monotonically increasing in y and we define its inverse by $f(\mathbf{p}, \hat{y}) = y$. Naturally f is continuous and strictly increasing and with (2.13) it is $\partial_y E(\mathbf{p}, \hat{y}) = \partial_y C(\mathbf{p}, f(\hat{y}))/(1 - \partial_y C(\mathbf{p}, f(\mathbf{p}, \hat{y})))$ for all $\hat{y} \geq 0$. Since $E(\mathbf{p}, y)$ can be written, using the fundamental theorem of calculus[11] as

$$E(\mathbf{p}, y) = E(\mathbf{p}, y - C(\mathbf{p}, y)) + \int_{y - C(\mathbf{p}, y)}^{y} \partial_y E(\mathbf{p}, \xi)\, d\xi,$$

we obtain the main result by a simple change of variables.

[11] See Rudin (1976), p. 134.

THEOREM 2.1 *Let $(\mathbf{p}, y) \geq 0$. The willingness to accept is related to the willingness to pay by*

$$E(\mathbf{p}, y) = C(\mathbf{p}, y) + \int_0^{C(\mathbf{p},y)} \frac{\partial_y C(\mathbf{p}, f(\mathbf{p}, \xi))}{1 - \partial_y C(\mathbf{p}, f(\mathbf{p}, \xi))} \, d\xi, \tag{2.14}$$

where $f : \mathbb{R}_+^{n+1} \to \mathbb{R}_+$, with $f(\mathbf{p}, \xi) = \{y : \xi = y - C(\mathbf{p}, y)\}$, is a continuous single-valued function.

As a direct consequence of the first mean-value theorem in integral calculus we obtain the following corollary.

COROLLARY 2.1 *There exists a constant $\mu(\mathbf{p}, y) \in [m, M]$, such that*

$$E(\mathbf{p}, y) = (1 + \mu(y))C(\mathbf{p}, y),$$

where $m = m(\mathbf{p}, y)$ and $M = M(\mathbf{p}, y)$ are the infimum and supremum of the set

$$L(\mathbf{p}, y) = \left\{ \frac{\partial_y C(\mathbf{p}, f(\mathbf{p}, \xi))}{1 - \partial_y C(\mathbf{p}, f(\mathbf{p}, \xi))} : \xi \in [0, C(\mathbf{p}, y)] \right\}$$

respectively.

COROLLARY 2.2 *(i) If $E(\mathbf{p}, y)$ is monotonically increasing (decreasing) in y, then $E(\mathbf{p}, y) \geq (\leq)\ C(\mathbf{p}, y)$. (ii) If $C(\mathbf{p}, y)$ is increasing (decreasing) in y, then $C(\mathbf{p}, y) \geq\ (\leq)\ E(\mathbf{p}, y)$.*

The dual relation to (2.14), formulated below, expressing $C(\mathbf{p}, y)$ as a function of $E(\mathbf{p}, y)$ can be obtained in a manner completely analogous to Theorem 2.1.

THEOREM 2.2 *The willingness to pay is related to the willingness to accept by*

$$C(\mathbf{p}, y) = E(\mathbf{p}, y) - \int_0^{E(\mathbf{p},y)} \frac{\partial_y E(\mathbf{p}, g(\mathbf{p}, \xi))}{1 + \partial_y E(\mathbf{p}, g(\mathbf{p}, \xi))} \, d\xi, \tag{2.15}$$

where $g : \mathbb{R}_+^{n+1} \to \mathbb{R}_+$, with $g(\mathbf{p}, \xi) = \{y : \xi = y + E(\mathbf{p}, y)\}$, is a continuous single-valued function.

The obtained relation between the equivalent and compensating variation directly relates WTA and WTP, so that from a normative viewpoint the question of which one is greater is resolved by determining the sign of the integral in either equation (2.14) or (2.15). In addition, the difference between the two can be bounded, if the concrete problem at hand allows specifying limits on the slope of either C or E. Another advantage is that the exact relation can help avoid solving the dual problem to (2.9), when the sole objective is to obtain a bound on the difference or obtain one from the other. In addition, equations (2.14) and (2.15) may be useful as an alternative way of measuring the difference between WTA and WTP by estimating the slope of either compensating or equivalent variation over the relevant interval.

2.4 References

BLACKWELL, D. 1953. Equivalent comparison of experiments. *Annals of Mathematical Statistics* **24**(2) 265–272.

BOHNENBLUST, H., SHAPLEY, L., SHERMAN, S. 1949. Reconnaissance in game theory. *Research memorandum RM-208*, RAND Corporation.

DEGROOT, M. 1962. Uncertainty, information, and sequential experiments, *Annals of Mathematical Statistics* **33**(2) 404–419.

HANEMANN, W.M. 1991. Willingness to pay and willingness to accept: how much can they differ? *American Economic Review* **81**(3) 635–647.

HICKS, J. 1939. *Value and Capital.* Clarendon Press, Oxford, UK.

KAHNEMAN, D., KNETSCH, J.L., THALER, R.H. 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives* **5**(1) 193–206.

MAS-COLELL, A., WHINSTON, M.D., GREEN, J.R. 1995. *Microeconomic Theory.* Oxford University Press, Oxford, UK.

RANDALL, A., STOLL, J.R. 1980. Consumer's surplus in the commodity space. *American Economic Review* **70**(3) 449–455.

RUDIN, W. 1976. *Principles of Mathematical Analysis.* McGraw-Hill, New York, NY.

WEBER, T.A. 2003. An exact relation between willingness to pay and willingness to accept. *Economics Letters*, to appear.

WILLIG, R.D. 1976. Consumer's surplus without apology. *American Economic Review* **66**(4) 589–597.

If we lose our Money, it gives us some Concern.
If we are cheated or robb'd of it, we are angry:
But Money lost may be found;
What we are robb'd of may be restored:
The Treasure of Time once lost, can never be recovered;
Yet we squander it as tho' 'twere nothing worth,
Or we had no Use for it.
— Benjamin Franklin

Chapter 3

Information Transfer and Strategic Manipulation

3.1 Introduction

We determine the sale value of information in an investment decision wherein payment for the information may occur both before and after investment returns are realized. The demand for this information comes from an investor who can allocate her finite wealth freely between an idiosyncratic investment opportunity of random return and a risk-free asset with zero return. More specifically, the risk-averse investor holds an exclusive nontransferable option on the risky investment opportunity, which one may think of as a privately-financed startup or a non-publicly-traded position in a productive asset whose absolute return is proportional to the amount invested in it and whose expected return is positive but subject to random shocks. To support her decision of how to divide her wealth between this risky investment opportunity and a zero-return risk-free asset, the investor can procure an informative signal about the actual return from an outside risk-neutral information seller.

In the economics literature (e.g., Lawrence 1999), the value of information that supports a decision is defined as the simple difference between the payoffs from the investor-optimal actions chosen with and without the information. While this common definition emphasizes the *substitution* effect of information – producing economic benefit to a decision maker who changes her actions as a result of acquiring it – it neglects the *wealth* effects that the payment for the information has upon the budget set and thereby on determining or constraining the feasible investment actions. Further, the wealth effect depends not only on the magnitude of this payment but also its timing; ex-ante payments made to the information seller cannot also be invested, whereas ex-post payment obligations may not be met if poor investment outcomes occur due to limited liability.

In this paper, we examine the time and use value of imperfect information in such an investment decision. The contractually agreed monetary transfer for a signal from the investor (the agent) to the seller (the principal) of the information, which determines the agent's investable net wealth, can take place *before* and/or *after* the realization of the uncertainty. First and foremost, information-sale contracts which do not require payment until after returns are realized offer a potential Pareto improvement over fully prepaid arrangements, as the full amount of the investor's wealth is available for deployment in favorable signal states which both the investor and the information seller value. Rather than simply dividing a fixed payment into two time periods, the information seller can rebalance his prices between ex ante and ex post to collect a higher amount (in expectation) after the returns are realized, leading to a higher expected revenue than insisting on collecting everything up front. Second, the seller generally suffers from default risk for a positive ex-post price component: if part of the payment is to occur after the return realization, then a situation may arise in which (due to poor investment results) the investor's remaining wealth does not cover the agreed payment to the information seller, resulting in partial or complete default. Under our assumption of limited liability, the seller's recovery is limited to the investor's remaining wealth. Third, the investor's choices may be altered by the strategic (ab-)use of this default possibility: payment deferral, combined with the implicit downside protection afforded by limited liability, may induce the investor to take higher risks than she would otherwise under an ex-ante payment plan, even when starting with the same amount of investable wealth. Such incentive effects resulting from unsecured debt have been identified by Jensen and Meckling (1976) as a source of moral hazard and earlier by Fama and Miller (1972) as a potential cause for conflict between the interests of debtholders and shareholders of private enterprises. Particularly under full payment deferral, although the information seller might well be said to still 'own' the information (having 'lent' it to the investor but not having yet received payment), the investor effectively has assumed 'control' of the information (Berle and Means 1932); she can use this information for any purpose, including the unintended and damaging purpose of crafting an investment policy so risky that it reduces the expected amount collected by the seller.

We argue that, in addition to optimally balancing the ex-ante and ex-post portions of the payment schedule, it is sometimes revenue-increasing for an information seller (acting as a principal in this agency problem) to garble the signal so as to induce the agent to take less risky actions when these actions increase the chances of payment default. The principal is able to reduce agency cost, and increase expected revenues, by “damaging” his information product in the sense of Deneckere and McAfee (1996). Unlike the situation described by Gjesdal (1982), who shows that a signal is strictly preferred to a nonsufficient garbling of it for *all* agency problems of a certain form, we obtain the inverse conclusion for situations in which the principal can influence the informational environment of the agent. We further suggest that for contracts with an ex-post payment provision including a portion of *unsecured* debt, the principal may be able to increase his expected surplus by properly garbling the signal.

The sale and purchase of information in an investment setting has a fifteen-year history in the financial economics literature. The value of information for agents that trade on a common market has been examined in a series of articles by Admati and Pfleiderer (1986,1987,1990), who provide conditions under which an insider would either trade and sell inside information, trade without disclosing any information, or exclusively sell information without trading on the asset market. Also in order to preserve the ability to trade and to be able to sell the information to many traders, the insider has an incentive to distort the signal by adding noise. Allen (1990) builds on this work and establishes the link to financial intermediation in a model where all agents have utility functions with constant absolute risk aversion (for which the parameters are sometimes not known to all players). The seller in this model cannot capture the full value of the information due to the reliability problem (any would-be seller can claim to have information), and intermediaries may be able to capture some of the residual value. Rather than examine a multibuyer market setting, we focus on the contract structure where there is only one investor that seeks information helping her to decide what proportion of her investable wealth she allocates to an idiosyncratic risky asset. We assume that the risk-neutral information seller can diversify the induced payoff default risk (applicable only if he chooses an ex-post payment contract) completely. Further we suppose that initially both the investor and the information seller have identical rational expectations about the return distribution of the risky asset. We then examine a number of different contracting forms for the sale of information (by, for instance, admitting the possibility of ex-post payment, which introduces moral hazard as noted above), absent from the market-interaction mechanisms considered by Admati-Pfleiderer and Allen.

We also provide a quite different structure than in the previous work on investment advice (e.g., Kihlstrom (1988), in which the investor would like to both control the effort of the security analyst (moral hazard) and to get the analyst to reveal whether the information obtained is reliable (adverse selection). The optimal payment contracts from the investor’s viewpoint (the “principal” in this context) make the analyst’s fee dependent on both realized and predicted return (to deal with the moral-hazard problem) and introduce risk for the analyst (to discourage an unreliable analyst from accepting the contract, which can achieve a separating equilibrium in the adverse selection scenario). In our model, the investor is effectively the agent, but knows the actual signal quality offered by the information seller. Moral hazard still arises because the investor, by being offered the opportunity to gamble with the deferred fees owed the security analyst, has an incentive to invest more aggressively than is in the interest of the seller (the principal).

It should be noted at this point that we relate the investor's willingness to pay (WTP) defined as Hicks' (1939) *compensating variation* that moves the investor back to her maximum expected utility achievable without information. In many models of information acquisition (Athey and Levin 2000 is a recent example) the funds used to purchase information are implicitly segregated from the wealth used to generate utility, effectively assuming additive separability in utility from the "wealth to be invested" and "wealth earmarked for purchasing information."[1] In these models the cash disbursed to acquire information does not constrain the feasible set of actions, nor does economizing on information or deferring its purchase settlement yield any investment benefit. Our treatment here is inclusive in that we explicitly take into account the wealth effects induced by the contraction of the budget set by the information expenditure, as introduced by Kihlstrom (1974) under a different setting.

The outline of this paper is as follows. In Section 3.2 we specify the basic investment model and examine the investor's asset allocation decision as a function of the prices charged by the information seller ex ante and ex post. We show that the information seller's optimal pricing problem always possesses a solution, and then characterize solution properties in certain special cases. As a benchmark we examine the sale of perfect information, to which we provide a general solution.[2] In Section 3.3 we then more closely examine the structure of optimal payment timing, if it is to be used both to take advantage of the co-investment opportunity and to mitigate the economic damage caused by investor moral hazard. We determine the optimal investment action from the information seller's point of view, which trades off gains from speculative co-investment and losses from default risk. A "moral hazard state" is then

[1]This applies to a risk-neutral agent. For risk-averse agents, the cost of an informative signal is typically modeled as a "utility cost," which is subtracted from the agent's full-information utility, *as if* the amount for the information was paid out of a separate budget. See for instance Raiffa and Schlaifer (1961), Marschak and Miyasawa (1968), Kihlstrom (1976) or Athey (2000).

[2]The value of perfect information has been examined earlier by Thon and Thorlund-Petersen (1993), but without using compensating variation as the relevant welfare measure and without varying the contractual form under which information can be procured.

defined to characterize signal realizations that induce the investor to deliberately overshoot these seller-optimal investment actions, taking advantage of her limited liability for the ex-post payments portion to capture the incremental upside from favorable outcomes without suffering losses from unfavorable ones. Section 3.4 introduces an additional tool to influence the noncontractible investor action: garbling of the information seller's signal. By systematically misreporting favorable signal outcomes, the seller can put a check on investor aggressiveness to limit moral hazard, resulting in lower default risk and therefore increased revenues. Section 3.5 concludes the paper with a description of research in progress and opportunities for future theory development.

3.2 The Model

3.2.1 The Investor's Asset Allocation Decision

The investor can allocate her wealth w between two assets: a safe asset and a risky asset. Without loss of generality we assume the return of the safe asset to be zero. The return of the risky asset is described by the realization of a random variable, $\tilde{r}$, with continuous probability density function $h : \mathbb{R} \to \mathbb{R}_+$ with support $\mathcal{R} = [\underline{r}, \bar{r}]$ where $-1 \leq \underline{r} < \bar{r}$. We assume that for simplicity the initial beliefs of both the investor and the information seller on the distribution of the returns are correct and thus given by the probability density function h on $\mathcal{R}$. Moreover, the ex-ante expected return[3] of the risky asset is assumed to be positive,

$$E\tilde{r} = \int_{\mathcal{R}} h(r) r dr > 0.$$

Let $u : \mathbb{R} \to \mathbb{R}$ be the strictly increasing and differentiably strictly concave utility function of the investor as a function of wealth, such that without loss of generality $u(0) = 0$. The investor has the option to buy from a risk-neutral monopolist information seller a partially informative signal $\tilde{s} \in \mathcal{S}$, where the set of possible signal realizations, $\mathcal{S}$, is Lebesgue-measurable. The characteristics of the signal are common knowledge: its continuous probability density conditional on $\tilde{r} = r$ is given by $f(\cdot|r)$. If the investor receives access to a realization, s, of this signal, she updates her initial beliefs according to Bayes' rule, obtaining the posterior return density

$$g(r|s) = \frac{f(s|r)h(r)}{\displaystyle\int_{\mathcal{R}} f(s|r)h(r)dr}.$$

[3]We refer to "ex ante" and "ex post" *only* with respect to the return realizations. Any signal is observed during an intermediate period between contracting and investing.

To ensure integrability we assume that the signal realizations and their ordering in $\mathcal{S}$ is such that the posterior density $g(r|\cdot)$ is a measurable function for any fixed $r \in \mathcal{R}$.[4] Based on the information available to her, the investor chooses an action $a \in [0, 1]$, namely the proportion of her investable wealth.[5] After the return realization has taken place, the investor's utility payoff π depends on her initial wealth w, her chosen investment action a, the realized return r, and the price vector $p = (p_1, p_2)$ that she agreed to pay to the information seller (of which p_1 is charged ex ante, and p_2 ex post under limited liability). Hence

$$\pi(a; p, w, r) = u\left([(w - p_1)(1 + ar) - p_2]_+\right) \geq 0,$$

where the $[\cdot]_+$-operator is defined for any real ξ by $[\xi]_+ = \max\{0, \xi\}$. Note that $\pi(a; p, w, r)$ vanishes, if for $0 < a \leq 1, p_1 < w$ the return r does not exceed the critical value

$$r_c(a; p, w) = \left(-1 + \frac{p_2}{w - p_1}\right)/a. \tag{3.1}$$

Given these payoff prospects, the investor's optimal action a_s^* after observing a signal realization $s \in \mathcal{S}$ is given by the (possibly set-valued) maximizer

$$a_s^*(p, w) = \arg\max_{a \in [0,1]} \pi_s(a; p, w), \tag{IC}$$

where π_s represents the expected utility payoff, $\pi_s(a; p, w) = E\left[\pi(a; p, w, \tilde{r})|\, s\right]$, or using (3.1) equivalently

$$\pi_s(a; p, w) = \int_{r_c(a;p,w)}^{\bar{r}} u\left((w - p_1)(1 + ar) - p_2\right) g(r|s)dr. \tag{3.2}$$

The indirect utility from observing the realization s is thus $v_s(p, w) = \pi_s(a_s^*(p, w); p, w)$. Similarly, the no-information utility is given by $v_0(w) = \max_{a \in [0,1]} E\left[\pi(a; 0, w, \tilde{r})\right]$. If there exists an interior maximizer $a_s^o \in a_s^* \cap (0, 1)$, it must satisfy the following first-order necessary optimality condition:[6]

$$\frac{\partial_a \pi_s(a)}{w - p_1} = \int_{r_c(a)}^{\bar{r}} u'\left((w - p_1)(1 + ar) - p_2\right) g(r|s) r\, dr = 0, \tag{3.3}$$

where naturally $p_1 < w$ (cf. Proposition 3.1, part (i)). Using the Leibniz Rule the associated second-order condition can be written as

$$\frac{\partial_a^2 \pi_s(a)}{(w - p_1)^2} = \int_{r_c(a)}^{\bar{r}} u''\left((w - p_1)(1 + ar) - p_2\right) g(r|s) r^2\, dr + \frac{u'(0) g\left(r_c(a)|s\right) r_c^2(a)}{(w - p_1)a} < 0. \tag{3.4}$$

[4]The choice of the particular Lebesgue-probability measure ds naturally depends on the cardinality of $\mathcal{S}$ (finite/countable/uncountable). In addition, there is no need for $\mathcal{S}$ to be convex.

[5]The information seller may or may not demand prepayment, which directly influences the remaining wealth. If the investor has to pay ex ante (i.e., before the returns of the investment are realized) for example, then her investable wealth is reduced by the price for the information.

[6]Here and in the following discussion we drop the notation's dependence on w and p where not necessary.

In the presence of default risk (i.e., if $r_c(a_s^o) > \underline{r}$) nonconcavities can arise naturally due to the second term in (3.4). This term vanishes as long as $r_c(a) \notin (\underline{r}, \bar{r})$, which is generally the case for small positive a. In that case the utility payoff π_s is strictly concave in a and the investor's problem (IC) has a unique solution. Thus, for a very risk averse investor (with all a_s^o small enough), the presence of limited liability has no effect on her investment action. This is true as long as investing all her wealth in the risky asset and thereby "going for broke" cannot make her better off in expectation. On the other hand, if default risk is significant, the second term in (3.4) can cause multiple local extrema of the investor's expected utility π_s to appear, and thus a_s^* may contain several elements.

Before considering the information seller's problem below, it is necessary to fully understand the investor's reaction to price changes. We therefore examine the comparative statics of the investor's optimal action set a_s^* in (IC) under any signal realization $s \in \mathcal{S}$.

Proposition 3.1 (Comparative Statics) *Let $s \in S$ be fixed and a_s^* be defined by (IC).*

(i) At any interior maximizer $a_s^o \in a_s^* \cap (0,1)$, the investor is indifferent between changes in p_1 and changes in p_2, i.e., $D_{p_1}\pi_s\left(a_s^o(p,w);p,w\right) = D_{p_2}\pi_s\left(a_s^o(p,w);p,w\right)$.

(ii) For any constant policy $\bar{a} \in (0,1]$, the investor strictly prefers an increase in p_2 (decrease in p_1) to an increase in p_1 (decrease in p_2), iff $\partial_a \pi_s(\bar{a}) > 0$. In particular, if $a_s^* = 1$ (singleton), then $D_{p_1}\pi_s(1;p,w) < D_{p_2}\pi_s(1;p,w)$. For $\bar{a} = 0$, the investor is indifferent between changes in p_1 and p_2.

(iii) The response of an optimal interior policy $a_s^o \in a_s^* \cap (0,1)$ to price changes is given by

$$\partial_{p_1} a_s^o = -\frac{\int_{r_c(a_s^o)}^{\bar{r}} \rho(\cdot)u'(\cdot)g(r|s)r\,dr}{\partial_a^2 \pi_s(a_s^o)} + p_2 \frac{\partial_{p_2} a_s^o}{w - p_1}, \tag{3.5}$$

and

$$\partial_{p_2} a_s^o = \frac{w - p_1}{\partial_a^2 \pi_s(a_s^o)} \int_{r_c(a_s^o)}^{\bar{r}} u''(\cdot)g(r|s)r\,dr + \frac{u'(0)g(r_c(a_s^o)|s)r_c(a_s^o)}{a_s^o \partial_a^2 \pi_s(a_s^o)}, \tag{3.6}$$

where $\rho(\cdot) = -u''(\cdot)/u'(\cdot)$ is the Arrow-Pratt coefficient of relative risk aversion and $(\cdot)$ abbreviates $((w-p_1)(1+a_s^o r) - p_2)$.

Proof: See the Appendix.

Part (i) of this proposition states that as long as the optimal investment action is not at the boundary (i.e., either zero or one), the investor is indifferent between price changes in p_1 and p_2. This finding is very general and does not depend on the curvature of the utility function or on the presence of limited liability as long as a_s^o is a *strict* maximizer, so that no "transitions" between neighboring maximizers occur resulting from small price changes. As stated in the second part of the proposition, the investor is generally not indifferent between changes in p_1 and p_2 for any nonzero *constant* investment policy $\bar{a}$. In particular if $a_s^* = 1$ (a singleton), then the investor strictly prefers a unit increase in p_2 to an increase in p_1. As we will detail below (cf. Proposition 3.4), this immediately implies an optimal price vector for the sale of *perfect* information: charge the lowest admissible amount p_1 (possibly negative) ex ante and then as much ex post (via p_2) as is permissible until *either* the investor's individual-rationality constraint binds *or* she becomes indifferent to choosing $a_s^* = 1$ and deliberately defaulting. Part (iii) details the behavior of interior maximizers with respect to price changes. It turns out that a_s^o increases with p_2 when news is "good" (i.e., s such that most of $g(\cdot|s)$ is concentrated in high returns) and the limited-liability constraint is present but not too restrictive (i.e., $\underline{r} < r_c(a_s^o) < 0$). If a_s^o increases in p_2, then a_s^o increases even more (less) in p_1 provided the investor's utility is of increasing (decreasing) relative risk aversion.

3.2.2 The Information Seller's Optimal Pricing Policy

We examine now the general problem where the information seller can choose a division of ex-ante and ex-post payments that maximizes his revenues. Thereby we assume that negative prices *per se* are admissible, but that those need to be bounded from below, as the seller's liquidity is finite. Negative ex-ante prices thus allow for the possibility of speculative co-investment. We furthermore assume that the investor's actions are not contractible since they are either not observable or not verifiable by a benevolent court of law.[7] The investment opportunity is idiosyncratic to the investor and he is assumed to be the only agent endowed with the inalienable option to invest in this venture. The information seller can be seen in this light as a trusted investment advisor or a consulting company that can generate an informative signal about the future returns of the venture, but does not have the option to invest directly. To collect as much as possible while still mitigating the possible default of the investor, the seller can spread payments for the information over time. Ex-ante payments are default-free, but naturally bounded from above by w. In addition, by charging a positive price ex-ante, the risk-neutral information seller foregoes a portion of the return prospects generated by speculative co-investment. However, to focus on the payment *timing* under a limited-liability contract, we assume – as indicated before – that the payment cannot be made contingent on

[7] We also exclude contractibility on the signal realizations and on the investor's ex-post wealth except for the zero-threshold that determines her limited liability.

the return or on the signal's realizations. Formally, the information seller seeks a price vector $p^* \in \mathcal{P} \subseteq \mathbb{R}^2$ (where $\mathcal{P}$ is convex, contains the origin and is bounded from below), such that

$$p^* \in \arg\max_{p\in\mathcal{P}} \left\{ p_1 + E_s\left[E_r\left[\min\{p_2, (w-p_1)(1+a^*_{\tilde{s}}(p,w)\tilde{r})\} | \,\tilde{s}\right]\right]\right\} \tag{3.7}$$

subject to the investor's incentive-compatibility constraint (IC) as well as her individual-rationality constraint,

$$v(p^*, w) \equiv \int_{\mathcal{S}} v_s(p^*, w)\sigma(s)ds \geq v_0(w), \tag{IR}$$

where $\sigma(\cdot)$ is the marginal probability density of the signal realizations,

$$\sigma(s) = \int_{\mathcal{R}} f(s|r)h(r)dr.$$

The existence of a solution to this principal-agent problem cannot generally be taken for granted,[8] and Grossman and Hart's (1983) existence result is not applicable here, since the agent's (i.e., the investor's) utility is *not* multiplicatively separable in her actions and the principal's (i.e., the information seller's) price.

Proposition 3.2 (Existence) *There exists a price vector $p^* \in \mathcal{P}$ that solves the seller's problem (3.7), subject to (IC) and (IR).*

Proof: See the Appendix.

The existence is mainly a consequence of the continuity of the underlying functions and their possessing by the Weierstrass' theorem extrema on compact sets. Essential here is the assumption that the set $\mathcal{P}$ of possible prices is bounded from below (i.e., existence of a seller liquidity constraint). This implies by the investor's individual-rationality constraint (IR) the existence of a bounded solution $p^* \in \mathcal{P}$. Having guaranteed the existence of a solution, let us now consider solving the seller's problem. The seller's expected revenues in (3.7) can be rewritten as

$$R(p,w) = p_1 + \int_{\mathcal{S}} \left\{ p_2 \int_{r_c}^{\bar{r}} g(r|s)dr + (w-p_1)\int_{\underline{r}}^{r_c} (1+a^*_s r)g(r|s)dr \right\} \sigma(s)ds. \tag{3.8}$$

The following proposition characterizes some general properties of revenue-maximizing price vectors.

Proposition 3.3 (Properties of Solutions to the Seller's Problem) *Assume that the price domain $\mathcal{P}$ is of the form $[\underline{p}, w] \times [\underline{p}, \infty)$ with $\underline{p} \leq 0$, and let a^*_s be defined by (IC). A solution $p^* = (p^*_1, p^*_2)$ to the seller's pricing problem (3.7) subject to (IC) and (IR) is such that*

[8] An interesting and disturbing nonexistence example has been given by Mirrlees (1974).

(i) $p_1^* < w$, *and* $p_k^* \in [\underline{p}, 0] \Rightarrow p_{3-k}^* \geq 0$ *for* $k = 1, 2$.

(ii) *If* $p^* \in \mathrm{int}(\mathcal{P})$ *is a strict maximizer and the investor's actions are interior (or zero),* $a_s^* \subset [0,1)$, *for all* $s \in \mathcal{S}$, *then necessarily* $\partial_{p_1} R = \partial_{p_2} R$ *and*

$$\int_{\mathcal{S}} \left\{ \left(\partial_{p_1} a_s^* - \partial_{p_2} a_s^* \right) \int_{\underline{r}}^{r_c} g(r|s) dr - \frac{a_s^*}{w - p_1} \int_{\underline{r}}^{r_c} g(r|s) r \, dr \right\} \sigma(s) ds = 0 \tag{3.9}$$

holds at $p = p^*$, *where for* $a_s^o \in a_s^* \cap (0,1)$:

$$\left(\partial_{p_1} - \partial_{p_2} \right) a_s^o = - \frac{\int_{r_c(a_s^o)}^{\bar{r}} \rho(\cdot) u'(\cdot) g(r|s) r \, dr}{\partial_a^2 \pi_s(a_s^o)} + r_c(a_s^o) \partial_{p_2} a_s^o. \tag{3.10}$$

On the other hand, if $a_s^* = 1$ *for some* s, *so that* $\mathcal{S}_+^* = \{ s \in \mathcal{S} : a_s^*(p^*, w) = 1 \}$ *is of positive measure* and *the investor's individual-rationality constraint (IR) is binding, then necessarily*

$$\frac{\partial_{p_1} R(p^*, w)}{\partial_{p_2} R(p^*, w)} = 1 + \frac{\int_{\mathcal{S}_+^*} \left\{ \int_{\frac{p_2^*}{w - p_1^*} - 1}^{\bar{r}} u'(\cdot) g(r|s) r \, dr \right\} \sigma(s) ds}{\int_{\mathcal{S}} \left\{ \int_{\frac{p_2^*}{w - p_1^*} - 1}^{\bar{r}} u'(\cdot) g(r|s) dr \right\} \sigma(s) ds}. \tag{3.11}$$

(iii) *If there is no default risk at* p^*, *i.e., if* $\max \{ -1, r_c(a_s^*(p^*, w); p^*, w) \} \leq \underline{r}$ *for all* $s \in \mathcal{S}$, *then a price vector* $p^* \in \partial \mathcal{P}$ *such that* $p_1^* = \underline{p}$ *and*

$$\int_{\mathcal{S}} v_s \left((\underline{p}, p_2^*), w \right) \sigma(s) ds = v_0(w) \tag{3.12}$$

is optimal.

Proof: See the Appendix.

The first part of Proposition 3.3 states that it is never optimal for the seller to charge all of the investor's wealth up front, since this essentially destroys the investor's idiosyncratic investment opportunity. The only way the investor would part with all his wealth before investing it would be for the seller to promise a compensation in the second period equal to the certainty equivalent of the expected no-information utility $v_0(w)$, which is necessarily greater than w; thus charging $p_1 = w$ is a negative revenue proposition for the seller. Also, if the seller subsidizes in one of the two payment periods, he naturally must charge a positive price in the other period. Part (ii) characterizes necessary optimality conditions for a solution

p^* in the interior of $\mathcal{P}$. As long as the optimal investor action given the price chosen by the information seller is interior to her action set (i.e., lies in the open interval $(0, 1)$), the slopes of the seller's expected revenue function with respect to p_1 and p_2 must be equal at the optimum. In the event of a corner solution to (IC) (with $a_s^* = 1$) for "good" signal realizations s, the difference between $\partial_{p_1} R$ and $\partial_{p_2} R$ at the optimum is negative, so that the information seller prefers to charge ex post, as this also corresponds to the investor's preferences according to Proposition 3.1, part (ii). The last part of Proposition 3.3 provides the optimal solution to the seller's problem in the special case when the information transferred is not too valuable and the lowest achievable return is $\underline{r}$ is above the critical return r_c below which the investor at her optimal policy may default for some signal-return realizations. Without default, a maximal ex-ante subsidy combined with an ex-post charge that extracts all of the investor's surplus (making (IR) binding) will maximize seller expected revenue.

In the next two sections, we will further examine the optimal pricing under potential default. Section 3.3 deals with mitigating the moral hazard induced by the investor's limited liability by payment timing and Section 3.4 examines the possibility of garbling information to deter the investor from too aggressively exploiting her "non-payment" option. Before turning to these issues of optimal timing and revenue-increasing information garbling, we briefly consider the case of selling perfect information, as a benchmark. Interestingly, the optimal pricing policy for perfect information is essentially the same as the one in Proposition 3.3 (iii), even though we do not explicitly exclude default.

3.2.3 Benchmark: Selling Perfect Information

Consider the situation in which the information seller is in a position to offer perfect information. Clearly, the revenues from the sale of perfect information constitute a natural upper bound for maximum extractable revenues. In addition, perfect information maximizes the investment return for the investor and also the seller's extractable revenues.[9] Without loss of generality perfect information about the investment returns can be communicated by setting $\mathcal{S} = \mathcal{R}$ and $s = r$. The investor's optimal policy given a signal realization $s = r$ is thus $a_r^* = [\mathrm{sgn}(r)]_+ \in \{0, 1\}$. Whenever the investment return r is positive, the investor invests her full wealth (possibly augmented by a subsidy via p_1), otherwise she invests nothing at all.[10] Thus, under perfect information any investor behaves exactly like a risk-neutral investor. In

[9] In particular, garbling perfect information is never in the seller's interest as Proposition 3.6 (iii) shows.

[10] Strictly speaking, in the case when $\tilde{r} = 0$ we have that $a_s^* = [0, 1]$ which is however of no importance for our discussion.

particular, we see that the investor's optimal policy a_s^* does not depend on the price vector charged by the seller, it is therefore constant with respect to changes in p. Thus, by part (ii) of Proposition 3.1 we obtain that the investor's utility is reduced less by increases in p_2 than by increases in p_1, as long as (IR) is not binding. In addition, the information seller is indifferent between revenues from p_1 and p_2; as long as the investor follows her optimal strategy given above, no default can occur if $p_1 + p_2 \leq w$. But this in turn implies – in view of making (IR) the least restrictive possible and maximizing the amount to be co-invested in this now riskless opportunity – that the information seller charges a minimum possible p_1, and then the maximal individually rational amount in p_2. If on the other hand default is a possibility with $p_1 + p_2 > w$, then the seller's revenues are given by

$$R(p, w) = p_1 + p_2 + (w - p_1 - p_2)H(r_c) + (w - p_1) \int_0^{r_c} h(r) r \, dr,$$

where $r_c = -1 + p_2/(w - p_1) > 0$ and $H(\cdot)$ is the probability distribution function associated with the density $h(\cdot)$. Similar to part (iii) of Proposition 3.3, we can then compute the difference between the slopes of R with respect to p_1 and p_2,

$$(\partial_{p_1} - \partial_{p_2}) \, R(p, w) = -\int_0^{r_c} h(r) r \, dr < 0,$$

so that with default the seller's revenues increase faster by charging ex post than by charging ex ante. Hence, we may conclude that the optimal price to any type of investor (essentially independent of the curvature of u) for the sale of perfect information is given by a maximal upfront subsidy of $p_1^* = \underline{p} \leq 0$ and an ex-post charge of p_2^*, such that (IR) is binding,

$$v_0(w) = \begin{cases} -H(0) u \left(w - \underline{p} - p_2^* \right) + \displaystyle\int_0^{\bar{r}} u\left(\cdot\right) h(r) dr, & \text{if } p_2^* < w - \underline{p}, \\ \displaystyle\int_{\frac{p_2^*}{w - \underline{p}} - 1}^{\bar{r}} u\left(\cdot\right) h(r) dr, & \text{otherwise}, \end{cases} \tag{3.13}$$

where $(\cdot)$ abbreviates $\big((w - \underline{p})(1 + r) - p_2^*\big)$. A seller of perfect information has thus an incentive to use speculative co-investment to the largest extent possible, i.e., up to his liquidity constraint. The following proposition summarizes our results on the sale of perfect information.

Proposition 3.4 (Optimal Pricing of Perfect Information) *Consider the sale of perfect information about the future return r. Given an admissible price domain of $\mathcal{P} = [\underline{p}, w] \times [\underline{p}, \infty)$, an optimal price vector is given by $p^* = (\underline{p}, p_2^*)$, where p_2^* is such that the investor's individual-rationality constraint (IR) (equivalent to (3.13)) is binding.*

Note that the optimal policy for the sale of perfect information specified in Proposition 3.4 is exactly the same as the no-default pricing in Propostion 3.3 (iii), even though we did not rule out default for the sale of perfect information. The key is that any risk-averse investor with perfect information effectively behaves as if risk-neutral, so that in view of the seller's speculative co-investment their incentives are aligned to the point of producing a first-best outcome.

3.3 Optimal Payment Timing

The monopoly seller has the choice not only of the amount to charge for his information but also of the form and timing that payment will take (subject to the buyer's voluntary participation and the buyer's capability to make such payments). In this section we consider the optimal payment timing schedule in terms of choosing the vector $p = (p_1, p_2)$ in the presence of investor moral hazard and associated default risk, and contrast this to the case of pure ex-ante payment that naturally eliminates any possibility for default. By deferring a portion or all of the payment, the information seller may be able to substantially increase his revenues. Investor default risk somewhat limits the seller's ability to indirectly participate in the investment opportunity: providing maximal ex-ante financing (by charging $p_1 = \underline{p} < 0$) as obtained in the previous section in the *absence* of default may prove detrimental to the information seller in its presence. This has interesting consequences for the optimal pricing under default risk: generally $p_1^* > \underline{p}$ and sometimes the investor's individual-rationality constraint (IR) will be slack at $p^* = (p_1^*, p_2^*)$.[11]

3.3.1 Pure Ex-Ante Payment

In Section 3.2, we have already provided a general optimal pricing policy when there is no possibility for the investor to default on the ex-post payment portion, p_2. If p_2 is constrained to be zero – for instance because the outcome of the investment cannot be observed by the information seller – then no investor default is possible. At the same time the information seller is unable to derive residual benefits from the information by co-investing along with the risk-averse investor.

[11]The fact that (IR) may be slack at the optimal prices can be seen as the investor obtaining an "efficiency wage" from the seller which discourages her from going for broke and overinvesting. The investor surplus is in this case an information rent that the investor can extract as a result of the noncontractibility of her (hidden) investment actions.

The resulting optimal price vector $(p_1^*, 0)$ can be implicitly determined by computing the investor's compensating variation for the use of the information,

$$\int_{\mathcal{S}} v_s\left((p_1^*, 0), w\right) \sigma(s) ds = v_0(w).$$

By varying the timing of the payment collection from ex ante to ex post, the information seller may, in addition to being able to co-invest, also increase the investor's overall WTP by allowing her to invest her full wealth. In order to deal with the resulting moral hazard we show that in general an interior payment policy is revenue-maximizing for the seller.

3.3.2 Payment Policies Under Default Risk

Before discussing optimal pricing policies under investor default, let us briefly focus on the "ideal" solution to the seller's revenue-maximization problem, as if he had complete control over the investor's investment actions:

$$\hat{a}_s = \arg\max_{a \in [0,1]} \int_{\mathcal{S}} \left\{ p_2 \int_{r_c}^{\bar{r}} g(r|s) dr + (w - p_1) \int_{\underline{r}}^{r_c} (1 + ar) g(r|s) r \, dr \right\} \sigma(s) ds. \tag{3.14}$$

This allows us to be clear about what exactly constitutes investor moral hazard. Such moral hazard creates extra default risk, which reduces the seller's expected payoff below what could be achieved if the investor had ignored the pending payment of p_2 when determining her optimal investment policy (e.g., treating the cost of information as sunk, as in the case of ex-ante payments).

PROPOSITION 3.5 *The solution to the seller-determined revenue-maximization problem (3.14) is given by*

$$\breve{a}_s = \left[0, \left(\frac{p_2}{w - p_1} - 1\right) / \underline{r}\right] \cap [0, 1], \tag{3.15}$$

for $p_1 + p_2 \leq w$, *and*

$$\breve{a}_s = \begin{cases} \left[0, \left(\frac{p_2}{w-p_1} - 1\right) / \bar{r}\right] \cap [0, 1], & \text{if } E[\tilde{r}|s] \leq 0, \\ \breve{a}_s^o, & \text{otherwise}, \end{cases} \tag{3.16}$$

for $p_1 + p_2 > w$, *where* $\breve{a}_s^o$ *is defined by*

$$\int_{\underline{r}}^{r_c(\breve{a}_s^o)} g(r|s) r \, dr = 0. \tag{3.17}$$

Proof: See the Appendix.

Definition. *If an element of the investor's action* a_s^* *is for a given signal realization* s *higher than any seller-revenue-maximizing action* $\breve{a}_s$ *in that state, then the difference,* $\max\{a_s^*\} - \max\{\breve{a}_s\}$, *is called the* moral hazard *(induced by the presence of unsecured default risk). The corresponding signal realization is termed a* moral-hazard state.

Note that all states $s \in \mathcal{S}_+^*$ in which the investor goes for broke, with $\mathcal{S}_+^*$ as defined in Proposition 3.3, are moral-hazard states. In the following section we concentrate on states exhibiting moral hazard and show that in some circumstances, it may be beneficial for the seller to garble the corresponding signal realization, if the investor's action can be reduced by enough such that the resultant reduction in default risk offsets her decrease in WTP for the garbled signal.

Let us now show an argument that leaving slack in the investor's individual rationality constraint (IR) may be optimal. For this let $H \in \mathcal{S}_+^* \neq \emptyset$ be a moral-hazard state (with associated return density $g(r|H)$) that occurs with positive probability. Then the change in the investor's propensity to invest following a change of either p_1 or p_2 is governed by $\partial_{p_1}\left(V_H^1 - V_H\right)$ and $\partial_{p_2}\left(V_H^1 - V_H\right)$ respectively, where we have set $V_H^1 = \int_{r_c(1)}^{\bar{r}} u\left((w-p_1)(1+r)-p_2\right) g(r|H)dr$ and $V_H = \int_{r_c(a_H^*)}^{\bar{r}} u\left((w-p_1)(1+a_H^* r)-p_2\right) g(r|H)dr$ to denote the investor's state-contingent utility from investing fully ($a_H^* = 1$) or at an interior optimal allocation $a_H^* \in (0,1)$ when H occurs. Note that as H is by assumption an element of $\mathcal{S}_+^*$ (so that $1 \in a_H^*$) we have that $V_H^1 \geq V_H$ at the optimal price vector p^*. Let us first compute the change of $V_H^1 - V_H$ with respect to changes in p_1:

$$\begin{aligned}
\partial_{p_1}\left(V_H^1 - V_H\right) &= -\int_{r_c(1)}^{\bar{r}} u'(\cdot)_1(1+r)g(r|H)dr + \int_{r_c(a_H)}^{\bar{r}} u'(\cdot)_H(1+a_H^* r)g(r|H)dr \\
&= -\int_{r_c(1)}^{\bar{r}} u'(\cdot)_1 r g(r|H)dr + \int_{r_c(1)}^{0} \left(u'(\cdot)_H - u'(\cdot)_1\right) g(r|H)dr \\
&\quad + \int_{r_c(a_H^*)}^{r_c(1)} u'(\cdot)_H g(r|H)dr - \int_0^{\bar{r}} \left(u'(\cdot)_H - u'(\cdot)_1\right) g(r|H)dr. \qquad (3.18)
\end{aligned}$$

Thereby the abbreviations $(\cdot)_1$ and $(\cdot)_H$ stand for $((w-p_1)(1+r)-p_2)$ and $((w-p_1)(1+a_H^* r)-p_2)$ respectively. The first and the last term of the preceding expression are negative while the other two are positive but in most cases of smaller magnitude, since generally the posterior return density $g(r|H)$ concentrates relatively more mass into the region of positive returns. The computation of the change of $V_H^1 - V_H$ with respect to changes in p_2 proceeds in an analogous manner to the above and yields:

$$\partial_{p_2}\left(V_H^1 - V_H\right) = \partial_{p_1}\left(V_H^1 - V_H\right) + \int_{r_c(1)}^{\bar{r}} u'(\cdot)_1 r g(r|H)dr \qquad (3.19)$$

an expression generally strictly larger than $\partial_{p_1}\left(V_H^1 - V_H\right)$. As a result of (3.18)–(3.19) we note that the investor's propensity to default in state H is generally higher following an increase of p_2 than following an increase in p_1. Conversely, a reduction in p_2 is a more effective deterrent from going for broke than a decrease in p_1. Hence, instead of decreasing p_1 even further, the information seller may choose to limit p_2 to a level (together with an appropriate interior p_1) that makes the investor indifferent between going for broke and investing at an interior optimal level. The consequence of this favorable switch in investor behavior towards an interior investment level generally has a *discontinuous* positive impact on the seller's revenue, so that leaving a modicum of slack in the investor's individual-rationality constraint to accomplish this shift may *improve* expected revenue in some situations. Note that if (IR) is slack at p^*, then there must be $1 \in a_s^*$ for a set of states $\mathcal{S}_+^*$ of positive measure. Otherwise, expected revenue could be strictly increased by a small increase of either p_1 or p_2 contradicting the assumption that the vector p^* was optimal. – In the next section we discuss how the information seller, by distorting his information, can extract this slack to his advantage.

3.4 Revenue-Increasing Garbling of Information

For the purpose of discussion we consider in this section signals with only two possible realizations, "high" and "low." As argued above, a revenue-maximizing payment plan (p_1^*, p_2^*) for the sale of an informative signal must account for the investor's propensity of going for broke investing all her money into the risky asset. Particularly if the ex-post component p_2^* is high, the investor's potential gains from full investment may be large whereas her downside is bounded at zero ex-post wealth because of limited liability. This misbehavior is particularly serious for the high signal state, in which the probability of a positive payoff from incremental investment is highest. Sometimes simply reducing prices will be required to induce more favorable investor behavior by making her prefer an interior investment level. As we argued above in Section 3.3, such a "preemptive" pricing policy must generally leave the investor with some surplus at the optimum set of prices (1983). In this section, we discuss how the seller can further increase her expected profit by using another technique – garbling the signal prior to its transmission – that similarly induces an interior optimum choice of investment but costs less than a price reduction.

3.4.1 Garbling Properties

Garbling a signal means strictly reducing its informativeness, be it through the adding of random noise or simply by systematically misreporting its realizations. The elements of the research literature related to endogenous garbling of information fall into three major categories: (1) the concept of signal garbling and relative informativeness of multiple signals (Blackwell 1953, Marschak and Miyasawa 1968, Holmstrom 1979, Gjesdal 1982), (2) the concept of deliberately degrading the quality of a good to be sold, for the purpose of encouraging buyers to separate by type (Deneckere and McAfee 1996, Varian and Shapiro 1998), (3) diverse analyses of "perverse" situations in agency contexts, wherein more accurate information or more intensive monitoring induces less effort or worse outcomes for the principal (Cowen and Glazer 1996, Dubey and Haimanko 2000, Jacobides and Croson 2001).

We believe that the analysis of the deliberate sabotage of investment information, treating the information seller as a strategic actor, is a novel application. Unlike in the damaged goods or versioning literatures, there is no intent to sell the full version of the information to any buyer; the intent is to alleviate moral hazard, not enable monopolistic screening. Precise garbling, the extent of which is chosen by the information seller, can supplement the relatively coarse tool of payment timing to encourage fine alterations in the investor's portfolio allocation choice.[12] Following Blackwell (1953) based on Bohnenblust et al. (1949), the garbled signal is less informative if the resulting signal $\tilde{x}$ is not a sufficient statistic for $\tilde{s}$, and therefore in DeGroot's (1962) sense not of more value when benchmarked over *all* decision problems.[13] In a one-person decision problem, the garbled signal could thus not possibly be of higher value (in the Blackwell-DeGroot sense) to the decision-maker. In a principal-agent setting, however, wherein a principal is searching for a method to distort the incentives of an agent to his advantage, the garbling of information may become a viable option even though it may destroy overall welfare (Lewis 2000, Weber and Croson 2002). Even though such garbling can shift payoffs internally between the principal and agent, the information is still less valuable to a social planner (in accordance with Blackwell and DeGroot's results.)

Proposition 3.6 (Fundamental Garbling Properties) *Let* (p_1^*, p_2^*) *be an optimal price*

[12] In addition, if the information for sale is of unknown quality, the seller of high-quality information may be forced to adopt return-contingent prices, as a credible signal of quality, to make any sales at all – even if inefficient risk-bearing results.

[13] For more specific situations, such as monotone decision problems with certain characteristics of the payoff function, there are stronger characterizations of the value of information available (Athey 2000), which allow for less restrictive criteria than statistical sufficiency.

vector for the informative signal.

(i) Garbling can increase the information seller's expected revenue R only if there is a default risk by the investor.

(ii) Garbling never increases overall welfare in the absence of moral hazard by the investor.

(iii) Garbling perfect information never increases revenue.

Proof: See the Appendix.

Proposition 3.3 describes the seller's optimal pricing policy under no default or in the absence of moral hazard states. Proposition 3.6 states that in both of these situations neither seller nor buyer can be made better off by decreasing the informativeness of the signal, so it must be welfare maximizing not to garble. By extension, since perfect information eliminates investor misbehavior and default at the optimal price vector (given in Proposition 3.4), the seller would never find it optimal to distort perfectly informative signals.

3.4.2 Existence of a Revenue-Improving Garbling

For simplicity consider the case in which the space of signal realizations consists solely of two elements, $\mathcal{S} = \{L, H\}$, so that $\text{Prob}(\tilde{s} = k) = \sigma_k \in (0, 1)$ for $k \in \{L, H\}$ and $\sigma_L + \sigma_H = 1$. The realization H represents "good news" and L "bad news" in the sense that $g_H(r) = g(r|H)$ first-order stochastically dominates $g_L(r) = g(r|L)$ on $\mathcal{R}$. The information seller can distort his informative signal by deliberately misreporting realizations. In particular, the seller will find it useful to claim that the received signal is H when it is actually L to induce the investor to be cautious even when H is reported. Instead of reporting all states truthfully, the seller thus misreports realizations L with probability $\alpha \in [0, 1]$ as H, whereas realizations H are always reported truthfully. The resulting probabilities of the signal reports observed by the investor are $\hat{\sigma}_H(\alpha) = \sigma_H + \alpha\sigma_L$ and $\hat{\sigma}_L(\alpha) = (1 - \alpha)\sigma_L$. Given such an information garbling policy (the nature of which is assumed to be common knowledge, just as if it were provably announced by the information seller), the investor adjusts her posterior beliefs given a report of H to

$$\hat{g}_H(r, \alpha) = \frac{\sigma_H g(r|H) + \alpha\sigma_L g(r|L)}{\sigma_H + \alpha\sigma_L}, \tag{3.20}$$

whereas for a report of L her beliefs are not influenced by the garbling: $\hat{g}_L(r, \alpha) = g(r|L)$ for all $\alpha \in [0, 1]$. Thus, the investor's investment action is influenced only for a report of "good news." We will assume that for $\alpha = 0$, H is a moral-hazard state exhibiting default risk (i.e.,

$H \in \mathcal{S}_+^*$). In this situation, there may exist a revenue-increasing garbling of information. The intuition proceeds as follows: if the signal H were always reported truthfully, the investor would choose $a = 1$ at the optimal prices derived in Section 3.3, above. Despite the investor's justified optimism given this information, there are still return realizations that fall short of the expected return given good news. If these returns are negative enough, the seller cannot collect all of p_2 (which may include a recapture of capital lent via a negative p_1) and thus receives lower revenue than if L had occurred (and the investor selected a more conservative investment policy based on that bad news). By occasionally misreporting L as H, the seller can temper the investor's enthusiasm to reduce the risk of default because a report of H is now less favorable than before. Of course, the investor's WTP for this adulterated information is no longer as high as for the truthfully reported information, but by judicious garbling the seller may be able to collect more by avoiding default than she must give up in a price discount for the damaged information good.

Proposition 3.7 (Existence of a Revenue-Improving Garbling) *If an optimal price vector, (p_1^*, p_2^*), subject to $\alpha = 0$, is such that the investor's individual-rationality constraint (IR) is not binding, and $H \in \mathcal{S}_+^*$ (so that $a_H^* = 1$ is an element of the investor's utility-maximizing choice in (IC)), then garbling the information (i.e., choosing $\alpha > 0$) will result in higher expected revenue than not garbling, and seller revenue will be maximized by a series of progressively more intense garblings and corresponding adjustments (p_1^*, p_2^*) until (IR) becomes binding.*

Note that in the two-state case, at an optimum with positive garbling α the investor must be indifferent between investing at an interior level and going for broke (i.e., $\hat{a}_H^* = \{\hat{a}_H^o, 1\}$ with $\hat{a}_H^o < 1$), the only possible reason for the seller to leave slack in the investor's individual-rationality constraint (IR). Indeed if (IR) was not binding and $\hat{a}_H^* = 1$ was not an element of the investor's optimal investment policy in (IC) at H, then a small increase in either p_1 or p_2 would yield higher seller revenues without triggering going-for-broke behavior on the part of the investor (from $\hat{a}_H^* < 1$ to $\hat{a}_H^* = 1$). The seller therefore responds to the investor's implicit of going for broke by leaving her with a positive surplus (i.e., slack in (IR)). As we show in the proof of Proposition 3.7, increasing α strictly reduces the investor's propensity to go for broke, as the difference between $\hat{V}_H^1$, her utility under $a = 1$, and $\hat{V}_H$, her utility under interior optimal $a_*^o \in (0, 1)$, strictly decreases, $\partial_\alpha \left(\hat{V}_H^1 - \hat{V}_H \right) < 0$. Thus a "discontinuous" switch in the investor's optimal policy may sometimes be achieved by garbling the information without the seller having to adjust prices downward accordingly. Continuing in this manner choosing increasingly positive α permits the information seller to extract all slack from the investor's individual-rationality constraint, strictly increasing his revenue.

Our conclusion is that garbling can thus expand the profitability frontier for the information seller, and will enter the optimal solution whenever zero garbling leads an investor to invest fully in the moral hazard state.

3.5 Discussion

We have focused in this paper on the value of information when the seller can split payments in ex-ante and ex-post components, and examined the challenges facing the seller who tries to collect payments partially ex post from a wealth-constrained investor under limited liability. We find that the deferral of collection can create value, even though it subjects the seller to default risk and potential moral hazard of overinvestment. The optimal pricing structure generally involves initial speculative co-investment followed by extraction of investor surplus through an ex-post charge. Furthermore, we show that garbling the signal can increase expected revenues by discouraging overinvestment in favorable signal states, thereby reducing the frequency and costs of default.

An intriguing analogy to the time value of money suggests itself. Unlike the timing of cash flows in which collecting earlier is better, collecting later is better here because substantially more (in expectation) can be collected. This value accrual through payment deferral is not driven by assumptions about the pure time value of money, interest payments from the buyer, or as a required return to compensate for default risk taken (all of which we have assumed away). Rather, the seller's deferral expands the total (expected) gain through placing more capital behind the informative signal's recommendations, as the seller can effectively invest the value of the information (plus any subsidy) alongside its buyer's wealth. In addition, deferring payment enables more accurate extraction of the buyer's value of the information, as the seller can choose an optimal combination of ex-ante and ex-post charges rather than being constrained to ex-ante contracts only. The time value of this information thus accrues not to its buyer (who receives value first but pays later), but to the seller (who even subsidizes the buyer at first, and collects value only later if at all.)

The obvious extension in the investment context is to examine contingent pricing of the information, either contingent on signal realizations (ex-ante contingency) or on return outcomes (ex-post contingency), even if investors' levels of final wealth remain noncontractible. Presumably, more favorable signals would result both in higher prices and the choice of higher investment levels. Our preliminary investigations show that ex-ante contingent pricing strictly outperforms ex-ante noncontingent pricing, and that ex-post contingent pricing strongly outperforms ex-post noncontingent pricing, but that ex-ante contingent pricing and ex-post noncontingent pricing are not strongly ordered.

We see the sale of quality-differentiated information as the first extension outside the investment context (cf. Allen 1986). While the results presented here make only minimal assumptions about investor characteristics, we can certainly calculate optimal prices and levels of garbling for specific forms of buyer utility. Given that different types of buyers differ in their WTP for the same information, a discriminating monopolist could screen them by charging different prices for different qualities of information (i.e., more- or less-garbled versions of the true signal). The optimal degrading of this information, however, interacts with the price received when pricing structures other than ex-ante pricing are employed. The spread of versions offered under various pricing structures may transcend simply vertically differentiated versions of the same information (Deneckere and McAfee 1996, Shapiro and Varian 1998) and the optimal combination may mix different revenue models in the same market.

We also hope to apply our theory of deliberate sabotage of information for sale to a corporate governance setting. Ever since Fama and Miller (1972), the moral-hazard effect of the presence of debt on equity holders' risk-taking incentives has been a staple of corporate finance research. Our approach suggests an efficient way to address the costs of this moral hazard in the selection of risky projects. Garbling (or censoring) the information given to managers may encourage them to take less risk, reducing the variation of the firm's cash flows (and the probability of default on the firm's debt) and thereby lowering the firm's weighted-average cost of capital - which is in the long-term interest of both bond- and stockholders. Of course, the decisions made by these managers will suffer from this lower-quality information, leading to a direct theoretical basis for the tradeoff between costs of incentive misalignment and decision errors posited in Jensen and Meckling (1992).

Finally, the application of the tools we develop in this paper, although developed in a financial context, need not stop in the area of financial economics. The revenue-maximization problem faced by the information seller, subject to both individual-rationality and incentive constraints by the buyer and the noncontractibility of investment intensity, strongly resembles a principal-agent problem. We hope to extend these results to a general principal-agent model, where the principal can use a two-part wage structure to induce higher actions but faces moral hazard because agent actions are noncontractible, and who cannot collect a flat fee in excess of the agent's initial wealth. Combining the relatively blunt tool of flat-rate wage schedules with the principal's ability to fine-tune the information conditions and risks faced by the agent can lead to a setting in which the principal may even deliberately garble his own information to induce higher agent effort, in the spirit of Cowen and Glazer (1996). Techniques to motivate wealth-constrained agents are currently at the frontiers of agency theory (e.g., Che and Gale 2000, Lewis and Sappington 2000) and financial decision-making is a logical research area in which to derive, extend, and apply these theoretical innovations.

3.6 References

Admati, A., Pfleiderer, P. 1986. A monopolistic market for information. *Journal of Economic Theory* **39**(2) 400–438.

Admati, A., Pfleiderer, P. 1987. Viable allocations of information in financial markets. *Journal of Economic Theory* **43**(1) 76–115.

Admati, A., Pfleiderer, P. 1990. Direct and indirect sale of information. *Econometrica* **58**(4) 901–928.

Allen, B. 1986. The demand for (differentiated) information. *Review of Economic Studies* **53**(3) 311–323.

Allen, F. 1990. The market for information and the origin of financial intermediation. *Journal of Financial Intermediation* **1**(1) 3–30.

Athey, S. 2000. Investment and information value for a risk averse firm. Working paper, Massachusetts Institute of Technology.

Athey, S., Levin, J. 2000. The value of information in monotone decision problems. Working paper, Massachusetts Institute of Technology.

Berge, C. 1959. *Espaces Topologiques et Fonctions Multivoques.* Dunod, Paris.

Berle, A., Means, G. 1932. *The Modern Corporation and Private Property.* Commerce Clearing House, Reprinted in 1993 by Gryphon Editions, New York, NY.

Blackwell, D. 1953. Equivalent comparison of experiments. *Annals of Mathematical Statistics* **24**(2) 265–272.

Bohnenblust, H., Shapley, L., Sherman, S. 1949. Reconnaissance in game theory. *Research memorandum RM-208*, RAND Corporation.

Che, Y., Gale, I. 2000. The optimal mechanism for selling to a budget-constrained buyer. *Journal of Economic Theory* **92**(2) 198–233.

Cowen, T., Glazer, A. 1996. More monitoring can induce less effort. *Journal of Economic Behavior and Organization* **30**(1) 113–123.

DeGroot, M. 1962. Uncertainty, information, and sequential experiments, *Annals of Mathematical Statistics* **33**(2) 404–419.

Deneckere, R., McAfee, R.P. 1996. Damaged goods. *Journal of Economics and Management Strategy* **5**(2) 149–174.

Dubey, P., Haimanko, O. 2000. Optimal scrutiny in multi-period promotion tournaments. Cowles foundation discussion paper, Cowles Foundation, Yale University.

Fama, E., Miller, M. 1972. *The Theory of Finance.* Dryden Press, Hinsdale, IL.

Franklin, B. 1795. *The Way to Wealth, or, Poor Richard Improved.* Printed for Ant. Aug. Renouard, Paris.

Gjesdal, F. 1982. Information and incentives: the agency information problem. *Review of Economic Studies* **49**(3) 373–390.

Grossman, S., Hart, O. 1983. An analysis of the principal-agent problem. *Econometrica* **51**(1) 7–45.

Hicks, J. 1939. *Value and Capital*, Clarendon Press, Oxford, UK.

HOLMSTROM, B. 1979. Moral hazard and observability. *Bell Journal of Economics* **10**(1) 74–91.

JACOBIDES, M., CROSON, D.C. 2001. Information policy: shaping the value of agency relationships. *Academy of Management Review* **26**(2) 202–223.

JENSEN, M., MECKLING, W. 1976. Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* **3**(4) 305–360.

JENSEN, M., MECKLING, W. 1992. Specific and general knowledge, and organizational structure, in L.Werin and H. Wijkander (eds.), *Contract Economics.* Blackwell, Oxford, UK, 251–274.

KIHLSTROM, R. 1974. A general theory of demand for information about product quality. *Journal of Economic Theory* **8**(4) 413–439.

KIHLSTROM, R. 1976. Firm demand for information about price and technology. *Journal of Political Economy* **84**(6) 1335–1341.

KIHLSTROM, R. 1988. Optimal contracts for security analysts and portfolio managers. *Journal of Banking and Finance* **5** 291–325.

LAWRENCE, D. 1999. *The Economic Value of Information*, Springer, Springer, NY.

LEWIS, T., SAPPINGTON, D. 2000. Motivating wealth-constrained actors. *American Economic Review* **90**(4) 944–960.

MARSCHAK, J., MIYASAWA, K. 1968. Economic comparability of information systems. *International Economic Review* **9**(2) 137–173.

MIRRLEES, J. 1974. Notes on Welfare Economics, Information, and Uncertainty, in M. Balch, D. McFadden and S.-Y. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam, NL.

Raiffa, H., Schlaifer, R. 1961. *Applied Statistical Decision Theory.* Division of Research, Graduate School of Business Administration, Harvard University, Boston, MA.

Sappington, D. 1983. Limited liability contracts between principal and agent. *Journal of Economic Theory* **29**(1) 1–21.

Shapiro, C., Varian, H.R. 1998. Versioning: the smart way to sell information. *Harvard Business Review* **76**(6) 107–114.

Thon, D., Thorlund-Petersen, L. 1993. The value of perfect information in a simple investment problem. *Information Economics and Policy* **5** 51–71.

Weber, T.A., Croson, D.C. 2002. Selling less information for more: garbling with benefits. Working paper, Wharton School, University of Pennsylvania.

3.7 Appendix: Proofs

Proof of Proposition 3.1: (i) Let $a_s^0 \in (0,1)$ be such that it satisfies (3.3)–(3.4). Differentiating the expression for π_s in (3.2) at $a = a_s^o$ with respect to p_2, we obtain by the envelope theorem (i.e., by using the first-order condition (3.3)), the Leibniz Rule, and our convention that $u(0) = 0$:

$$D_{p_2}\pi_s(a_s^o) = \partial_{p_2}\pi_s(a_s^o) = -\int_{r_c(a_s^o)}^{\bar{r}} u'(\cdot)\, g(r|s)dr. \tag{3.21}$$

Similarly, differentiating (3.2) with respect to p_1 yields

$$D_{p_1}\pi_s(a_s^o) = \partial_{p_1}\pi_s(a_s^o) = -\int_{r_c(a_s^o)}^{\bar{r}} (1 + a_s^o r)u'(\cdot)\, g(r|s)dr, \tag{3.22}$$

and thus by using (3.3) the same right-hand side as in (3.21). We see that this proof does not depend on the curvature properties of u as long as a_s^o is a strict local maximizer. (ii) If $\bar{a} \in (0,1]$ is a constant investment policy, then expression (3.21) remains unchanged.[14] Differentiating with respect to p_1 we obtain

$$D_{p_1}\pi_s(a_s^o) = -\int_{r_c(\bar{a})}^{\bar{r}} u'(\cdot)g(r|s)dr - \bar{a}\int_{r_c(\bar{a})}^{\bar{r}} u'(\cdot)g(r|s)r\, dr, \tag{3.23}$$

[14] Clearly we cannot use the envelope theorem, since the first-order condition is not satisfied in general. However, we have $D_{p_1}\pi_s(\bar{a};p,w) = \partial_{p_2}\pi_s(\bar{a};p,w) + \partial_a\pi_s(\bar{a};p,w)\partial_{p_2}\bar{a} = \partial_{p_2}\pi_s(\bar{a};p,w)$.

so that the equivalence statement in (ii) obtains. If $a_s^* = 1$, then necessarily $\partial_a \pi_s(1) > 0$, so that $D_{p_1}\pi_s(1; p, w) < D_{p_2}\pi_s(1; p, w)$. If $\bar{a} = 0$, then the two derivatives are equal as a direct consequence of (3.21) and (3.23). (iii) Differentiating the first-order condition (3.3) at $a = a_s^o \in a_s^* \cap (0, 1)$ on both sides with respect to p_2 (totally) using the Leibniz Rule we obtain (3.6) directly. On the other hand, differentiating (totally) with respect to p_1 yields

$$\frac{\partial_a^2 \pi_s(a_s^o)}{w - p_1}\partial_{p_1} a_s^o - \int_{r_c(a_s^o)}^{\bar{r}} u''(\cdot)g(r|s)r\,dr - \frac{u'(0)g(r_c\,(a_s^o)|s)\,p_2 r_c(a_s^o)}{(w - p_1)^2 a_s^o} = 0.$$

After adding and subtracting $p_2 \int_{r_c(a_s^o)}^{\bar{r}} u''(\cdot)g(r|s)r\,dr$ and then multiplying both sides by $(w - p_1)/(\partial_a^2 \pi_s(a_s^o)) < 0$, we obtain for $\partial_{p_1} a_s^o$ the expression

$$\frac{\int_{r_c(a)}^{\bar{r}} u''(\cdot)g(r|s)r\,dr}{\partial_a^2 \pi_s(a)} + p_2 \left(\frac{w - p_1}{\partial_a^2 \pi_s(a_s^o)} \int_{r_c(a_s^o)}^{\bar{r}} u''(\cdot)g(r|s)r\,dr + \frac{u'(0)g(r_c(a_s^o)|s)r_c(a_s^o)}{a_s^o \partial_a^2 \pi_s(a_s^o)} \right),$$

which – after substituting the definition of the relative risk aversion $\rho(\cdot)$ – is identical to the right-hand side of (3.6). ■

Proof of Proposition 3.2: We first note that for any $p \in \mathcal{P}$, $w > 0$ and $s \in \mathcal{S}$, due to the continuity of $E\left[\pi(a; p, w, \tilde{r})|\, s\right]$ in a on the compact set $[0, 1]$, there exists a maximizer a_s^* solving (IC). By the maximum theorem (Berge 1959) the maximizer a_s^* is compact-valued and upper-hemicontinuous in p. In addition, the maximum theorem asserts that $v_s(\cdot, \cdot)$ is continuous for any $s \in \mathcal{S}$, which implies that the investor's expected indirect ex-ante utility from observing the signal, $v(\cdot, \cdot)$ as defined in (IR), is also continuous. As a consequence, the seller's objective function in (3.7) is continuous in p, but with an image of $\mathcal{P}$ that is may be unbounded, if $\mathcal{P}$ is unbounded. Introducing the investor's individual-rationality constraint (IR) we can restrict attention to price vectors $p \in \mathcal{P}$ that achieve at least her reservation utility $v_0(w)$, which together with the fact that $\mathcal{P}$ is bounded from below, defines a compact subset $\mathcal{P}_{\text{IR}}$ of $\mathcal{P}$. The image of $\mathcal{P}_{\text{IR}}$ under the mapping $v(\cdot, w)$ is compact for all $w > 0$, and hence it contains an element p^* solving (3.7) subject to (IC) and (IR). ■

Proof of Proposition 3.3: (i) If $p_1 = w$, then the investor's investable wealth is zero and her investment opportunity, which by assumption has positive expected value cannot yield any return. Thus in order for the price for the information to be individually rational for the investor, p_2 needs to be negative and its absolute value equal to the certainty equivalent of the investment opportunity, $p_2 = -\text{CE}(w) = -u^{-1}\,(v_0(w)) < -w$. Thus, for any signal this yields negative revenues for the information seller, since $p_1 + p_2 < 0$. Hence necessarily $p_1^* < w$.

Similarly, if any one of the two components p_k^*, $k = 1, 2$, is nonpositive, the other component must be nonegative in order for total revenues to be nonnegative. (ii) The Lagrangian for the constrained optimization problem (3.7) subject to (IC) and (IR) with the investor's individual-rationality constraints binding is $\mathcal{L}(p,w) = R(p,w) + \lambda \Big(v(p,w) - v_0(w) \Big)$ with Lagrange multiplier $\lambda \geq 0$. Provided that there exists an optimal price in the interior of $\mathcal{P}$, the first-order necessary optimality conditions are thus

$$
\begin{aligned}
1 + \int_{\mathcal{S}} \Bigg\{ \Big(w - p_1 - p_2 + (w - p_1) a_s^* r_c \Big) g(r_c|s) \partial_{p_1} r_c - \int_{\underline{r}}^{r_c} g(r|s) dr \\
-a_s^* \int_{\underline{r}}^{r_c} g(r|s) r dr + (w - p_1) \left(\partial_{p_1} a_s^* \right) \int_{\underline{r}}^{r_c} g(r|s) dr \Bigg\} \sigma(s) ds + \lambda \partial_{p_1} v(p,w) &= 0, \quad (3.24) \\
\int_{\mathcal{S}} \Bigg\{ \Big(w - p_1 - p_2 + (w - p_1) a_s^* r_c \Big) g(r_c|s) \partial_{p_2} r_c + \int_{r_c}^{\bar{r}} g(r|s) dr \\
+(w - p_1) \left(\partial_{p_2} a_s^* \right) \int_{\underline{r}}^{r_c} g(r|s) dr \Bigg\} \sigma(s) ds + \lambda \partial_{p_1} v(p,w) &= 0, \quad (3.25)
\end{aligned}
$$

where

$$
\begin{aligned}
\partial_{p_1} r_c &= \frac{p_2}{(w-p_1)^2 a_s^*} - r_c \partial_{p_1} \ln \left(a_s^* \right), & (3.26) \\
\partial_{p_2} r_c &= \frac{1}{(w-p_1) a_s^*} - r_c \partial_{p_2} \ln \left(a_s^* \right). & (3.27)
\end{aligned}
$$

By subtracting (3.25) from (3.24) and using (3.1) as well as (3.26)–(3.27) we obtain

$$
\begin{aligned}
\frac{\left(\partial_{p_1} - \partial_{p_2} \right) R}{w - p_1} &= \int_{\mathcal{S}} \left\{ \Big(\partial_{p_1} a_s^* - \partial_{p_2} a_s^* \Big) \int_{\underline{r}}^{r_c} g(r|s) dr - \frac{a_s^*}{w - p_1} \int_{\underline{r}}^{r_c} g(r|s) r \, dr \right\} \sigma(s) ds, \\
&= -\lambda \left(\partial_{p_1} - \partial_{p_2} \right) v(p,w) \qquad (3.28)
\end{aligned}
$$

i.e., condition (3.9) as long as $\partial_{p_1} v = \partial_{p_2} v$. But from part (ii) of Proposition 3.1 we can immediately conclude that $\partial_{p_1} v = \partial_{p_2} v$ whenever $1 \notin a_s^*$ on $\mathcal{S}$. For interior optima $a_s^o \in a_s^* \cap (0,1)$ by part (iii) of Proposition 3.1 relation (3.10) also holds. If on the other hand $\mathcal{S}_+$ (as defined in the proposition) is of positive measure, then relation (3.11) follows directly from (3.22)–(3.23) after noting that $(\partial_{p_1} R)/(\partial_{p_2} R) = (\partial_{p_1} v)/(\partial_{p_2} v)$. (iii) No default risk means that no return realization is below the critical return r_c in (3.1), given the investor's choice at the seller's optimal price. If the information is not very valuable and the investor sufficiently risk-averse, there will be no default risk. In that case, the sellers revenue function degenerates to $R = p_1 + p_2$ and thus he is naturally indifferent between amounts collected ex ante and amounts collected ex post. Hence the optimal price will be such that the investor's compensating variation is highest, which according to part (i) of Proposition 3.1 is (weakly) achieved if the least possible amount is charged ex ante (i.e., $p_1^* = \underline{p}$) and subsequently an amount ex post that makes the investor indifferent between procuring the partially informative signal or not, so that (3.12) determines p_2^*. ■

Proof of Proposition 3.5: Differentiating the maximand $\breve{R}(a;p,w)$ in the seller-determined revenue-maximization problem (3.14) with respect to a we obtain using the Leibniz Rule

$$\begin{aligned} \partial_a \breve{R}(a;p,w) &= \frac{p_2 g(r_c|s) r_c}{a} + (w-p_1)\int_{\underline{r}}^{r_c} g(r|s) r\, dr - (w-p_1) g(r_c|s) \frac{(1+ar_c)r_c}{a} \\ &= (w-p_1)\int_{\underline{r}}^{r_c(a)} g(r|s) r\, dr, \end{aligned} \tag{3.29}$$

where we have substituted the definition of $r_c(a)$ in (3.1). At this point it is useful to distinguish between r_c nonpositive and positive. (i) If $r_c(1) \le 0$ (equivalent to $p_1 + p_2 \le w$), then it is clear that $\partial_a \breve{R} < 0$ and an optimal action is zero investment or any action $\breve{a}_s$ for which $r_c(\breve{a}_s) \le \underline{r}$, i.e., (3.15) holds. (ii) If on the other hand $r_c(1) > 0$ (equivalent to $p_1 + p_2 > w$), then $r_c(a) > 0$ for all $a \in [0,1]$ and the integrand in (3.29) becomes positive on the interval $[0, r_c(a)] \cap [0,1]$. If in addition $E[\tilde{r}|s] \le 0$, then similar to the previous case, an optimal action is zero investment or any action $\breve{a}_s$ for which $r_c(\breve{a}_s) \ge \bar{r}$. If $E[\tilde{r}|s] > 0$, then there is only one optimal action a_s^o, so that (3.16)–(3.17) hold. ■

Proof of Proposition 3.6: (i) If the investor's default risk is zero, then the optimal price vector is given by Proposition 3.3 (iii). In other words, the investor is liable for all of her investment losses incurred by investing aggressively, including the agreed upon payment to the information seller. The buyer's individual-rationality constraint (IR) is binding and the seller's expected payoff is $R = p_1 + p_2$. Any decrease in the investor's WTP as a result of garbling the signal would be borne directly by the information seller. Thus garbling cannot increase the seller's revenues without investor default.
(ii) In the absence of investor moral hazard, i.e., if $\mathcal{S}_+^* = \emptyset$, the seller's optimal price vector is first-best, since prices according to Proposition 3.3 are chosen such that investor utility decreases as little as possible, the seller being indifferent between charging p_1 or p_2. In fact this leads to a maximum upfront subsidy ($p_1^* = \underline{p}$), and a expected full extraction of ex-post suplus ((IR) binding). Thus, if there were a garbling that increased overall welfare, then the resulting garbled signal would have to be more informative for the one-person social planner's decision problem of maximizing overall welfare than the original signal. Hence, garbling information in the absence of investor moral hazard cannot increase revenues. (iii) As pointed out in Section 3.2.3, perfect information yields signal realizations equal to the actual investment returns,

$s = r$, and an optimal investor policy $a_s^* = [\text{sgn}(s)]_+ \in \{0, 1\}$. As the amount invested is zero whenever $r < 0$, there can be no default risk. Thus by (i) garbling cannot increase revenue.[15]

■

Proof of Proposition 3.7: Consider the moral-hazard state H, in which the investor and seller's posterior probability density is given by $\hat{g}_H(r, \alpha)$. For $\alpha = 0$ the moral-hazard state remains ungarbled.[16] Since by assumption g_H first-order stochastically dominates g_L, we have that the corresponding investor actions a_H^* and a_L^* are such that $a_H^*(\alpha; p, w) \geq a_L^*(p, w)$ for all $\alpha \in [0, 1]$. Increasing the garbling generates progressively worse gambles in the sense of FOSD, because of the shift of posterior probability mass from higher return states to lower ones. Hence such a garbling smoothly reduces the investor's action a_H^* as a function of α. The reduction is *strict* for all $\alpha > 0$ if without any garbling $a_H^* < 1$ (i.e., no "saturation"), which we assume henceforth. Let $\hat{V}_H^1 = \int_{r_c(1)}^{\bar{r}} u\left((w - p_1)(1 + r) - p_2\right) \hat{g}_H(r, \alpha) dr$ and $\hat{V}_H = \int_{r_c(\hat{a}_H^o)}^{\bar{r}} u\left((w - p_1)(1 + \hat{a}_H^o r) - p_2\right) \hat{g}_H(r, \alpha) dr$, where $\hat{a}_H^o \in (0, 1)$ is an interior maximizer of the investor's problem under garbling, analogous to (IC). Correspondingly we define V_H^1, V_L^1, V_H as the investor utilities under a policy $a = 1$ in high/low signal state and $a = a_H^o \in (0, 1)$ (where solves (IC)) respectively. Assume that $V_H^1 > V_H$, so that without garbling (for $\alpha = 0$) the investor strictly prefers to go for broke allocating all her wealth (including the seller subsidy if any) to the risky asset. We will show that the propensity to go for broke, $\hat{V}_H^1 - \hat{V})_H$, strictly decreases in α. For this, let us first compute $\partial_\alpha \hat{V}_H^1$. Using the definition of $\hat{g}_H(r, \alpha) = \left(\sigma_H g(r|H) + \alpha \sigma_L g(r|L)\right) / \hat{\sigma}_H$, we can write

$$\hat{V}_H^1 = \frac{\sigma_H V_H^1 + \alpha \sigma_L V_L^1}{\hat{\sigma}_H},$$

so that

$$\partial_\alpha \hat{V}_H^1 = \frac{\sigma_L \sigma_H \left(V_L^1 - V_H^1\right)}{\hat{\sigma}_H^2}.$$

Thereby we have taken into account that for $a = 1$ both V_H^1 and V_L^1 do not depend on α. On the other hand, using the envelope theorem, the Leibniz Rule, as well as the definition of r_c we have that

$$\begin{aligned} \partial_\alpha \hat{V}_H &= \partial_\alpha \int_{r_c(\hat{a}_H^*)}^{\bar{r}} u\left((w - p_1)(1 + \hat{a}_H^* r) - p_2\right) \hat{g}_H(r, \alpha) dr \\ &= \int_{r_c(\hat{a}_H^*)}^{\bar{r}} u\left((w - p_1)(1 + \hat{a}_H^* r) - p_2\right) \partial_\alpha \hat{g}_H(r, \alpha) dr, \end{aligned}$$

[15] Note also that there is no moral hazard in any state, as $a_s^* = \hat{a}_s$ for all $s \in \mathcal{R}$.

[16] This mechanism is assumed to be common knowledge and irrevocably committed to by the seller before the signal results are observed.

whereby

$$\partial_\alpha \hat{g}_H(\alpha, r) = \frac{\sigma_L \sigma_H \left(g(r|L) - g(r|H)\right)}{\hat{\sigma}_H^2}.$$

Thus

$$\partial_\alpha \left(\hat{V}_H^1 - \hat{V}_H\right) = \frac{\sigma_L \sigma_H}{\hat{\sigma}_H^2} \left[\left(V_L^1 - \int_{r_c(\hat{a}_c^H)}^{\bar{r}} u(\cdot)_{\hat{H}} g(r|L) dr\right) - \left(V_H^1 - \int_{r_c(\hat{a}_c^H)}^{\bar{r}} u(\cdot)_{\hat{H}} g(r|H) dr\right)\right], \tag{3.30}$$

where $(\cdot)_{\hat{H}}$ abbreviates $((w - p_1)(1 + \hat{a}_H^* r) - p_2)$. Note first that

$$V_L^1 < \int_{r_c(\hat{a}_c^H)}^{\bar{r}} u(\cdot)_{\hat{H}} g(r|L) dr,$$

since given a true signal realization L with associated probability density $g(r|L)$, it is strictly better for the investor to invest $\hat{a}_H^*$ which is closer to the unique maximizer a_L^* than is investing the full wealth under $a = 1$. The second term component of (3.30) is also negative, since by assumption $V_H^1 > V_H$ ("going for broke" in the absence of garbling when $\alpha = 0$) and

$$V_H > \int_{r_c(\hat{a}_c^H)}^{\bar{r}} u(\cdot)_{\hat{H}} g(r|H) dr,$$

by optimality of V_H for an unfiltered realization of H with associated probability density of $g(r|H)$. As a result, we obtain from (3.30) that

$$\partial_\alpha \left(\hat{V}_H^1 - \hat{V}_H\right) < 0,$$

implying that *any* garbling reduces the investor's propensity to go for broke. Thus, if the point of indifference is "close" enough, or in other words, if $V_H^1 - V_H > 0$ is small for $\alpha = 0$, then garbling will tend to induce a change in investor behavior for the high signal state from $a_H^* = 1$ to an interior optimum $a_H^* \in (0, 1)$. The resulting sharp decline in the investor's misbehavior entails a finite positive impact on the seller's revenues. In addition to this, the seller does not have to adjust prices downward for small enough α, for by assumption there is slack in the investor's individual-rationality constraint (IR) at the outset when $\alpha = 0$. We have thus shown that garbling may provide a way to increase the seller's revenue, whenever (IR) is slack at the optimal price vector subject to $\alpha = 0$. In fact, by adjusting prices up to investor indifference between going for broke and choosing an interior investment level, the seller can successively extract all of the investors initial slack in her individual-rationality constraint (IR). The minimal α that is necessary to achieve no default in equilibrium (with $\mathcal{S}$ restricted to two states) is optimal.[17] ■

[17] For a support $\mathcal{S}$ containing more states, the seller may find it optimal to tolerate going-for-broke behavior for some signal realizations.

The greatest loss of time is delay and expectation,
Which depend upon the future.
We let go the present, which we have in our power,
And look forward to that which depends upon chance,
And so relinquish a certainty for an uncertainty.
— SENECA

Chapter 4

Information Sharing

4.1 Introduction

In the current global networked economy information sharing has become an imperative. Still, the *value of the information* to be shared is often poorly defined and ill-assessed. Much less clear is *how much each beneficiary of such shared information ought to contribute* financially in order to at least offset the expenses for the creation and dissemination of the information. A widespread inability of dotcom's to profitably offer purely informational content over the Internet speaks for itself. This is particularly striking in the light of the often-repeated argument that "information is nonrival," i.e., due to its generally very low cost of reproduction more than one agent is able to observe at the same time essentially the same realization of the same informative signal (e.g., the report of a news event or a current stockmarket price). From such an observation each agent then draws conclusions relating to her respective project and adjusts her actions accordingly. Hence, depending on the differences in the projects the agents pursue, their levels of confidence in the observed signal in terms of providing decision-relevant information about an uncertain future payoff are generally varied.

One can think of a *shared information service* as an organization that is run by its members for the sole purpose of producing information that is being used in a *nonstrategic* but *excludable* way.[1] In particular, we assume that the individual members of the shared information service pursue *independent* projects; for instance the involved agents comprise only a small

[1] This means that the group using the information can prevent others from using this same information source, but each member does not use the information in a strategic way against another member. In Section 4.3.4 the assumption of nonstrategic agent interaction is somewhat relaxed through the explicit consideration of positive and negative network effects.

portion of a common market, in which they do not have market power or they act in different fields. There are many concrete examples of nonstrategic shared information services, such as the Aviation Weather Forecast, the Insurance Services Office,[2] networked medical expert and knowledge-base systems, any type of member-operated information- or news-aggregation networks, and national agencies with a common set of objectives.[3] We allow for the possibility of customization or versioning of the information (Shapiro and Varian 1998): the precise observation of the signal may be different from agent to agent; moreover, even the signal itself may be different across users. For instance, queries to a common database input by different users corresponding to their individual needs yield different informative signals, even though the underlying information resource is shared.

The *economic incentives* for setting up and using shared information services stem from economies of scale and scope in assembling the raw information as well as economies of specialization in generating the informative signal. The fixed cost for the formation of an information service capable of producing signals of acceptable reliability ("confidence") is often too high for any single firm, which calls for group action and a contractual bargaining solution to avoid the tragedy of the commons (Liebowitz and Margolis 1994). We thereby assume that the firms, once they decide to enter the bargaining negotiations, act rational in the sense of Harsanyi (1966) and that the coordination of the negotiations does not pose particular problems.

The pioneering contributions in strategic information sharing have been made by Novshek and Sonnenschein (1982) and Clarke (1983) for oligopolistic firms producing homogeneous goods (no incentive for information sharing), as well as Vives (1984) for a differentiated Cournot duopoly (incentive for information sharing exists if the goods are complements or weak substitutes). Raith (1996) unifies and generalizes this and most subsequent work on strategic information sharing with market interaction. As far as we know, there exists no prior treatment of information sharing in a *cooperative* context in the Economics literature. In the literature on information systems, there is some concern investments in cooperative inter-organizational information systems. For instance Wang and Seidmann (1995) consider the investment in an electronic data interchange (EDI) system and show that a "supplier's adoption of EDI can generate a positive externality for the buyer and negative externalities for

[2]The Insurance Services Office (http://www.iso.com) acts as a supplier of statistical, actuarial, underwriting, and claims information to the property and casualty insurance industry.

[3]As a current example emphasizing the need for information sharing between federal agencies, consider Informationweek's call that "[s]haring data is key to antiterrorism efforts" (Stahl 2001) in the wake of the September-11 attack on the World Trade Center.

other suppliers" (p. 401). These externalities play a role in the adoption dynamics of interorganizational systems (1994). A question untouched by Wang and Seidmann and other authors in this area is precisely *how* the investment for shared information systems should be distributed between the different users, based on the *decision value* of the information. Our goal here is to derive *efficient* investment sharing rules, based on cooperative bargaining. We thereby model explicitly the information to be shared as a signal that is imperfectly correlated with a decision maker's uncertain investment payoff. Nash (1953) founded the field of "cooperative bargaining theory," showing that based on a number of axioms (cf. Section 4.3.3) a unique solution could be obtained that implements a Pareto-optimal allocation. A good overview is provided by Roth (1979). In particular, we focus on the following two research questions:

1. Given that a number of economic agents (e.g., firms, federal agencies or individuals) may experience a benefit from setting up a shared information service, how much should each one contribute financially, and how much is the overall venture worth? An answer to this question can serve as an upper bound for what could be charged by, say, a monopolist information provider to that same group of agents, and we will compare both cases. In the first case, firms are assumed to cooperatively bargain about the share of the implementation cost; in the second case, a monopolist information provider maximizes his profits if he decides to enter. Naturally, the cooperative bargaining solution outperforms the monopolistic solution, if of course its efficiency in actually providing and disseminating the information is the same as for the monopolist.

2. What are the conditions for existence of shared information services? We provide explicit expressions for the *critical mass* of users that is necessary to form an information network, both in the cooperative and the monopolistic situation. We introduce the notion of a "critical Nash network size" which counts the smallest subset of agents that can create a shared information network subject to each agent being at least as well off as without the information.

The outline of this paper is as follows. In Section 4.2 we give an exact definition of the value of information in terms of the "compensating variation" of the agent's utility that makes her indifferent between acquiring the informative signal or not. A definition of the value of information as compensating variation has first been introduced by Kihlstrom (1974) in a product-consumption setting and by Treich (1997) for a standard portfolio investment problem. Although not widely spread, this definition is superior to what is commonly used in the Economics literature, since it takes full account of all wealth effects. We also compare the general effect of the number of participants on the cost of a shared information service, depending on its cost characteristics. Section 4.3 then determines what each participant should pay as a result of a cooperative bargaining process and how the cost for setting up a shared information service varies with the size of its membership body and associated network externalities. In Section 4.4 we compare this to what a monopolist can charge in terms of an "ultimatum offer," and we discuss the sources of inefficiency. Section 4.5 concludes with directions for future research.

4.2 The Value of Information

In the traditional Economics literature the value of information to a particular agent is assessed, depending on the agent's risk posture, using two different methods (Raiffa and Schlaifer 1961, Marschak and Miyasawa 1968, Athey and Levin 2000): (i) for risk-neutral agents the information value is determined as difference between the optimal payoff with information and the optimal "default" (or "no-information") payoff; (ii) for risk-averse agents, the cost of an informative signal is typically modeled as a "utility cost," which is subtracted from the agent's full-information utility, *as if* the amount for the information was paid out of a separate budget. These two common approaches to computing the value of information typically neglect the effect the payment itself has on the agent's budget set and thereby on constraining her feasible actions. Such "wealth effects" are, as should be expected, of considerable relevance at least in a portion of the wealth domain, irrespective of the assumed risk-aversion characteristics. Also, wealth effects *cannot* generally be eliminated by an "appropriate" choice of a class of utility functions possessing multiplicative or additive separability properties (such as those of constant absolute risk aversion (CARA) as is often claimed). We demonstrate below, using CARA utilities, that wealth effects are significant when the agent's risk aversion, her endowed wealth, or her project risk is relatively small or her confidence in the information is extremely large.

As a consequence, the value of information needs to be carefully defined taking into account the exact contractual terms governing the transfer of information from the seller to the agent. We define the monetary value or willingness to pay (WTP) for information in terms of Hicks' (1939) measure of welfare change, as the *compensating variation* that moves the agent's expected utility with information to the same level as without information. To simplify the analytical treatment and segregate the "pure" value of information from any specific decision situation as far as possible, we will assume here that a noncontingent fixed payment p for an informative signal $\tilde{s}$ (correlated with the exogenous payoff-relevant random event $\tilde{x}$) is to be made *before* any uncertainty is resolved, and that neither the agent's actions nor her ex-post wealth are contractible.

More specifically, we consider an agent that has the option to invest a portion $a \in [0,1]$ of her endowed wealth $w > 0$ into a project of risky return $\tilde{x}$. The agent's preferences are supposed to be representable by a smooth utility function $u : \mathbb{R} \to \mathbb{R}$ that is strictly increasing and concave. The agent's optimal "default" or "no-information" strategy,

$$\hat{a} \in \arg\max_{a\in[0,1]} E_x \left[u \left(w(1 + a\tilde{x}) \right) \right], \tag{4.1}$$

exists by virtue of u being continuous over the compact action set $[0,1]$ (applying Weierstrass' Theorem) and is unique, since u is concave and strictly monotonic (i.e., one-to-one). The agent can now improve her *default utility*,

$$d(w) = E_x \left[u \left(w(1 + \hat{a}\tilde{x}) \right) \right], \tag{4.2}$$

from this project at least weakly by observing a random signal $\tilde{s}$, correlated with $\tilde{x}$, before choosing her optimal state-contingent action,

$$\hat{a}_s \in \arg\max_{a\in[0,1]} E_x \left[u \left((w-p)(1 + a\tilde{x}) \right) \middle| \, s \right], \tag{4.3}$$

where p represents the price of the signal.[4] The agent's expected utility *with* information, v, as a function of her investable wealth $w-p$ is therefore $v(w-p) = E_s \left[E_x \left[u \left((w-p)(1 + \hat{a}_s\tilde{x}) \right) \middle| \, \tilde{s} \right] \right]$. Hence we *define the value of information* for this agent as the compensating variation $\bar{p}$ that makes her indifferent between observing the signal or not,

$$v(w - \bar{p}) = d(w). \tag{4.4}$$

In other words, $\bar{p}$ is the highest amount that this agent would be willing to pay for obtaining the right to observe a realization of the signal $\tilde{s}$ prior to choosing her optimal action.

4.3 Cooperative Information Services

We consider now the situation where $N+1$ agents bargain about their respective contributions, $t_0, \ldots, t_N$, to a fixed net present investment $F(N)$ that is needed to set up a shared member-operated information service.[5] This shared service would allow each agent to observe a signal

[4] The the return of the agents' investments is assumed to be conditionally independent.

[5] To simplify the analysis, we assume that this investment has to be fully committed to ex ante, even though in reality it could be spread over time, realizing a real option value of flexibility through "chunkification" of the investment.

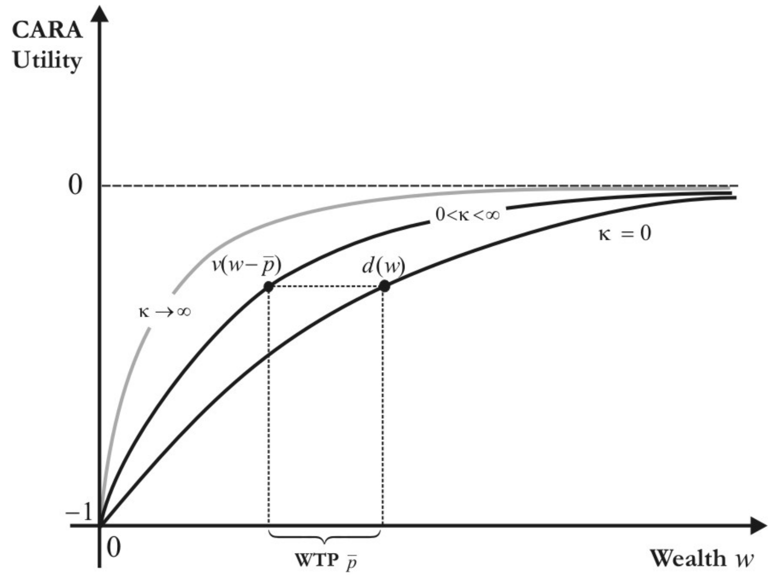

Figure 4.1: WTP is defined as the compensating variation $\bar{p}$ in (4.4).

correlated with her respective uncertain future payoff. We examine at first the value of the information service to a single user with CARA utility. Then we analyze the cost structure of a shared information service, which may give rise to both a "critical mass" and a maximal "carrying capacity," imposing a lower and (possibly) upper bound on the size of the information network. At last we provide a characterization of the Nash bargaining solution and examine the resulting *efficient* cost sharing and membership policies.

4.3.1 Value of the Information Service to a Single User

Let us focus on the case, where a given user has a CARA utility for wealth w of the form $u(w) = -\exp(-\rho w)$ and faces the decision to allocate a fraction $a \in [0,1]$ of her wealth into a risky project of return $\tilde{x}$. This return is assumed to be normally distributed with mean μ and variance σ^2, and her expected utility, given that she invests aw and retains $(1-a)w$, is

$$\begin{aligned} E\left[u\left(w(1+a\tilde{x})\right)\right] &= -\int_{-\infty}^{\infty} \exp\left[-\rho w(1+a\xi) - \frac{(\xi-\mu)^2}{2\sigma^2}\right] \frac{d\xi}{\sqrt{2\pi}\sigma} \\ &= -\exp\left[-\left(\rho w(1+a\mu) - \frac{(a\sigma\rho w)^2}{2}\right)\right], \end{aligned} \tag{4.5}$$

so that we obtain the unique no-information maximizer[6]

$$\hat{a} = \left[\frac{\mu}{\sigma^2 \rho w}\right]_{[0,1]}, \tag{4.6}$$

and consequently for the expected utility of the agent's default option

$$d(w) = \begin{cases} -\exp\left[-\left(\rho w + \mu^2/(2\sigma^2)\right)\right], & \text{if } w \geq \mu/\left(\sigma^2 \rho\right), \\ -\exp\left[-\left(\rho w(1+\mu) - (\sigma \rho w)^2/2\right)\right], & \text{otherwise.} \end{cases} \tag{4.7}$$

Now the agent observes the signal $\tilde{s} = \tilde{x} + \tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is uncorrelated with $\tilde{x}$ and normally distributed with mean zero and variance $1/\kappa$. The positive constant κ, corresponding to the inverse of the variance of $\tilde{\varepsilon}$, represents in Bayesian terms the "confidence" (or "precision") with which the signal informs the decision-maker about the realization of the exogenous random variable $\tilde{x}$. Using Bayes' rule one can compute the probability density of $\tilde{x}|s$,[7] which is again normally distributed with mean

$$E\left[\tilde{x}|\, s\right] = \left(\frac{\kappa\sigma^2}{1+\kappa\sigma^2}\right) s + \left(\frac{1}{1+\kappa\sigma^2}\right)\mu, \tag{4.8}$$

and variance

$$\mathrm{Var}(\tilde{x}|s) = \left(\kappa + \frac{1}{\sigma^2}\right)^{-1}. \tag{4.9}$$

The posterior mean $E[\tilde{x}|s]$ is a weighted average of signal realization s and prior mean μ. The relative weight of s is increasing in the confidence of the signal, κ. The posterior expected utility is then

$$E\left[u\left(w(1+\tilde{x}a)\right)\middle|\, s\right] = -\sqrt{\frac{1+\kappa\sigma^2}{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\rho w(1+a\xi) - \frac{\left(1+\kappa\sigma^2\right)\left(\xi - E[\tilde{x}|s]\right)^2}{2\sigma^2}\right) d\xi,$$

which can be maximized with respect to a as before in (4.5) to yield the optimal state-contingent policy

$$\hat{a}_s = \left[\frac{\kappa\sigma^2 s + \mu}{\sigma^2 \rho w}\right]_{[0,1]}. \tag{4.10}$$

Note that $\hat{a}_s$ adjusts the default strategy $\hat{a}$ in (4.6) with the additive state-contingent term $\kappa s/(\rho w)$, so that the investment intensity is increased, iff the observation s is positive.[8] Hence the investor only reduces her investment in the risky asset if she receives unambiguous "bad" news,

[6]The $[\cdot]_{[0,1]}$ operator is defined for any $\alpha \in \mathbb{R}$ as $[\alpha]_{[0,1]} = \max\{0, \min\{1, \alpha\}\} = (\min\{1,\alpha\})_+$, truncating the value of α to the admissible interval $[0,1]$.

[7]We denote by s a particular realization of the signal $\tilde{s}$.

[8]This adjustment is additive but *not* linear, since it occurs *inside* the $[\cdot]_{[0,1]}$ operator.

i.e., iff $s < 0$, which has a relatively low probability of occurrence. The correction of the default action $\hat{a}$ is proportional to the investor's confidence in the signal, κ, and inversely proportional to her Arrow-Pratt risk-aversion parameter ρ. As her risk aversion increases, the investor gets more reluctant to adjust her default action, and consequently the information contained in the observation loses value.[9] The agent's indirect utility contingent on the observation s is therefore (using (4.10))

$$v_s(w) = \begin{cases} -\exp\left[-\left(\rho w\left(1+\dfrac{\kappa\sigma^2 s+\mu}{1+\kappa\sigma^2}\right) - \dfrac{(\sigma\rho w)^2}{2\left(1+\kappa\sigma^2\right)}\right)\right], & \text{if } s \geq \bar{s}, \\ -\exp\left[-\left(\rho w + \dfrac{\left(\kappa\sigma^2 s+\mu\right)^2}{2\sigma^2(1+\kappa\sigma^2)}\right)\right], & \text{if } s \in [\underline{s}, \bar{s}], \\ -\exp\left[-\rho w\right], & \text{if } s \leq \underline{s}, \end{cases}$$

where we have set $\underline{s} = -\mu/\left(\kappa\sigma^2\right)$ and $\bar{s}(w) = \left(\sigma^2\rho w - \mu\right)/\left(\kappa\sigma^2\right)$ as the two critical observations, above and below which the agent invests respectively all or none of her wealth in the risky asset. In expectation, the agent obtains

$$\begin{aligned} E_s v_s(w) &= -e^{-\rho w}\left(F(\underline{s}) + \int_{\underline{s}}^{\bar{s}} f(\varsigma)\exp\left[-\frac{\left(\kappa\sigma^2\varsigma+\mu\right)^2}{2\sigma^2(1+\kappa\sigma^2)}\right]d\varsigma\right. \\ &\qquad \left. + \exp\left[-\frac{\mu+\sigma^2\rho w/2}{1+\kappa\sigma^2}\right]\int_{\bar{s}}^{\infty} f(\varsigma)\exp\left[-\frac{\kappa\sigma^2\rho w\,\varsigma}{1+\kappa\sigma^2}\right]d\varsigma\right), \end{aligned}$$

where

$$f(s) = \frac{\exp\left[-\dfrac{\kappa(s-\mu)^2}{2\left(1+\kappa\sigma^2\right)}\right]}{\sqrt{2\pi\left(\sigma^2+1/\kappa\right)}}$$

is the probability density of the signal $\tilde{s}$ based on (4.8)–(4.9), and F is the associated cumulative distribution function. This yields

$$\begin{aligned} v(w) &= -e^{-\rho w}\left\{F(\underline{s}) + \nu\exp\left(-\frac{\mu^2}{2\sigma^2}\right)\left[F\left(\frac{\bar{s}}{\nu}+\mu\right) - F\left(\frac{\underline{s}}{\nu}+\mu\right)\right]\right. \\ &\qquad \left. + \exp\left(-\frac{\mu+\sigma^2\rho w/2}{1+\kappa\sigma^2} - \frac{\sigma^2\rho w\left(\mu+\sigma^2\rho w/2\right)}{\sigma^2+1/\kappa}\right)\left[1-F\left(\bar{s}+\sigma^2\rho w\right)\right]\right\}, \end{aligned} \tag{4.11}$$

where we have set $\nu = \sqrt{1+\kappa\sigma^2}$ and $v(w) = E_s v_s(w)$. Provided that both μ and w are large enough, there is a very *simple approximate solution*, since then the interval $[\underline{s}, \bar{s}]$ comprises most of the probability mass. When concentrating on observations s plus or minus three standard

[9] In fact, the signal is most valuable to a risk-neutral investor who aggressively adjusts the investment policy to $\hat{a}_s \in \{0, 1\}$.

deviations, $\pm 3\sqrt{\sigma^2 + 1/\kappa}$, around the prior mean μ, more than 99 percent of the probability mass is concentrated in the critical interval $[\underline{s}, \bar{s}]$ and we can approximate,

$$v(w) \approx \frac{d(w)}{\sqrt{1 + \kappa\sigma^2}}, \tag{4.12}$$

and from there compute the agent's WTP, $\bar{p}$, for the signal by comparing $v(w - \bar{p})$ to her utility of the no-information strategy, $d(w)$, which yields

$$\bar{p} \approx \frac{\log\left(1 + \kappa\sigma^2\right)}{2\rho}. \tag{4.13}$$

A *sufficient condition* for the validity of this approximation is

$$\mu \in \left[\frac{3\sigma^2}{\sqrt{1 + \kappa\sigma^2}}, \frac{\rho w - 3\kappa\sqrt{\sigma^2 + 1/\kappa}}{1 + \kappa}\right], \tag{4.14}$$

guaranteeing an error in probability weight of less than one percent. This approximate solution is *not* valid if (i) endowed wealth w is small, (ii) signal confidence κ is very high, or (iii) the agent is almost risk-neutral (ρ close to zero), since then $\underline{s}$ and $\bar{s}$ will be close, i.e., a small variation in the observation within the interval $[\underline{s}, \bar{s}]$ induces a large change in the investment behavior. For such cases, the exact computation of the expected utility with information needs to take into account all terms as determined in (4.11).[10]

Special case: Perfect Information ($\kappa \to \infty$). Then $\underline{s} = \bar{s} = 0$ and the approximation (4.12) is *not* valid. From (4.11) we obtain the value of perfect information,

$$v(w)\Big|_{\kappa\to\infty} = -e^{-\rho w}\left[F(0) + \left(1 - F(\sigma^2\rho w)\right)\exp\left(-\rho w\mu - (\sigma\rho w)^2/2\right)\right], \tag{4.15}$$

which defines an *upper bound for the value of any signal*. It depends strongly on the wealth of the investor and is naturally zero for $w = 0$. No information service can rationally charge more for its information than is implied by (4.15) combined with (4.4).

4.3.2 Cost of Shared Information Provision

We suppose that the cost of the information network $F(N)$ contains a *fixed* portion F_0, independent of the number of users participating in the shared information service, and a *variable* component $c(N)$ that is strictly increasing in N.[11] If we further assume for simplicity that the

[10] Note that in this approximation we have effectively eliminated the dependence on the agent's wealth w in the expression for her WTP (4.13), but under condition (4.14) this approximation is justified.

[11] The cost function $c : \mathbb{R} \to \mathbb{R}$ is assumed to be smooth with $c' > 0$ and $c(0) = 0$.

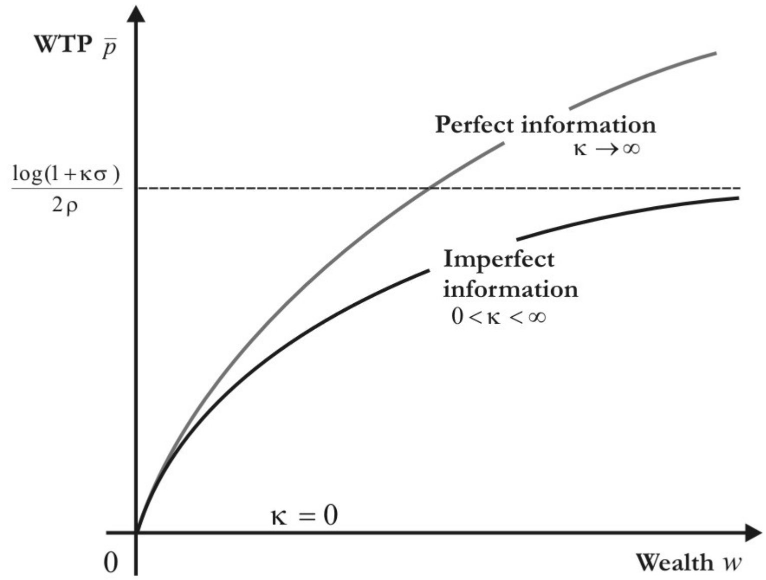

Figure 4.2: The WTP $\bar{p}$ can be approximated for w large enough by (4.13) (not possible for perfect information).

users of the information service are *identical* and each have a WTP of $\bar{p} < F_0$, as determined in (4.4), then for a *rational* creation of the service we need that

$$F(N) = F_0 + c(N) \leq (N+1)\bar{p}. \tag{4.16}$$

If the variable component is linear, $c = \gamma N$ for some nonnegative constant $\gamma < \bar{p}$, then the *critical mass* of users N_c can be computed explicitly,[12]

$$N_c = \frac{F_0 - \bar{p}}{\bar{p} - \gamma}. \tag{4.17}$$

If the variable cost component c is strictly concave in N (i.e., $c'' < 0$, due for instance to economies of scale), then there always exists a finite critical mass. The critical mass decreases if $\bar{p}$ is increasing in N due to positive network externalities (cf. footnote 12).[13] On the other hand, if c contains convexities (i.e., $c'' > 0$ somewhere), these may be balanced by the growth of $\bar{p}(N)$ in N so that even with high fixed cost and increasing complexity in maintaining a shared

[12] If $\gamma \geq \bar{p}$, then $N_c = \infty$, i.e. the shared information service stands no chance of creation, unless in some way (through additional assumptions) $\bar{p}$ increases in N. This could be the case as additional members create *positive* network externalities on other members by contributing information valuable to everyone.

[13] There may also exist *negative* network externalities, where the average confidence in the signal decreases as more and more agents join the shared information network, because of aggregation effects that decrease the customization value of the signal or because of *strategic* interaction between agents (cf. Section 4.3.4).

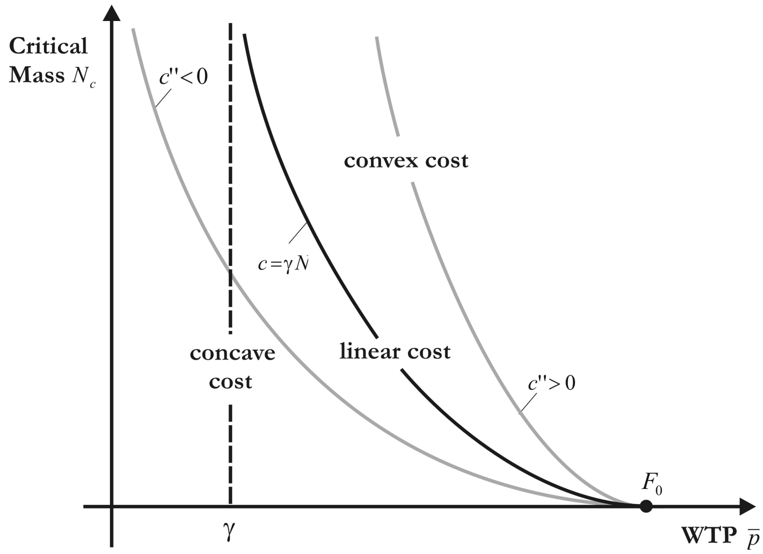

Figure 4.3: Critical Mass of the Shared Information Network as a Function of Agents' Willingness to Pay and Different Marginal Cost Functions.

information network, there may still exist a finite critical mass above which the investment in this shared venture is viable. Figure 4.3 provides an overview. A maximum "carrying capacity" can arise in the case of convex costs ($c'' > 0$) if $c > \bar{p}$ from some N on, since then each new member of the information service is willing to contribute less than the community has to pay for adding her.

4.3.3 The Bargaining Problem

In this section we derive an efficient apportioning of the cost of the shared information service among its members. Assume that there are $N + 1$ agents that are willing to bargain about their contributions $t_0, t_1, \ldots, t_N$, where

$$\sum_{i=0}^{N} t_i = F(N). \tag{4.18}$$

The bargaining solution is chosen here as the benchmark of what can be achieved using a coordinated approach to the shared allocation of costs to the members of a shared information network. Harsanyi (1966) notes that such "[a] cooperative game can always be replaced by a noncooperative game if we incorporate promises and threats in the strategies available to the players" (p. 616). The threats in the cooperative negotiation game correspond to the players' ability to interrupt the bargaining process by forcing the default outcome, an extreme action that is not generally individually rational (cf. Roth 1977) given that entry into the bargaining round was free.[14] More specifically, we assume that each agent $i \in \mathcal{I} = \{0, 1, \ldots, N\}$ has an endowed wealth of $w_i > 0$, is of constant absolute risk aversion with coefficient $\rho_i > 0$, and pursues a risky project of normal return $\tilde{x}_i$ with mean μ_i and variance σ_i^2. Furthermore, she has confidence $\kappa_i > 0$ in the personalized signal $\tilde{s}_i = \tilde{x}_i + \tilde{\varepsilon}_i$, where the error term $\tilde{\varepsilon}_i$ has mean zero and is statistically independent from the project return $\tilde{x}_i$.[15] If the agent disagrees with a proposed outcome she can force the default outcome $d_i(w_i) \in \mathbb{R}^{N+1}$ as previously determined in (4.2) and substituting agent i's parameters. The compact, convex and nonempty set $\mathcal{S} \subset \mathbb{R}^{N+1}$ contains all the feasible utility payoff vectors for the agents in $\mathcal{I}$.[16] Nash (1953) defined a solution to the bargaining game as a function $\mathbf{b} : \mathcal{S} \to \mathbb{R}^{N+1}$, such that $\mathbf{b}(\mathcal{S}, \mathbf{d}) \in \mathcal{S}$ for all possible bargaining situations $(\mathcal{S}, \mathbf{d})$, where $\mathbf{d} = (d_0, \ldots, d_N)$. Given the four axioms (cf. Roth (1979), pp. 6–8) (i) *Independence of Equivalent Utility Representations*, (ii) *Symmetry*, (iii) *Independence of Irrelevant Alternatives*, and (iv) *Pareto Optimality*,[17] there exists a unique solution $\mathbf{v} = \mathbf{b}(\mathcal{S}, \mathbf{d})$, such that[18]

$$\mathbf{v} = \arg\max_{\hat{\mathbf{v}} \geq \mathbf{d}} \prod_{i=0}^{N} (\hat{v}_i - d_i), \tag{4.19}$$

where $\mathbf{v} = (v_0, v_1, \ldots, v_N)$ and $v_i = v_i(w_i - t_i) = E_s\left[v_{is}(w_i - t_i)\right]$ is defined in the same manner as the expected utility for the individual agent in (4.11). Note that the bargaining set $\mathcal{S}$ is defined by (4.18).

Proposition 4.1 (Cooperative Allocation) *Consider the bargaining problem* $(\mathcal{S}, \mathbf{d})$.

(i) The solution to (4.18)–(4.19) is uniquely characterized by

$$\frac{v_j'(w_j - t_j)}{v_j(w_j - t_j) - d_j(w_j)} = \frac{v_0'(w_0 - t_0)}{v_0(w_0 - t_0) - d_0(w_0)}, \quad j = 1, \ldots, N, \tag{4.20}$$

[14] It may appear unrealistic that any of the players could carry so much weight as to stop the formation of the shared information service. However, the preceding remarks should clarify that this does not generally happen, given the assumption that the players have agreed to bargain about their contribution in the first place. Individual rationality will generally prevent firms from carrying out their threat, which is a mere modeling device to support the efficiency achieved by the Nash bargaining solution (cf. also footnote 17). If a particular member indeed decided to exit the N-firm bargaining, negotiations would subsequently resume with $N - 1$ firms.

[15] Each agent $i \in \mathcal{I}$ is only observing her own customized signal $\tilde{s}_i$.

[16] We further assume that all the projects' payoffs are uncorrelated and that the agents do not form coalitions.

[17] The Pareto-optimality axiom can be relaxed to mere "individual rationality" as shown by Roth (1977). In that case the solution comprises the Nash solution *and* the default option $\mathbf{d}$.

[18] Nash (1953) shows this only for two players. The generalization to N players (without coalitions) is trivial as Roth (1979) points out.

together with

$$t_0 = F(N) - \sum_{i=1}^{N} t_i. \tag{4.21}$$

(ii) If all agents have the same risk aversion ρ and the approximation (4.12)–(4.13) holds for all $i \in \mathcal{I}$, then the optimal cost sharing for the common information service is given by

$$t_j(N) = \frac{1}{N}\left(F(N) - \frac{\log Q}{\rho}\right) + \frac{1}{\rho}\log\left(\frac{q_j}{q_0}\right), \quad j = 1, \ldots, N, \tag{4.22}$$

and $t_0(N) = (\rho F(N) - \log Q)/(\rho N)$, where $Q = q_1 q_2 \cdots q_N / q_0^N$ and $q_i = \sqrt{1 + \kappa_i \sigma_i^2}$ for all $i \in \mathcal{I}$.

(iii) If the risk aversion ρ_i of agent $i \in \mathcal{I}$ increases, then – all else being equal – this agent's share t_i increases, iff

$$\frac{\partial \log\left(v_i(w_i - t_i) - d_i(w_i)\right)}{\partial \rho_i} < \frac{\partial \log v_i'(w_i - t_i)}{\partial \rho_i}. \tag{4.23}$$

Proposition 4.1 characterizes the efficient Nash bargaining solution both generally (in terms of a first-order condition), and explicitly for an important special case where all the agents' risk-aversion parameters ρ_i are identical and the approximation condition (4.14) holds. For differing risk-aversion parameters a general result by Kihlstrom et al. (1981) states that bargaining performance decreases with increasing risk aversion. We find that the agent's absolute monetary contribution to the shared information service does *not* need to increase as a consequence. Indeed, the situation of information procurement produces a countervailing effect, since the value of information generally *decreases* with increasing risk aversion.[19] The *net* change of an agent's contribution is positive if and only if – as a response to an increase of her risk aversion – her demand for information decreases faster than her bargaining power relative to the other agents.[20]

4.3.4 Effect of Network Externalities

Externalities in information networks do occur as a result of (possibly mediated) interactions of different network members. These externalities can have both positive and negative effects

[19] For instance, a risk-averse agent will respond less aggressively to a (imperfect) signal than a risk-neutral investor, and thus is willing to pay less for the information.

[20] "Bargaining power" is meant here in the sense of Kihlstrom et al. (1981) who show – as pointed out in the main text – that agents with lower risk aversions obtain a larger surplus in ccoperative bargaining situations and therefore have a higher implicit bargaining power. For an *explicit* modeling of bargaining power, see footnote 27.

on the members' perceived utility. Riggins et al. (1994) examine the growth of interorganizational systems which can be both hampered and facilitated by differences in perceived network externalities across the different users. *Positive* externalities may be a result of the benefit that comes from sharing information in a nonstrategic way insofar as it increases the signal quality of individual users. On the other hand, *negative* externalities can be a consequence of high user traffic and network capacity limitations that may it harder to retrieve usable signals and thus indirectly reduce the confidence in these signals. Negative externalities may also result from a *strategic* use of information by different members of the same industry against each other.[21] To incorporate such effects in our model, we assume that the firms' respective confidence κ in an informative signal retrieved from the shared information service depends on the number of users N of the information service. For simplicity we thereby assume that $\kappa(N) \geq 0$ is quasiconcave in N or in other words that $\kappa(N)$ is either monotonic or single-peaked in N. For any given network size N, the results of the previous sections remain unaffected; however, incorporating (possibly strategic) network effects may have a dramatic effect on the critical mass and maximum carrying capacity of the shared information network. Under the assumption that (4.14) holds we obtain, by combining (4.13) with (4.16), for the critical mass under network externalities:[22]

$$N_c^e = \min \left\{ N \geq 1 : F_0 + c(N) \leq \frac{(N+1) \log \left(1 + \kappa(N)\sigma^2\right)}{2\rho} \right\}. \tag{4.24}$$

For the simple case of a linear network cost expansion path with $c(N) = \gamma N$ one can directly see that any strict increase $\hat{\kappa} > \kappa$ yields a (weak) decrease in the critical mass, as the right-hand side of (4.17) is strictly decreasing in $\bar{p}$ which in turn is strictly increasing in κ. This naturally

[21]Such "strategic externalities" are not necessarily negative. They tend to be positive – increasing the incentive of industry participants to share information – when firms' best responses to their actions are strategic complements (cf. Vives 1984, Raith 1996).

[22]To keep the formalism simple, we have here implicitly assumed that the confidences and perceived network externalities are uniform across all firms. Of course there is no reason for this to be true in any particular situation. In addition to a more complicated form of the expressions for the critical mass due to aggregation of the heterogeneous firms with different quasiconcave confidences $\kappa_i(N)$, $i = 0, \ldots, N$, the fundamental difference is that the feasibility of the shared investment and the surplus of the different players will depend on exactly *who* is bargaining and not just on the number of participants. Firm heterogeneity in perceived network effects generally results in nonmonotonic surplus in the number of participants. To deal with full firm heterogeneity in a proper fashion it is necessary to endogenize firm entry into the cooperative bargaining situation, an extension that is for space constraints beyond the scope of this paper. It naturally involves determining the fulfilled-expectations equilibria of which there may be several so that coordination and preannouncements may become important dimensions of the problem (cf. Farrell and Saloner 1986).

generalizes to any network cost function considered in the previous section. For convex c, the quasiconcavity of $\kappa(\cdot)$ ensures that the critical mass in (4.24) is unique.

4.3.5 Efficient Cost Sharing and Membership Policies

The Nash bargaining solution is efficient provided that all of the bargaining members need to *either* agree on a common resource allocation *or* force the disagreement outcome $\mathbf{d}$ (also sometimes appropriately referred to as the "threat point," cf. footnote 14). A sequential enlargement of the membership body does not fulfill this requirement; and naturally the existing members wish to enlarge their user base such that each of their contributions are nonincreasing, which may result in the exclusion of certain agents, if they did not belong to the original network.

4.4 Monopolist Information Services

4.4.1 Pricing and Network Size

Consider now a monopolist that decides about entering a market to sell customized information to a potential client base of M agents. We assume that each agent $m \in \mathcal{M} = \{1, \dots, M\}$ disposes about enough wealth and is sufficiently risk averse, so that (4.14) holds and each agent's reservation price is given by (4.13). Agents are heterogeneous in their risk aversions, $\rho_1, \rho_2, \dots, \rho_M$, where $\rho_m = m\bar{\rho}/M$ for $m \in \mathcal{M}$, and $\bar{\rho}$ is a positive constant.[23] Corresponding to these different risk aversions are the reservation prices $p_1 < p_2 < \cdots < p_M$, where by (4.13) and by the definition of ρ_m we have

$$p_m = \frac{\log\left(1 + \kappa_m \sigma_m^2\right)}{2\rho_m} = \frac{M \log\left(1 + k\right)}{2m\bar{\rho}}.$$

Thereby we have assumed that the customization of the information is such that the "risk-adjusted confidence" $\kappa_m \sigma_m^2$ is constant and equal to k for all $m \in \mathcal{M}$.[24] If (for simplicity) the marginal cost of providing one additional agent with the customized information is *constant* (i.e., $c(N) = \gamma N$), we can provide closed-form solutions for the monopolist's optimal price, optimal profits, and market penetration.

[23] In other words, we assume here that there is an upper bound for the risk aversion, and that the M agents are uniformly distributed on the interval $[\bar{\rho}/M, \bar{\rho}]$.

[24] In that sense the information is horizontally versioned in such a way that the product of confidence in the signal and variance of the risky project is the same across all agents (i.e., if a project is more risky, then the agent can be less sure about the relevance of the signal realization, which corresponds to the notion that the monopolist has "constant" marginal cost to satisfy each additional user of the information network).

Proposition 4.2 (Monopolistic Allocation) *Assume that the monopolist rationally enters the market. Then the optimal price is $p^* = 2\gamma$ and optimal profits are*

$$\Pi^* = \frac{M \log(1+k)}{4\bar{\rho}} - F_0, \tag{4.25}$$

with a market penetration (and resulting network size) of

$$m^* = M \cdot \min\left\{1, \frac{\log(1+k)}{4\gamma\bar{\rho}}\right\}. \tag{4.26}$$

The monopolist rationally enters the market iff the critical mass M_c of *potential* users is reached ($M \geq M_c$), so that optimal profits Π^* in (4.25) are nonnegative, where $M_c = 4\bar{\rho}F_0/\log(1+k)$. If the monopolist enters the market (the likelihood of which is inversely proportional to the dispersion $\bar{\rho}$ in the agents' risk aversions), then he in general tends to *restrict* the size of the information network to maximize his profits. Those agents with high risk aversions also have a low WTP and are therefore excluded from the use of the network. Especially when fixed costs F_0 are high, this can lead to market failure, i.e., the monopolist does not enter the market, even though a bargaining solution (which cross-subsidizes the low-value agents) exists, which is the case whenever M is larger than the minimal network size, which we term the "critical Nash network size."[25]

Proposition 4.3 (Welfare Differentials and Critical Mass) *Assume that the approximation condition (4.14) is satisfied for each agent.*

(i) The total welfare achieved in a non-discriminating monopolistic environment is[26]

$$W = \begin{cases} \dfrac{M \log(1+k)}{2\bar{\rho}} \left[\dfrac{1}{2} + \displaystyle\sum_{m=1}^{m^*} \dfrac{1}{m}\right] - 2\gamma m^* - F_0, & \text{if } M \geq M_c, \\ 0, & \text{otherwise.} \end{cases} \tag{4.27}$$

whereas the welfare achieved in the Pareto-optimal Nash bargaining solution is given by

$$\bar{W} = \begin{cases} \dfrac{N \log(1+k)}{2\bar{\rho}} \left(\displaystyle\sum_{n=1}^{N} \dfrac{1}{n}\right) - \gamma N - F_0, & \text{if } N \geq N_c, \\ 0, & \text{otherwise,} \end{cases} \tag{4.28}$$

[25] The *critical Nash network size* is simply the smallest subset of a given universe of agents that can individually rationally set up an information network (of a given cost structure).

[26] The finite sum can also be expressed in "closed form," using the relation $\sum_{k=1}^{K} 1/k = \Psi(K+1) + \gamma_e$, where Ψ is the Digamma function, related to the standard Gamma function Γ by $\Psi(x) = d \log \Gamma(x)/dx$, and $\gamma_e \approx .5772$ is Euler's constant.

(ii) *The critical number of potential users of the monopoly shared information service, M_c, and the critical Nash network size N_c are related as follows:*

$$N_c = \frac{2\bar{\rho}F_0}{\log(1+k)\left(\sum_{n=1}^{N_c} 1/n\right) - 2\bar{\rho}\gamma} = \frac{M_c}{2\left(\sum_{n=1}^{N_c} 1/n\right) - \gamma\bar{\rho}/\log(1+k)}. \tag{4.29}$$

(iii) *The critical Nash network size N_c is the smallest possible network size of any shared information network that guarantees individually rational membership, in particular $N_c \leq m^* \leq M_c$.*

For small marginal costs γ, the critical Nash network size is less than half as large as M_c, since using (4.29) we have then:

$$N_c \leq \frac{M_c}{2 - \gamma\bar{\rho}/\log(1+k)} \approx \frac{M_c}{2}. \tag{4.30}$$

However, M_c just represents the consumer base necessary for the monopolist to serve this heterogeneous market. The actual size of the monopolist's network, m^*, is smaller than that. But still joining the network has to be for each user individually rational, not allowing for a cross-subsidization of low-value agents by high-value agents, as is generally the case with a Pareto-optimal bargaining outcome.

4.4.2 Sources of Inefficiency

The monopolistic solution is market based, unlike the bargaining equilibrium, which presupposes that each player is telling the truth about her confidence in the signal and that there is common knowledge about the utility functions. In practical situations, an efficiency loss due to asymmetries in the information can be expected, but nevertheless *if* all parties can agree on their respective parameter vectors $(\rho, \kappa, \mu, \sigma)$ and costs for the service are known to be of the form (4.16), then any disagreements about splitting the investment burden F is resolved by the cooperative allocation described in Proposition 4.1.[27] Thus, even though there is a *nominal* efficiency loss by implementing the market-based solution it may actually be preferable, whenever the expected differences in welfare are small.

[27] An additional hidden assumption is that the agents all have equal (explicit) bargaining power (cf. also footnote 20). Situations with asymmetric bargaining power can be dealt with in a manner already suggested by Nash (1953) by introducing positive exponents β_i, $i \in \mathcal{I}$ into the Nash function (4.19) proportional to the agent's bargaining power, such that $\beta_1 + \beta_2 + \ldots + \beta_N = N$. This modification makes the model's prescribed sharing rule potentially contentious whenever not all β_i's are equal to one.

4.5 Discussion

We have examined shared information services from both a cooperative bargaining viewpoint and the perspective of a monopolist supplier. To approximate the *efficient* outcome generated by Nash bargaining, the pricing of information should be contingent on the agent's confidence in the signal, her project risk, and her risk aversion. Even though each agent's demand for information decreases as her risk aversion increases, in a cooperative bargaining situation her contribution share may actually increase, an effect that a market-based mechanism implemented by a monopolistic information supplier cannot exploit. We have also shown that wealth effects are significant for low-wealth agents, due to the strong impact of a prepayment on their investable wealth. This suggests that pricing should be significantly different for agents with a low endowed wealth, but should be essentially independent of wealth after a certain threshold (given by condition (4.14)) is reached. We also discussed the *critical mass* of users necessary for the creation of shared information services in the presence of network externalities. Clearly any prospective information provider or cooperating group of agents should verify that they meet these minimum requirements before setting up the information service in order to avoid failure. We identified the ingredients of an efficient pricing scheme and a (future) manager of a shared information service network, given that it is somehow possible to isolate the salient agent characteristics (e.g., their confidence in the informative signal), knows *how* an efficient pricing scheme can be structured to maximize welfare or – in the case of monopolistic information provision – profits.

Future theoretical contributions could evolve along the following three axes: (a) to endogenize entry in the presence of uncertainty about the costs of the shared information network and examine the associated fulfilled-expectations equilibria; (b) to examine monopolistic information selling mechanisms that involve various screening techniques based on versioning; (c) to introduce informational asymmetries in the bargaining solution to capture parameter uncertainty and differences in bargaining power among the agents.

4.6 References

Athey, S. 2000. Investment and information value for a risk averse firm. Working Paper, Massachusetts Institute of Technology.

Clarke, R.N. 1983. Collusion and the incentives for information sharing. *Bell Journal of Economics* **14**(2) 383–394.

FARRELL, J., SALONER, G. 1986. Installed base and compatibility: innovation, product preannouncements, and predation. *American Economic Review* **76**(5) 940–955.

HARSANYI, J.C. 1966. A general theory of rational behavior in game situations. *Econometrica* **34**(3) 613–634.

HICKS, J. 1939. *Value and Capital*. Clarendon Press, Oxford, UK.

KIHLSTROM, R.E. 1974. A general theory of demand for information about product quality. *Journal of Economic Theory* **8**(4) 413–439.

KIHLSTROM, R.E., A.E. ROTH, SCHMEIDLER, D. 1981. Risk aversion and solutions to Nash's bargaining problem, in O. Moeschlein and D. Pallaschke, eds., *Game Theory and Mathematical Economics*. North-Holland, Amsterdam, NL.

LIEBOWITZ, S.J., MARGOLIS, S.E. 1994. Network externalities: an uncommon tragedy. *Journal of Economic Perspectives* **8**(2) 133–150.

MARSCHAK, J., MIYASAWA, K. 1968. Economic comparability of information systems. *International Economic Review* **9**(2) 137–173.

NASH, J.F. 1953. Two-person cooperative games. *Econometrica* **21**(1) 128–140.

NOVSHEK, W., SONNENSCHEIN, H. 1982. Fulfilled expectations Cournot duopoly with information acquisition and release. *Bell Journal of Economics* **13**(1) 214–218.

RAIFFA, H., SCHLAIFER, R. 1961. *Applied Statistical Decision Theory*, Boston, MA: Division of Research, Graduate School of Business Administration, Harvard University.

RAITH, M. 1996. A general model of information sharing in oligopoly. *Journal of Economic Theory* **71**(1) 260–288.

RIGGINS, F.J., KRIEBEL, C.H., MUKHOPADHYAY, T. The growth of interorganizational systems in the presence of network externalities. *Management Science* **40**(8) 984–998.

Roth, A.E. 1977. Individual rationality and Nash's solution to the bargaining problem. *Mathematics of Operations Research* **2**(1) 64–65.

Roth, A.E. 1979. *Axiomatic Models of Bargaining*, Springer, Berlin.

Shapiro, C., Varian, H.R. 1998. Versioning: the smart way to sell information. *Harvard Business Review* **76**(6) 107–114.

Stahl, S. 2001. Sharing data is key to antiterrorism efforts. *Informationweek*, 4th October, **858** 8.

Treich, N. 1997. Risk tolerance and value of information in the standard portfolio model. *Economics Letters* **55**(3) 361–363.

Vives, X. 1984. Duopoly information equilibrium, Cournot and Betrand. *Journal of Economic Theory* **34**(1) 71–94.

Wang, E.T.G., Seidmann, A. 1995. Electronic data interchange: competitive externalities and strategic implementation policies. *Management Science* **41**(3) 401–418.

4.7 Appendix: Proofs

Proof of Proposition 4.1. (i) Define $\bar{u}_i(t_i) = v_i(w_i - t_i)$ for all $i \in \mathcal{I}$. Then using (4.18) the first-order optimality conditions for the problem (4.19) can be written as[28]

$$-\bar{u}_0'\left(F - \sum_{i=1}^{N} t_i\right)\prod_{i=1}^{N}(\bar{u}_i(t_i) - d_i) + \bar{u}_j'(t_j)\prod_{(j\neq)i=0}^{N}(\bar{u}_i(t_i) - d_i) = 0, \quad j = 1,\ldots,N,$$

which is equivalent to (4.20)–(4.21). (ii) From (4.12)–(4.13) and (4.20) we obtain

$$\frac{-\rho d_j(w_j - t_j)}{d_j(w_j - t_j) - \sqrt{1+\kappa_j\sigma_j^2}\,d_j(w_j)} = \frac{-\rho d_0(w_0 - t_0)}{d_0(w_0 - t_0) - \sqrt{1+\kappa_0\sigma_0^2}\,d_0(w_0)},$$

[28] The second-order conditions are satisfied as a consequence of the concavity of $\bar{u}_i$ for all $i = 0,\ldots,N$.

which, after using the multiplicative separability of the CARA utilities implying that $d_j(w_j - t_j) = e^{\rho t_j} d_j(w_j)$ for all $i \in \mathcal{I}$, is equivalent to

$$q_j e^{-\rho t_j} = q_0 e^{-\rho t_0}, \quad j = 1, \ldots, N,$$

where we have set the q_i's as in the proposition. From the last relation we can conclude using (4.21) that

$$t_j = t_0 + \frac{1}{\rho} \log \frac{q_j}{q_0}, \quad j = 1, \ldots, N. \tag{4.31}$$

Summing up the t_j's from 1 through to N one obtains t_0 as given in the proposition. Substituting this expression for t_0 into (4.31) we obtain (4.22). (iii) Let $i, j \in \mathcal{I}$ and, without loss of generality, $i \neq 0$. If we substitute (4.21) into the first-order condition (4.20) and differentiate both sides with respect to ρ_i we obtain (using the abbreviation $\bar{u}_i(t_i) = v_i(w_i - t_i)$ as above) [29]

$$\left(\frac{\bar{u}_j''}{\bar{u}_j - d_j} - \left(\frac{\bar{u}_j'}{\bar{u}_j - d_j} \right)^2 \right) t_j' + \Delta_{ij} = \left(\frac{\bar{u}_0''}{\bar{u}_0 - d_0} - \left(\frac{\bar{u}_0'}{\bar{u}_0 - d_0} \right)^2 \right) t_0', \tag{4.32}$$

where we have set $t_j' = \partial t_j / \partial \rho_i$, $\Delta_{ij} = 0$ for $j \neq i$, and

$$\Delta_{ii} = \frac{\partial_{\rho_i} \bar{u}_i'}{\bar{u}_i - d_i} - \bar{u}_i' \, \frac{\partial_{\rho_i} (\bar{u}_i - d_i)}{(\bar{u}_i - d_i)^2}.$$

Differentiating (4.21) with respect to ρ_i it is $t_0' = -\sum_{k=1}^N t_k'$, and thus after summation of (4.32) for $j = 1, 2, ..., N$ we find (using (4.20)) that

$$\left(\frac{\bar{u}_0''}{\bar{u}_0 - d_0} - \left(\frac{\bar{u}_0'}{\bar{u}_0 - d_0} \right)^2 \right) t_0' = - \frac{t_i'}{\sum_{k \neq i} \left[\frac{\bar{u}_k''}{\bar{u}_k - d_k} - \left(\frac{\bar{u}_0'}{\Delta_0} \right)^2 \right]^{-1}}.$$

Substituting this last relation into (4.32) for $j = i$ we obtain an explicit expression for t_i':

$$\frac{\partial t_i}{\partial \rho_i} = -\Delta_{ii} / \left\{ \frac{\bar{u}_0''}{\bar{u}_0 - d_0} - \left(\frac{\bar{u}_0'}{\bar{u}_0 - d_0} \right)^2 + 1 / \left(\sum_{k \neq i} \left[\frac{\bar{u}_k''}{\bar{u}_k - d_k} - \left(\frac{\bar{u}_0'}{\Delta_0} \right)^2 \right]^{-1} \right) \right\}. \tag{4.33}$$

Since $\bar{u}_j$ is concave and strictly increasing, the bracketed expression in (4.33) is negative, so that t_i' is positive iff $\Delta_{ii} > 0$. In other words, t_i' is positive iff (4.23) holds, which completes the proof of Proposition 4.1. ∎

[29] Here we omit the arguments for simplicity. Note that the arguments in for utilities with the "zero" index are $(F - t_1 - t_2 - \cdots - t_N)$.

Proof of Proposition 4.2. The monopolist's profit function is

$$\Pi(p) = \left(\sum_{m=1}^{M} \mathbf{1}_{\{p_m \geq p\}} \right) (p - \gamma) - F_0,$$

where $\mathbf{1}$ denotes the indicator function. Since there are only a finite number of agents, maximizing profits is equivalent to finding $m^* \in \mathcal{M}$, such that $\Pi(p_{m^*}) = \max_p\{\Pi(p)\}$, i.e.,

$$m^* \in \arg\max_{m \in \mathcal{M}} \left\{ m \left(1 - \frac{\gamma}{p_m} \right) \right\}, \tag{4.34}$$

which yields (4.26) as the unique solution. Hence the optimal price is $p^* = 2\gamma$ and profits are given by (4.25). ■

Proof of Proposition 4.3. (i) Let us first compute the consumer surplus in the monopolistic situtation,

$$\mathrm{CS} = \sum_{m=1}^{m^*} (p_m - p_{m^*}) = \frac{M \log(1+k)}{2\bar{\rho}} \left(\sum_{m=1}^{m^*} \frac{1}{m} \right) - 2\gamma m^* \geq 2\gamma m^* \left(\sum_{m=2}^{m^*} \frac{1}{m} \right), \tag{4.35}$$

where the last inequality binds, whenever the *unconstrained* solution of (4.34) lies inside $[0, M]$, i.e., whenever the monopolist actively *excludes* potential users from the network. Adding CS in (4.35) and (4.25) we obtain the expression (4.27). On the other hand, the welfare $\bar{W}$ (in monetary terms) for the Pareto-optimal case can be – given that the approximation condition (4.14) holds – simply determined by adding up the WTPs (4.13) over all agents $i = 1, \ldots, N$, and subtracting the cost of the network $F_0 + \gamma N$,

$$\bar{W} = \log(1+k) \sum_{i=1}^{N} \frac{1}{2\rho_m} - \gamma N - F_0 = \frac{N \log(1+k)}{2\bar{\rho}} \sum_{n=1}^{N} \frac{1}{n} - \gamma N - F_0,$$

as long as $N \geq N_c$ and zero otherwise, which yields (4.28). (ii) The critical number M_c of potential users of the monopolistic network have been computed in the main text exactly offsetting the monopolist's total cost. The Nash network size is analogously determined by setting $\bar{W}$ to zero. From this, equation (4.29) follows directly. (iii) Clearly, since the Nash bargaining solution is welfare maximizing, and at N_c each agent is as low in her utility that removing a single agent would prompt at least one user to force the default outcome $\mathbf{d}$. ■

"One and one make two" assumes that the changes in the shift of circumstance are unimportant. But it is impossible for us to analyse this notion of "unimportant change." We have to rely upon common sense.
— Alfred North Whitehead

Chapter 5

Multiattribute Product Differentiation

5.1 Introduction

Information goods such as computer software or electronic newspapers can be provided by firms at a low marginal cost, while in many cases large capital outlays are required to produce their first unit. The substantial setup cost is thereby mainly driven by the cost of developing the top quality product. Having established this "flagship" product a firm can degrade it or in other ways modify it, and in this way create a multitude of products with different attributes at a small "versioning" cost. Information goods are not the only products that enjoy such cost complementarities in development. Component standardization and design based on common platforms can be found in a wide range of industries (Ulrich 1995). In addition, information goods are often bundled in with physical goods determining in part the functionality of the underlying physical product. Therefore finding *optimal* versioning policies is becoming increasingly important, as often lower distribution costs and newfound customer intimacy render intricate second-degree price discrimination strategies feasible. We present a model of a firm that chooses a two-product portfolio sequentially, when products can differ with respect to both horizontal and vertical nonprice attributes. The firm determines its second, or *versioned product* after it has completed the development of its *flagship product*, and uncertainty over the market acceptance of the latter has resolved.

It is well known that component sharing or in a broader sense cost complementarities between products can induce a firm to increase its product variety (Fisher et al. 1999) by keeping the number of components necessary for their assembly manageable. On the other hand, over time the creation of product variety should be commensurate with the information available about demand, leading to the practice of delayed differentiation (Lee and Tang 1997, Anand and Mendelson 1998, Swaminathan and Tayur 1998). Cost complementarities and opportunities for delayed differentiation are especially large for information goods: provided a sufficiently modular product design, features can be easily disabled or rebundled leading to vertically and horizontally different versions of the initial flagship product (Shapiro and Varian 1998). Since most of the development effort goes into the design of the top quality product, the optimal initial research effort and therefore ultimately the choice of the flagship product's quality level depends on the firm's options of creating a versioned product once the market demand for the first product has been observed.[1] While this paper is not so much concerned with the timing of product introductions (Moorthy and Png 1992), we are interested in finding an optimal segmentation of a heterogeneous customer base of multiple characteristics both with and without delayed differentiation. We derive the option value for delaying the versioning decision and show that investment in product quality does not not have to be higher when this option is available.

All consumers are initially endowed with a wealth level and a taste. Each consumer of a certain (*taste*,*wealth*)-type has a utility, which depends on price, product quality, and distance of the horizontal product attribute from the own taste. Quality is thereby a vertically differentiating instrument and can be thought of as performance or product breadth. If a product is characterized as a bundle of features (Lancaster 1966), then a product including more features than another is of higher quality. On the other hand, two products with different bundles of features, not distinguished in terms of performance or overall product breadth, can be seen as horizontally differentiated. Each of them appeals to consumer tastes "located" sufficiently close. Given a distribution of consumer types (here assumed to be uniform for simplicity), generated for instance by usage preferences or existing standards, the firm can choose product attributes so as to attract as many of the consumers with sufficiently high wealth (i.e., potential buyers) as possible. More specifically, the goal of the risk-neutral multiproduct firm is to create a certain number of product offerings of the form *(price, quality, horizontal product attribute)*, such that expected profits are maximized. Each product offering generally targets a certain subset or *segment* of the (*taste*,*wealth*)-space, into which the consumers self-select by choosing the product that maximizes their respective utilities. However, the firm's ability

[1] As an example, Microsoft's Office 2000 software was available in at least four different versions (Premium, Professional, Small Business, and Standard) essentially to segment the market and thereby screen the heterogeneous consumer base. These versions contained both vertically (i.e., more vs. less features) and horizontally (i.e., disjoint features) differentiating attributes.

to optimize the choice of its portfolio is limited by imperfect information about the consumer characteristics at the outset. We assume that this subjective uncertainty gets resolved after the creation of its flagship product, which can typically be used to evaluate demand. This creates an option value of delaying differentiation to reduce the firm's partially irreversible commitment to a particular product portfolio.

The underlying problem of multiattribute product differentiation is highly nonconvex. We show that both pure horizontal and pure vertical versioning are *locally* optimal, whereas *mixed versioning*, i.e., the simultaneous differentiation along both horizontal and vertical nonprice attributes, is never optimal in our quasilinear homogeneous setting. Pure vertical versioning is *globally* optimal for relatively low development costs, whereas for high development costs pure horizontal versioning is superior. Under delayed differentiation the optimal policy is contingent on the demand realization ("state"). A consequence of this added flexibility might be that the firm's optimal ex-ante investment in product development may drastically drop, instead of increase as a result of diminished (strategic) irreversibility of the upfront sunk cost.

5.1.1 Literature Review

Beginning with Hotelling's (1929) seminal paper on horizontal competition, numerous contributions have been made to product differentiation. The corresponding literature can be divided into locational models in the tradition of Hotelling, where each firm is attributed an "address" in product space, and into so-called "non-address" models in the spirit of Chamberlin's (1933) monopolistic competition, where a representative consumer exhibits (probabilistic) preferences for different products.[2] An important distinction between the two groups of models is that in the latter group, *each* product is competing with each other, while in the former consumers are truly heterogeneous in their preferences, and some products may have no overlap, i.e., may never be in direct competition.[3] In this paper we adopt the locational approach that in our view better captures consumer heterogeneity and allows the explicit consideration of participation constraints that inevitably arise when dealing with a spatial distribution of endowed unobservable consumer characteristics. Our model is thereby inspired by Salop's (1979) "circular city" that we extend to a "cylinder" by adding a vertical product characteristic.

Mussa and Rosen (1978), based on earlier work by Mirrlees (1971), essentially founded the line of work on second-degree price discrimination of multiproduct firms with quality as

[2]For a good bibliography see Anderson et al. (1992), Beath and Katsoulacos (1991), as well as Tirole (1988).

[3]A notable exception in this dichotomy is the model by Perloff and Salop (1985) that combines characteristics from both locational and non-address models, driven by symmetry assumptions in the preferences of a representative consumer that is faced with localized products.

the differentiating instrument. Cremer and Thisse (1991) demonstrate that in many setups horizontal differentiation can be seen as a special case of vertical differentiation. However, as our model contains both horizontal and vertical features it can not be generally mapped into either a pure horizontal or a pure vertical one, as we shall see below. Jones and Mendelson (1998) show that for quality-differentiated information goods and a uniform distribution of consumer types, no differentiation is optimal.[4] Our results in this paper do not confirm these findings interpreting one consumer characteristic as reservation price, which induces an additional type-dependent participation (or "feasibility") constraint (no consumer can pay more than her reservation price), and this in turn gives rise to a stable separation of the consumer base.

In addition to allowing the multidimensional screening of a heterogeneous consumer base (Laffont et al. 1987, Rochet and Choné 1998), product variety and broader product lines based on a modular product design[5] can help to delay differentiation along the supply chain (Lee and Tang 1997, Swaminathan and Tayur 1998). The real-option value of delaying irreversible commitment to a full product line can be seen in analogy to financial decision making (McDonald and Siegel 1986): delaying irreversible investments can carry a significant value, sometimes comparable to the investment volume itself (Pindyck 1988). This option value is generally increasing in the magnitude of the uncertainty unless, as Huchzermeier and Loch (2001) show, imperfectly correlated risks from different sources are pooled, potentially decreasing the value of managerial flexibility. We restrict ourselves to what they term "market payoff variability," a single risk class, and thus circumvent this effect. Examining the real-option value of delaying the product versioning until – with release of the flagship product – overall demand has been observed, we find that it is generally nonmonotonic in the cost of quality β. Naturally, the option value depends in an important manner on the performance of the default ex-ante versioning policy, and is therefore not necessarily monotonic in the amount of uncertainty over all β's. This points to resource flexibility as a driver of this option value (Van Mieghem 1998). Delaying differentiation does not generally allow to reverse the upfront quality investment. Being able to make product-line-extension and pricing decisions *contingent* on the observed demand carries significant value depending on the firm's default policy. For instance for development costs, β, close to zero when the irreversibility of the investment in flagship quality is not very

[4]These results have been recently generalized by Bhargava and Choudhary (2001). They give sufficient conditions for goods with *nonzero* marginal costs and general utility functions under which a stable, incentive-compatible separation of the consumer base into segments can be reached.

[5]Krishnan and Ulrich (2001) provide an excellent survey of the recent literature on product development decisions.

significant, the option to delay differentiation carries a high value. In fact we show that it is locally maximal at $\beta = 0$. The effects of increased versioning flexibility on upfront investment in product development as well as on consumer surplus are ambiguous. We provide an example where delayed differentiation leads to a lower flagship quality as a result of a state-contingent policy where the mode of differentiation depends on the observed demand realization, which is generally the case if the dispersion in the firm's prior beliefs about demand is large. For very little demand uncertainty, the versioning mode is generally "locked in" ex ante, and the delay in the differentiation allows to make slight improvements on the pricing and positioning decisions. Consumer surplus depends both on the actual product development costs as well as on the chosen mode of product differentiation and thus may vary in both directions when introducing an option to delay the versioning decision. To the best of our knowledge our paper provides a first complete analytical treatment of multiattribute versioning combining horizontal and vertical differentiation. We further extend this new approach to account for demand uncertainty by deriving optimal state-contingent policies under delayed differentiation. Information goods are chosen here for mere convenience. Everything in this paper equally applies to goods with positive marginal costs (cf. Section 5.4.4).

5.1.2 Outline

In the following section we introduce the basic model together with some structural simplifications that can be made without loss of any generality. In Section 5.3 we examine the ex-ante versioning case in which both flagship and versioned product are chosen simultaneously. We show that mixed versioning is never optimal and that investment in quality increases with estimated demand. Section 5.4 covers the case when the firm possesses the option to delay product versioning until after demand has realized. We determine the optimal versioning policy and compute the option value of delayed differentiation in this context. We also show that investment in product quality under delayed differentiation may both increase or decrease using a state-contingent horizontal-vertical versioning policy. Section 5.5 discusses the results and concludes with directions for further research.

5.2 The Model

Let all consumers be distributed uniformly on $\mathcal{V} \times \mathcal{W}$, where $\mathcal{V} = \mathbb{R}/2\mathbb{Z}$ and $\mathcal{W} = [0, W]$ denote the sets of tastes and wealths respectively, with $W > 0$ the maximum reservation price. The quotient space $\mathcal{V}$ is topologically nothing else than the interval $[0, 2]$, where the points

0 and 2 have been identified; one might think of $\mathcal{V} \times \mathcal{W}$ as a cylinder of radius $1/\pi$ and of height W.[6] The parameter W represents the maximum reservation price and is initially only imperfectly known, in terms of a measurable probability density function f with compact support $\mathcal{S} \subset \mathbb{R}_+$. We assume that any information product offered can be uniquely described in terms of its attributes, price $p \in \mathcal{W}$, quality $q \in \mathbb{R}_+$, and taste $z \in \mathcal{V}$. The *utility* of a consumer of type $(v, w) \in \mathcal{V} \times \mathcal{W}$ that buys a product of attributes (p, q, z) is assumed to be additively separable of the form $w - p + \kappa q - |v - z|$, and equal to w if she does not buy the product.[7] The positive constant κ defines the marginal valuation of quality relative to the other product attributes, price and taste.[8] This consumer can only afford the product, if her reservation price (disposable wealth) w is at least equal to the price p of the product,

$$p \leq w. \tag{5.1}$$

For any given maximum reservation price W there is a total of $2W$ consumers, which is therefore also the normalization constant for the consumer density.[9]

Change of Variables. To simplify some of the discussions that follow, we introduce the new variable $u = \kappa q - p$, so that consumer utility can be written in the form $w + u - |v - z|$. The new variable u represents the net utility gain at perfect horizontal fit. Using this simple change of variables consumer utility becomes independent of price. The participation constraint (i.e., for any consumer the utility obtained for the product has to weakly exceed utility for wealth w) can be written as

$$|v - z| \leq u. \tag{5.2}$$

Any optimization of the product portfolio at the firm level needs to account for the *feasibility* and *participation* constraints (5.1)–(5.2) as well as *incentive compatibility* in the sense that

[6] This model generalizes the common setup of "circular city" models in the spirit of Salop (1979), which are used to describe pure horizontal differentiation, typically to avoid the endpoint effects Hotelling's "linear city." We have chosen the cylinder topology for this mixed differentiation model with precisely this in mind.

[7] General conditions for the additive separability of utility functions in wealth have been provided by Gorman (1968). Instead of κq it is possible to incorporate any nonlinear strictly increasing quality measure $\phi(q)$ which simply corresponds to a rescaling (cf. also equation (5.4)).

[8] Without loss of generality the marginal utility of taste differences $|v - z|$ can be normalized to one, even when the cylinder circumference $2V$ (for some $V > 0$) is not precisely known. Varying V changes the relative marginal valuation of taste mismatch. By substituting $W' = W/V$ for W and using an appropriate prior probability density g of W' (instead of f) one obtains the same results as the ones derived for the unit cylinder $\mathcal{V} \times \mathcal{W}$, up to a constant factor. Thus, the results in this paper do also hold when both maximum reservation price *and* extent of the taste domain are ex-ante uncertain.

[9] Total wealth in the set $\mathcal{V} \times \mathcal{W}$ is $W/2$, independent of the diameter of the cylinder.

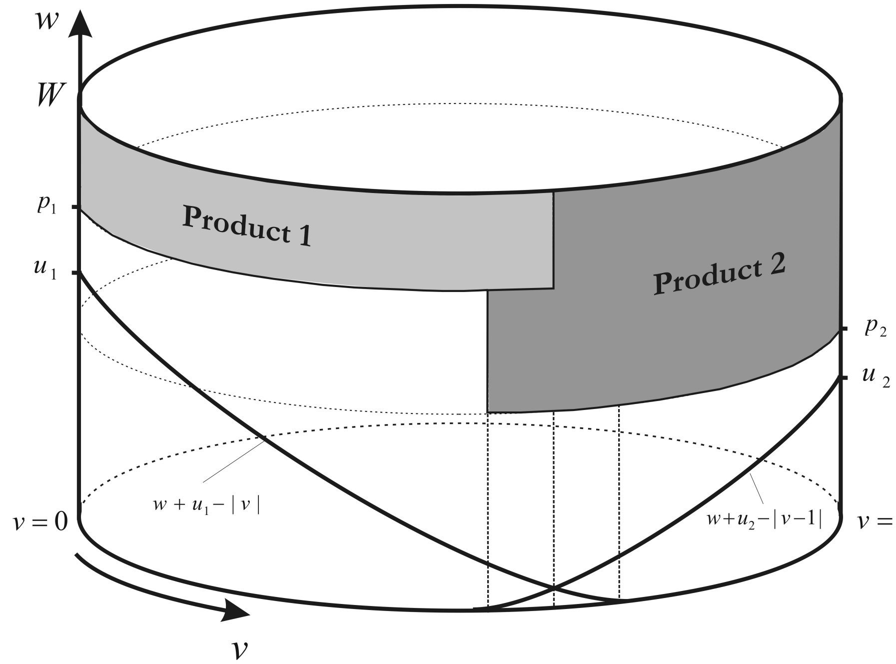

Figure 5.1: Consumer Utilities and Resulting Market Segments ($q_1 > q_2$).

each consumer's choice is utility maximizing.

Cost Structure. Throughout most of this paper we consider information goods with zero marginal cost which simplifies the closed-form solutions. This assumption is not crucial and all results can be easily generalized to goods with positive marginal cost as is shown in Section 5.4.4. For information goods, the costs of reproduction and distribution are indeed very small, so that this has become a standard assumption in much of the literature.[10] The cost βq of creating the flagship product depends linearly on the quality q, and can be written in the new variables as

$$C(p, u) = \frac{\beta}{\kappa}(p + u), \tag{5.3}$$

where β is a positive constant. Without loss of generality it is possible to set

$$\kappa = 1, \tag{5.4}$$

as the results for any arbitrary value of $\kappa > 0$ can be recovered by substituting everywhere $\beta' = \beta/\kappa$ for β. Indeed, the weight κ in the utility function has the sole effect of scaling the

[10]Information goods are merely chosen here to keep analytical complexity as low as possible. Of course, there are settings for which the small, but nonzero, nature of the marginal costs matters, such as in deriving asymptotic properties of large bundles of information goods (Bakos and Brynjolfsson 1999).

quality cost of the firm. In everything that follows we therefore assume that (5.4) holds.

Versioning Problem. After having created the flagship product the risk-neutral firm may have the option to create a versioned product by varying product attributes in the following admissible ways:

(a) *horizontal version* (i.e., same quality, different taste) at a cost $\alpha_H \geq 0$,

(b) *vertical version* (i.e., lower quality, same taste) at a cost $\alpha_V \geq 0$,

(c) *mixed version* (i.e., lower quality, different taste) at a cost $\alpha_M \in [\max\{\alpha_H, \alpha_V\}, \alpha_H + \alpha_V]$.[11]

This cost structure exhibits strong cost complementarities or economies of scope (Panzar and Willig 1981) as development costs are only incurred once to create the flagship product.[12] The firm maximizes profits subject to the consumers' choice and participation. Solutions to this profit-maximization problem exhibit *translation invariance*: profits will not change by just horizontally translating all products, as a direct consequence of the circular symmetry. One can therefore arbitrarily fix the horizontal location of the first product to zero. Any implied solution then naturally represents an entire equivalence class of solutions to the more general problem with arbitrary horizontal location of the first product.

Demand Estimation. As has been pointed out above, the demand parameter W is initially only imperfectly known to the firm (cf. also footnote 8). For any product of price p the profit function depends on W through the marginal revenue density $p(p-W)/(2W)$. By replacing the random value W with a nonlinear estimator $\hat{W} = (E[1/W])^{-1}$ we obtain the *expected* marginal revenue density. The following technical lemma will be used throughout the paper without special mention.

LEMMA 5.1 *(i)* $E\left[p(p-W)/(2W)\right] = p\left(p-\hat{W}\right)/\left(2\hat{W}\right)$. *(ii)* $\hat{W} < \bar{W}$, *where* $\bar{W} = E(W)$.

The first part of this result is trivial but useful, as it allows replacing the imperfectly known W with the nonrandom $\hat{W}$ when computing expected profits. Part (ii) indicates that when maximizing expected profits, uncertainty in W prompts the firm to use as decision-relevant

[11] This condition can also be written in terms of the r-norm of the row vector (α_H, α_V), i.e., $\alpha_M = ||(\alpha_H, \alpha_V)||_r = (\alpha_H^r + \alpha_V^r)^{1/r}$ for $r \in [1, \infty)$ and $\alpha_M = ||(\alpha_H, \alpha_V)||_\infty = \max\{\alpha_H, \alpha_V\}$ for $r = \infty$.

[12] A multiproduct cost structure of this form has also been proposed by Shaked and Sutton (1987).

"certainty equivalent" of W an estimator $\hat{W}$ that is always strictly (and sometimes substantially) lower than the expected maximum reservation price $\bar{W}$.[13]

5.3 Ex-Ante Versioning

5.3.1 Benchmark: Single-Product Firm

Before solving the multiproduct case we briefly discuss the single-product monopoly as a benchmark. Without loss of generality we can set $z_1 = 0$, in view of the translation invariance discussed above. With this the expected single-product monopoly profit $\bar{\Pi}_1(p_1, u_1) = E\Pi_1(p_1, u_1; W)$ as a function of (p_1, u_1) becomes

$$\bar{\Pi}_1(p_1, u_1) = \frac{p_1(\hat{W} - p_1)}{\hat{W}} \min\{u_1, 1\} - \beta(p_1 + u_1).$$

It is clear that $u_1^* \leq 1$, since $\partial_{u_1}\bar{\Pi}(p_1, u_1) < 0$ for any $(p_1, u_1) \in (0, \hat{W}) \times (1, \infty)$. Furthermore, there are no interior extrema, since the determinant of the corresponding Hessian matrix $H = \partial^2\bar{\Pi}(p_1, u_1)/\partial^2(p_1, u_1)$ is negative,[14] for all $p_1 \neq \hat{W}/2$. Therefore, maximum profits are achieved on the boundary of $\mathcal{V} \times \mathcal{W}$,

$$(p_1^*, u_1^*) = \begin{cases} (\hat{W}(1-\beta)/2, 1), & \text{if } \beta \leq 1, \\ (0, 0), & \text{if } \beta > 1. \end{cases} \tag{5.5}$$

Thus, the single-product firm would like to achieve "full horizontal market coverage" (i.e., $q_1^* = 1 + p_1^*$), whenever it decides to enter the market.[15]

PROPOSITION 5.1 *(i) The optimal expected profit of the single-product firm is*

$$\bar{\Pi}_1^* = \frac{\hat{W}}{4}(1-\beta)^2 - \beta, \tag{5.6}$$

increasing in $\hat{W}$ and decreasing in β. (ii) The single-product firm enters the market, iff

$$\hat{W} > \frac{4\beta}{(1-\beta)^2}. \tag{5.7}$$

[13] The difference between $\bar{W}$ and $\hat{W}$ can become very large. As an example consider a uniform distribution of W on an interval $[1, \lambda]$ with $\lambda > 1$. Then $\bar{W} = (1+\lambda)/2$ and $\hat{W} = (\lambda - 1)/\ln\lambda$, so that $\bar{W}/\hat{W} \to \infty$ for $\lambda \to \infty$.

[14] In particular, this means that interior critical points can only be saddle points.

[15] This result is driven by the linearity of the cost function. In the case when the cost function is strictly convex, say, of the form βq^2, there may exist interior solutions making it impossible to derive closed-form solutions in the multiproduct case.

Relation (5.7) is a *viability* condition that imposes a lower bound on the maximum-reservation-price estimator $\hat{W}$ as a function of β. In what follows we will use the optimal expected single-product monopoly profits as a benchmark for the performance of the various differentiation modes in a multiproduct setting.

5.3.2 Ex-Ante Product Differentiation

As for the single-product firm it is, due to the symmetry of the problem, without loss of generality possible to set $z_1 = 0$, and z_2 can be restricted to the interval $[0, 1]$ (otherwise just switch products and translate). Any solution under these restrictions defines an equivalence class of solutions to the versioning problem. There is one further simplification, which directly results from the linearity of the problem in z_2. Indeed, the optimal value of z_2 must be at an extremity of $[0, 1]$, i.e., $z_2^* \in \{0, 1\}$. The induced two canonical cases of vertical and horizontal differentiation have to be distinguished in the solution to the versioning problem. Furthermore we assume that the quality of the second product does not exceed the quality of the first product ($q_1 \geq q_2$), which after linear variable transformation ($u = \kappa q - p$) can be written equivalently as

$$q_1 = p_1 + u_1 \geq p_2 + u_2 = q_2. \tag{5.8}$$

This assumption correlates with the fact that in order for a "versioning" policy to be meaningful, the first (flagship) product has to be of higher quality than the derived product. Figure 5.2 provides an overview of the different modes of differentiation that can arise: (1) no differentiation ($z_2 = 0, q_1 = q_2$), which is *de facto* equivalent to the single-product monopoly; (2) pure vertical differentiation ($z_2 = 0, q_1 > q_2$); (3) mixed differentiation ($z_2 = 1, q_1 > q_2$); and, (4) pure horizontal differentiation ($z_2 = 1, q_1 = q_2$). To simplify the exposition, we will first set versioning costs to zero as their effect can be analyzed separately, once all the main expressions for the zero-versioning cost base case have been established.

Pure Vertical Differentiation. The profit function $\bar{\Pi}_V$ is linear in u_1 and u_2, so that we can concentrate on the situation when $u_1 = u_2 = 1$, since otherwise the firm would only enter with one product ($u_2 = 0$, cf. Section 5.3.1). The maximization problem can be written in the form

$$\max_{p_1,p_2} \bar{\Pi}_V(p_1, p_2) = \max_{p_1,p_2} \left\{ \frac{1}{\hat{W}} \Big(p_1(\hat{W} - p_1) + p_2(p_1 - p_2) \Big) - \beta(p_1 + 1) \right\}, \tag{5.9}$$

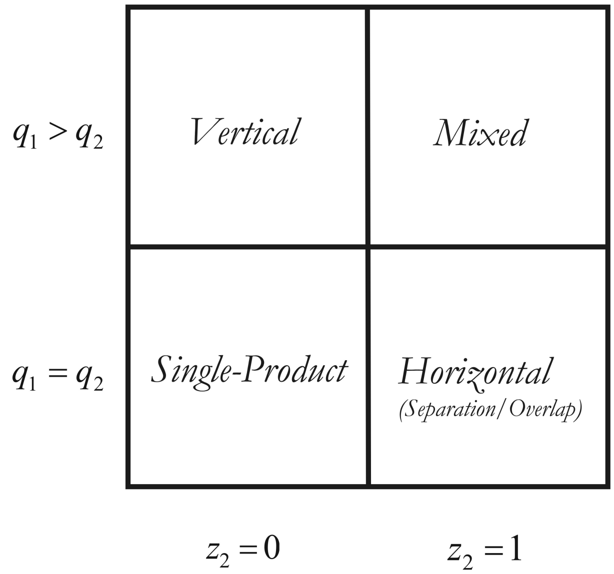

Figure 5.2: Modes of Differentiation.

the solution of which yields the unique interior maximizers, $p_1^* = 2\hat{W}(1-\beta)/3$ and $p_2^* = \hat{W}(1-\beta)/3$, provided that $\beta < 1$. Optimal expected profits are then

$$\bar{\Pi}_V^* = \frac{\hat{W}}{3}(1-\beta)^2 - \beta = \bar{\Pi}_1^* + \frac{\hat{W}}{12}(1-\beta)^2, \tag{5.10}$$

strictly greater than single-product monopoly profits $\bar{\Pi}_1^*$. Thus, pure vertical differentiation strictly dominates a single-product monopoly, given that versioning costs α_V are small enough. Pure vertical differentiation is ex-ante viable, iff

$$\hat{W} > \frac{3\beta}{(1-\beta)^2}. \tag{5.11}$$

It is important to note that since $u_1^* = u_2^*$ at the optimum in pure vertical differentiation, a consumer of type (v, w) obtains equal surplus from both products. This follows from the latent assumption that *in equilibrium consumers buy the higher-price product if this yields exactly the same utility as the lower-price product* (e.g., since resale value is higher also). If this was not the case, the firm could technically achieve stable separation by letting $u_1^* = u_2^* + \varepsilon$ for any (arbitrarily small) $\varepsilon > 0$ without, in the limit, changing any of the results. The fundamental reason for the stable segmentation of the consumer population lies in the feasibility constraint (5.1): consumers never spend more than their wealth w (reservation price) permits.

Mixed Differentiation. In analogy to the single-product monopoly, profits $\bar{\Pi}_M$ only depend linearly on u_1 and u_2 (whence results full horizontal coverage). Since the constraint (5.8) is by assumption not binding, it is $u_1 = u_2 = 1$. The profit-maximization problem thus takes the form

$$\max_{p_1,p_2} \bar{\Pi}_M(p_1,p_2) = \max_{p_1,p_2} \left\{ \frac{1}{2\hat{W}} \Big(p_1(\hat{W}-p_1) + p_2(\hat{W}-p_2) + p_2(p_1-p_2) \Big) - \beta(p_1+1) \right\}. \tag{5.12}$$

Straightforward unconstrained maximization yields the optimal prices $p_1^* = \hat{W}(5-8\beta)/7$ and $p_2^* = \hat{W}(3-2\beta)/7$, which are well defined (i.e., in $[0,\hat{W}]$ and such that $p_1^* > p_2^*$) for any positive $\beta < 1/3$. Optimal expected profits are

$$\bar{\Pi}_M^* = \frac{\hat{W}}{7}\left(\frac{5}{4} - 2\beta\right)^2 + \frac{\hat{W}}{16} - \beta = \bar{\Pi}_1^* + \frac{9\hat{W}}{28}\left(\frac{1}{3} - \beta\right)^2, \tag{5.13}$$

strictly increasing in $\hat{W}$ and strictly decreasing in β. In the mixed-differentiation mode, the multiproduct monopolist enters, iff

$$\hat{W} > \frac{\beta}{\frac{1}{7}\left(\frac{5}{4} - 2\beta\right)^2 + \frac{1}{16}}, \textbf{ and } \beta < \frac{1}{3}. \tag{5.14}$$

Given entry in mixed-differentiation mode, quality dispersion does in general occur, since by construction for $\beta < 1/3$ that $q_1^* = \hat{W}(5-8\beta)/7+1 \geq \hat{W}(3-2\beta)/7+1 = q_2^*$. However, vertical differentiation strictly dominates mixed differentiation. Indeed,

$$\bar{\Pi}_V^* - \bar{\Pi}_M^* = \hat{W}\left(\frac{(1-\beta)^2}{3} - \frac{1}{7}\left(\frac{5}{4} - 2\beta\right)^2 - \frac{1}{16}\right) = \hat{W}\left(\frac{1}{20} - \frac{5}{21}\left(\beta - \frac{1}{10}\right)^2\right) > 0$$

for all $\beta \in (0, 1/3)$. Therefore, the mixed-differentiation mode does not play a role, which by itself is an interesting fact.

Proposition 5.2 *For any given viable market configuration $(\beta, \hat{W})$ either pure horizontal or pure vertical differentiation is optimal.*

Mixed versioning is dominated by pure vertical differentiation. As soon as goods are horizontally differentiated with ($z_2 = 1$), it is best to reduce the quality of the flagship product so as to economize on providing costly horizontal coverage with each product. Hence leaving slack in constraint (5.8) does not fully utilize the potential of the flagship product, whose quality will ex ante be reduced until $q_1 = q_2$, i.e., until pure horizontal differentiation is obtained.

Pure Horizontal Differentiation. In this mode constraint (5.8) is binding, so that $q_1 = q_2$ or in other words

$$u_1 = u_2 + p_2 - p_1. \tag{5.15}$$

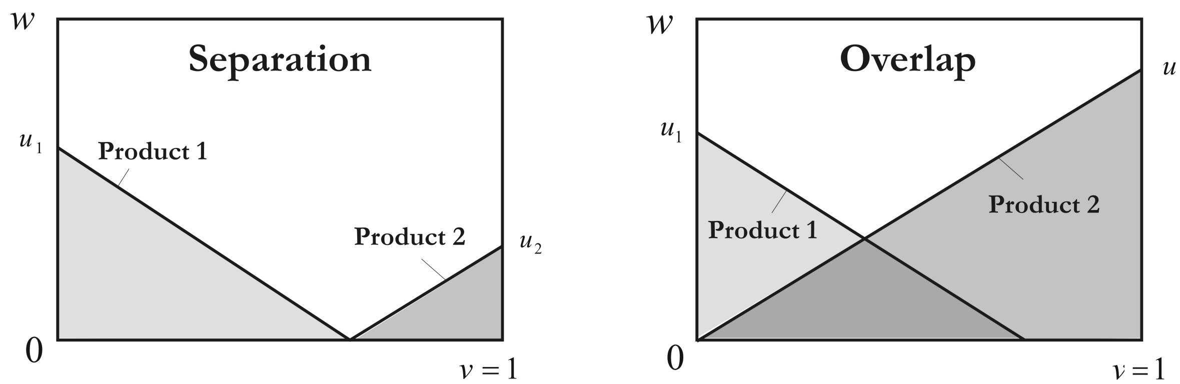

Figure 5.3: Horizontal Separation and Overlap in Terms of Consumer Surplus.

Expected profits are $\bar{\Pi}_H = \frac{1}{2\hat{W}}\Big(p_1(\hat{W}-p_1)(1+u_1-u_2)+p_2(\hat{W}-p_2)(1-u_1+u_2)+\min\{p_1,p_2\}$ $|p_2-p_1|(u_1+u_2-1)\Big)-\beta(p_1+u_1)$ or, using (5.15),

$$\begin{aligned}\bar{\Pi}_H &= \frac{1}{2\hat{W}}\Big(p_1(\hat{W}-p_1)(1-p_1+p_2)+p_2(\hat{W}-p_2)(1+p_1-p_2) \\ &\quad +\max\{p_1(p_2-p_1),p_2(p_1-p_2)\}(2u_2-1-p_1+p_2)\Big)-\beta(p_2+u_2), \quad (5.16)\end{aligned}$$

subject to $p_2 \in [p_1-1,p_1+1]\cap[0,\hat{W}]$ and $u_2 \in [(1+p_1-p_2)/2,1]$. Thus, depending on the sign of $\partial_{u_2}\bar{\Pi}_H = \frac{1}{\hat{W}}\max\{p_1(p_2-p_1),p_2(p_1-p_2)\}-\beta$, either $u_2 = 1$ (for $\partial_{u_1}\bar{\Pi}_H > 0$) or $u_2 = (1+p_1-p_2)/2$ (for $\partial_{u_1}\bar{\Pi}_H < 0$). In the latter case there is no overlap between the two products, which we term "separation" (cf. Figure 5.3). In case when $u_2 = 1$, there clearly is "overlap," and relation (5.15) entails that then $p_1 \geq p_2$. It is clear that overlap in practice means horizontal "product cannibalization," for some consumers are offered an individually rational choice (satisfying (5.1)–(5.2)) between both products. It turns out that under ex-ante versioning horizontal overlap is never optimal as the firm can do better by *either* eliminating cannibalization *or* purely vertically differentiating. This finding does not hold under delayed differentiation where horizontal cannibalization may be optimal as the best possible "compromise" for certain intermediate demand realizations.

PROPOSITION 5.3 *Let the parameters $0 < \beta < 1$ and $\hat{W} > 0$ be given.*

(i) Horizontal differentiation with separation yields optimal expected profits of

$$\bar{\Pi}^*_{H_1} = \frac{\hat{W}}{4}(1-\beta)^2 - \frac{\beta}{2} = \bar{\Pi}^*_1 + \frac{\beta}{2}, \quad (5.17)$$

with symmetric product portfolio $(p^*_k, q^*_k) = \Big(\hat{W}(1-\beta)/2, \Big(1+\hat{W}(1-\beta)\Big)/2\Big)$ for $k = 1, 2$.

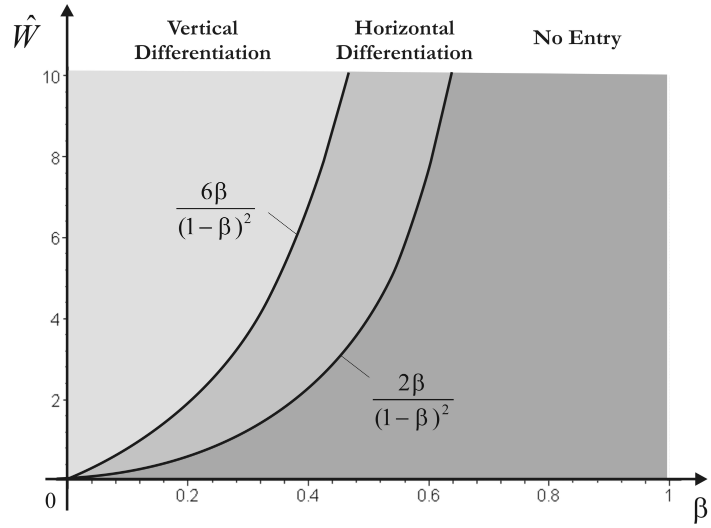

Figure 5.4: Optimal Ex-Ante Product Versioning.

(ii) There is no horizontal product cannibalization, i.e., optimal expected profits $\bar{\Pi}^*_{H_2}$ *under horizontal differentiation with overlap are dominated,*

$$\bar{\Pi}^*_{H_2} < \max\{\bar{\Pi}^*_{H_1}, \bar{\Pi}^*_V\} = \bar{\Pi}^*. \tag{5.18}$$

When only considering horizontal differentiation, Proposition 5.3 implies that product cannibalization is not inherently ruled out, which may seem counterintuitive in our simple setting. In fact, for lower unit cost of quality β overlap (i.e., product cannibalization) performs better than separation, which only yields a constant improvement of $\beta/2$ over single-product optimal profits $\bar{\Pi}^*_1$. Cannibalization implies an overinvestment in quality in the sense that more quality is provided than necessary for full horizontal coverage of the consumer base. Exploiting the consumers' feasibility constraint it is possible (for low β) to achieve a *vertical* separation of the consumer base offering one product at a lower price than the other. The high-price product is then solely preferred by high-w consumers of very good horizontal fit. This feature of the model explains the possibility of *price dispersion* in a purely horizontally differentiated market when the marginal cost of creating additional quality is sufficiently low.

The following proposition summarizes the firm's entry decision into the market in the absence of an option to delay product differentiation.

PROPOSITION 5.4 *Consider a firm that has the option to enter the market, given* $(\beta, \hat{W}) \in (0,1) \times \mathcal{S}$*, with one or two information goods.*

(i) For any $(\beta, \hat{W})$ such that

$$2\beta/(1-\beta)^2 < \hat{W}, \tag{5.19}$$

the firm's entry is viable, if versioning costs are small enough.

(ii) In the case that (5.19) does not hold, the firm's entry is not viable, even if versioning costs are zero.

(iii) If entry is viable and $\hat{W} < 4\beta/(1-\beta)^2$, the firm needs to enter the market with two products.

An interesting implication from the last proposition is that *in some markets for information goods that are unprofitable for any single product, a firm may still be able to enter profitably with a versioned portfolio of products.*

Effect of Versioning Costs. We have seen above that under perfect information mixed differentiation is strictly dominated by pure vertical differentiation. Thus we can limit ourselves to adjusting and comparing expressions (5.6), (5.10), and (5.29). As a consequence of the earlier discussion, write $\bar{\Pi}^*_H$ (instead of $\bar{\Pi}^*_{H_1}$) for the optimal profits under pure horizontal differentiation (separation). Following (5.29), pure horizontal differentiation is dominated by single-product monopoly, iff $\alpha_H > \beta/2$. On the other hand, it is dominated by pure vertical differentiation, iff $\alpha_V < \alpha_H + \frac{\hat{W}}{12}(1-\beta)^2 - \frac{\beta}{2}$. The resulting partition of the (α_H, α_V)-plane is depicted in Figure 5.5.

5.4 Delayed Differentiation

Assume that the risk-neutral firm can delay the versioning decision until after the demand uncertainty has resolved through introduction of the flagship product. The exact timing proceeds over two stages ($t \in \{1, 2\}$): at $t = 1$ the firm develops its flagship product of fixed[16] attributes (q_1, z_1) and presents it to the market without collecting revenues. In the second stage ($t = 2$) firm learns the true value for the maximum reservation price W, decides about introducing a versioned product of nonprice attributes (q_2, z_2), and prices both products at p_1, p_2 respectively. As a result, the firm needs to choose its product portfolio such that its

[16] Quality upgrades are not possible for $t > 1$, since e.g., the development team has been disbanded or upfront costs to restart significant quality improvements are prohibitive. The firm has never an incentive to change the horizontal nonprice attribute of the flagship product, $z_1 = 0$.

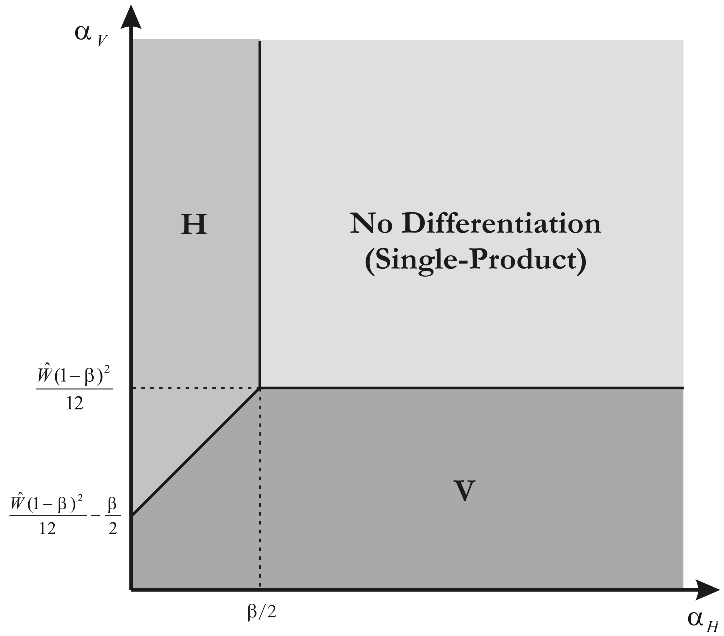

Figure 5.5: Optimal Modes of Ex-Ante Differentiation for Versioning Cost $(\alpha_H, \alpha_V) \geq 0$.

ex-ante expected profits achieve

$$\bar{\Pi}^{**} = \max_{q_1} \left\{ E \left[\max \left\{ \Pi_V^*(q_1; W), \Pi_H^*(q_1; W) \right\} \right] \right\} \tag{5.20}$$

instead of the $\bar{\Pi}^* = \max\{\bar{\Pi}_V^*, \bar{\Pi}_H^*\}$ it obtains under ex-ante versioning according to Proposition 5.3. We solve this problem in stages, beginning with the second stage.

5.4.1 Optimal Ex-Post Versioning

Assume the firm has observed the realization of W and has decided about the quality q_1 of its flagship product. As in Section 5.3, it is possible to differentiate the versioned product either vertically ($z_2 = 0$) or horizontally ($z_2 = 1$), provided that its quality does not exceed the flagship product's, i.e., $q_1 \geq q_2$ (cf. footnote 16). Alternately the firm has the option not to introduce a vertically versioned product at all.[17]

[17] The firm could also reprice the first product out of the market ($p_1 \geq W$) while vertically versioning a second, which is, however, strictly inferior if versioning cost α_V is positive. If the firm does not introduce the second product at all, then it needs to solve $\max_{p_1' \in [0,W]} p_1'(W - p_1') \min\{1, q_1 - p_1'\}$, with solution $p_1'^*$ that can be obtained by replacing W with $3W/4$ in (5.21). Note that the firm does not incur versioning cost, so that at least for large α_V not introducing a second product is optimal, given that vertical differentiation is considered. If on the other hand $\alpha_V = 0$ (or $\alpha_V > 0$, but small enough), then introducing the versioned product in the second round is clearly dominant, since it can always be achieved as a special case by setting $p_2 = 0$.

Pure Vertical Differentiation. Consider the case when $z_2 = 0$ and the resulting portfolio contains two products. Without loss of generality we can assume that $u_1^* = u_2^* = q_1 - p_1^*$, since – as long as they are free – larger values for u weakly increase profits. The profit-maximization problem becomes

$$\max_{p_1,p_2} \Pi_V(p_1, p_2; q_1, W) = \max_{p_1,p_2} \left\{ \frac{p_1(W - p_1) + p_2(p_1 - p_2)}{W} \min\{1, q_1 - p_1\} - \beta q_1 \right\}$$

with solution equivalent to that of the ex-ante versioning problem, as long as q_1 is large enough so that full horizontal coverage holds. If on the other hand q_1 is small compared to the realization W, then horizontal coverage will not be full and product breadth is traded for the ability to charge high prices. In other words, "transportation costs" for consumers with significant misfit will be too high to buy any of the products. More specifically we obtain $(p_2^*, q_2^*) = (p_1^*/2, q_1 - p_1^*/2)$, and

$$p_1^* = \begin{cases} 2W/3, & \text{if } q_1 \geq 1 + 2W/3, \\ q_1 - 1, & \text{if } q_1 \in [\bar{q}_V(W), 1 + 2W/3], \\ \left(4W + 3q_1 - \sqrt{(4W - 3q_1)^2 + 12Wq_1}\right)/9, & \text{otherwise,} \end{cases} \tag{5.21}$$

where $\bar{q}_V(W) = \left(6 + 2W - \sqrt{9 + 4W^2}\right)/3 \in [1, \min\{2, 1 + 2W/3\}]$. From this one can derive an expression for the optimal profits as a function of q_1,

$$\Pi_V^*(q_1; W) = \begin{cases} W/3 - \beta q_1, & \text{if } q_1 \geq 1 + 2W/3, \\ (q_1 - 1)\left(1 - \dfrac{3(q_1 - 1)}{4W}\right) - \beta q_1, & \text{if } q_1 \in [\bar{q}_V(W), 1 + 2W/3], \\ p_1^*\left(1 - \dfrac{3p_1^*}{4W}\right)(q_1 - p_1^*) - \beta q_1. & \text{otherwise,} \end{cases} \tag{5.22}$$

Note that due to the cost complementarities a versioned product will always be introduced as long as versioning cost α_V are small enough. The development cost for the flagship product is already sunk and a second product can then only increase profits (at least as long as α_V, α_H are small enough). It is straightforward to show that for any $W > 0$ the profit function $\Pi_V^*(q_1; W)$ is continuously differentiable in $q_1 > 0$ with $\partial_{q_1}\Pi_V^*(0) = \partial_{q_1}\Pi_V^*(1 + 2W/3)) = -\beta < \partial_{q_1}\Pi_V^*(\bar{q}_V) = \left(\sqrt{4W^2 + 9} - 3\right)/(2W) - \beta$. In addition, the state-contingent profit function Π_V^* has at most one maximum in $q_1 > 0$ but is generally not quasiconcave.[18]

Horizontal Differentiation (Including Mixed). When deciding to horizontally differentiate, the firm chooses $z_2 = 1$ and then needs to decide about p_1, p_2 and $q_2 = u_2 + p_2 \leq$

[18]It is also locally maximal at $q_1 = 0$, as $\partial_{q_1}\bar{\Pi}_V^*|_{q_1=0} = -\beta < 0$.

$u_1 + p_1 = q_1$. As in the ex-ante versioning problem (cf. Section 5.3) there are two general cases, "separation" and "overlap." The details of the overlap case, which is synonymous with "horizontal product cannibalization," are discussed in the Appendix, where analytical solutions are provided. Under delayed differentiation it is indeed sometimes best to horizontally cannibalize products, whereas this is never optimal under ex-ante versioning (Proposition 5.5). Indepedent of the versioning costs it is *never* optimal to version along both horizontal and vertical attributes at the same time.[19]

There are four possible configurations that all exhibit horizontal separation, depending on how large the quality q_1 is with respect to the demand realization W: (i) if $q_1 \geq 1 + W/2$ (i.e., $q_1 - p_1 \geq 1$), given separation there is just one product and $p_1^* = W/2$ analogous to the single-product firm. (ii) If $1/2 \leq q_1 - p_1 = 1 - (q_2 - p_2) < 1$, then

$$(p_1^*, p_2^*) = \arg\max_{p_1,p_2} \left\{ \frac{p_1(W-p_1)(q_1-p_1)}{2W} + \frac{p_2(W-p_2)(1-q_1+p_1)}{2W} - \beta q_1 \right\} = \left(\frac{W}{2}, \frac{W}{2} \right),$$

for $q_1 \in [(1+W)/2, 1+W/2]$. Note that in this case $q_2^* = 1 + W - q_1$ resulting in *mixed* differentiation without price dispersion. (iii) If $q_1 \in [\bar{q}_H(W), (1+W)/2]$, then $p_1^* = p_2^* = q_1 - 1/2$, and $q_2^* = q_1$ (pure horizontal differentiation). Thereby $\bar{q}_H(W) = \left(2 + W - \sqrt{1+W^2}\right)/2 \in [1/2, \min\{1, (1+W)/2\}]$. (iv) If $q_1 - p_1 < 1 - (q_2 - p_2)$ (equivalent to $q_1 - p_1 < 1/2$), then $p_1 = p_2$ and $q_1 = q_2$, since the situation is symmetric (pure horizontal differentiation). We have

$$p_1^* = \arg\max_{p_1} \left\{ \frac{p_1(W-p_1)(q_1-p_1)}{W} - \beta q_1 \right\} = \left(W + q_1 - \sqrt{(W-q_1)^2 + Wq_1} \right)/3, \quad (5.23)$$

for $q_1 < \bar{q}_H(W)$. Summarizing these results we obtain

$$\Pi_{H_1}^*(q_1; W) = \begin{cases} W/4 - \beta q_1, & \text{if } q_1 \geq (1+W)/2, \\ (q_1 - 1/2)\left(1 - (q_1 - 1/2)/W\right) - \beta q_1, & \text{if } q_1 \in [\bar{q}_H, (1+W)/2], \\ 2p_1^*(1 - p_1^*/W)(q_1 - p_1^*) - \beta q_1, & \text{otherwise.} \end{cases} \quad (5.24)$$

Note that in the case of realizations W such that $q_1 \in [(1+W)/2, 1+W/2]$ the firm would use mixed versioning; however, the following proposition shows that this is never optimal. Analogous to the pure-vertical-versioning case, it can be shown that the profit function Π_{H_1} is continuously differentiable and has at most one maximum in $q_1 > 0$.

[19]This finding is somewhat driven by the quasi-linearity of the model in preferences and the assumption that the consumers are uniformly distributed. Including nonlinearities in the consumers' preferences interior solutions favoring simultaneous differentiation along more than one attribute cannot be ruled out.

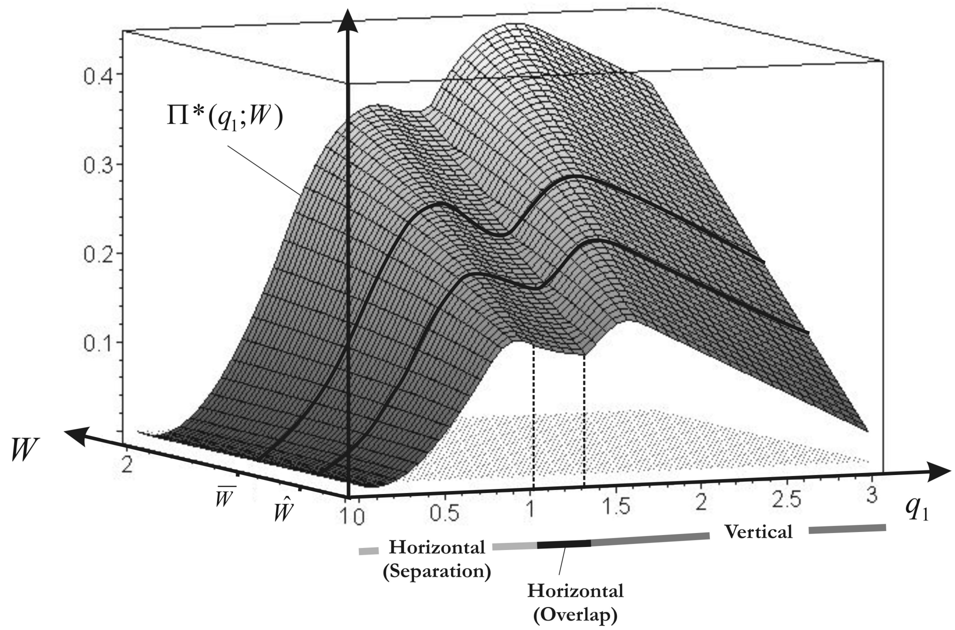

Figure 5.6: Optimal Cumulative Profits as a Function of q_1 and W (for $\beta = .1$, $\mathcal{S} = [1, 2]$).

PROPOSITION 5.5 *(i) For any given realization of W in $\mathcal{S}$, either pure horizontal or pure vertical differentiation is optimal. (ii) Horizontal product cannibalization (overlap) may be optimal for certain "intermediate" realizations of W.*

Horizontal overlap in the product portfolio is most likely to be useful if the support $\mathcal{S}$ of demand realizations is such that for low states W vertical differentiation is optimal while for high states horizontal separation is best. Note that in the presence of horizontal product cannibalization, there generally is price dispersion (cf. the details on the overlap case in the Appendix).

Effect of Versioning Costs. As a consequence of Proposition 5.5 mixed versioning cost, $\alpha_M \geq \alpha_H, \alpha_V$, will never be incurred. With *perfect* ex-ante information about W, the effect of versioning cost is of course adequately displayed in Figure 5.5 with $\hat{W}$ replaced by W. In the presence of uncertainty, the optimal policy depends on the magnitudes and the difference between α_V and α_H as these influence the attractiveness of one versioning mode over the other and of product-line extensions overall.

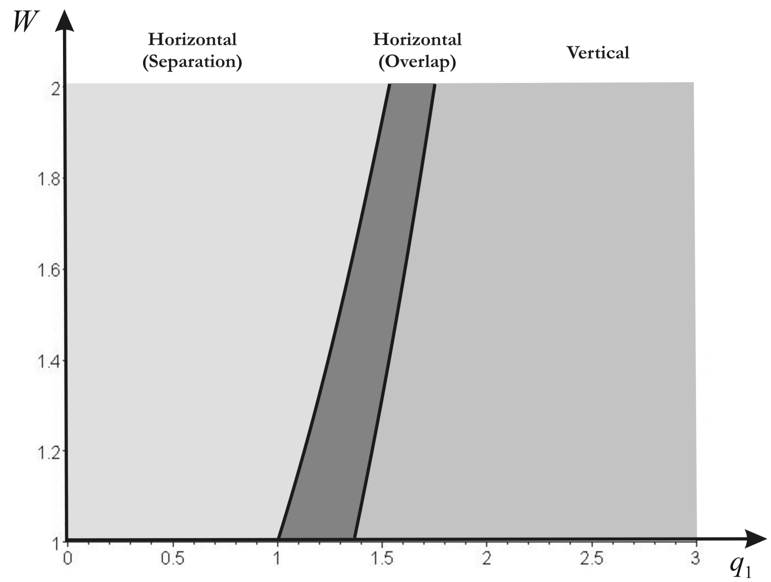

Figure 5.7: Optimal Modes of Ex-Post Differentiation (for $\beta = .1$, $\mathcal{S} = [1, 2]$).

5.4.2 The Product Development Decision

In the first stage of the versioning problem with delayed differentiation the firm decides about the appropriate development effort (i.e., investment in product quality q_1^{**}) by solving problem (5.20). As pointed out in Section 5.1, irreversibilities in investments lead firms to value the option of delaying commitment (Weisbrod 1964, McDonald and Siegel 1986, Pindyck 1988). Under delayed differentiation in the presence of demand uncertainty the marginal value of quality may be higher than under ex-ante versioning where an adaptation of the product portfolio to the realization of the uncertain demand is not possible.

Proposition 5.6 *If the optimal ex-post versioning policy is not state-contingent, i.e., the mode of differentiation does not depend on W, then the optimal ex-ante investment q_1^{**} is characterized by*[20]

$$q_1^{**} = 1 + \frac{2}{3} \frac{\int_{W_1(q_1^{**})}^{W_2(q_1^{**})} f(w)dw - \beta + \int_{W_2(q_1^{**})}^{\bar{s}} p_1^*(w) \left(1 - \frac{3p_1^*(w)}{4w}\right) f(w)dw}{\int_{W_1(q_1^{**})}^{W_2(q_1^{**})} \frac{f(w)}{w} dw} \tag{5.25}$$

[20] It is clear that horizontal differentiation with overlap can never be the result of a non-state-contingent versioning policy.

for pure vertical differentiation (with $\bar{s} = \max\{\mathcal{S}\}$*,* $W_1 = 3(q_1^{**} - 1)/2$*,* $\bar{q}_V(W_2) = q_1^{**}$*, and* p_1^* *as in (5.21) for* $w = W$*), and*

$$q_1^{**} = \frac{1}{2} + \frac{\displaystyle\int_{W_1(q_1^{**})}^{W_2(q_1^{**})} f(w)dw - \beta + 2\int_{W_2(q_1^{**})}^{\bar{s}} p_1^*(w)\left(1 - \frac{p_1^*(w)}{w}\right) f(w)dw}{\displaystyle 2\int_{W_1(q_1^{**})}^{W_2(q_1^{**})} \frac{f(w)}{w}dw} \tag{5.26}$$

for horizontal differentiation without overlap (with $\bar{s}$ *as above,* $W_1 = 2q_1^{**} - 1$*,* $\bar{q}_{H_1}(W_2) = q_1^{**}$*, and* p_1^* *as in (5.23) for* $w = W$*).*

If the uncertainty about the demand characteristic W is low, then $[W_1, W_2] \cap \mathcal{S} = \mathcal{S}$, so that the expressions (5.25) and (5.26) are identical to the optimal ex-ante quality choice q_1^* under vertical and horizontal differentiation respectively (cf. Section 5.3.2). If the development cost β is zero, then clearly it is optimal to vertically version and choose $q_1^{**} = 1 + 2\bar{s}/3$, a quality larger than the ex-ante optimal quality of $q_1^* = 1 + 2\hat{W}/3$ for $\beta = 0$, since uncertainty about demand implies uncertainty about the optimal pricing of the product portfolio. Hence, even if quality can be provided at an arbitrarily small positive cost, a firm lacking the flexibility to version and to reprice according to observed demand will generally limit its flagship product's quality.

Example. For *state-contingent* versioning policies a firm's investment in product development does not necessarily increase. Consider the case when the support $\mathcal{S}$ is finite containing only the positive elements $W_L = 2$ and $W_H = 4$, which correspond to a "low" and "high" demand realization respectively. The probability of high demand is denoted by π_H and the development cost is $\beta = 1/4$. Thus the ex-ante demand estimator is $\hat{W} = 4/(2 - \pi_H)$, which satisfies the viability condition (5.19) for all $\pi_H \in [0, 1]$. For $\hat{W}$ less than $6\beta/(1-\beta)^2 = 8/3$ (i.e., $\pi_H < 1/2$) the firm decides ex-ante to horizontally version, whereas for $\pi_H > 1/2$ vertical versioning is optimal. Under delayed differentiation the firm prefers to not commit to any particular mode of differentiation for a large range of intermediate π_H's. In the case of a "surprisingly" high realization of demand the ex-ante quality choice is relatively too small so that horizontal versioning is optimal. The converse is true for a low demand realization, when it is better to vertically version in order to make use of the "unexpectedly" high flagship quality. Figure 5.8 shows that the firm's development effort (proportional to q_1) may go both up or down with delayed differentiation as a result of actively using its added flexibility in the form of a state-contingent horizontal-vertical versioning policy.

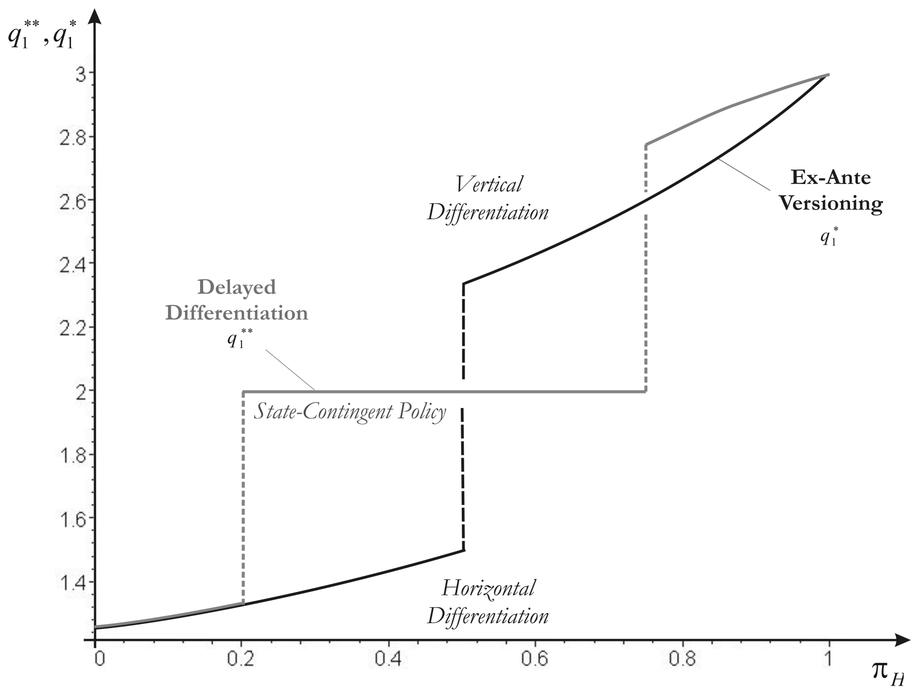

Figure 5.8: Optimal Choice of q_1 under Ex-Ante Versioning and under Delayed Differentiation (for $\beta = 1/4$, $\mathcal{S} = \{2, 4\}$).

Value of Perfect Information. The expected value of *perfect* information (EVPI) about the demand under ex-ante versioning is simply the difference between ex-ante optimal profit evaluated at W (in expectation) and evalutated at $\hat{W}$ respectively: EVPI(ex-ante versioning) = $E\Pi^*(W) - \Pi^*(\hat{W})$, where

$$\Pi^*(W) = \begin{cases} W(1-\beta)^2/3 - \beta, & \text{if } W \geq 6\beta/(1-\beta)^2, \\ W(1-\beta)^2/4 - \beta/2, & \text{if } W \in \left[2\beta/(1-\beta)^2, 6\beta/(1-\beta)^2\right], \\ 0, & \text{otherwise.} \end{cases}$$

Clearly EVPI is an upper bound for the value of the option to delay differentiation (Conrad 1980). It is generally *not* monotonic in the development cost β, as the way in which information alters the versioning decision depends on the firm's default no-information versioning policy.

Value of the Option to Delay Differentiation. The option value of not having to commit to the product portfolio ex ante is given by $V_d = \bar{\Pi}^{**} - \bar{\Pi}^*$. This value must be zero for large values of β, since both $\bar{\Pi}^{**}$ and $\bar{\Pi}^*$ are tightly bounded from below by zero and are strictly decreasing (and continuous) in β.

Proposition 5.7 *(i) The value of the option to delay differentiation, $V_d(\beta)$, is locally maximal at $\beta \in \left\{0, \left(3+\hat{W}-\sqrt{9+6\hat{W}}\right)/\hat{W}\right\}$. (ii) Its slope with respect to the development cost β is given almost everywhere by the difference of the optimal flagship qualities without and with delayed differentiation,*

$$\partial_\beta V_d \stackrel{a.e.}{=} q_1^* - q_1^{**}. \tag{5.27}$$

(iii) The option value is zero for large β's.

Development costs in which the firm is ex ante indifferent between vertical and horizontal versioning make the option to delay decisions about the composition of the product portfolio particularly valuable. A state-contingent policy specifying the mode of differentiation as a function of the realized demand clearly outperforms the noncontingent ex-ante decision. In particular, the value of flexibility is generally non-monotonic in quality cost, it depends on the way in which a deviation from the firm's default ex-ante strategy influences expected profits. Varying the development cost, the option value varies proportional to the difference $q_1^* - q_1^{**}$, which can be positive and negative. Figure 5.9 depicts the option value our previous two-state example (with $\mathcal{S} = \{W_L, W_H\}$) together with the expected value of perfect information as a function of β.

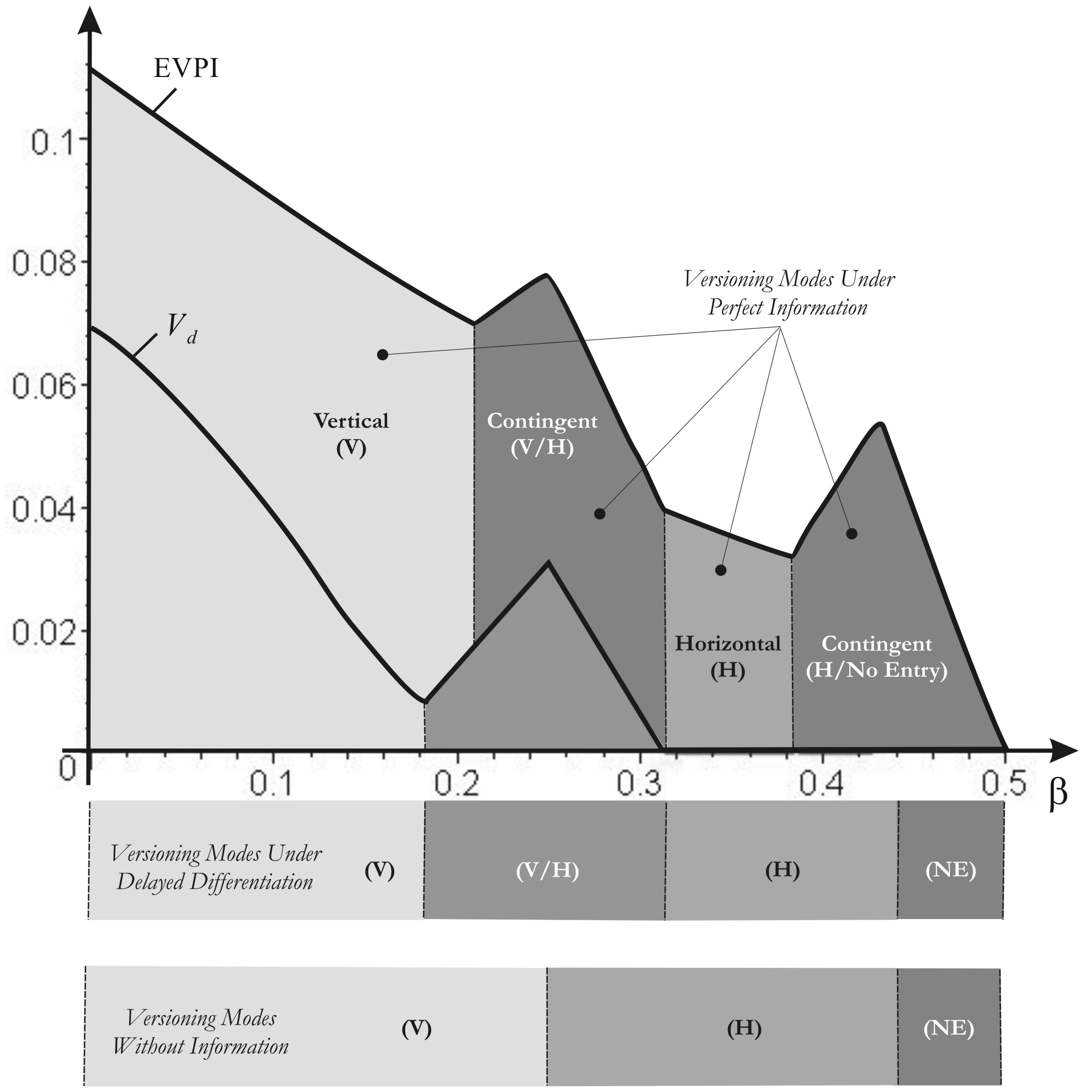

Figure 5.9: Expected Value of Perfect Information (EVPI) and Value of the Option to Delay Differentiation (V_d) as a Function of the Development Cost β (for $\mathcal{S} = \{2, 4\}$, $\pi_H = .5$).

Mode of Differentiation	*Consumer Surplus* $[u_k = q_k - p_k,\ \delta = p_1 - p_2]$
Vertical	$\frac{W-p_2}{W}\left(u_1 - \min\{u_1, 1\}/2\right)\min\{u_1, 1\}$
Horizontal (Separation)	$\frac{W-p_1}{W}\left(u_1 - \min\{u_1, \frac{1}{2}\}/2\right)\min\{u_1, \frac{1}{2}\}$
Horizontal (Overlap)	$\frac{W-p_1}{W}\left(u_1 - \frac{1-\delta}{4}\right)\frac{1-\delta}{2} + \frac{W-p_2}{W}\left(u_2 - \frac{1+\delta}{4}\right)\frac{1+\delta}{2} + \frac{\delta}{W}\left(u_2 - \frac{\min\{u_2,1\}}{2}\right)\min\{u_2, 1\}$

Table 5.1: Expected Consumer Surplus Given a Two-Element Product Portfolio Ω.

5.4.3 Effects on Consumer Surplus

Consumer surplus depends directly on the firm's product portfolio and can be directly determined from the set $\Omega = \{(p_k, q_k, z_k)\}_{k=1,2}$.

Proposition 5.8 *(i) Given a demand realization W and the firm's product portfolio Ω, the consumer surplus is given in Table 5.1. (ii) Under ex-ante optimal versioning, the expected consumer surplus is given by*

$$CS = \begin{cases} (2+\beta)/6, & \text{if } \beta \leq \beta', \\ (1+\beta)/8, & \text{if } \hat{W} \in [\beta', \beta''], \\ 0, & \text{otherwise}, \end{cases} \tag{5.28}$$

where $\beta' = \left(3 + \hat{W} - \sqrt{9 + 6\hat{W}}\right)/\hat{W}$ *and* $\beta'' = \left(1 + \hat{W} - \sqrt{1 + 2\hat{W}}\right)/\hat{W}$.

The first part of Proposition 5.8 describes how consumer surplus can be determined from the results in Sections 5.4.1 and 5.4.2. The second part shows that under ex-ante differentiation consumer surplus is piecewise linearly increasing in β. At the points of discontinuity β' and β'', where the first mode of differentiation changes and then the firm decides not to enter the market, consumer surplus drastically decreases. For each given versioning mode, consumer surplus is increasing in β because the firm needs compensates its decreasing quality investment by reducing price so as to ensure a large participation. In fact, expected market penetration[21] increases between two and four times faster than consumer surplus.

As shown above, the firm generally values versioning flexibility which may both increase or decrease the optimal ex-ante quality investment depending on the no-information default strategy. Interestingly, the increased profitability of the firm's investment does not always imply a decrease in consumer surplus. For low development cost, the benefits of the higher flagship quality may outweigh the firm's improved ability to screen the consumer base, leading

[21] Expected market penetration under ex-ante versioning is $(1+\beta)/2 \in [1/2, 1]$ for a horizontally and $(2+\beta)/3 \in [2/3, 1]$ for a vertically differentiated optimal product portfolio.

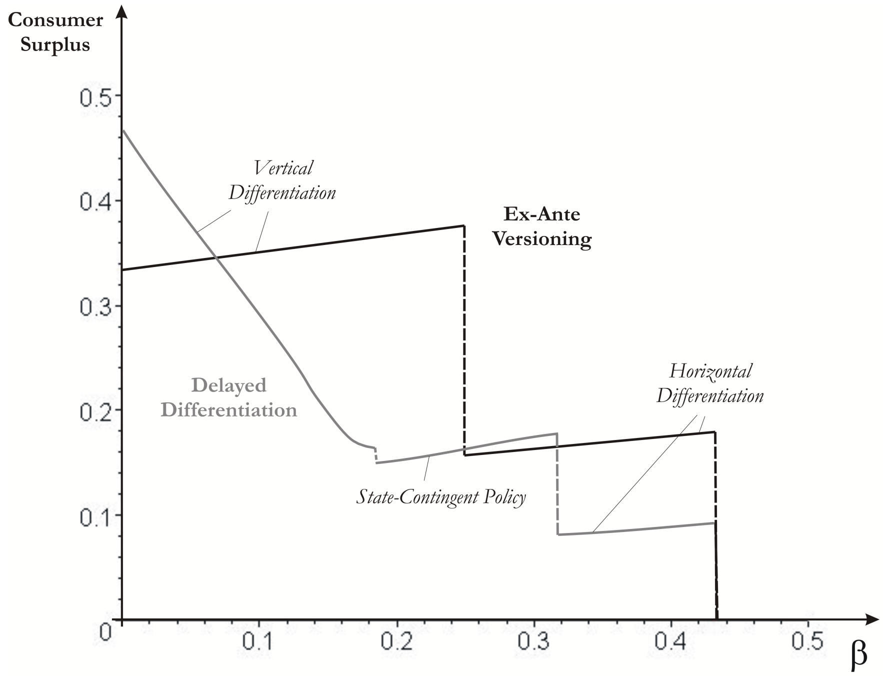

Figure 5.10: Consumer Surplus (CS) as a Function of the Development Cost β (for $\mathcal{S} = \{2, 4\}$, $\pi_H = .5$).

for low β's to a higher consumer surplus under delayed differentiation than under ex-ante versioning (cf. Figure 5.10).

5.4.4 Generalization to Goods with Positive Marginal Costs

A generalization of virtually all results in this paper to goods with positive marginal costs $c > 0$ can be achieved via a simple variable transformation. The only restrictive assumption is that the marginal costs for both goods be the same. This is realistic for many products, and especially those where the actual versioning decision can be taken far downstream in the supply chain (Anand and Mendelson 1998), e.g., by the retailer. Replacing the prices p_k in Section 5.3 by the markups $m_k = p_k - c$ (for $k = 1, 2$), together with using the constants $\hat{W}_c = \hat{W} - c$ and $\beta_c = \beta\hat{W}/\left(\hat{W} - c\right)$ is enough for all profit-maximization problems under ex-ante versioning to carry over to the case with $c > 0$. Under delayed differentiation the same change of variables can be used, only that the estimator $\hat{W}$ needs to be replaced by the actual demand realization W. In addition, the firm needs to consider $q_1' = q_1 - c$ for its second-stage versioning decision. Again all of the maximizers can be simply recovered by substituting the new constants. The overall effect of these transformations is that positive marginal costs are equivalent to increasing the development costs and also add an additional fixed cost of entry changing the firm's viability requirements for market entry.

5.5 Conclusion and Further Research

Even when a company can extend its product line by varying horizontal *and* vertical attributes of a flagship product, it is – given our assumptions of quasilinear preferences and homogeneous consumer distribution – not optimal to do both at the same time.[22] This finding is independent of the versioning cost structure, as long as $\alpha_M \geq \alpha_V, \alpha_H$. Nevertheless, under delayed differentiation the firm's optimal versioning policy contingent on the demand realization generally incorporates both modes of differentiation. For low demand realizations it is best to differentiate vertically adding a product of degraded quality to the then relatively high-performance flagship product to more adequately segment the consumer base by maintaining full horizontal market coverage. This can be accomplished by deliberately "damaging" the flagship product (Deneckere and McAfee 1996). On the other hand, for (unexpectedly) large demand realizations horizontal differentiation is generally superior, since the effective flagship product quality may not be high enough to guarantee full horizontal market coverage. For intermediate realizations of demand horizontal product cannibalization may be the best compromise. The latter generally implies price dispersion for products of equal quality. Despite the quite general insight from option theory that added flexibility usually leads to an increase of a firm's ex-ante investment which has become "less irreversible," we find that for multiattribute versioning the opposite may be true. Optimal upfront product development efforts may decrease, if due to the possibility of delayed differentiation a policy contingent on the demand state becomes optimal. This result is driven by the nonconvexity of the problem, horizontal and vertical differentiation yielding locally optimal profits.[23]

Examining the option value of delayed differentiation further we show that it contains a local maximum at the point of ex-ante indifference between the versioning modes, and that it naturally vanishes for very high product development costs. The effect of delayed differentiation on consumer surplus is mixed. Under ex-ante versioning consumer surplus heavily depends on

[22]This finding cannot be expected to hold for product portfolios with more than two products, the treatment of which in our framework poses significant analytical difficulties. It is also not necessarily what is observed in practice (cf. footnote 1).

[23]From a technical point of view, the maximizer (i.e., the "optimal policy") is upper hemicontinuous, essentially as a consequence of the continuity of the objective function. This is guaranteed by Berge's (1959) maximum theorem.

the mode of differentiation used and for each mode is monotonically increasing in the firm's development cost, as the firm tends to compensate an increase in quality cost β by lowering price and thereby enlarging the consumer base. If differentiation is delayed the versioning modes switch at different points so that effects on consumer surplus are ambiguous. Overall it tends to decrease as the firm's ability to make effective pricing and product-line-extension decisions greatly increases.

Future research could proceed along the following three promising axes: (1) extend the approach to more than two products. One of the main challenges here is that the potentially multiple overlaps are hard to capture analytically; (2) admit more general cost functions and demand distributions; and (3) incorporate competitive dynamics into the multiattribute versioning decision.

5.6 References

Anand, K.S., Mendelson, H. 1998. Postponement and information in a supply chain. Working Paper, Kellogg Graduate School of Management, Northwestern University.

Anderson, S., de Palma, A., Thisse, J.-F. 1992. *Discrete-Choice Theory of Product Differentiation.* MIT Press, Cambridge, Massachusetts.

Bakos, Y., Brynjolfsson, E. 1999. Bundling information goods: pricing, profits, and efficiency. *Management Science* **45**(12) 1613–1630.

Beath, J., Katsoulacos, Y. 1991. *The Economic Theory of Product Differentiation.* Cambridge University Press, New York.

Berge, C. 1959. *Espaces Topologiques et Fonctions Multivoques.* Dunod, Paris.

Bhargava, H.K., Choudhary, V. 2001. Information goods and vertical differentiation. *Journal of Management Information Systems* **18**(2) 89–106.

Bronshtein, I.N., Semendyayev, K.A. 1997. *Handbook of Mathematics.* Springer, New York.

Chamberlin, E. 1933. *The Theory of Monopolistic Competition.* Harvard University Press, Cambridge, Massachusetts.

Conrad, J.M. 1980. Quasi-option value and the expected value of information. *Quaterly Journal of Economics* **94**(4) 813–820.

Cremer, H., Thisse, J.-F. 1991. Location models of horizontal differentiation: a special case of vertical differentiation models. *Journal of Industrial Economics* **39**(4) 383–390.

Deneckere, R.J., McAfee, R.P. 1996. Damaged goods. *Journal of Economics and Management Strategy* **5**(2) 149–174.

Fisher, M., K. Ramdas, Ulrich, K.T. 1999. Component sharing in the management of product variety: a study of automotive braking systems. *Management Science* **45**(3) 297–315.

Gorman, W.M. 1968. Conditions for additive separability. *Econometrica* **36**(3-4) 605–609.

Hotelling, H. 1929. Stability in competition. *Economic Journal* **39**(153) 41–57.

Huchzermeier, A., Loch, C.H. 2001. Project management under risk: using the real options approach to evaluate flexibility in R&D. *Management Science* **47**(1) 85–101.

Jones, R., Mendelson, H. 1998. Product and price competition for information goods. Working Paper, Graduate School of Business, Stanford University.

Krishnan, V., Ulrich, K.T. 2001. Product development decisions: a review of the literature. *Management Science* **47**(1) 1–21.

Laffont, J.-J., Maskin, E., Rochet, J.-C. 1987. Optimal nonlinear pricing with two-dimensional characteristics. T. Groves, R. Radner, S. Steiger, Eds. *Information, Incentives, and Economic Mechanisms: Essays in Honor of Leonid Hurwicz.* University of Minnesota Press, Minneapolis, Minnesota, 255–266.

Lancaster, K.J. 1966. A new approach to consumer theory. *Journal of Political Economy*

74(2) 132–157.

Lee, H.L., Tang, C.S. 1997. Modeling the costs and benefits of delayed product differentiation. *Management Science* **43**(1) 40–53.

McDonald, R., Siegel, D. 1986. The value of waiting to invest. *Quarterly Journal of Economics* **101**(4) 707–728.

Mirrlees, J.A. 1971. An exploration in the theory of optimal income taxation. *Review of Economic Studies* **38**(2) 175–208.

Moorthy, K.S., Png, I. 1992. Market segmentation, cannibalization, and the timing of product introductions. *Management Science* **38**(3) 345–359.

Mussa, M., Rosen, S. 1978. Monopoly and product quality. *Journal of Economic Theory* **18**(2) 301–317.

Panzar, J.C., Willig, R.D. 1981. Economies of scope. *American Economic Review* **71**(2) 268–272.

Perloff, J., Salop, S. 1985. Equilibrium with product differentiation. *Review of Economic Studies* **52**(1) 107–120.

Pindyck, R.S. 1988. Irreversible investment, capacity choice, and the value of the firm. *American Economic Review* **78**(5) 969–985.

Rochet, J.-C., Choné, P. 1998. Ironing, sweeping and multidimensional screening. *Econometrica* **66**(4) 783–826.

Salop, S. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* **10**(1) 141–156.

Shaked, A., Sutton, J. 1987. Product differentiation and industrial structure. *Journal of Industrial Economics* **36**(2) 131–146.

Shapiro, C., Varian, H. 1998. Versioning: the smart way to sell information. *Harvard Business Review* **76**(6) 107–114.

Swaminathan, J.M., Tayur, S.R. 1998. Managing broader product lines through delayed differentiation using vanilla boxes. *Management Science* **44**(12-2) S161–S172.

Tirole, J. 1988. *The Theory of Industrial Organization.* MIT Press, Cambridge, Massachusetts.

Ulrich, K.T. 1995. The role of product architecture in the manufacturing firm. *Research Policy* **24**(3) 419–440.

Van Mieghem, J.A. 1998. Investment strategies for flexible resources. *Management Science*. **44**(8) 1071–1078.

Weisbrod, B. 1964. Collective-consumption services of individual consumption goods. *Quaterly Journal of Economics* **78**(3) 471–477.

5.7 Appendix: Proofs

Proof of Lemma 5.1. (i) This part is trivial. (ii) The probability density of W is by assumption given by the Lebesgue-measurable function $f \in L_1(\mathcal{S})$. Thus we have using Fubini's theorem:

$$\begin{aligned}\frac{\bar{W}}{\hat{W}} &= \left(\int_{\mathcal{S}} x f(x)\ dx\right)\left(\int_{\mathcal{S}} \frac{f(y)}{y} dy\right) \\ &= \int_{\mathcal{S}\times\mathcal{S}} \frac{x}{y} f(x) f(y)\ d(x,y) \\ &= \int_{\mathcal{S}\times\mathcal{S}} \frac{1}{2}\left(\frac{x}{y}+\frac{y}{x}\right) f(x) f(y)\ d(x,y) \\ &> \int_{\mathcal{S}\times\mathcal{S}} f(x) f(y)\ d(x,y) = 1,\end{aligned}$$

since $(x/y + y/x)/2 > 1$ for all $x, y > 0$ with $x \neq y$. Note that the set of points $(x,y) \gg 0$ where $x = y$ is of measure zero in $\mathcal{S} \times \mathcal{S}$, so that the strict inequality is warranted. ■

Proof of Proposition 5.3. (i) Since there is no overlap, the expected profits under horizontal differentiation (5.16) are under separation

$$\bar{\Pi}_{H_1}(p_1, p_2) = \frac{1}{2\hat{W}} \Big(p_1(\hat{W} - p_1)(1 - p_1 + p_2) + p_2(\hat{W} - p_2)(1 + p_1 - p_2) \Big) - \beta(1 + p_1 + p_2)/2.$$

Note that $\bar{\Pi}_{H_1}(p_1, p_2) = \bar{\Pi}_{H_1}(p_2, p_1)$ and therefore, if there is an asymmetric maximizer, there must be at least two of them. Hence, without loss of generality, assume that $p_1 \geq p_2$. Let us now introduce the variable $\delta = p_1 - p_2 \in [0, 1]$, which leads to the problem of maximizing

$$\bar{\Pi}_{H_1}(p_2 + \delta, p_2) = -\frac{p_2^2}{\hat{W}} + \left(1 - \beta - \frac{\delta(1-\delta)}{\hat{W}}\right) p_2 + \frac{1}{2}\left(\delta(1-\delta)\left(1 - \frac{\delta}{\hat{W}}\right) - \beta(1+\delta)\right)$$

with respect to $p_2 \in [0, \hat{W}]$ and $\delta \in [0, 1]$, subject to the additional constraint $0 \leq p_2 + \delta \leq \hat{W}$. Neglecting the latter constraint, the first-order condition with respect to p_2 is $\partial_{p_2}\bar{\Pi}_{H_1}(p_2 + \delta, p_2) = -\frac{2p_2}{\hat{W}} + 1 - \beta - \frac{\delta(1-\delta)}{\hat{W}} = 0$, which yields $p_2^*(\delta) = \frac{\hat{W}}{2}(1-\beta) - \frac{\delta(1-\delta)}{2} \leq \frac{\hat{W}}{2}(1-\beta)$. Thus,

$$\bar{\Pi}_{H_1}(p_2^*(\delta) + \delta, p_2^*(\delta)) = \frac{\hat{W}}{4}(1-\beta)^2 - \frac{\beta}{2} - \frac{\delta^2}{4\hat{W}}\left(1 - \delta^2 + 2\beta\hat{W}\right) \leq \frac{\hat{W}}{4}(1-\beta)^2 - \frac{\beta}{2}.$$

The equal sign in the last inequality holds, iff $\delta = 0$. The unique maximizers of the original maximization in the "separation" case, now satisfying *all* of the above constraints, are $p_1^* = p_2^* = \frac{\hat{W}}{2}(1-\beta)$, yielding optimal profits

$$\bar{\Pi}_{H_1}^* = \frac{\hat{W}}{4}(1-\beta)^2 - \frac{\beta}{2} = \bar{\Pi}_1^* + \frac{\beta}{2}. \tag{5.29}$$

These profits are by the amount $\beta/2$ larger than profits for the single-product monopoly and are positive (i.e., viable), iff

$$\hat{W} > \frac{2\beta}{(1-\beta)^2}. \tag{5.30}$$

In particular, a multiproduct monopolist firm can substitute horizontal differentiation for a quality increase, and thereby save on cost (provided that $\alpha_H < \beta/2$).

(ii) When there is overlap, the expected profit function (5.16) becomes

$$\begin{aligned} \bar{\Pi}_{H_2}(p_1, p_2) \;&=\; \frac{1}{2\hat{W}} \Big(p_1(\hat{W} - p_1)(1 - p_1 + p_2) + p_2(\hat{W} - p_2)(1 + p_1 - p_2) \\ &\quad + \max\{p_1(p_2 - p_1), p_2(p_1 - p_2)\}(1 - p_1 + p_2) \Big) - \beta(1 + p_2). \end{aligned}$$

Note that the single-product monopoly profits $\bar{\Pi}_1^*$ are attained for $p_1 = p_2$ and $p_1 = p_2 + 1$. As pointed out above, $p_1 \geq p_2$. Let us again introduce $\delta = p_1 - p_2 \in [0, 1]$. Then

$$\bar{\Pi}_{H_2}(p_2 + \delta, p_2) = -\frac{p_2^2}{\hat{W}} + \left(1 - \beta - \frac{\delta(1-\delta)}{2\hat{W}}\right) p_2 + \frac{\delta(1-\delta)}{2}\left(1 - \frac{\delta}{\hat{W}}\right) - \beta.$$

The first-order condition with respect to p_2, $\partial_{p_2}\bar{\Pi}_{H_2}(p_2+\delta, p_2) = 0$, yields $p_2^*(\delta) = \frac{\hat{W}}{2}(1-\beta) - \frac{\delta(1-\delta)}{4}$, and therefore

$$\begin{aligned}\bar{\Pi}_{H_2}(p_2^*(\delta)+\delta, p_2^*(\delta)) &= \frac{\hat{W}}{4}(1-\beta)^2 - \beta - \frac{\delta(1-\delta)}{16\hat{W}}\left(\delta^2 + 7\delta - 4\hat{W}(1+\beta)\right) \\ &= \bar{\Pi}_1^* + \frac{\delta(1-\delta)}{16\hat{W}}\left(\delta - \delta_-\right)\left(\delta_+ - \delta\right), \end{aligned} \tag{5.31}$$

where $\delta_\pm = \left(-7 \pm \sqrt{49 + 16\hat{W}(1+\beta)}\right)/2$. It is $\delta_- < 0 < \delta_+$, and therefore $\bar{\Pi}_{H_2}^*$ in expression (5.31) is greater than $\bar{\Pi}_1^*$, iff $0 < \delta < \min\{\delta_+, 1\}$. The following lemma helps establishing that the optimal profit $\bar{\Pi}_{H_2}^*$ is strictly inferior to $\max\{\bar{\Pi}_{H_1}^*, \bar{\Pi}_V^*\}$. In other words, if pure horizontal differentiation is superior to pure vertical differentiation, then separation will maximize profits.

Lemma 5.2 *Let the parameters $0 < \beta < 1$ and $\hat{W} > 0$ be given. Then for any $\delta \in [0, 1]$:*

$$\delta(1-\delta)(\delta - \delta_-)(\delta_+ - \delta) < \max\{8\beta\hat{W}, (2\hat{W}(1-\beta))^2/3\}. \tag{5.32}$$

Proof. Let $\bar{\delta} = \min\{1, \delta_+\}$. As the RHS of (5.32) is always positive, we can limit ourselves to $\delta \in I = [0, \bar{\delta}]$ for which the LHS is nonnegative. Set $f(\delta)$ to be equal to the LHS of (5.32). Since f is a polynomial with a set of roots $\{\delta_-, 0, \delta_+, 1\}$, where $\delta_- < 0 < \bar{\delta}$, it is

$$\max_{\delta \in I} |f'(\delta)| = \max\{f'(0), -f'(\bar{\delta})\}.$$

More specifically, by the mean value theorem and the fact that $f''|_I < 0$, for all $\delta \in I$:

$$f(\delta) \le \min\{f'(0)\cdot\delta, -f'(\bar{\delta})(\bar{\delta} - \delta)\} \le -\frac{f'(\bar{\delta})f'(0)\cdot\bar{\delta}}{f'(0) - f'(\bar{\delta})}. \tag{5.33}$$

Thereby $f'(0) = -\delta_-\delta_+$ and $-f'(\bar{\delta}) = \max\{\delta_+(\delta_+ - \delta_-)(1-\delta_+), (1-\delta_-)(\delta_+ - 1)\}$. Let us first consider the case when $\bar{\delta} = 1$, or equivalently

$$\hat{W}(1+\beta) \ge 2. \tag{5.34}$$

Then by (5.33)

$$f(\delta) \le \frac{-\delta_-\delta_+(1-\delta_-)(\delta_+ - 1)}{-\delta_-\delta_+ + (1-\delta_-)(\delta_+ - 1)} = \frac{16\hat{W}(1+\beta)\left(\hat{W}(1+\beta) - 2\right)}{4\hat{W}(1+\beta) + 4\left(\hat{W}(1+\beta) - 2\right)} = \frac{2\hat{W}(1+\beta)\left(\hat{W}(1+\beta) - 2\right)}{\hat{W}(1+\beta) - 1}. \tag{5.35}$$

One can verify by straightforward manipulations that the RHS of (5.35) is less than $8\beta\hat{W}$, iff

$$\hat{W}(1 - 3\beta) < 2\frac{1-\beta}{1+\beta},$$

which establishes inequality (5.32) for all

$$(\beta, \hat{W}) \in D_1 = \left\{ (\breve{\beta}, \breve{W}) : 0 < \breve{\beta} < \frac{1}{3},\ \breve{W} < \frac{1-\breve{\beta}}{(1+\breve{\beta})(1-3\breve{\beta})} \right\} \cup \Bigg([1/3, 1) \times (0, \infty) \Bigg).$$

On the other hand, one can verify that the RHS of (5.35) is less than $\left(2\hat{W}(1-\beta)\right)^2/3$, iff

$$\frac{2}{1+\beta} + \frac{2}{3}\frac{(1-\beta)^2}{1+\beta}\hat{W}^2 - \left(\frac{2}{3}\left(\frac{1-\beta}{1+\beta}\right)^2 + 1 \right)\hat{W} > 0. \tag{5.36}$$

Restricting β to the interval $(0, 1/3]$, inequality (5.36) is satisfied, if

$$3+(1-\beta)^2\hat{W}^2 - \left(\frac{(1-\beta)^2}{1+\beta} + \frac{3}{2}(1+\beta) \right)\hat{W} \geq 3 + \left[\frac{2}{3}(1-\beta)\hat{W} - \left(\frac{(1-\beta)^2}{1+\beta} + \frac{3}{2}(1+\beta) \right) \right]\hat{W} > 0.$$

The expression in square brackets can be further minorized by setting $\beta = 1/3$, based on the fact that the derivative of that expression with respect to β,

$$-\frac{4}{3}\frac{1}{1+\beta} - \frac{4}{3}\frac{1-\beta}{(1+\beta)^2} + 2\frac{1-\beta}{1+\beta} + \frac{(1-\beta)^2}{(1+\beta)^2} - \frac{3}{2}$$

is negative on $(0, 1/3]$. In particular,

$$3 + \hat{W}\left[\frac{2}{3}(1-\beta)\hat{W} - \left(\frac{(1-\beta)^2}{1+\beta} + \frac{3}{2}(1+\beta) \right) \right]_{\beta=1/3} = \frac{1}{9}(\hat{W}-3)(4\hat{W}-9) > 0$$

implies (5.32) for all $(\beta, \hat{W}) \in D_2 = (0, 1/3] \times ((0, 9/4) \cup (3, \infty))$. Alternatively one can minorize the RHS of (5.36) by substituting

$$\hat{W} = \frac{2(1-\beta)^2 + 3(1+\beta)^2}{4(1-\beta)^2(1+\beta)},$$

its unique minimizer. Relation (5.36) is satisfied for a given $\beta \in (0, 1/3)$ and all $\hat{W} > 0$, if

$$23\beta^4 - 20\beta^3 - 150\beta^2 - 20\beta + 23 > 0.$$

The last polynomial has the set of approximate roots $\{-2.0221, -.4945, .3269, 3.0593\}$, so that (5.32) follows for $(\beta, \hat{W}) \in D_3 = (0, 3/10) \times (0, \infty)$. Since $D_1 \cup D_2 \cup D_3 = (0, 1) \times (0, \infty)$, we have indeed shown that (5.32) holds for all relevant[24] $(\beta, \hat{W})$ whenever $\bar{\delta} = 1$. Let us now examine the case when $\bar{\delta} = \delta_+ < 1$ or in other words

$$0 < \hat{W} < \frac{2}{1+\beta}. \tag{5.37}$$

[24]Note that $\bar{\delta} = 1$ is equivalent to (5.34), which slightly restricts the relevant domain of $(\beta, \hat{W})$ to a subset of $(0, 1) \times (0, \infty)$.

Using (5.33) write

$$\begin{aligned} f(\delta) &\leq \frac{-\delta_-\delta_+^3(\delta_+-\delta_-)(1-\delta_+)}{-\delta_-\delta_+ + \delta_+(\delta_+-\delta_-)(1-\delta_+)} < \frac{\delta_+^2(\delta_+-\delta_-)(1-\delta_+)}{2-\delta_+} \\ &= \frac{1}{4}\frac{\sqrt{49+16\eta}\left(-9+\sqrt{49+16\eta}\right)\left(-7+\sqrt{49+16\eta}\right)^2}{-11+\sqrt{49+16\eta}}, \end{aligned} \tag{5.38}$$

where $\eta = \hat{W}(1+\beta)$ and by (5.37) it is $0 < \eta < 2$. Straightforward computations yield that expression (5.38) is less than $8\beta\hat{W} = 8\eta(1 - 1/(1+\beta))$, iff

$$(\beta, \eta) \in E_1 = \left\{(\breve{\beta}, \breve{\eta}) : 0 < \breve{\beta} < R_1(\breve{\eta}), 0 < \breve{\eta} < 2\right\}, \tag{5.39}$$

where

$$\begin{aligned} R_1(\eta) \;=\; & \left(2\eta^3 - 153\eta^2 - 537\eta - 294\right)^{-1}\Bigg(2\eta^3 + 279\eta^2 + 306\eta - 1764 + \sqrt{49+16\eta}\left(15\eta + 12\eta^2 - 294\right) \\ & + \Big[4032\eta^5 + 70620\eta^4 - 91110\eta^3 - 1744848\eta^2 - 444528\eta + 7260624 \\ & + \sqrt{49+16\eta}\left(48\eta^5 + 6756\eta^4 + 14538\eta^3 - 197208\eta^2 - 232848\eta + 1037232\right)\Big]^{1/2}\Bigg). \end{aligned}$$

On the other hand, expression (5.38) is less than $(2\hat{W}(1+\beta))^2/3 = 4\eta^2\left(1 - 4/(1+\beta) + 4/(1+\beta)^2\right)/3$, iff

$$(\beta, \eta) \in E_2 = \left\{(\breve{\beta}, \breve{\eta}) : R_2(\breve{\eta}) < \breve{\beta} < 1, 0 < \breve{\eta} < 2\right\}, \tag{5.40}$$

where

$$R_2(\eta) = \frac{\sqrt{49+16\eta}\,(23\eta+98) - 16\eta^2 - 273\eta - 686}{16\eta^2 + 295\eta + 686 - \sqrt{49+16\eta}\,(25\eta+98)}.$$

Since $R_1(\eta) > R_2(\eta)$ for all $0 < \eta < 2$, it is $E_1 \cup E_2 = (0,1) \times (0,2)$, and therefore (5.32) holds also for the case that $\bar{\delta} = \delta_+$, which concludes the proof of Lemma 5.2. □

As mentioned before, based on Lemma 5.2 as well as (5.10), (5.29), and (5.31) pure horizontal differentiation in the overlap mode is strictly dominated by *either* pure vertical differentiation *or* pure horizontal differentiation in the separation mode. This concludes the proof of Proposition 5.3. ■

Proof of Proposition 5.4. (i) For zero versioning costs, condition (5.19) is necessary and sufficient, since the viability domains of single-product, pure vertical differentiation, and mixed differentiation are subsets of the region of positive profits for pure horizontal differentiation in separation mode, given precisely by (5.19). Thus, if the latter condition is satisfied there exists

$\varepsilon > 0$ such that $\alpha_H \in [0, \varepsilon]$ is small enough for the resulting profits to be positive. (ii) This is an immediate consequence of (5.19) being necessary and sufficient for viability under zero versioning costs. (iii) From $\hat{W} \in (2\beta/(1-\beta)^2, 4\beta/(1-\beta)^2)$ it follows that a single-product monopoly is not viable according to (5.7), while at the same time (5.19) is satisfied. ■

Details on the Horizontal-Overlap Case under Delayed Differentiation. Since $q_1 \geq q_2$ we have that necessarily $p_1 \geq p_2$, which generally implies price dispersion for goods of equal quality. In the following it it useful to distinguish three cases. (i) If $q_1 - p_2 < 1$, then $q_1 - p_1 < 1$ and $q_2 - p_2 < 1$. We can conclude that $p_1, p_2 \in [q_1 - 1, q_1]$ or in other words that $p_1 - p_2 < 1$. Thus $q_1 = q_2$ (pure horizontal differentiation). The profit-maximization problem is

$$(p_1, p_2) = \arg\max_{p_1, p_2} \left\{ \frac{W - p_1}{2W} \left(p_1 + p_2 - (p_1 - p_2)^2 \right) + \frac{p_2(p_1 - p_2)(q_1 - p_2)}{W} \right\}. \tag{5.41}$$

Setting $\delta = p_1 - p_2 \in [0, 1]$ it is possible to rewrite the profit function as

$$\Pi_{H_2}(p_2 + \delta, \delta) = \frac{(W - p_2 - \delta)\left(2p_2 - \delta(1 - \delta)\right) + 2p_2\delta(q_1 - p_2)}{2W} - \beta q_1.$$

Any interior extremum satisfies the necessary optimality condition for p_2, from which we obtain

$$p_2^*(\delta) = \frac{2W + \delta\left(2(q_1 - 1) - (1 - \delta)\right)}{4(1 + \delta)}.$$

By substituting this expression, the profit function Π_{H_2} can be expressed as a function of δ only, and the first-order condition with respect to δ is

$$a\delta^4 + b\delta^3 + c\delta^2 + d\delta + e = 0, \tag{5.42}$$

with $a = 27/16$, $b = (q_1 + 3)/2 - W$, $c = (q_1^2 - 5W)/4 - 17/16$, $d = (q_1^2 - 3q_1 + W)/2 + 1/8$, and $e = -W(W + 1 - 2q_1)/4$. To analytically find the roots of this fourth-order polynomial we follow Bronshtein and Semendyayev (1997). Substituting $y = \delta + b/(4a)$, equation (5.42) can be written equivalently in the reduced form

$$y^4 + Py^2 + Qy + R = 0, \tag{5.43}$$

where $P = (c/a) - 3b^2/\left(8a^2\right)$, $Q = d/a - cb/a^2 + b^3/\left(8a^3\right)$, and $R = e/a + \left(cb^2/16 - db/4\right)/a^2 - 3b^4/\left(256a^4\right)$. The solution behavior of (5.43) depends on the behavior of its *cubic resolvent*,

$$z^3 + 2Pz^2 + (P^2 - 4R)z - Q^2 = 0,$$

the roots of which are all real[25] and given by $z_k = 2\rho\cos(\varphi/3 + 2(k-1)\pi/3)$, for $k = 1, 2, 3$, with $\rho = \sqrt{4P^2 - 3Q}/3$ and $\varphi = \arccos\left(\left(2P(P^2-4R)/3 + Q^2 - 16P^3/27\right) / \left(2\rho^3\right)\right)$. From this, the roots of the reduced-form polynomial in (5.43) can be computed to

$$\begin{aligned} y_1 &= \left(\sqrt{z_1} + \sqrt{z_2} - \sqrt{z_3}\right)/2, \\ y_2 &= \left(\sqrt{z_1} - \sqrt{z_2} + \sqrt{z_3}\right)/2, \\ y_3 &= \left(\sqrt{z_1} + \sqrt{z_2} + \sqrt{z_3}\right)/2, \\ y_4 &= \left(\sqrt{z_1} - \sqrt{z_2} - \sqrt{z_3}\right)/2, \end{aligned}$$

and a local maximizer of $\Pi_{H_2}(p_2^*(\delta) + \delta, p_2^*(\delta))$ is

$$\delta^* = y_1 - \frac{b}{4a}, \tag{5.44}$$

as the second-order condition is satisfied at that point.[26]

(ii) If $q_1 - p_2 \geq 1$, then the optimization problem is

$$\begin{aligned} (p_1, p_2) &= \arg\max_{p_1,p_2} \left\{ \frac{p_1(W-p_1)(1-p_1+p_2)}{2W} + \frac{p_2(p_1-p_2)}{W} + \frac{p_2(W-p_1)(1+p_1-p_2)}{2W} \right\} \\ &= \arg\max_{p_1,p_2} \left\{ \frac{W-p_1}{2}\left(p_1 + p_2 - (p_1-p_2)^2\right) + p_2(p_1-p_2) \right\}. \end{aligned} \tag{5.45}$$

The quality q_2 is actually indeterminate in $[1+p_2, q_1]$, but without loss of generality we set $q_2 = q_1$ (which is true for even very small versioning cost). Note that $p_1 - p_2 \leq 1$, otherwise there is only one product. Also, this case is equivalent to the corresponding ex-ante versioning situation for $\beta = 0$. Introducing $\delta = p_1 - p_2 \in [0, 1]$, we obtain

$$\Pi_{H_2}(p_2 + \delta, \delta) = \frac{(W - p_2 - \delta)\left(2p_2 + \delta - \delta^2\right) + 2p_2\delta}{2W} - \beta q_1.$$

Assuming that there is an interior extremum the first-order condition with respect to p_2 is

$$2W - 4p_2 - \delta(1-\delta) = 0,$$

so that $p^*(\delta) = W/2 - \delta(1-\delta)/4$. The resulting expression for the profits under horizontal versioning with overlap is

$$\Pi_{H_2}(\delta) = \frac{W}{4} - \beta q_1 + \frac{\delta(1-\delta)(\delta-\delta_-)(\delta_+-\delta)}{16W}, \tag{5.46}$$

[25] This follows from the fact that the *discriminant* $\Delta = \left((P^2 - 4R - 4P^2/3)/3\right)^3 + \left(16P^3/27 - 2P(P^2-4R)/3 - Q^2\right)^2/4$ is negative in the relevant domain, which can be verified numerically.

[26] An analytical test of the second-order condition is complicated, but it can be readily verified numerically.

where $\delta_{\pm} = \left(-7 \pm \sqrt{16W+49}\right)/2$. The first-order condition with respect to δ is then

$$\delta^3 + \frac{9}{2}\delta^2 - \left(2W + \frac{7}{2}\right)\delta + W = 0. \tag{5.47}$$

Using the cardanic formula to solve the cubic equation (5.47) it turns out that for any W there are three real roots, of which only one is associated with an interior maximum of the fourth-order polynomial at (5.46),

$$\delta^*(W) = \sqrt{\frac{8W+41}{3}} \sin\left(\frac{\pi}{6} - \frac{1}{3}\arccos\left(\frac{48(W+3)}{\sqrt{(8W+41)^3/3}}\right)\right) - \frac{3}{2}. \tag{5.48}$$

It is $\delta^*(W) \in [0, 1/2]$ strictly increasing in W with $\delta^*(0) = 0$ and $\lim_{W\to\infty} \delta^*(W) = 1/2$.

(iii) In the intermediary case with binding constraint, $q_1 - p_2 = 1$, the optimization problem is identical to (5.45), subject to $p_2 = q_1 - 1$,

$$\Pi_{H_2}(\delta) = \frac{(W - (q_1 - 1))(q_1 - 1)}{W} - \beta q_1 + \frac{\delta(1-\delta)(A-\delta)}{2W}$$

with the abbreviation $A = W - (q_1 - 1)$. The first-order necessary optimality condition is $3\delta^2 - 2(A+1)\delta + A = 0$, and the maximizer is therefore

$$\delta^* = \frac{A + 1 - \sqrt{A^2 - A + 1}}{3}. \tag{5.49}$$

The second-order condition is satisfied here, since $\Pi''_{H_2}(\delta^*) = -2\sqrt{A^2 - A + 1} < 0$. This concludes our analytical discussion of the horizontal-overlap case under delayed differentiation. □

Proof of Proposition 5.5. (i) It is sufficient to show that the profit function under mixed ex-post differentiation is dominated by pure horizontal differentiation with overlap. Let (p_1, q_1) be such that $1/2 \leq q_1 - p_1 < 1$. Under horizontal separation (in which the local mixed-differentiation optimum occurs) this implies that $q_1 - p_1 = 1 - (q_2 - p_2) < 1$, and furthermore, as shown in Section 5.4.1, then $p_1^* = p_2^* = W/2$. Thus, also $q_1 - p_2^* < 1$ so that we can compare the profit function under horizontal separation to the maximand in (5.41) discussed earlier in this Appendix for the case that $q_1 - p_2 < 1$. Subtracting the latter from the former yields $-p_1(W - p_1)(1 - (q_1 - p_2))/2 - p_2(W - p_2)(q_1 - p_2) - p_2(p_1 - p_2)(q_1 - p_2)(q_1 - p_2) < 0$, so that mixed differentiation is strictly dominated. (ii) The claim that horizontal cannibalization is sometimes optimal can be shown by example and thus Figure 5.7 is sufficient as a proof. It is not optimal for large q_1 (compared to W), since it can be shown analogously to the computation in part (i) that pure vertical differentiation is globally better. For q_1 small compared to W horizontal overlap cannot be achieved. Thus, horizontal overlap can only be best for "intermediate" realizations of W. This concludes the proof of Proposition 5.5. ■

Proof of Proposition 5.6. Assume first that ex post pure vertical differentiation is optimal, independent of the demand realization W. Then differentiating the expression (5.22) for $\Pi_V^*(q_1; W)$ and then taking the expectation with respect to W yields[27]

$$-\beta + \int_{W_1}^{W_2} \left(1 - \frac{3(q_1^{**} - 1)}{2w}\right) f(w)dw + \int_{W_2}^{\bar{s}} p_1^* \left(1 - \frac{3p_1^*}{4w}\right) f(w)dw = 0,$$

where W_1, W_2, p_1^*, and $\bar{s}$ are as the corresponding part of Proposition 5.6. The last equation is equivalent to (5.25). Next assume that ex post pure vertical differentiation is optimal, independent of the demand realization W. As before we determine the first-order condition by first taking the derivative of $\Pi_{H_1}(q_1; W)$ in (5.24) with respect to q_1 and then taking the expectation with respect to W. This yields

$$-\beta + \int_{W_1}^{W_2} \left(1 - \frac{2q_1^{**} - 1}{w}\right) f(w)dw + 2\int_{W_2}^{\bar{s}} p_1^* \left(1 - \frac{p_1^*}{w}\right) f(w)dw = 0,$$

which is equivalent to (5.26) using the definitions of W_1, W_2, p_1^{**}, and $\bar{s}$ as in the corresponding part of Proposition 5.6. ■

Proof of Proposition 5.7. (ii) By Berge's (1959) maximum theorem the option value V_d is continuous in β as difference of two continuous functions $\bar{\Pi}^{**}$ and $\bar{\Pi}^*$. Also as a consequence of the maximum theorem both maximizers q_1^{**} and q_1^* are generally set-valued and upper hemicontinuous in the parameter β. The maximizers are thereby set-valued, iff the firm is indifferent between several different q_1-values. Let $\hat{W} > 0$ be given. Following our developments in Section 5.3, $q_1^*(\beta)$, $0 \le \beta \le 1$, is given by

$$q_1^*(\beta) = \begin{cases} 1 + 2\hat{W}(1-\beta)/3, & \text{if } \beta \ge \beta', \\ 1/2 + \hat{W}(1-\beta)/2, & \text{if } \beta \in [\beta', \beta''], \\ 0, & \text{otherwise,} \end{cases}$$

where $\beta' = \left(3 + \hat{W} - \sqrt{9 + 6\hat{W}}\right)/\hat{W}$ and $\beta'' = \left(1 + \hat{W} - \sqrt{1 + 2\hat{W}}\right)/\hat{W}$. Only for $\beta \in \{\beta', \beta''\}$ is q_1^* possibly set-valued; on $[0, \beta'') \setminus \{\beta'\}$ it is strictly monotonically decreasing. Similarly, the optimal flagship quality under delayed differentiation, q_1^{**}, can only be set-valued at

[27] This means to switch differentiation and integration when determining the first-order conditions of the optimization problem (5.20). This naturally presupposes sufficient regularity of the probability density function f.

$0 < \beta_1, \beta_2, \beta_3 < 1$, where $\beta_1 \leq \beta' < \beta_2 < \beta'' \leq \beta_3$, corresponding to the vertical-contingent (β_1), contingent-horizontal (β_2), and horizontal-no entry (β_3) mode transitions. We have noted in Section 5.4 that expressions (5.22) and (5.24) are continuously differentiable in q_1. The same is true for the corresponding expression in the overlap mode. In fact, these expressions are twice continuously differentiable ("smooth") almost everywhere, and thus the envelope theorem can be applied in the smooth portions, so that we obtain relation (5.27). (i) At $\beta = \beta'$ the firm is indifferent between horizontal and vertical versioning modes. But if the firm can observe demand $W \neq \hat{W}$ *before* taking the versioning decision,[28] then it generally will not be indifferent between horizontal and vertical differentiation, so that a state-contingent policy of intermediate ex-ante quality will at least weakly increase profits at $\beta = \beta'$. As a consequence, $q_1^{**}(\beta') \in (q_1^*(\beta'+), q_1^*(\beta'-))$ so that with (5.27) we obtain a local maximum of V_d at $\beta = \beta'$, in other words a state-contingent policy will generally improve profits around β'. Consider now $\beta = 0$. Then $q_1^{**}(0) = 1 + 2\bar{s}/3$ with $\bar{s} = \max\{\mathcal{S}\}$. On the other hand, it is following Section 5.3.2: $q_1^*(0) = 1 + 2\hat{W}/3 \leq q_1^{**}(0)$ because generally $\bar{s} \geq \hat{W}$. (iii) This part is trivial: for $\beta \geq \beta_3, \beta''$, the firm does not enter the market and thus the option value of being able to delay differentiation is not worth anything. ■

Proof of Proposition 5.8. (i) Consider first the case of pure vertical differentiation. Consumer surplus CS_V is then

$$\begin{aligned} CS_V &= \frac{1}{W} \int_0^{\min\{u_1,1\}} \left[(u_1 - v)(w - p_1) + (u_1 - v)\right] dv \\ &= \frac{W - p_2}{W} \left[u_1 v - \frac{v^2}{2}\right]_0^{\min\{u_1,1\}}, \end{aligned}$$

an expression equivalent to the first entry in Table 5.1. The consumer surplus for pure horizontal differentiation with separation follows in the same manner. We now turn to the case to horizontal differentiation with overlap, for which the consumer surplus can be computed as follows:

$$CS_{H_2} = \frac{W - p_1}{W} + \frac{W - p_2}{W} \int_0^{\frac{1+p_1-p_2}{2}} (u_2 - v) dv + \frac{p_1 - p_2}{W} \int_0^{\min\{u_1,1\}} (u_2 - v) dv.$$

Using the abbreviation $\delta = p_1 - p_2$, we obtain the third expression in Table 5.1 by straightforward integration. (ii) Substituting the ex-ante optimal product portfolio into the first two

[28] If $W = \hat{W}$ (which can only happen with positive probability for certain "pathological" discrete distributions of W), then the firm might be indifferent between horizontal and vertical versioning modes, even under delayed differentiation.

expressions for consumer surplus in Table 5.1, one immediately obtains (5.28). ■

Since it is Reason which shapes and regulates all other things, it ought not itself to be left in disorder.
— EPICTETUS

Chapter 6

Discussion, Conclusion, and Further Research

6.1 Discussion and Conclusion

Let me briefly return to the three research questions posed in Chapter 1 and discuss some of the results obtained in this dissertation. I will also mention some specific directions for future research where appropriate. More general directions for research are outlined in Section 6.2.

QUESTION 1 *How can the value of information be generally defined in terms of an agent's willingness to pay and willingness to accept, and what is the precise relationship between the two welfare measures?*

In Chapter 2, I have provided the definition of information value as compensating variation (and in some cases equivalent variation) in the spirit of, for instance, LaValle (1968) or Kihlstrom (1974). I have argued that the loss of monetary wealth to the payment for information often entails a contraction of the decision maker's action set. Thus, there is a natural tradeoff between engaging in productive actions and procuring information to enhance these actions (Arrow 1974, p. 49), which are then generally bound to lie in a subset of the action set available under full wealth.

In certain cases information is rival, in the sense that once the seller has transferred the information to the agent, the seller loses any right to use the information himself (such as for a patent). In that case, the seller's WTA for the information would have to be below the buyer's WTP for trade to take place. In Section 2.3, I provide an exact relation between WTA and WTP, which amounts to a relation between compensating and equivalent variation. Thus, in effect by requiring that WTA $\leq$ WTP as a condition for trade of a unique object between individuals of equivalent utility representation U and wealth endowment, we have effectively in relations (2.14) and (2.15) an "endogenous" criterion for bilateral trade between agents of identical preferences to occur.[1]

As pointed out in Chapter 2, the value of *transferred* information generally depends on the precise details of the payment to an information seller and its contingencies upon actions, payoffs and/or signal realizations.[2] Only in the *transfer* of information is the value of information actually realized by an information seller. The amount that the information seller can extract from a given agent (for a given purpose) constitutes the true *realizable* value of information, which naturally has to depend on the precise context in which the information transfer is to take place. If the payment for the information depends in some way on a hidden action by the buyer ("agent"), which may include the use of the information itself (in case payment is deferred), then the problem of maximizing the value of a given signal to the information seller ("principal") becomes one of designing an optimal contract (possibly subject to certain limitations in contractibilities). A resulting optimal contract can naturally include the strategic modification of the information itself.

QUESTION 2 *Consider a principal-agent environment.*

(i) What are the incentives for the distortion of information in a principal-agent relationship where information exerts an externality?

(ii) What are the consequences for organizational design for an informed principal?

In Chapter 3, I have shown that consciously damaging information before its transfer may allow the information seller to mitigate moral hazard that arises through limited liability of the information buyer in the presence of an ex-post payment component. Limiting the buyer's downside risk might induce "going-for-broke" behavior in the sense that the buyer-investor tends to invest more wealth into the risky asset than desired by the information seller. The moral hazard can be sometimes reduced by misreporting "bad" signal states as "good" ones, thereby "poisoning" the good news, which may lead to a more conservative investment behavior of the information buyer. We have seen in Chapter 3 that information garbling can only be beneficial for imperfect information and in the presence of signal states that result in significant agent moral hazard.

[1]More precisely, for trade to occur, the integral expressions in (2.14) and (2.15) have to be *negative*.

[2]This aspect is not entirely new, but not very well articulated in the literature.

In organizations (or hierarchies) it is often the "center" (i.e., the principal) that holds a large portion of the information relevant to a decentralized decision maker (i.e., the agent). Typically the agent's actions are only imperfectly observable to the principal and standard incentive theory (e.g., Laffont and Martimort 2002) describes how to set *monetary* incentives such as to maximize the benefits to the principal, subject to securing the agent's voluntary participation and incentive compatibility in her actions. If information is included in the principal's transferrable assets, then – as a consequence of the discussion in Chapter 3 – it may be best for the principal not to transmit the full information to the agent. Anticomplementarities between effort and information can be taken advantage of by the principal by withholding possibly mission-critical information in order to induce higher agent effort (Croson and Weber 2002).

Information transfer and use in *multi*agent environments is considered in Chapters 4 and 5 of this dissertation that address the following question.

Question 3 *Consider a multiagent environment.*

(i) How should a heterogeneous group of agents share cooperatively an investment for a common source of information?

(ii) What is the option value of being able to wait for information that helps a firm to screen a heterogeneous consumer base?

By jointly financing the generation of non-rival information a number of heterogeneous agents may be able to overcome the critical mass for the creation of a common information source and thereby utilize the increasing returns to scale from the fixed-cost investment.[3] Nash bargaining provides by construction an efficient mechanism for the joint allocation of resources. The bargaining outcome thereby depends on the project parameters (volume, riskiness), the agents' characteristics (respective wealth, risk aversion), and the signal's fidelity ("confidence" in a Bayesian framework). Due to network externalities the size of the fixed-cost block for the shared information source may also depend on the number of agents participating in the common investment. In Chapter 4, I provide a general framework for analyzing the sharing problem in a nonstrategic context. It is shown that for heterogeneous agents, due to the differences in their evaluations of the size of the bargaining "pie" (the value of a shared information source can be quite different for each agent), the comparative statics may bear surprising re-

[3] The variable cost of adding members may of course introduce further increases in the returns, or – as in the case of convex costs – possible decreases in the returns, cf. (4.16).

sults, such as *increasing* risk aversion possibly leading to *lower* contributions as a result of the concomitant decrease in the decision value of the received signal.[4]

If a multiproduct monopolist firm is faced with a heterogeneous base of consumer-agents, then it is generally profit maximizing to offer a *differentiated* product portfolio, if the cost for creating additional products is sufficiently low. In particular for information products it is relatively easy to create additional versions of products once a flagship version with a complete set of features has been established. Delaying the differentiation until after uncertain demand has been observed effectively amounts for the firm to procuring a signal about demand, the value of which depends on its default (i.e., no-information) versioning policy. In Chapter 5, I present a model of multiattribute product differentiation for two products, which can be horizontally and/or vertically differentiated to screen a consumer base of heterogeneous wealth and horizontal brand preference. The option value of delaying information is composed of the value of improved screening (i.e., product positioning and pricing) abilities as well as the value of flexibility.

Limitations. The treatment of the endogenous value of information in this dissertation is limited as it only pertains to certain situations. The general conclusions obtained from the research undertaken so far suggests that by endogenously varying the information structure in agency situations it is sometimes possible to significantly shift welfare between agents, especially in information-rich environments. The finding that actively damaging information may help an informed principal to manipulate agents that depend on this information to take actions better suited for the principal suggests there can be *upper bounds* for the value of investment in information systems that are not generated by the typical cost-efficiency tradeoff, but are due to the negative marginal utility of information past a certain threshold that is an endogenous result of the agency relationships of the problem. The results suggested by my dissertation should be seen as a starting point of a longer research agenda in this direction.

6.2 Future Research

This dissertation has shown, using a number of examples, that the realizable value of transferred information, indeed is very sensitive to the precise conditions under which this transfer takes place. There are a number of promising directions for future research in the area of

[4] Agents with higher risk aversion sometimes tend to act less aggressively on a received signal and therefore their WTP for the information may be lower.

endogenous information value. We found that the information value depends on its direct use and its "co-use" in the sense that the information seller may derive side benefits (such as favorable agent behavior) from the information, which go beyond its direct value with respect to the agent's optimal decision. In that sense, the value of information may be nonmonotonic in its informativeness. More generally, the principal is often "endowed" with a number of "complementary" assets (such as information) that he can use to pay "in kind" rather than using direct (monetary) wages. The question of how a principal should design general compensation schemes, using complementary assets in contracting with his agent seems a fruitful area for further research. Complementary assets, by potentially influencing agent behavior favorably, may alleviate the inefficiencies resulting from noncontractible actions.

In the arena of information sharing, the absence of strategic interactions between the agents was limiting the discussion in this dissertation, and further research could be directed towards integrating the strategic use of information and the incentives for sharing it. As mentioned in Chapter 4, there exists an extensive literature on information sharing in a standard oligopoly context, but there is very little work pertaining to situations in which agents modify their own internal capabilities as a result of information sharing, such as for instance in competitive benchmarking.

The value of marginal information in a situation of asymmetric information about agent types ("screening") is another area for future research. An interesting question is, How should a firm allocate its resources towards acquiring information about agents (e.g., consumers) and how should he design the corresponding experiments to maximize the value of the incremental information (e.g., for the purposes of product differentiation)?

6.3 References

Arrow, K.J. 1974. *Limits of Organization*. Norton, New York, NY.

Croson, D.C., Weber, T.A. 2002. The mushroom treatment: information suppression and misrepresentation in organizations. *Mimeo*, The Wharton School, University of Pennsylvania.

Kihlstrom, R.E. 1974. A general theory of demand for information about product quality. *Journal of Economic Theory* **8**(4) 413–439.

Laffont, J.-J., Martimort, D. 2002. *The Theory of Incentives*. Princeton University Press, Princeton, NJ.

LaValle, I.H. 1968. On cash equivalents and information evaluation in decisions under uncertainty. *Journal of the American Statistical Association* **63**(321) 252–290.

A room without books
is like a body without a soul.
— Cicero

Bibliography

[1] Admati, A., Pfleiderer, P. 1986. A monopolistic market for information. *Journal of Economic Theory* **39**(2) 400–438.

[2] Admati, A., Pfleiderer, P. 1987. Viable allocations of information in financial markets. *Journal of Economic Theory* **43**(1) 76–115.

[3] Admati, A., Pfleiderer, P. 1990. Direct and indirect sale of information. *Econometrica* **58**(4) 901–928.

[4] Allen, B. 1986. The demand for (differentiated) information. *Review of Economic Studies* **53**(3) 311–323.

[5] Allen, F. 1990. The market for information and the origin of financial intermediation. *Journal of Financial Intermediation* **1**(1) 3–30.

[6] Anand, K.S., Mendelson, H. 1998. Postponement and information in a supply chain. Working Paper, Kellogg Graduate School of Management, Northwestern University.

[7] Anderson, S., de Palma, A., Thisse, J.-F. 1992. *Discrete-Choice Theory of Product Differentiation*. MIT Press, Cambridge, Massachusetts.

[8] Armstrong, M. 1996. Multiproduct nonlinear pricing. *Econometrica* **64**(1) 51–75.

[9] Armstrong, M., Rochet, J.-C. 1999. Multidimensional screening: a user's guide. *European Economic Review* **43**(4-6) 959–979.

[10] Arrow, K.J. 1974. *Limits of Organization*. Norton, New York, NY.

[11] Athey, S. 2000. Investment and information value for a risk averse firm. Working paper, Massachusetts Institute of Technology.

[12] Athey, S., Levin, J. 2000. The value of information in monotone decision problems. Working paper, Massachusetts Institute of Technology.

[13] Bakos, Y., Brynjolfsson, E. 1999. Bundling information goods: pricing, profits, and efficiency. *Management Science* **45**(12) 1613–1630.

[14] Barua, A., Kriebel, C.H., Mukhopadhyay, T. 1991. An economic analysis of strategic information technology investments. *MIS Quaterly* **15**(3) 313–327.

[15] Beath, J., Katsoulacos, Y. 1991. *The Economic Theory of Product Differentiation*. Cambridge University Press, New York.

[16] Benaroch, M., Kauffman, R.J. 1999. A case for using real options pricing analysis to evaluate information technology investments. *Information Systems Research* **10**(1) 70–86.

[17] Berge, C. 1959. *Espaces Topologiques et Fonctions Multivoques*. Dunod, Paris.

[18] Berle, A., Means, G. 1932. *The Modern Corporation and Private Property*, Commerce Clearing House, Reprinted in 1993 by Gryphon Editions, New York, NY.

[19] Bhargava, H.K., Choudhary, V. 2001. Information goods and vertical differentiation. *Journal of Management Information Systems* **18**(2) 89–106.

[20] Blackwell, D. 1953. Equivalent comparison of experiments. *Annals of Mathematical Statistics* **24**(2) 265–272.

[21] Bohnenblust, H., Shapley, L., Sherman, S. 1949. Reconnaissance in game theory. *Research memorandum RM-208*, RAND Corporation.

[22] Bronshtein, I.N., Semendyayev, K.A. 1997. *Handbook of Mathematics.* Springer, New York.

[23] Brynjolfsson, E., Hitt, L.M. 1996. Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Science* **42**(4) 541–558.

[24] Brynjolfsson, E., Hitt, L.M. 2000. Beyond computation: information technology, organizational transformation and business performance. *Journal of Economic Perspectives* **14**(4) 23–48.

[25] Chamberlin, E. 1933. *The Theory of Monopolistic Competition.* Harvard University Press, Cambridge, Massachusetts.

[26] Chambers, R.G., Quiggin, J. 2000. *Uncertainty, Production, Choice, and Agency*, Cambridge University Press, Cambridge, UK.

[27] Champsaur, P., Rochet, J.-C. 1989. Multiproduct duopolists. *Econometrica* **57**(3) 533–557.

[28] Che, Y., Gale, I. 2000. The optimal mechanism for selling to a budget-constrained buyer. *Journal of Economic Theory* **92**(2) 198–233.

[29] Chwelos, P., Benbasat, I., Dexter, A.S. 2001. Research report: empirical test of an EDI adoption model. *Information Systems Research* **12**(3) 304–321.

[30] Clarke, R.N. 1983. Duopolists do not wish to share information. *Economics Letters* **11**(2) 33–36.

[31] Clarke, R.N. 1983b. Collusion and the incentives for information sharing. *Bell Journal of Economics* **14**(2) 383–394.

[32] Clemons, E.K. 1991. Evaluation of strategic investments in information technology. *Communications of the ACM* **34**(1) 22–36.

[33] Clemons, E.K., Kleindorfer, P.R. 1991. An economic analysis of interorganizational information technology, *Decision Support Systems* **8**(5) 431–446.

[34] Clemons, E.K., Reddi, S.P., Row, M.C. 1993. The impact of information technology on the organization of economic activity: the "move to the middle" hypothesis. *Journal of Management Information Systems* **10**(2) 9–35.

[35] Conrad, J.M. 1980. Quasi-option value and the expected value of information. *Quaterly Journal of Economics* **94**(4) 813–820.

[36] Cowen, T., Glazer, A. 1996. More monitoring can induce less effort. *Journal of Economic Behavior and Organization* **30**(1) 113–123.

[37] Cremer, H., Thisse, J.-F. 1991. Location models of horizontal differentiation: a special case of vertical differentiation models. *Journal of Industrial Economics* **39**(4) 383–390.

[38] Croson, D.C., Weber, T.A. 2002. The mushroom treatment: information suppression and misrepresentation in organizations. *Mimeo*, The Wharton School, University of Pennsylvania.

[39] DeGroot, M. 1962. Uncertainty, information, and sequential experiments, *Annals of Mathematical Statistics* **33**(2) 404–419.

[40] DeGroot, M.H. 1970. *Optimal Statistical Decisions*, McGraw-Hill, New York, NY.

[41] Deneckere, R.J., McAfee, R.P. 1996. Damaged goods. *Journal of Economics and Management Strategy* **5**(2) 149–174.

[42] Dewan, S., Min, C.-K. 1997. The substitution of IT for other factors of production: a firm-level analysis. *Management Science* **43**(12) 1660–1675.

[43] Dewan, S., Michael, S.C., Min, C.-K. 1998. Firm characteristics and investments in information technology: scale and scope effects. *Information Systems Research* **9**(3) 219–232.

[44] Dixit, A., Stiglitz, J. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* **67**(3) 297–308.

[45] Dixit, A.K., Pindyck, R.D. 1994. *Investment Under Uncertainty*, Princeton University Press, Princeton, NJ.

[46] Doherty, N., Mahul, O. 2002. Mickey Mouse, moral hazard and insurance securitization: decomposing informative triggers. Working Paper, Wharton School, University of Pennsylvania.

[47] Dubey, P., Haimanko, O. 2000. Optimal scrutiny in multi-period promotion tournaments. *Cowles Foundation Discussion Paper*, Cowles Foundation, Yale University.

[48] Fama, E., Miller, M. 1972. *The Theory of Finance*. Dryden Press, Hinsdale, IL.

[49] Farrell, J., Saloner, G. 1986. Installed base and compatibility: innovation, product preannouncements, and predation. *American Economic Review* **76**(5) 940–955.

[50] Fisher, M., Ramdas, K., Ulrich, K.T. 1999. Component sharing in the management of product variety: a study of automotive braking systems. *Management Science* **45**(3) 297–315.

[51] Franklin, B. 1795. *The Way to Wealth, or, Poor Richard Improved.* Printed for Ant. Aug. Renouard, Paris.

[52] Gal-Or, E. 1985. Information sharing in oligopoly. *Econometrica* **53**(2) 329–344.

[53] Gal-Or, E. 1986. Information transmission: Cournot and Bertrand equilibria. *Econometrica* **53**(1) 85–92.

[54] Gal-Or, E. 1987. 1st mover disadvantages with private information. *Review of Economic Studies* **54**(2) 279–292.

[55] Gal-Or, E. 1988. The advantages of imprecise information. *Rand Journal of Economics* **19**(2) 266–275.

[56] Galbraith, J.R. 1974. Organization design: an information processing view. *Interfaces* **4**(3) 28–36.

[57] Gallager, R.G. 1968. *Information Theory and Reliable Communication*, Wiley, New York, NY.

[58] Gebauer, J., Buxmann, P. 2000. Assessing the value of interorganizational systems to support business transactions. *International Journal of Electronic Commerce* **4**(4) 61–82.

[59] Gjesdal, F. 1982. Information and incentives: the agency information problem. *Review of Economic Studies* **49**(3) 373–390.

[60] Gorman, I.E. 1985. Conditions for economies of scope in the presence of fixed costs. *Rand Journal of Economics* **16**(3) 431–436.

[61] Gorman, W.M. 1959. Separable utility and aggregation. *Econometrica* **27**(3) 469–481.

[62] Gorman, W.M. 1968. Conditions for additive separability. *Econometrica* **36**(3-4) 605–609.

[63] Grossman, S., Hart, O. 1983. An analysis of the principal-agent problem. *Econometrica* **51**(1) 7–45.

[64] Hanemann, W.M. 1991. Willingness to pay and willingness to accept: how much can they differ? *American Economic Review* **81**(3) 635–647.

[65] Harsanyi, J.C. 1966. A general theory of rational behavior in game situations. *Econometrica* **34**(3) 613–634.

[66] Hart, O.D., Holmstrom, B. 1987. The theory of contracts, in T. Bewley, ed., *Advances in Economic Theory*, Fifth World Congress, Cambridge University Press, New York, NY.

[67] Hicks, J. 1939. Value and Capital. Clarendon Press, Oxford, UK.

[68] Hirshleifer, J. 1971. The private and social value of information and the reward to inventive activity. *American Economic Review* **61**(4) 561–574.

[69] Hirshleifer, J., Riley, J.G. 1992. *The Analytics of Uncertainty and Information*, Cambridge University Press, Cambridge, UK.

[70] Hitt, L.M. 1999. Information technology and firm boundaries. *Information Systems Research* **10**(2) 134–149.

[71] Hitt, L.M., Brynjolfsson, E. 1996. Productivity, business profitability and consumer surplus: three different measures of information technology value. *MIS Quaterly* **20**(2) 121–142.

[72] Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* **10**(1) 74–91.

[73] Hotelling, H. 1929. Stability in competition. *Economic Journal* **39**(153) 41–57.

[74] Huchzermeier, A., Loch, C.H. 2001. Project management under risk: using the real options approach to evaluate flexibility in R&D. *Management Science* **47**(1) 85–101.

[75] Hviid, M. 1989. Risk-averse duopolists and voluntary information transmission. *Journal of Industrial Economics* **38**(1) 49–64.

[76] Ingersoll, J.E. 1987. Theory of Financial Decision Making, Savage, ML: Rowman and Littlefield.

[77] Innes, R. 1990. Limited liability and incentive contracting with ex ante action choices. *Journal of Economic Theory* **52**(1) 45–67.

[78] Jacobides, M., Croson, D.C. 2001. Information policy: shaping the value of agency relationships. *Academy of Management Review* **26**(2) 202–223.

[79] Jensen, M., Meckling, W. 1976. Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* **3**(4) 305–360.

[80] Jensen, M., Meckling, W. 1992. Specific and general knowledge, and organizational structure, in L.Werin and H. Wijkander (eds.), *Contract Economics*. Blackwell, Oxford, UK, 251–274.

[81] Jones, R., Mendelson, H. 1998. Product and price competition for information goods. Working Paper, Graduate School of Business, Stanford University.

[82] Judd, K.L. 1985. Credible spatial preemption. *Rand Journal of Economics* **16**(2) 153–166.

[83] Kahneman, D., Knetsch, J.L., Thaler, R.H. 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives* **5**(1) 193–206.

[84] Kalai, E., Smorodinsky, M. 1975. Other solutions to Nash's bargaining problem. *Econometrica* **43** 513–518.

[85] Kambil, A., Henderson, J., Mohsenzadeh, H. 1993. Strategic management of information technology investments: an options perspective, in R.D. Banker, R.J. Kauffmann, and M.A. Mehmod (eds.) *Strategic Information Technology Management: Perspectives on Organizational Growth and Competitive Advantage*. Idea Publishing, Harrisburg, PA.

[86] Kihlstrom, R.E. 1974. A general theory of demand for information about product quality. *Journal of Economic Theory* **8**(4) 413–439.

[87] Kihlstrom, R.E., Mirman, L.J. 1974. Risk aversion with many commodities. *Journal of Economic Theory* **8**(2) 361–388.

[88] Kihlstrom, R.E. 1976. Firm demand for information about price and technology. *Journal of Political Economy* **84**(6) 1335–1341.

[89] Kihlstrom, R.E., Roth, A.E., Schmeidler, D. 1981. Risk Aversion and Solutions to Nash's Bargaining Problem, in O. Moeschlein and D. Pallaschke (eds.) *Game Theory and Mathematical Economics*. North-Holland, Amsterdam, NL.

[90] Kihlstrom, R.E. 1984. A Bayesian exposition of Blackwell's theorem on the comparison of experiments, in M. Boyer and R.E. Kihlstrom (eds.) *Bayesian Models in Economic Theory*. North-Holland, Amsterdam, NL, 13–31.

[91] Kihlstrom, R.E. 1988. Optimal contracts for security analysts and portfolio managers. *Journal of Banking and Finance* **5** 291–325.

[92] Kihlstrom, R.E., Vives, X. 1989. Collusion by asymmetrically informed duopolists. *European Journal of Political Economy* **5**(2-3) 372–402.

[93] Kirby, A.J. 1988. Trade associations as information exchange mechanisms. *Rand Journal of Economics* **19**(1) 138–146.

[94] Krishnan, V., Ulrich, K.T. 2001. Product development decisions: a review of the literature. *Management Science* **47**(1) 1–21.

[95] Laffont, J.-J., Maskin, E., Rochet, J.-C. 1987. Optimal nonlinear pricing with two-dimensional characteristics. T. Groves, R. Radner, S. Steiger (eds.) *Information, Incentives, and Economic Mechanisms: Essays in Honor of Leonid Hurwicz*. University of Minnesota Press, Minneapolis, MN, 255–266.

[96] Laffont, J.-J. 1989. *The Economics of Uncertainty and Information*. MIT Press, Cambridge, MA.

[97] Laffont, J.-J., Martimort, D. 2002. *The Theory of Incentives*. Princeton University Press, Princeton, NJ.

[98] Lancaster, K.J. 1966. A new approach to consumer theory. *Journal of Political Economy* **74**(2) 132–157.

[99] Launhardt, W. 1885. *Mathematische Begründung der Volkswirtschaftslehre*. B.G. Teubner, Leipzig. Reprinted by Scientia, Aalen, 1963.

[100] LaValle, I.H. 1968. On cash equivalents and information evaluation in decisions under uncertainty. *Journal of the American Statistical Association* **63**(321) 252–290.

[101] Lawrence, D. 1999. *The Economic Value of Information*. Springer, Springer, NY.

[102] Lee, H.L., Tang, C.S. 1997. Modeling the costs and benefits of delayed product differentiation. *Management Science* **43**(1) 40–53.

[103] Lewis, T., Sappington, D. 2000. Motivating wealth-constrained actors. *American Economic Review* **90**(4) 944–960.

[104] Liebowitz, S.J., Margolis, S.E. 1994. Network externalities: an uncommon tragedy. *Journal of Economic Perspectives* **8**(2) 133–150.

[105] Marschak, J., Miyasawa, K. 1968. Economic comparability of information systems. *International Economic Review* **9**(2) 137–173.

[106] Mas-Colell, A., Whinston, M.D., Green, J.R. 1995. Microeconomic Theory. Oxford University Press, Oxford, UK.

[107] Matthews, S., Moore, J. 1987. Monopoly provision of quality and warranties: an exploration in the theory of multidimensional screening. *Econometrica* **55**(2) 441–467.

[108] McDonald, R., Siegel, D. 1986. The value of waiting to invest. *Quarterly Journal of Economics* **101**(4) 707–728.

[109] Merton, R. 1982. On the microeconomic theory of investment under uncertainty, in K. Arrow and M.D. Intriligator (eds.) *Handbook of Mathematical Economics*. North-Holland, New York, NY.

[110] Millon, M.H., Thakor, A.V. 1985. Moral hazard and information sharing: a model of financial information gathering agencies. *Journal of Finance* **40**(5) 1403–1421.

[111] Mirrlees, J.A. 1971. An exploration in the theory of optimal income taxation. *Review of Economic Studies* **38**(2) 175–208.

[112] Mirrlees, J. 1974. Notes on Welfare Economics, Information, and Uncertainty, in M. Balch, D. McFadden and S.-Y. Wu (eds.) *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam.

[113] Moorthy, K.S., Png, I. 1992. Market segmentation, cannibalization, and the timing of product introductions. *Management Science* **38**(3) 345–359.

[114] Mukhopadhyay, T., Kerke, S., Kalathur, S. 1995. Business value of information technology: a study of electronic data interchange. *MIS Quaterly* **19**(2) 137–156.

[115] Mukhopadhyay, T., Lerch, F.J., Mangal, V. 1997. Assessing the impact of information technology on labor productivity. *Decision Support Systems* **19**(2) 109–122.

[116] Mukhopadhyay, T., Rajiv, S., Srinivasan, K. 1997b. IT impact on process output and quality. *Management Science* **43**(12) 1645–1659.

[117] Mussa, M., Rosen, S. 1978. Monopoly and product quality. *Journal of Economic Theory* **18**(2) 301–317.

[118] Nash, J.F. 1953. Two-person cooperative games. *Econometrica* **21**(1) 128–140.

[119] NAULT, B.R., TYAGI, R.K. 2001. Implementable mechanisms to coordinate horizontal alliances. *Management Science* **47**(6) 787–799.

[120] NOVSHEK, W., SONNENSCHEIN, H. 1982. Fulfilled expectations Cournot duopoly with information acquisition and release. *Bell Journal of Economics* **13**(1) 214–218.

[121] NOVSHEK, W., THOMAN, L. 1998. Information disaggregation and incentives for non-collusive information sharing. *Economics Letters* **61**(3) 327–332.

[122] PANZAR, J.C., WILLIG, R.D. 1981. Economies of scope. *American Economic Review* **71**(2) 268–272.

[123] PERLOFF, J., SALOP, S. 1985. Equilibrium with product differentiation. *Review of Economic Studies* **52**(1) 107–120.

[124] PINDYCK, R.S. 1988. Irreversible investment, capacity choice, and the value of the firm. *American Economic Review* **78**(5) 969–985.

[125] QUINZII, M., ROCHET, J.-C. 1985. Multidimensional signalling. *Journal of Mathematical Economics*. **14** 261–284.

[126] RADNER, R. 1993. The organization of decentralized information processing. *Econometrica* **62**(5) 1109–1146.

[127] RAIFFA, H., SCHLAIFER, R. 1961. *Applied Statistical Decision Theory*. Division of Research, Graduate School of Business Administration, Harvard University, Boston, MA.

[128] RAIFFA, H., PRATT, J., SCHLAIFER, R. 1995. *Introduction to Statistical Decision Theory*, MIT Press, Cambridge, MA.

[129] RAITH, M. 1996. A general model of information sharing in oligopoly. *Journal of Economic Theory* **71**(1) 260–288.

[130] RANDALL, A., STOLL, J.R. 1980. Consumer's surplus in the commodity space. *American Economic Review* **70**(3) 449–455.

[131] RIGGINS, F.J., KRIEBEL, C.H., MUKHOPADHYAY, T. 1994. The growth of interorganizational systems in the presence of network externalities. *Management Science* **40**(8) 984–998.

[132] Rochet, J.-C., Choné, P. 1998. Ironing, sweeping and multidimensional screening. *Econometrica* **66**(4) 783–826.

[133] Ross, S.A. 1973. The economic theory of agency: the principal's problem. *American Economic Review* **63**(2) 134–139.

[134] Roth, A.E. 1977. Individual rationality and Nash's solution to the bargaining problem. *Mathematics of Operations Research* **2**(1) 64–65.

[135] Roth, A.E. 1979. *Axiomatic Models of Bargaining*. Springer, Berlin.

[136] Rudin, W. 1976. *Principles of Mathematical Analysis*. McGraw-Hill, New York, NY.

[137] Salanié, B. 1999. *The Economics of Contracts*. MIT Press, Cambridge, MA.

[138] Salop, S. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* **10**(1) 141–156.

[139] Sappington, D. 1983. Limited liability contracts between principal and agent. *Journal of Economic Theory* **29**(1) 1–21.

[140] Shaked, A., Sutton, J. 1982. Relaxing price competition through product Differentiation. *Review of Economic Studies* **49** 3–13.

[141] Shaked, A., Sutton, J. 1987. Natural oligopolies. *Econometrica* **51**(5) 1469–1484.

[142] Shaked, A., Sutton, J. 1987. Product differentiation and industrial structure. *Journal of Industrial Economics* **36**(2) 131–146.

[143] Shaked, A., Sutton, J. 1990. Multiproduct firms and market structure. *Rand Journal of Economics* **21**(1) 45–62.

[144] Shannon, C. 1948. A mathematical theory of communication. *Bell System Technical Journal* **27** 379–423, 623–656.

[145] Shapiro, C. 1986. Exchange of cost information in oligopoly. *Review of Economic Studies* **53**(3) 433–446.

[146] Shapiro, C., Varian, H.R. 1998. Versioning: the smart way to sell information. *Harvard Business Review* **76**(6) 107–114.

[147] Shin, N. 1997. The impact of information technology on coordination cost. *Proceedings of the Eighteenth International Conference on Information Systems*, Atlanta, GA, 400–438.

[148] Stahl, S. 2001. Sharing data is key to antiterrorism efforts. *Informationweek*, 4th October, **858** 8.

[149] Stole, L. 1995. Nonlinear pricing and oligopoly. *Journal of Economics and Management Strategy* **4**(4) 529–562.

[150] Swaminathan, J.M., Tayur, S.R. 1998. Managing broader product lines through delayed differentiation using vanilla boxes. *Management Science* **44**(12-2) S161–S172.

[151] Thon, D., Thorlund-Petersen, L. 1993. The value of perfect information in a simple investment problem. *Information Economics and Policy* **5** 51–71.

[152] Tirole, J. 1988. *The Theory of Industrial Organization*. MIT Press, Cambridge, Massachusetts.

[153] Treich, N. 1997. Risk tolerance and value of information in the standard portfolio model. *Economics Letters* **55**(3) 361–363.

[154] Trigeorgis, L. 1996. *Real Options*. MIT Press, Cambridge, MA.

[155] Ulrich, K.T. 1995. The role of product architecture in the manufacturing firm. *Research Policy* **24**(3) 419–440.

[156] Van Mieghem, J.A. 1998. Investment strategies for flexible resources. *Management Science*. **44**(8) 1071–1078.

[157] VanZandt, T. 1997. Decentralized information processing in the theory of organizations, in M. Sertel, ed., *Contemporary Economic Development Reviewed, Volume 4*, Macmillan, London, UK.

[158] Vives, X. 1984. Duopoly information equilibrium, Cournot and Betrand. *Journal of Economic Theory* **34**(1) 71–94.

[159] Vives, X. 1990. Trade association disclosure rules, incentives to share information, and welfare. *Rand Journal of Economics* **21**(3) 409–430.

[160] Wang, E.T.G., Seidmann, A. 1995. Electronic data interchange: competitive externalities and strategic implementation policies. *Management Science* **41**(3) 401–418.

[161] Weber, T.A., Croson, D.C. 2002. Selling less information for more: garbling with benefits. Working paper, Wharton School, University of Pennsylvania.

[162] Weber, T.A. 2003. An exact relation between willingness to pay and willingness to accept. *Economics Letters*, to appear.

[163] Weisbrod, B. 1964. Collective-consumption services of individual consumption goods. *Quaterly Journal of Economics* **78**(3) 471–477.

[164] Whitehead, A.N. 1941. Immortality, in P.A. Schilpp, ed., *The Philosophy of Alfred North Whitehead*. Banta Publishing, Menaha, WI, 682–700.

[165] Willig, R.D. 1976. Consumer's surplus without apology. *American Economic Review* **66**(4) 589–597.

[166] Wilson, R. 1985. Multi-dimensional signalling. *Economics Letters* **19** 17–21.

[167] Wilson, R. 1993. *Nonlinear Pricing*. Oxford University Press, Oxford, UK.