Year: 2005

Four essays in economics

Schneider, Yves

Other titles: 1. How much internalization of nuclear risk through liability insurance? 2. Spatial effects in willingness-to-pay: the case of nuclear risks 3. Switching costs, firm size, and market structure 4. Can conglomerates ease credit constraints?

Four Essays in Economics

Dissertation
der Wirtschaftswissenschaftlichen Fakultät
der Universität Zürich

zur Erlangung der Würde
eines Doktors der Ökonomie

vorgelegt von
Yves Schneider Bonassi
von Seftigen BE

genehmigt auf Antrag von

Prof Dr. P. Zweifel
Prof Dr. A. Schmutzler

Die Wirtschaftswissenschaftliche Fakultät der Universität Zürich gestattet hierdurch die Drucklegung der vorliegenden Dissertation, ohne damit zu den darin ausgesprochenen Anschauungen Stellung zu nehmen.

Zürich, den 9. November 2005

Der Dekan: Prof. Dr. H. P. Wehrli

Vorwort

Meine Zeit am Sozialökonomischen Institut (SOI) der Universität Zürich begann mit der Betreuung eines Projektes im Auftrag des Bundesamtes für Energie. Prof. Dr. Peter Zweifel ermöglichte es mir, direkt nach Abschluss der Lizenziatsarbeit ein interessantes, empirisches Projekt zu betreuen. Ich danke ihm herzlichst für das mir entgegengebrachte Vertrauen und die lehrreiche gemeinsame Arbeit an diesem Projekt. Die ersten beiden Aufsätze dieser Dissertation sind das Resultat dieser Arbeit. Natürlich durfte ich mich auch nach Abschluss dieses Projektes jederzeit über die Unterstützung durch Prof. Dr. Peter Zweifel erfreuen. Ich danke auch Prof. Dr. Armin Schmutzler für seine Gesprächsbereitschaft und hilfreichen Anregungen zu meiner Arbeit. Nicht weniger hilfreich waren die Denkanstösse und Kommentare von Dr. Michael Breuer zu diversen Arbeiten von mir.

Die Idee zum dritten Papier entstand während einem Besuch in New York. Gemeinsam mit Dr. Simon Lörtscher dachte ich über die Rolle von Supermärkten nach: Wie viel komplizierter ist es doch, seinen geliebten Coop-Warenkorb im Whole Foods Supermarkt zusammenzustellen. Es wäre doch viel praktischer, wenn gleich um die Ecke 24th Street / 7th Avenue eine Coop-Filiale stünde. Das Resultat dieser Überlegungen ist Aufsatz Nummer drei dieser Dissertation. Die Arbeit mit Simon an diesem Aufsatz sowie die Diskussionen über andere Forschungsprojekte und Ideen waren immer sehr bereichernd. Ich freue mich, noch lange gemeinsam mit Simon Ökonomie zu betreiben. Hoffentlich war es nicht das letzte gemeinsame Papier.

Während den ersten Jahren am SOI profitierte ich viel von den z.T. ausgedehnten (Streit)gesprächen mit “der alten Garde” Dr. Hansjörg Lehmann, Dr. Markus König, Dr. Lukas Steinmann und Dr. Harry Telser. Es freut mich sehr, dass der Kontakt zu Ihnen bis heute anhält. Dass auch meine zweite Halbzeit am SOI nicht allzu trostlos verlief, verdanke ich zu einem grossen Teil Karolin Becker und Patrick Eugster. Die spannenden Diskussionen über Ökonomie und vieles

andere bleiben mir in bester Erinnerung. Respekt gebührt auch meinen Bürokollegen Chantal Grandchamp, Boris Krey und Christian Wyss, die meine Anwesenheit meist ohne Klage ertragen haben.

Eine wesentliche Erleichterung für meine Recherchen war die Hilfsbereitschaft der SOI-Bibliothek. Die Mitarbeiterinnen der SOI-Bibliothek haben mich jederzeit unbürokratisch und effizient unterstützt. Zeitschriftenartikel suchen und Bücher ausleihen war nie einfacher!

Natürlich sass ich in den letzten Jahren nicht nur im Institut am Schreibtisch. Dank gebührt deshalb auch der Andorra Bar in Zürich für die fortwährend herausragende Gastfreundschaft und die vielen gratis Pistazien. Spezieller Dank geht auch an Daniel Gasser, Stephan Illi und Christoph Wälti. Die drei haben immer wieder Zeit gefunden, mit mir das schöne Leben zu feiern. Besonders erholsam waren und sind die gemeinsamen Go-Abende mit Stephan.

Da man sich den grössten Dank für den Schluss des Vorworts vorbehält, verbleibt es mir meiner engsten Familie zu danken. Während der gesamten Zeit konnte ich auf die Unterstützung durch meine Frau Tamara Bonassi zählen. Sie hat mich oft ermuntert weiterzumachen. Schliesslich Danke ich meinen Eltern Fritz und Kathrin Schneider dafür, dass sie mir diese Ausbildung ermöglicht haben und jederzeit für mich da waren.

Zürich, im November 2005

Table of Contents

Chapter 1

Introduction

The four essays offered in this book all cover different topics in economics. The first two essay are empirical investigations on consumers' willingness-to-pay for increased financial security with regard to nuclear power plant accidents. The third essay addresses an industrial organization topic. It presents a model explaining the advantage of chain stores compared to single outlet stores when consumers face costs of switching their suppliers. The forth essay, finally, belongs to the realm of finance. It develops a theoretical model explaining the circumstances that mitigate credit constraints if projects are financed as divisions of a conglomerate rather than as independent firms.

The first essay is the result of a study financed by the Swiss Federal Energy Agency. The agency was interested in the optimal level of internalization of external costs caused by nuclear accidents. Certainly, a major nuclear accident will have financial consequences which are not covered by nuclear energy suppliers and thus will have to be borne by the public. Since the Swiss government already mandates operators of nuclear power plants to purchase a liability insurance covering roughly 1 bn. CHF, it seemed obvious to ask how much more coverage by liability insurance should be mandated in order to optimally internalize these costs. This requires estimation of marginal costs and marginal benefits of an increase in mandatory liability insurance. The essay presented investigates the benefit side of this question and aims at estimating consumers' willingness-to-pay (WTP) for additional mandatory insurance coverage.

Estimated WTP for increased financial security concerning nuclear accidents can be complemented with a spatial dimension. How does marginal willingness-to-pay change with respect to distance of domiciles from the power plants? This aspect is addressed in chapter 3. Arguably, there are two effects working in opposite directions; viz. a risk effect and a selection effect. According to the risk effect, people located near a nuclear power plant are expected to show higher WTP values than people located farther away. According to the selection effect, people living closer to nuclear power plants are expected to be more favorable to nuclear energy in general and less averse to their risk in particular. Therefore, their WTP should be smaller than that of people located at greater distance. Indeed, the empirical evidence is inconclusive with regard to the overall relationship between WTP and distance from the nearest nuclear power plant. However, when the selection effect is statistically controlled for, WTP values significantly decrease with distance.

Although these two essays seem almost unrelated to the latter two essays, there is nonetheless a connection. In order to analyze optimal internalization of external costs by governments, one must – in principle – explain why individuals and firms cannot contract with each other and deal with the externality themselves. What superior ability should a government have in solving the problem of externalities? This is a controversial and long-standing question in economics and ultimately relates to differences in transaction costs resulting from different governance regimes. As firms might improve upon arm's-length markets, governments might improve upon firms. It is not the aim of this short introduction to elaborate on such an extensive topic, but rather to point out that the question of optimal internalization of externalities also relates to the question of optimal governance structures. One reason for the existence of firms lies in the fact that they are a means for internalizing externalities in arm's-length markets (see for instance Holmstrom, 1999). A government which is able to force its citizens to obey its laws must have some superior ability to internalize certain externalities compared to firms, which do not have this authority.

The first two essays do not address these theoretical questions of optimal governance but deal with a specific instrument for internalizing a particular externality, viz. the risk of nuclear power plant accidents. The latter two essays neither address optimal governance in such a general sense, but are both concerned with rather specific questions relating to the boundaries of firms.

The third essay, entitled "Switching Costs, Firm Size, and Market Structure", takes a technological perspective in discussing the advantage of "large" over "small" firms. Whenever consumers shop at a new supplier for the first time, they incur some fixed costs for getting acquainted with the particular characteristics and products of this supplier. This non-convex consumption technology generates an advantage for multi-outlet (chain) stores. If consumers are mobile and therefore change locations, a chain store in fact insures them against the risk of incurring these fixed costs again at the new location. If instead they patronize local stores, they again must incur fixed costs of getting accustomed to the local store at the new location. It is shown that if consumers are heterogenous with respect to these fixed costs of consumption, chain stores and local stores coexist in equilibrium and that local stores charge lower prices than chain stores.

The forth essay addresses the question of optimal firm size in a contract theoretical framework. Firms are modeled as being credit constrained, i.e. not all profitable ventures can attract the outside investments necessary to undertake them. The essay seeks to answer the question whether these credit constraints are mitigated if two firms finance their projects jointly. The model presented motivates credit constraints by contractual incompleteness and moral hazard. The quality of a firm or project is assumed to be non-verifiable and thus not contractible. In addition, asymmetric information between investors and entrepreneurs is assumed. As pointed out in the literature (see, e.g., Shin and Stulz, 1998), asymmetric information is a key ingredient in explaining the tripartite structure of financing envisaged in this model, viz. outside investor contracting with headquarters and headquarters contracting with project managers. Otherwise it is difficult to imagine why actors in external capital markets (investors and headquarters) are not capable of using the same contracts as actors in internal capital markets (headquarters and project managers). However, actors in internal capital markets enjoy an informational advantage because they are bound by an informal employment contract that serves to limit moral hazard. Therefore, they can simultaneously use this informal employment contract and a formal financial contract with the investor to finance their investments jointly. It is shown that investors prefer to finance some types of projects as conglomerate in order to take advantage the superior information. On the other hand, they prefer to finance other types of projects as independent firms.

Note that Peter Zweifel co-authored chapter 2 and 3, Christian Wyss co-authored chapter 3 and Simon Lörtscher co-authored chapter 4. While the undersigned author was at least equally responsible for the intellectual input to chapters 2 and 4 as his co-authors, the main contribution to chapter 3 is by Peter Zweifel. Chapter 2 appeared in the Journal of Risk and Uncertainty, chapter 3 is under review with the American Economic Review and the chapter 4 will be submitted to an industrial economics journal soon. The last chapter has working paper status and requires additional effort to make it suitable for submission.

Yves Schneider
Zürich, September 2005

Bibliography

Holmstrom, B. (1999). The firm as a subeconomy. *Journal of Law, Economics and Organization 15*, 74–102.

Shin, H.-H. and R. M. Stulz (1998). Are internal capital markets efficient? *The Quarterly Journal of Economics 113*, 531–552.

How Much Internalization of Nuclear Risk Through Liability Insurance?

Yves Schneider and Peter Zweifel [*] [†]

[*]This article has been previously published in the Journal of Risk and Uncertainty, volume 29, number 3, p.219-240.

[†]We would like to acknowledge helpful comments and criticism by participants of the Economic Research Seminar of the university of Basel, participants of the 2002 meeting of the Verein für Socialpolitik, as well as two anonymous referees. Financial support came from the Swiss Federal Office of Energy (Berne) and the Swiss Institute of Technology.

Chapter 2

How Much Internalization of Nuclear Risk Through Liability Insurance?

2.1 Introduction

Nuclear power plants provoke conflicts in many countries. While many voters and politicians are committed on this issue, others will gauge the advantages and disadvantages of the nuclear option. On the downside, an important consideration is that with a very small probability, an accident causing billions of Dollars of damage may occur.

This paper seeks to estimate the willingness-to-pay (WTP) of Swiss citizens for relief from the financial consequences of a severe nuclear accident, to be provided by an extension of liability insurance coverage mandated to nuclear power operators. This statement of objective calls for two clarifications. First, the risk to be considered needs to be defined. In the production process of nuclear energy, at least ten stages can be identified, each with its proper risks (Hirschberg et al., 1998). This paper deals exclusively with the risks of nuclear energy produced in Switzerland. Second, it focuses on mandatory liability insurance as an instrument for risk internalization. The many norms that govern the production of nuclear power, monitored by the Swiss Nuclear Safety Inspectorate (HSK), are simply taken as given.

This regulation in combination with the legal norm of liability still leaves room for what Shavell (1986) calls the judgement proof problem. The judgement proof problem consists of the possibility that nuclear power plant operators may fail to pay compensation for the damage caused, due to lack of assets. As shown by Shavell, mandatory liability insurance serves to avoid this shortcoming. This finding justifies considering mandatory liability insurance as an instrument of nuclear risk internalization.

In Switzerland, the current coverage of CHF 0.7 billion (bn.) (approx. US$ 0.47 bn. at 2002 exchange rates) written by private insurers will hardly be sufficient to compensate the victims of an accident. However, an extension of coverage will result in a higher outlay for nuclear power. No market for individually contracted supplementary coverage has developed, quite likely because there are no contractual relationships between nuclear insurers and consumers to build upon. Therefore, determining the importance accorded to more comprehensive relief from the financial consequences of a severe nuclear accident amounts to estimating the increase in the price of electricity that would be accepted by consumers in return for extended liability coverage to be bought by plant operators.

'Stated choice' rather than conventional 'contingent valuation' experiments were conducted to measure WTP. In 'stated choice', different attributes of electric power are distinguished and varied from one scenario to the next. This should avoid fixation on the price attribute which in 'contingent valuation' may cause respondents to forget about alternate uses of their income, thus resulting in excessive estimates of WTP. Moreover, respondents are not asked to provide rankings but only to choose between the status quo and one alternative at a time, which serves to bring the experimental situation close to everyday decision making.

The basic hypothesis is that opting for the alternative must entail a utility gain that can be related to the levels of the attributes pertaining to the scenario. Therefore, the utility associated with an attribute can be inferred from the observed choice sequences using a Probit model. Since the surcharge per kwh is one of the attributes, it is possible to calculate WTP values for the other attributes (in particular, extension of insurance coverage).

Indeed, respondents valued five attributes of power in a way that is consistent with economic theory, and median estimated WTP amounts to a realistic 0.14 US cents per kwh. Since quite probably the extra cost of a fivefold extension of present liability coverage is below this figure, the benefits of such an extension exceed its cost, suggesting an improvement of efficiency.

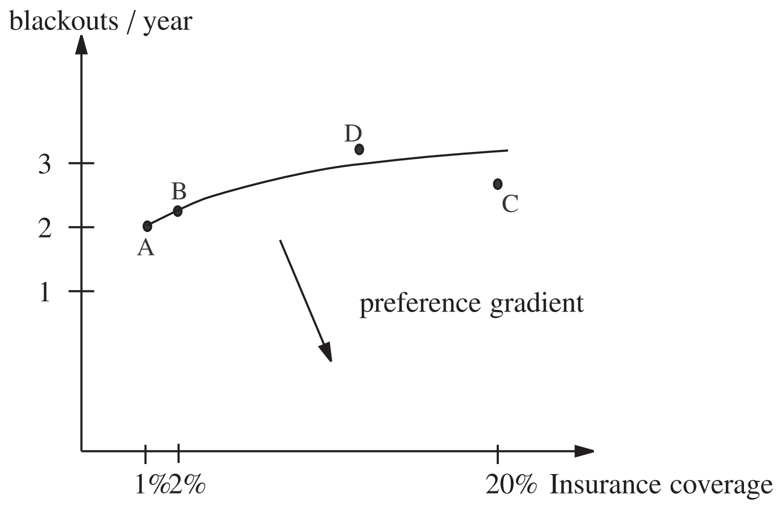


Figure 2.1: Trading off different product attributes.

2.2 Determining Willingness to Pay Through 'Stated Choice'

For goods and services traded on markets, there is no need to measure willingness to pay (WTP). By accepting a price the customer reveals that his WTP is at least as high as the price. The safety of nuclear power plants is not yet traded on markets (though it is possible to think of nuke-bonds which mature in case of a prespecified accident). Neither does individually contracted insurance against nuclear risks exist which would reveal potential victims' willingness to pay for nuclear safety.

2.2.1 'Stated Choice' Methodology

In the absence of opportunities to observe revealed preferences, it is necessary to measure stated preferences through experiments. Earlier attempts to measure WTP for nonmarket goods used the contingent valuation variant of stated preference (see e.g. Mitchell (1989) and Hausman (1993)). This methodology can also be applied to risk reduction, see e.g. Krupnick et al. (2002).

In the present context, respondents would have to state the maximum amount per kwh they would be prepared to pay for the increased financial security achieved by an extension of mandatory liability insurance. The difficulty with this direct approach is that in real life people hardly ever ask themselves such questions. Rather, they compare the attributes of a good and its price and then decide to buy it or not.

The 'stated choice' approach, developed by Louviere and Hensher (1982), seeks to simulate this everyday decision making. Rather than directly asking for willingness to pay, it seeks to elicit it indirectly through 'accept'/'reject' choices. Its basic assumption is that individuals derive utility from the attributes of products and are willing to trade them off against each other (Lancaster, 1966). The first application of stated choice to risk reduction seems to be Telser and Zweifel (2002), which extends the paired comparison approach used in Magat et al. (1988) to a framework of multiple product attributes and ordinal utility.

In the present context, the 'stated choice' method allows individuals to choose among different types of electricity. During the decision process, the attributes (among them price) of electricity are traded off against each other. By observing several similar decisions it is possible to estimate how much income (through higher electricity prices) respondents are ready to give up in return for an increased amount of some other desired attribute.

One particular tradeoff is illustrated in figure 2.1. Utility increases with insurance coverage in the event of an accident (as a percentage of maximum possible loss) and decreases with the average number of power blackouts; both attributes were found to be relevant to the persons interviewed (see section 2.3). Assume that the status quo is given by combination A (1% insurance coverage, 2 blackouts/year). Point B indicates that the individual is willing to accept a slightly higher number of blackouts if insurance coverage is raised to 2%. The corresponding sacrifice in terms of security of supply is the marginal WTP for an increased insurance coverage.

Now participants in the experiment are asked to evaluate additional combinations, for example point C. If C is accepted, then the individual's indifference curve must lie above C. Next, it must lie below point D if D is rejected. Proceeding in this manner, it is possible to approximate the indifference curve. Finally, the marginal willingness to pay measured in money terms (MWP) can be estimated in the same way, by introducing the increase in the price of electricity as an additional product attribute. The corresponding increase in the outlay on electricity is a sacrifice of income which would otherwise be available for spending on other goods.

Summing up, using 'stated choice' the experimenter is not limited to varying only price and insurance coverage but can introduce other attributes that influence real life decisions concerning electricity. Failing this, he runs the risk of causing respondents to associate with variations in price or insurance coverage attributes not explicitly included in the experiment (and therefore assumed as fixed).

2.2.2 Problems With Risk Assessment

The ultimate objective of this work is to provide guidance to public regulation. For this purpose it would be preferable to have experiments based on objective probabilities. However, in the present context objective probabilities are not available since experts differ in their estimates of the probability of a catastrophic nuclear accident (Zweifel and Umbricht, 2002). In this situation, respondents' own subjective estimates become crucially important. Specifically the additional ambiguity caused may affect respondents' level of expected utility as well as their tradeoffs between attributes such as the one shown in figure 2.1 (Viscusi, 1998, ch. 2).

Ganderton et al. (2000) provide a possible solution to the problem of ambiguity. They conduct laboratory experiments where subjects had to decide about insurance against a low probability - high consequence event. Several draws from a loss function known to participants were revealed to subjects, who made choices largely consistent with predictions of the expected utility theory. However, the loss function cannot be claimed to be known in the present context.

Another problem is low probabilities. Here, Kunreuther et al. (2001) found that supplying a reference point in terms of a more probable and familiar risk and providing a good deal of contextual information helped respondents in dealing with low probabilities. In the present study, experts' average estimated probability of the highest possible loss served as the reference point. However, respondents were not asked to adopt this value, but to indicate their own probability estimate relative to that of the experts. The required contextual information comes from two sources, viz. the introduction to the experiment (see appendix 2.6.1) and the remaining attributes characterizing types of power, which moreover are allowed to vary during the 'stated choice' experiment. In sum, this setup serves to reduce the complexity of the choice situations by only incorporating the loss dimension of risk, while still allowing the effect of different beliefs about the loss probabilities to be estimated.

Figure 2.2 shows the frequencies of deviations from the reference point (the scale used to elicit this information can be found in appendix 2.6.2). More pessimistic beliefs clearly outweigh optimistic deviations, in keeping with findings reported by Camerer and Kunreuther (1989).

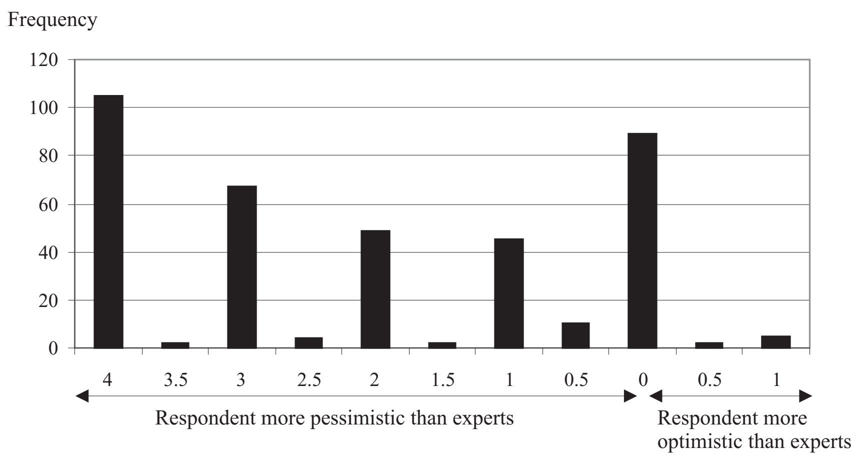


Figure 2.2: Beliefs about loss probabilities relative to experts' estimates. X-axis shows probability deviations in powers of 10 (see appendix 2.6.2).

2.3 Experimental Design

In the context of a 'stated choice' experiment several issues must be addressed. What product attributes are to be included? How many levels should be distinguished in an attribute? How should the level of attributes vary between choice scenarios? How many choice scenarios should be presented to each respondent? An extensive literature is devoted to these and other issues (see e.g. Louviere et al. (2000) and Hedayat et al. (1999)).

2.3.1 Determination of Attributes and Their Levels

First of all, the relevant attributes of electric power need to be identified. Their number must be kept low for the decision problem to remain manageable. In a separate survey, approximately 500 persons were asked in spring 2001 to rate 15 different attributes, assigning them ranks between 1 (not important) and 10 (very important). Among the most important were: secure and sustainable waste disposal (9.26), size of area exposed to hazard (8.91), reliability (low frequency of black-outs) (8.68), financial compensation of the victims in case of an accident (8.78), and average price per kwh (7.69).

Next, levels have to be assigned to attributes. Again they must be few in number in order to avoid long interviews. But then, the levels must reach sufficiently extreme values to cause respondents to switch from "accept" to "reject" and vice

Table 2.1: Levels of attributes.

Attribute	Levels (Coding[c])	Unit	Status quo
Price	0; 10; 30; 60 (**0**;...;60)	percent	0
Blackouts	2; 14 (**0**;1)	numb./year	2
Waste	unresolved problems (**1**); no unresolved problems (0)		unresolved problems
Damage[a]	0.1; 10; 100; 200 (0.1;...;**200**)	CHF bn.	200
Coverage[b]	1; 20; 50; 100 (**1**;...;100)	percent	1

[a] Values in US$ bn: 0.065; 6.5; 65; 130 (at 2002 exchange rates)
[b] Coverage in percent of loss
[c] Bold for status quo

versa.

To test the questionnaire, six persons were interviewed in a first pretest. Without exception they understood the questions and were able to process the choice scenarios without problems. However, the attribute 'insurance coverage' was regarded as relatively unimportant.

A second pretest comprising 20 persons was conducted, with the maximum price hike boosted to 60 percent in order to induce a sufficient frequency of rejected scenarios (compared to the status quo). The most important attribute was again safe waste disposal, this time followed by insurance coverage and price.

For the final survey the five product attributes 'price', 'blackout', 'waste disposal', 'damage' and 'insurance coverage' were used. Table 2.1 gives an overview and the next section describes the attributes in detail.

2.3.2 Description of Attributes

Price. The attribute 'Price' is the percentage increase caused by the extension of liability insurance coverage over the status quo. It read, e.g., "30 percent more expensive than at present (this is equivalent to a surcharge of CHF 285 on an annual electricity bill of CHF 950)". To obtain an absolute value for willingness to pay (WTP), this value was later multiplied by the actual annual electricity bill as indicated by respondents. Throughout the analysis it was assumed that respondents do not change the quantity of energy consumed. With this assumption, the higher price can be translated into a higher total outlay for electricity, and hence a lower disposable income which thus becomes a product attribute in the econometric analysis (see section 2.4.2).

Blackout. The attribute 'Blackout' indicates if the scenario considered has a high incidence of blackouts (14 per year, coded 1) or a low incidence (2 per year, coded 0). It is a proxy of service reliability. The scenario description read, "... In Switzerland, blackouts are rare and mainly caused by environmental effects (storms) or maintenance work..."

Waste. Since 'Waste' was an important attribute in both pretests, it had to be included in the final experiment. This variable takes on two values: Either there are unresolved problems with waste disposal (=1), or there are no unresolved problems with waste disposal (=0). 'Waste' was described as follows, "... Disposal of waste occasions problems and risks of variable magnitude. This holds in particular for nuclear waste, where these problems are not resolved yet."

Damage. This indicates that electricity generation may cause a (hypothetical) maximum loss amounting to e.g. CHF 100 bn. (appr. US$ 65 bn.) in the event of an accident. In order to make this amount more comprehensible, it was also expressed as an average damage per household. The attribute description said, "... All types of generating facility can cause accidents. Large scale accidents are rare. The magnitude of an accident cannot be calculated with precision but strongly depends on the type of facility. "

Coverage. This indicates the part of maximum loss which would be covered by liability insurance. The text said, e.g., "one percent of the financial damage is insured". See figure 2.3 for an example.

Note that the probability of an accident is not among the product attributes, as explained in section 2.2.2.

Figure 2.3 shows one of the choice scenarios. Type A power is always associated with the status quo scenario to simplify decision making. Type B power has the same five attributes as type A, but with levels changed with regard to four of them. Thus, by simply deciding between types A and B, respondents implicitly trade off attributes. Since this choice is repeated several times with varying attribute levels for type B, these tradeoffs become estimable.

2.3.3 Scenario Selection

The design summarized in table 2.1 gives rise to 256 possible scenarios.[1] Obviously this is an excessive number for the questionnaire, calling for a reduced

[1] $4 \cdot 2 \cdot 2 \cdot 4 \cdot 4 = 256$.

Decision No. 4209		
	Type A power	**Type B power**
Price	A kilowatt hour costs the same as today	A kilowatt hour is 60 percent more expensive than today
Blackouts	2 blackouts per year on average	2 blackouts per year on average
Waste	There are unresolved problems with waste disposal	There are no unresolved problems with waste disposal
Damage	A large scale accident can cause losses up to a maximum of Swiss francs 200 bn. (This amounts to Swiss francs 70,000 per household on average)	A large scale accident can cause losses up to a maximum of Swiss francs 100 mn. (This amounts to Swiss francs 35 per household on average)
Insurance Coverage	1 percent of this maximum damage is covered	100 percent of this maximum damage is covered
Your Choice	☐ Type A	☐ Type B
	☐ cannot decide	

Figure 2.3: Example of a choice scenario.

design. This was constructed using 'Gosset', a general purpose program for designing experiments (Hardin and Sloane, 1993) (see appendix 2.6.3 for the program code). 'Gosset' selects the vectors of a regressor matrix X in a way as to minimize or maximize a function of the covariance matrix of the parameters to be estimated, viz. $(X'X)^{-1}$ in the case of OLS. The D-optimality criterion which was used in the present experiment, maximizes the determinant of the Fisher information matrix. While 'Gosset' solves this maximization problem for linear models (such as OLS) only, it is a reasonable approximation for nonlinear models such as probit (Kanninen, 2002). The reduced design was chosen to enable estimation of all quadratic as well as all interaction terms between product attributes. In this way, the number of scenarios to be included in the survey was reduced to 42.

Since it is not possible for a respondent to evaluate all 42 scenarios, 14 scenarios were chosen at random. This resulted in a unique 'stated choice' questionnaire for each person interviewed.

2.3.4 Design of the Questionnaire

The questionnaire was divided into four parts. The first part consisted of some general warm-up questions concerning energy use. In addition the yearly electricity outlay of the respondent was asked for.

The second part contained information about the consequences of severe accidents in power production, focusing on nuclear and hydro (see appendix 2.6.1 for details). In Switzerland fossil fuels are negligible in electricity generation. In addition, the limited coverage of current nuclear liability insurance was evoked. Two solutions to this problem were sketched. First, the government could raise tax to compensate victims in the event of an accident. Second, mandated insurance coverage could be extended, which however would result in higher electricity prices. Respondents were asked to state their subjective probability of a severe accident relative to that of experts (see appendix 2.6.2). Finally the attributes used in the stated choice part of the questionnaire were explained (see section 2.3.2 above).

The third part contained the scenarios for the 'stated choice' experiment.

In the fourth part, socioeconomic data and information concerning the understanding of the 'stated choice' experiment were collected. Some 20 percent of respondents reported difficulties with the questionnaire. Roughly 73 percent stated they considered one of the attributes to be of overriding importance, which could be interpreted as an indication of lexicographic preferences. However, the econometric analysis failed to produce evidence suggesting that these individuals traded off attributes less frequently or less consistently.

2.4 Econometric Analysis

2.4.1 Data

Face to face interviews were performed in the German speaking part of Switzerland during September and October 2001. In total, 391 persons were interviewed. Average income of respondents is CHF 42,000 (US$ 28,000) having a yearly outlay on electricity of CHF 940 (US$ 630). 80 percent had a medium level of education (either vocational school, community college, technical college or equivalent education) and the average age was 42. Moreover, the sample was designed to have an equal proportion of men and women.

With 391 persons each evaluating 14 (out of a total of 42) choice scenarios, a total of 5,474 decisions were recorded. Respondents who felt unable to decide could always choose the option "not able to decide". This served to prevent choices made at random by individuals who in fact were indifferent or unable to decide.

In 819 cases (15 percent), no choice was stated, resulting in 4,655 usable observations. Only 90 percent of these (4,154) were used for estimation, while 10 percent were put aside for an out-of-sample test (section 2.4.5). In 27 percent of choices, the status quo was preferred. Missing values for socioeconomic information resulted in a final reduction to 4,119 observations.

2.4.2 Theoretical Background and Specification

No attempt was made to anchor specification in expected utility theory. Instead a general utility function in the retained attributes with linear, quadratic and mixed terms is used. To allow for heterogeneous preferences, socioeconomic characteristics of the respondents are interacted with income net of electricity outlay (i.e. disposable income), in accordance with Johnson and Desvousges (1997). This leads to the following utility function for individual i and power type j:

$$U_{ij} = \alpha X_j + \beta_1 w_{ij} + \beta_2 w_{ij}^2 + \gamma_1 z_i w_{ij} + \gamma_2 z_i w_{ij}^2 , \qquad (2.1)$$

where X_j includes all linear, quadratic and mixed terms of electricity attributes, except disposable income w_{ij}. Disposable income is given by $w_{ij} := \left(m_i - outlay_{ij}\right)$

with m_i for income and $outlay_{ij}$ for total outlay on electricity, z_i denotes the vector of socioeconomic variables, and $\alpha, \beta_1, \beta_2, \gamma_1, \gamma_2$ are to be estimated.

According to the random utility model (McFadden, 2001), the respondent evaluates the utility of the two scenarios and chooses the one with the higher utility. If 'A' denotes the attribute values of the status quo scenario, then individual i chooses the alternative (B) of choice set j if $U_{iB_j} > U_{iA_j}$, i.e. if

$$\begin{aligned}
&\alpha(X_{B_j} - X_{A_j}) + \beta_1(1 - d_i)(w_{iB_j} - w_{iA_j}) + \beta_2(1 - d_i)(w^2_{iB_j} - w^2_{iA_j}) \\
&+\gamma_1 z_i(1 - d_i)(w_{iB_j} - w_{iA_j}) + \gamma_2(1 - d_i)z_i(w^2_{iB_j} - w^2_{iA_j}) \\
&+\delta_1 d_i(w_{iB_j} - w_{iA_j}) + \delta_2 z_i d_i(w_{iB_j} - w_{iA_j}) \\
&+\delta_3 d_i(w^2_{iB_j} - w^2_{iA_j}) + \delta_4 d_i z_i(w^2_{iB_j} - w^2_{iA_j}) + \eta_i + \epsilon_{ij} > 0\ .
\end{aligned} \tag{2.2}$$

In addition to equation (2.1), equation (2.2) contains a dummy variable d_i to reflect the fact that income is missing with 43 percent of all individuals. It takes on the value of zero if income was revealed and one if missing, in which case the difference in disposable income reduces to the (inverse) difference of electricity outlay.

The error term appearing in this comparision has an individual-specific (η_i) and a general component (ϵ_{ij}) that also varies with the choice set presented. The two components are assumed to conform to the usual random effects specification (Greene, 1997, ch. 14), with $\rho = var(\eta_i)/var(\eta_i + \epsilon_{ij})$.

The dependent variable y_{ij} (choice of power type B) is given by

$$y_{ij} = \begin{cases} 1 & \text{if inequality (2.2) holds} \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

From this expression it becomes clear that the variables used in estimation are the differences between the attribute levels of scenario B and scenario A. For instance *damage* is defined as the maximum level of loss in scenario B minus the maximum level of loss in scenario A. For example, an individual opting for type B power in decision no. 4209 of figure 2.3 and having revealed his or her income has the following observation vector, in keeping with table 2.1: (y; *blackout*; *waste*; *damage*; *coverage*; …; $blackout^2$; $waste^2$; $damage^2$; $coverage^2$; …) = (1; 0 − 0; 0 − 1; 0.1 − 200; 100 − 1; …; $0^2 - 0^2$; $0^2 - 1^2$; $0.1^2 - 200^2$; $100^2 - 1^2$; …). Personal characteristics, which do not change between scenarios, drop out of the regression unless they interact with electricity attributes.

Table 2.2: Descriptive statistics for variables used in estimation.

	Mean	**Median**	**Unit**
blackouts[a)]	0.48	0	dummy
waste[a)]	-0.49	0	dummy
damage[a)]	-128	-190	CHF bn.
coverage[a)]	43	49	percent
disposable income[a)]	-152	0	CHF
damage^2 [b)]	-28,175	-39,900	CHF bn.
coverage^2 [b)]	3,360	2,500	percent
outlay on electricity	940	840	CHF
income (if revealed)	41,890	36,000	CHF / year
income not revealed	0.43	0	dummy
age	43	42	years
sex is female	0.51	1	dummy
pessimistic beliefs	0.72	1	dummy
medium level of education[c)]	0.82	1	dummy
high level of education	0.08	0	dummy

[a)] Denotes difference between status quo (A) and alternatives (B_j), see section 2.4.2

[b)] Difference of squared values, e.g. $damage^2 = damage_B^2 - damage_A^2$

[c)] Vocational school, community college, technical college

2.4.3 Explanatory variables

Table 2.2 shows descriptive statistics for some of the variables used. The product attributes were already explained in section 2.3. Note that outlay on electricity rather than price was used as an explanatory variable. In this way, an increase in outlay (occasioned by a higher price) can be interpreted as a reduction of disposable income. Of course, this holds only if outlay and price move in fixed proportions, i.e. if the quantity of power consumed stays constant. Since price elasticities of the household demand for electricity are low in Switzerland (Bonomo et al., 1998), this assumption is justifiable.

Product attributes appear in linear, quadratic, and mixed form. The dummy variables for *blackout* and *waste* cannot be squared because of multicollinearity. In keeping with section 2.4.2, only the difference in the attributes between scenario B (the alternative scenario) and scenario A (the status quo) are relevant for estimation. Therefore all product attribute variables in table 2.2 are differences between A and B, e.g. *damage* is defined as $damage_B - damage_A$ and $damage^2$ as $damage_B^2 - damage_A^2$ (note that $damage^2 \neq (damage_B - damage_A)^2$).

Table 2.3: Estimation of utility function.

Variables (selection)	**Coefficient**	**S.E.**
many blackouts (blackouts)	-0.26656**	0.09293
unresolved waste disposal problems (waste)	-0.67369**	0.09702
damage in 100 CHF bn. (damage)	0.06612	0.15924
insurance coverage in percent (coverage)	0.01416**	0.00329
disposable income in CHF 000s (income)	2.9755*	1.4302
income^2	0.0228	0.0142
damage^2	-0.2320**	0.0794
coverage^2	-0.00009**	0.00003
damage*coverage	0.0342**	0.0106
coverage*blackouts	-0.00363*	0.00164
(pessimistic beliefs) * income	-1.0153	0.6523
(medium level education)* income^2	-0.0239+	0.0136
(high level education)*income	-6.7719**	1.6609
age*income	0.0480*	0.0211
age*income^2	-0.0002+	0.0001
constant	0.22319+	0.13173
For complete estimation results see appendix 2.6.4		
Observations	4119	
Number of individuals	375	
Log likelihood	-1959.67	
Log likelihood constant only	-2326.77	
ρ	0.5425	
σ_η	1.0890	
+ significant at 10%,* significant at 5%, ** significant at 1%.		

2.4.4 Estimation Results

General Findings

The model from equation (2.3) was estimated using a random-effects probit specification. Estimation result for the utility function are displayed in table 2.3 as well as in appendix 2.6.4.

The linear forms of four out of five product attributes have the expected sign; note that (disposable) *income* has become a function of outlay and thus constitutes a product attribute. While *damage* has an unexpected (but insignificant) positive coefficient, the negative sign of *damage*^{2} changes the overall impact from positive to negative for a large subset of values of explanatory variables.

Most importantly, the coefficients for *coverage* and *income* have both a pos-

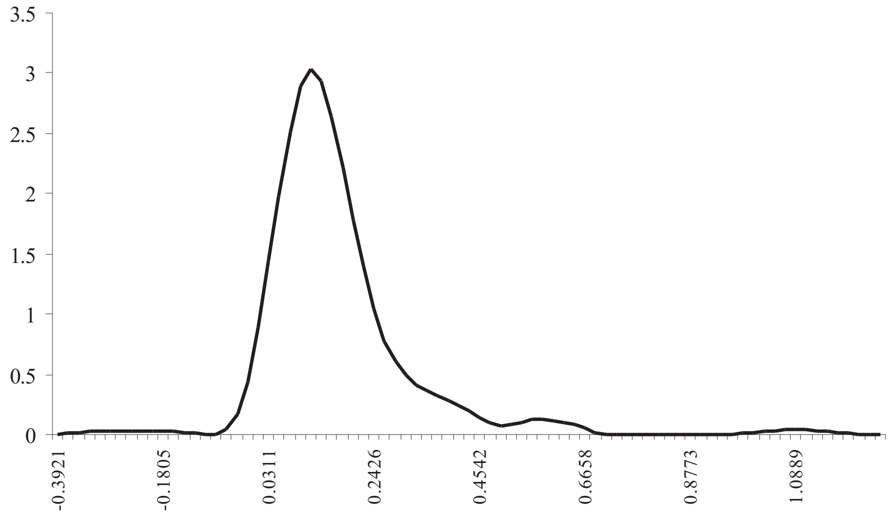


Figure 2.4: Density of MWP in US cents/kwh. *damage* = 200 (US$ 130 bn.), *waste* = 1, *blackouts* = 0 and *coverage* = 1. (Epanechnikov kernel with a bandwidth of 0.66)

itive and significant effect on utility and hence choice probability. Finally, there is some indication of respondents who hold a more pessimistic belief than experts with regard to the probability of an accident having a lower marginal utility of income over a large range of income (see the not quite significant coefficient of *pessimistic beliefs*income*). In keeping with equation (2.4) below, this should result in a higher marginal willingness to pay for additional coverage. The same holds for individuals with a high education level, however, the negative basis effect evidenced in table 2.3 is so strong that their WTP remains negative over the whole range of coverage.

Finally, the goodness of fit, measured by a pseudo-R^2 of 0.16, is satisfactory for a random-effects specification. The significantly positive value of $\rho = 0.54$ shows that 54 percent of the variance of the error term can be attributed to individual-specific effects.

Calculation of MWP for Financial Security

The marginal willingness to pay (MWP) for additional coverage is given by

$$MWP = \frac{\partial \hat{U}/\partial coverage}{\partial \hat{U}/\partial income}, \tag{2.4}$$

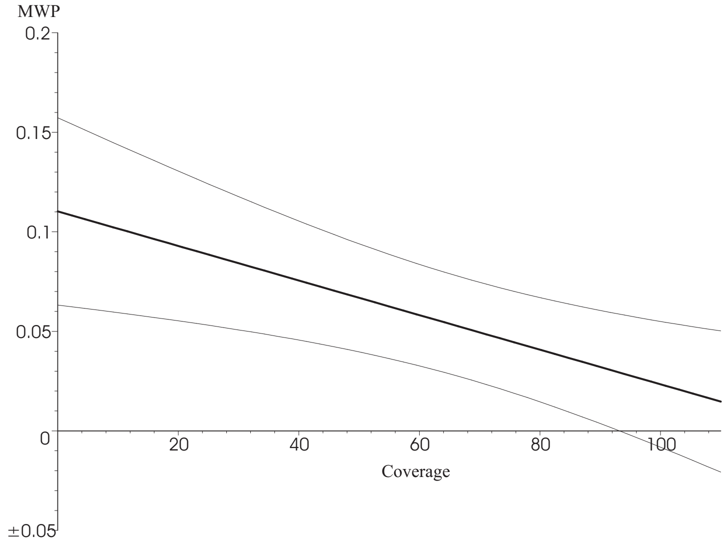


Figure 2.5: MWP in US cents/kwh including 95 percent confidence intervals; 42 year old woman, medium level education, pessimistic beliefs, income US$ 24,000, *damage* = 200 (US$ 130 bn.), *waste* = 1, *blackouts* = 0, *outlay* = 840 (US$ 560).

with *income* defined as income net of outlay on electricity so that $\partial\hat{U}/\partial income$ is the estimated marginal utility of disposable income. The ratio between the marginal utility of coverage and the marginal utility of income defines the MWP for additional financial security through increased insurance coverage. In order to express the MWP in terms of US cents per kwh, the quantity of power consumed must be known. This was calculated as the annual outlay devided by the average price of electricity at the household level during the year 1999 (16.2 Swiss cents/kwh, i.e. US 11 cents/kwh), obtaining

$$\text{MWP[US cents/kWh]} := \frac{\text{MWP[US\$/year]} \cdot 100}{\text{Outlay in US\$/0.11}} \tag{2.5}$$

In order to get an impression of the MWP across the entire sample, the MWP for increased coverage was calculated for each person based on his or her specific socioeconomic characteristics and a power type which has few blackouts (*blackouts* = 0), unresolved problems with waste disposal (*waste* = 1), a maximum possible loss of CHF 200 bn. (US$ 130 bn.) and initial coverage amounting to a mere 1 percent (*coverage* = 1 percentage). The resulting density function is shown in figure 2.4.

For the calculation of this density function, only the subsample of individuals who stated their income was used. Average MWP for this group amounts to 0.16 US cents/kwh, median MWP to 0.14 US cents/kwh. Compared to the average price of 11 cents paid in 1999, this corresponds to 1.3 percent of the electricity price. This does not seem excessive a priori, especially when taking the decline of MWP with increasing coverage into account (see figure 2.5).

In order to check whether these MWP values are significantly different from zero, standard errors were calculated, using the delta-method although it may result in an underestimate (see e.g. Polsky et al. (1997) and Telser (2002, ch. 4)). The alternatives would have been the Filler method, which only works for a simple ratio of coefficients, and bootstrapping.

Figure 2.5 shows the MWP of a person with median characteristics (woman aged 44 with medium level of education and pessimistic beliefs) along with its 95 percent confidence intervals. Estimated MWP declines, becoming indistinguishable from zero near an initial coverage rate of 95 percent. At a coverage level of 100 percent at the latest, MWP should theoretically be zero. However, this restriction was in no way built into the experiment. Figure 2.5 thus may be considered as providing preliminary evidence for the experiment's validity.

Plausibility tests of estimated MWP

A first plausibility test derives from the influence of income on MWP. If financial security is a normal economic good, MWP should be higher than average among individuals with high income. This prediction is borne out in figure 2.6. However, these differences lack statistical significance.

Second, due to the income effect, MWP should decline with increasing outlay on electricity. This is indeed the case without exception in both tables 2.4 and 2.5. Moreover, MWP again decreases with initial coverage for a given value of electricity outlay, confirming figures 2.5 and 2.6.

Comparison of tables 2.4 and 2.5 (with maximum loss doubled) shows that MWP increases systematically with maximum possible loss. This too corresponds with theoretical considerations if risk aversion is assumed (Chambers and Quiggin, 2000, ch. 3).

Finally, under very general conditions, the MWP of a risk averse individual is predicted to increase with increasing probability of an accident. Evaluation of the equation for MWP with regard to all relevant levels of the attributes shows that the MWP of individuals with pessimistic beliefs (probability of accident is higher

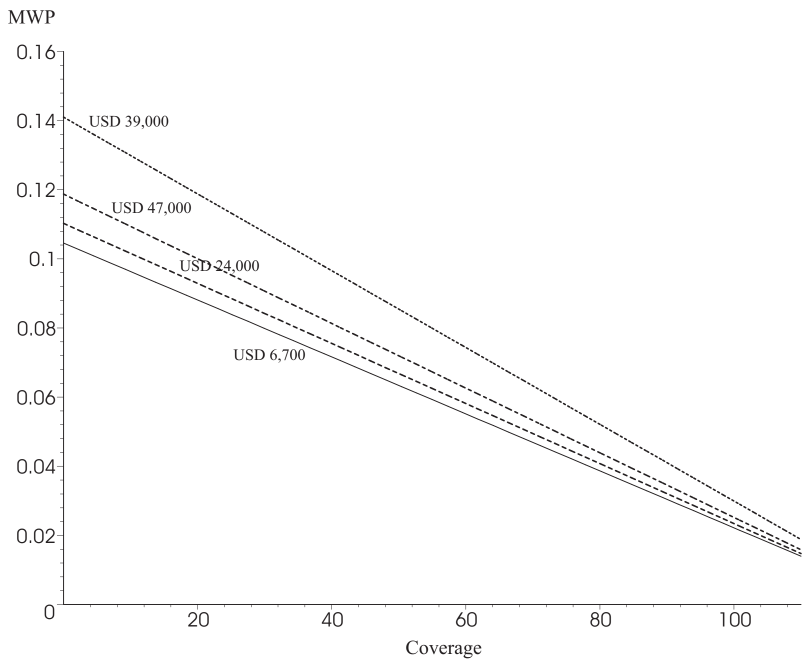


Figure 2.6: MWP in US cents/kwh of a 42 year old woman with medium level education, pessimistic beliefs and outlay on electricity of CHF 840 (US$ 560) per year. Maximum loss is CHF 200 bn. (US$ 130 bn.). Incomes at US$ 6,700; 24,000; 47,000; 93,000 respectively.

Coverage	**Electricity outlay (US$ per year)**						
(percent)	**400**	**530**	**670**	**800**	**930**	**1070**	**1200**
0	0.1306	0.0979	0.0783	0.0652	0.0559	0.0489	0.0434
20	0.1063	0.0797	0.0637	0.0531	0.0455	0.0398	0.0353
40	0.0820	0.0615	0.0491	0.0409	0.0351	0.0307	0.0273
60	0.0576	0.0432	0.0346	0.0288	0.0247	0.0216	0.0192
80	0.0333	0.0250	0.0200	0.0166	0.0142	0.0125	0.0111
100	0.0090	0.0067	0.0054	0.0045	0.0038	0.0034	0.0030

Table 2.4: MWP in US cents/kwh of a 42 year old woman with medium level education, pessimistic beliefs and an income of US$ $24,000$. Maximum loss is CHF 100 bn. (US$ 65 bn.), $waste$ = 1, $blackouts$ = 0.

Coverage	**Electricity outlay (US$ per year)**						
(percent)	**400**	**530**	**670**	**800**	**930**	**1070**	**1200**
0	0.1545	0.1158	0.0926	0.0771	0.0661	0.0578	0.0514
20	0.1301	0.0976	0.0780	0.0650	0.0557	0.0487	0.0433
40	0.1058	0.0793	0.0634	0.0528	0.0453	0.0396	0.0352
60	0.0815	0.0611	0.0488	0.0407	0.0349	0.0305	0.0271
80	0.0571	0.0428	0.0343	0.0285	0.0244	0.0214	0.0190
100	0.0328	0.0246	0.0197	0.0164	0.0140	0.0123	0.0109

Table 2.5: MWP in US cents/kwh of a 42 year old woman with medium level of education, pessimistic beliefs and an income of US$ $24,000$. Maximum loss is CHF 200 bn. (US$ 130 bn.), $waste$ = 1, $blackouts$ = 0.

than experts' estimate) is indeed greater than with beliefs congruent with experts' estimate. This effect, however, is not significant.

2.4.5 Out-of-Sample Test

The two performance criteria of an empirical investigation are its reliability and validity (Singleton and Straits, 1999). Reliability refers to the stability and consistency of the operational definition (here: MWP); validity refers to the goodness of fit between the operational definition and the concept that it is supposed to measure (here: valuation of nuclear risks).

Since effective choices cannot be observed and compared with the stated choices in the experiment, it is not possible to check validity directly, i.e. whether estimated MWP for a reduction of financial risk is a good proxi for the valuation of financial consequences of nuclear risks.

However, there are studies suggesting that the 'stated choice' method leads to results that are in line with corresponding hedonic price estimations (Gegax and Stanley (1997), Louviere et al. (1999), and Haener et al. (2000)). This suggests validity.

As to reliability, an out-of-sample test can be performed on the 10 percent of observations that were not used for estimation. The model predicts the probability of choosing the alternate scenario. For a calculated probability of more than 50 percent, the individual is assumed to choose the alternate scenario. It turns out that out of sample, roughly 70 percent of all decisions were predicted correctly. This share has to be compared to the share of correct decisions which would result from a random process. In the sample used for estimation, the alternate scenario was chosen 63 percent of the time. Now, a random process that generates choice of the alternate scenario in 63 percent of all cases and of the status quo scenario in 27 percent would predict correctly in 47 percent of cases. This value is the sum of the probability that the random process predicted the alternative and that the alternative was actually chosen (0.63^2) plus the probability that it predicted the status quo and that the status quo was actually chosen (0.27^2).

The estimated utility model thus serves to increase the share of correct predictions by 23 percentage points (= 70 − 47) over a random process. This points to a measure of reliability.

2.5 Conclusions

Measurement of willingness to pay for an increased internalization of the risks emanating from nuclear power plants is important for energy policy. One instrument of internalization is extending coverage provided by mandatory liability insurance for plant operators. For all its popular appeal, such a proposal will face opposition in parliament and by consumers since higher insurance premiums lead to higher electricity prices.

This study seeks to determine how much Swiss citizens value increased financial security through increased insurance coverage in case of an accident by using the economic concept of marginal willingness to pay for (financial) security.

Since additional coverage is not available to individual consumers, a 'stated choice' experiment was carried out, in which respondents decide in favor of or against an alternative to the status quo, characterized by several attributes of electricity. These attributes are varied throughout the experiment, in contradistinction to conventional 'contingent valuation' approaches. The relevant attributes were established by means of three pretests and turned out to be electricity price, frequency of blackouts, waste disposal, maximum possible loss in case of an accident, and insurance coverage. The econometric analysis confirms this selection, since all attributes are estimated to be statistically significant arguments of the underlying utility function. Average marginal willingness to pay for an additional percentage point of compensation for losses in excess of the status quo amounts to some 0.16 US cents per kwh (median value 0.14 cents), approaching zero when insurance coverage goes towards 100 percent.

Specifically, an increase of mandated liability insurance coverage from today's CHF 0.7 bn. (US$ 0.47 bn.) to CHF 4 bn. (US$ 2.7 bn.) would command a WTP amounting to 0.40 US cents/kwh. This can be compared to an estimate of additional cost. In a companion study, a log-logistic density function for nuclear damages (i.e. the loss function for nuclear insurers) was calibrated. According to that study, an increase of liability insurance from today's CHF 0.7 bn. (US$ 0.47 bn.) to CHF 4 bn. (US$ 2.7 bn.) would result in an increase in the price of electricity of 0.008 US cents/kwh (Zweifel and Umbricht, 2002, table 4.16). Therefore, quintuplicating current insurance coverage could lead to a welfare gain for the majority of Swiss citizens.

This proposition has to be qualified in several ways. On the cost side, the choice of the distribution law can be criticized. Indeed, a different choice (Gamma e.g.) would entail somewhat changed marginal cost estimates. On the benefits side investigated here, one has to accept the fact that no thought experiment can simulate the actual decision environment completely. In particular there is no guarantee that participants take described damages seriously and do not speculate on the government providing financial assistance to victims in case of a major accident[2]. On the other hand, estimated values of marginal willingness to pay do exhibit theoretically plausible variations in several dimensions, thus providing a measure of support for the validity of the experiment.

[2]In fact it is almost certain that the government will step in, as recent much less severe events have shown (e.g. the bailout of Swiss Airlines Ltd in 2002).

2.6 Appendix

2.6.1 Contextual Information Provided for the Interview

Pros and Cons of Nuclear Power

In the current debate, the following advantages of nuclear power are often mentioned:

- produces a great deal of electric power
- no environmental pollution
- low accident risk compared to other sources of energy

Disadvantages mentioned are:

- degradation of natural scenery
- disposal of nuclear waste
- long-term consequences of accidents (nuclear radiation).

Damages Caused by Accidents in Power Plants

Power plants can cause great damage in case of an accident. In case of a severe accident in a **nuclear power plant**, a large part of the resident population needs to be evacuated. Acute disease and death may occur in the vicinity of the power plant. However, only few deaths are expected as a rule. With considerable delay, an accident may affect remote areas in that entire regions suffer from radiation and may partially become uninhabitable for years.
The probability of such an accident is very small.

Breaches of hydro dams are somewhat more frequent. However, their probability is low as well.
A **sudden breach of a dam** releases a huge flood wave. Affected individuals cannot be warned in time, and thousands of them may die. In most cases however, there is enough time to warn residents, permitting the great majority of them to be evacuated. The immediate risk is limited to people on the downstream side of the dam. Most damages are repaired within a year after the accident. However, the natural environment may take several years to recover. At any rate, damages to man and nature are massive.

Today already, operators of nuclear power plant are mandated to buy insurance that pays in case of damage. However, insurance covers only part of the possible loss. It is an open question as to who would be responsible for the uncovered remainder. Federal parliament would decide on the issue. Therefore, there is no guarantee that victims are fully compensated. The uncovered remainder of the damage may easily exceed Switzerland's annual tax revenue.

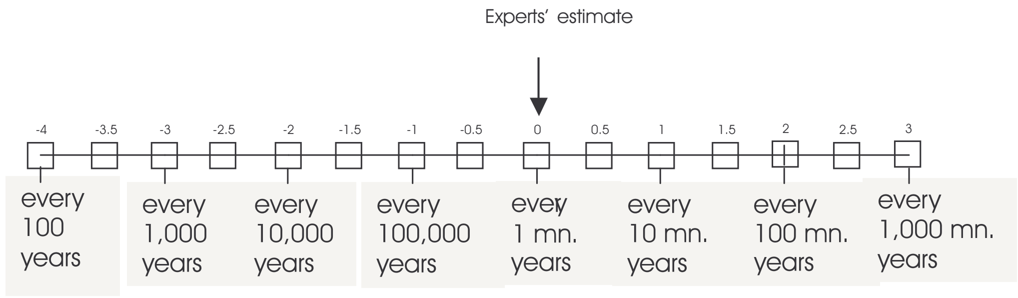
In order to better secure compensation of victims the law could stipulate an extension of **insurance coverage**. In this way a greater part of possible damage would be paid by insurers covering power plants. This would have the advantage of providing **improved financial protection** to victims. In addition, this could be an incentive for power plant operators to invest even more in the safety of their plant.

On the other hand, an extension of insurance coverage would increase the cost of operation of power plants, resulting in **higher prices for electricity** to consumers.

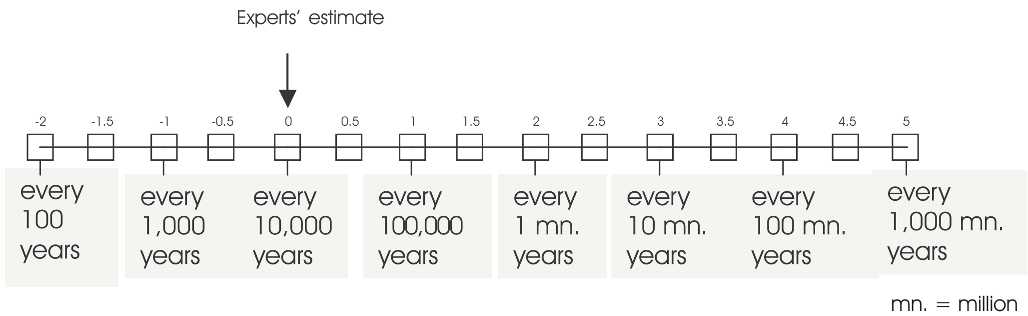
2.6.2 Risk Estimate

Method used in the questionnaire to elicit the subjective risk estimates for nuclear and hydro accidents respectively. The respondents had to mark their own risk estimate in one of the boxes.

Subjective probability of large-scale nuclear accident:



Subjective probability of dam failure (hydro):



2.6.3 Gosset Code Used to Construct Reduced Design

The experimental design was constructed using the program 'Gosset' by Hardin and Sloane. For more information on 'Gosset' see Hardin and Sloane (1993) and Hardin and Sloane (1994).

```
10 discrete price 0 10 30 60 20 discrete blk 0 1 30 discrete wst 0 1 40
discrete dam1 dam2 dam3 cov1 cov2 cov3 0 1 50 range dam4 cov4 0 1
         60 constraint
dam1+dam2+dam3+dam4=1 70 constraint cov1+cov2+cov3+cov4=1 80 model
(1+price+blk+wst+dam1+dam2+dam3+dam4+cov1+cov2+cov3+cov4)^2
         + price^3-blk^2-wst^2-dam1^2-dam2^2-dam3^2-dam4^2
```

```
    - cov1^2-cov2^2-cov3^2-cov4^2

design type=D n=50
```

2.6.4 Complete Estimation Results

Variables	Coefficient	S.E.
many blackouts (blackouts)	-0.26656**	0.09293
unresolved waste disposal problems (waste)	-0.67369**	0.09702
damage in 100 bn. CHF (damage)	0.06612	0.15924
insurance coverage in percent (coverage)	0.01416**	0.00329
disposable income in CHF 000s (income)	2.9755*	1.4302
$income^2$	0.0228	0.0142
$damage^2$	-0.2320**	0.0794
$coverage^2$	-0.00009**	0.00003
damage*coverage	0.0342**	0.0106
damage*waste	0.1145	0.0771
damage*blackouts	-0.0364	0.0718
coverage*waste	0.0012	0.0015
coverage*blackouts	-0.0036*	0.0016
noincome (= -outlay in CHF 000s)	-4.0500	2.8300
no $income^2$ (equal to $outlay^2$)	0.4740	1.0200
(pessimistic beliefs) * income	-1.0153	0.6523
(pessimistic beliefs) * $income^2$	0.0051	0.0031
(pessimistic beliefs) * noincome	0.0813	0.3305
(medium level of education)*income	-1.2353	1.1717
(medium level of education)*$income^2$	-0.0239+	0.0136
(medium level of education)*noincome	-1.0342	2.1487
(medium level of education)*$noincome^2$	0.5970	0.7770

continued...

...concluded

Variables	**Coefficient**	**S.E.**
(high level of education)*income	-6.7719**	1.6609
(high level of education)*income^2	-0.0066	0.0146
(high level of education)*noincome	1.0015	2.2493
(high level of education)*noincome^2	0.0724	0.7700
age*income	0.0480*	0.0211
age*income^2	-0.0002+	0.0001
age*noincome	-0.0041	0.0382
age*noincome^2	0.00000	0.0134
female*income	0.0338	0.6020
female*income^2	0.0029	0.0032
female*noincome	0.4417	1.1221
female*noincome^2	-0.4410	0.3820
constant	0.22319+	0.13173
Observations	4119	
Number of individuals	375	
Log likelihood	-1959.67	
Log likelihood constant only	-2326.77	
ρ	0.5425	
σ_η	1.0890	

+ significant at 10%,* significant at 5%, ** significant at 1%.

Note: Since *income* is equal to stated income minus *outlay* on electricity, *noincome* $=-$*outlay*.

Bibliography

Bonomo, S., M. Filippini, and P. Zweifel (1998). Neue aufschlüsse über die elektrizitätsnachfrage der schweizerischen haushalte (new evidence on electricity demand by swiss households). *Zeitschrift für Volkswirtschaft und Statistik 134*, 415–430.

Camerer, C. F. and H. Kunreuther (1989). Decision processes for low probability events: Policy implications. *Journal of Policy Analysis and Management 8*, 565–92.

Chambers, R. G. and J. Quiggin (2000). *Uncertainty, Production, Choice and Agency*. Cambridge University Press.

Ganderton, P. T., D. S. Brookshire, M. McKee, S. Stewart, and H. Thurston (2000).

Buying insurance for disaster-type risks: Experimental evidence. *Journal of Risk and Uncertainty 20*, 271–289.

Gegax, D. and L. Stanley (1997). Validating conjoint and hedonic preference measures: Evidence from valuing reductions in risk. *Quarterly Journal of Business and Economics 36*, 31–54.

Greene, W. H. (1997). *Econometric Analysis*. Prentice-Hall International.

Haener, M., P. Boxall, and W. Adamowicz (2000). Modeling recreation site choice: Do hypothetical choices reflect actual behavior? *Working Paper Staff Paper 00-01, Departement of Rural Economy, University of Alberta, Edmonton, Canada 49*, 1–15.

Hardin, R. H. and N. J. A. Sloane (1993). A new approach to the construction of optimal designs. *Journal of Statistical Planning and Inference 37*, 339–369.

Hardin, R. H. and N. J. A. Sloane (1994). Operating manual for gosset: A general purpose program for constructing experimental designs (second edition). http://www.research.att.com/~njas/gosset/.

Hausman, J. (Ed.) (1993). *Contingent Valuation - A Critical Assessment*. North Holland.

Hedayat, A., N. Sloane, and J. Stufken (1999). *Orthogonal Arrays - Theory and Applications*. Springer.

Hirschberg, S., G. Spiekerman, and R. Dones (1998). *Severe Accidents in the Energy Sector*. Paul Scherrer Insitute, Report 98-16.

Johnson, R. F. and W. H. Desvousges (1997). Estimating stated preferences with rated-pair data: Environmental, health, and employment effects of energy programs. *Journal of Environmental Economics and Management 34*, 79–99.

Kanninen, B. J. (2002). Optimal design for multinomial choice experiments. *Journal of Marketing Research 39*, 214–227.

Krupnick, A., A. Alberini, M. Cropper, N. Simon, B. O'Brien, R. Goeree, and M. Heintzelman (2002). Age, health and the willingness to pay for mortality risk reductions: A contingent valuation survey of ontario residents. *Journal of Risk and Uncertainty 24*, 161–186.

Kunreuther, H., N. Novemsky, and D. Kahneman (2001). Making low probabilities useful. *The Journal of Risk and Uncertainty 23*, 103–120.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy 74*, 132–57.

Louviere, J., R. Meyer, and D. Bunch (1999). Combining sources of preference data for modeling complex decision processes. *Marketing Letters 10*, 205–217.

Louviere, J. J. and D. A. Hensher (1982). On the design and analysis of simulated or allocation experiments in travel choice modelling. *Transportation Research Record 890*, 11–17.

Louviere, J. J., D. A. Hensher, and J. D. Swait (2000). *Stated Choice Methods - Analysis and Applications*. Cambridge University Press.

Magat, W. A., W. K. Viscusi, and J. Huber (1988). Paired comparison and contingent valuation approaches to morbidity risk valuation. *Journal of Environmental Economics and Management 15*, 395–411.

McFadden, D. (2001). Economic choices. *American Economic Review 91*, 351–378.

Mitchell, R. C. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Ressources for the Future.

Polsky, D., H. Glick, R. Willke, and K. Schulman (1997). Confidence intervals for cost-effectivness ratios: A comparison of four methods. *Health Economics 6*, 243–252.

Shavell, S. (1986). The judgement proof problem. *International Review of Law and Economics 6*, 45–58.

Singleton, R. A. J. and B. C. Straits (1999). *Approaches to Social Research, 3rd edition*. Oxford University Press.

Telser, H. (2002). *Nutzenmessung im Gesundheitswesen - Die Methode der Discrete Choice Experimente (Preference Elicitation in Health Care - The Technique of Discrete Choice Experiments)*. University of Zurich, PhD dissertation.

Telser, H. and P. Zweifel (2002). Measuring willingness-to-pay for risk reduction: An application of conjoint analysis. *Health Economics 11*, 129–139.

Viscusi, W. K. (1998). *Rational Risk Policy*. Oxford University Press.

Zweifel, P. and R. D. Umbricht (2002). Verbesserte deckung des nuklearrisikos - zu welchen bedingungen? (extended coverage of nuclear risk - on what conditions). *Forschungsprogramm Energiewirtschaftliche Grundlagen, Bundesamt für Energie*.

Spatial Effects in Willingness-to-Pay: The Case of Nuclear Risks

Peter Zweifel, Yves Schneider and Christian Wyss [*] [†]

[*]Submitted to the American Economic Review.

[†]Financial support came from the Swiss Federal Office of Energy (Berne) and the Swiss Institute of Technology.

Chapter 3

Spatial Effects in Willingness-to-Pay: The Case of Nuclear Risks

3.1 Introduction

This paper examines the spatial dimension of responses to the external effects of nuclear power. Most people would agree that these effects decrease - ceteris paribus - with distance from a nuclear power plant. They would therefore expect willingness-to-pay (WTP) for coverage against the financial risks of a nuclear accident to decrease with distance from plant. However, 'ceteris paribus' may not hold because homeowners and tenants can choose where to locate. If those who believe to be little affected by the externality settle in the vicinity of the plant, WTP for financial safety may well increase rather than decrease with distance. Still another spatial relationship is predicted for waste disposal, a second externality associated with nuclear power. As long as the (national) site of waste disposal is not determined yet but could be anywhere in the country where geological conditions are favorable, putting more distance between the plant and one's residence does not make much of a difference. Thus, WTP for getting rid of this externality is predicted to have no relation with distance (except in the very neighborhood of the plant, where all eligible train lines and truck routes necessarily originate).

The present study purports to test these predictions using data from a stated choice experiment (SCE) with Swiss individuals. It is of particular interest for two reasons. First, most of the existing published work concerned with the effect of distance on the WTP for reducing a negative externality focuses on housing prices.

By way of contrast, this contribution reports on WTP values derived from market experiments yielding measurements of demand uncontaminated by supply conditions. Second, the evidence comes from Switzerland, a country where residents have been having full opportunity to choose their location in response to nuclear externalities. As noted above however, such a relocation usually proves ineffective when it comes to escaping nuclear waste; in fact, Switzerland does not yet have a long-term repository for spent nuclear fuel and high-level radioactive waste . A large fraction of nuclear material is kept at the respective power plants before being shipped to either France or Great Britain.

The SCE was conducted in 2001. Some 370 respondents were asked to choose between the status quo and an alternative type of electricity, defined by a changed number of blackouts, problems with the disposal of nuclear waste, severity of (financial) damage in case of accident, and degree of coverage through mandatory liability insurance carried by nuclear plant operators.

At present, operators are obliged to insure for CHF 1 billion (bn.) (approx. US$ 0.8 bn. [US$ 1 ≈ CHF 1.25 at 2003 exchange rates]), an amount that will hardly be sufficient to compensate the victims of a major accident. However, an extension of coverage will c.p. result in higher consumer prices for electricity. This also holds true for any conceivable solution to the problem of nuclear waste, since in Switzerland 35 percent of electricity is provided through nuclear power (most of the remainder comes from hydro power).

The remainder of this paper is structured as follows. After a review of the literature dealing with the spatial effects of externalities associated with nuclear power, section 3.3 describes the SCE that was applied to measure WTP of the Swiss population for reducing these externalities. Section 3.4 develops the hypotheses to be tested, mainly with regard to the effect of distance on measured WTP. The econometric specification, based on the Random Utility Model, is presented in section 3.5. Estimation results and hypothesis tests follow in section 3.6, while concluding remarks are offered in section 3.7.

3.2 Review of the Literature

Choosing one's optimal location with respect to the risk of being affected by an externality can be viewed as self-insurance. By locating farther away from a nuclear power plant, individuals reduce their losses e.g. due to radiation in case of a severe accident. Ehrlich and Becker (1972) analyze the effect of the simultaneous availability of self-insurance and market insurance and conclude that both 'technologies' are substitutes as long as the price of market insurance is independent of the amount of self-insurance.

Since up to now it has not been possible for Swiss citizens to buy insurance against nuclear risks, the only option available is self-insurance. In the absence of an insurance market, the optimal level of self-insurance thus increases with increased risk aversion (see for instance Dionne and Eeckhoudt, 1985). Therefore, more strongly risk-averse individuals are predicted to spend more resources on loss mitigation than less risk-averse consumers.

The experiment conducted in the present study introduces a hypothetical insurance market. Respondents in this experiment evidently state their WTP for nuclear insurance after having set their optimal level of self-insurance through their residential choice. Although there is no need to account for strategic interaction between market insurance and self-insurance (Kelly and Kleffner, 2003), it is necessary to account for self-insurance that has taken place prior to the experiment. Thus, estimated WTP for insurance coverage is expected to vary systematically with the degree of self-insurance, i. e. residential location in the present context.

There is a large body of empirical work estimating the effect of proximity to a source of disamenity on property values. The case of nuclear power plants was first studied by Nelson (1981) and Gamble and Downing (1982). In the wake of the 1979 incident at Three Mile Island, they find weak or even reversed distance effects, viz. higher property values in the vicinity of the plant. Folland and Hough (2000) extend their focus beyond a single power plant, analyzing a panel data set of broad market areas across the United States. Their evidence points to a negative impact of nuclear power plants on land prices, with distance again having an ambiguous effect.

However, Gawande and Jenkins-Smith (2001) find that being five miles away from a nuclear waste shipment route was associated with a 3 percent increase of average house value compared to property on the route.

While important, nuclear power is only one of several sources of disamenities. In their review, Gawande and Jenkins-Smith (2001) conclude that a wide range of disamenities such as superfund sites and polluted water negatively influence the value of residential property. More specifically Faber (1998), collecting evidence on the effects of distance, finds that their magnitude depends on the type of facility, community characteristics, and setting (rural or urban). Chemical refineries and nuclear power plants seem to have roughly comparable (positive) gradients, amounting to $200-300 per mile of distance (in 1993 dollars). Compared to other facilities, this is a rather small effect, as a proposed radioactive waste disposal site was associated with a gradient of $4,440 per mile. As Clark and Allison (1999) found in their study, the distance effect weakens over time, suggesting that relocation of individuals may replenish demand for property close to the source of the externality by those who believe to be little affected, a consideration taken up in section 3.4.

Most of these studies rely on hedonic modeling, linking price data to a set of characteristics of real estate property. Davis (2005) applies this method to estimate marginal WTP to avoid pediatric leukemia risk. As he points out, the heterogeneity of individuals (with respect to income or preferences in general) contaminates housing price data. Furthermore, the cost of changing location, which constitutes the cost of this particular self-insurance technology, is arguably not trivial. Moreover, market prices also depend on supply which in turn is affected by zoning laws and building regulations. For these reasons, estimates of individual WTP derived from analyzing the compensating differentials contained in market data are potentially distorted and incomplete. Experimental evidence may thus complement information gleaned from market data.

For example, Smith and Desvousges (1986) analyze the impact of a waste disposal facility on the subjective value of a residential site using a contingent valuation experiment. For hazardous waste, they obtain a positive distance gradient of $330-$495 per mile. A study related to the present paper is by Riddel et al. (2003), who estimate the effect of several planned nuclear waste transportation routes from power plants to the Yucca Mountain (Nevada) repository. They find evidence that perceived risk decreases with distance to the planned transportation route and that higher perceived risk results in a higher probability of moving away from the route.

As opposed to Riddel et al. (2003), this study does not consider a planned and not yet effected change in the environment. Rather, it seeks to measure WTP for a reduction of two risks emanating from nuclear power plants that have been effective for at least two decades. Respondents had ample opportunity to relocate according to their preferences regarding nuclear power plants. The present investigation therefore estimates the net disutility caused by nuclear power plants, given the respondents' amount of self-insurance through locational choice.

Attribute	Levels (Coding[c])	Unit	Status quo
PRICE	0; 10; 30; 60 (**0**;...;60)	percent	0
BLACKOUTS	2; 14 (**0**;1)	numb./year	2
NOWASTE	unresolved problems (**1**); no unresolved problems (0)		unresolved problems
DAMAGE[a]	0.1; 10; 100; 200 (0.1;...;**200**)	CHF bn.	200
COVERAGE[b]	1; 20; 50; 100 (**1**;...;100)	percent	1

[a] Values in US$ bn: 0.065; 6.5; 65; 130 (at 2002 exchange rates)

[b] Coverage in percent of loss

[c] Bold for status quo

Table 3.1: Levels of attributes.

3.3 The Stated Choice Experiment

3.3.1 Methodology

In stated choice experiments (SCE), respondents are confronted with hypothetical choice situations where they have to decide whether they prefer the status quo or some alternative product that is allowed to differ in all product attributes. For each such choice set, respondents have to indicate their preferred choice, which requires them to trade off one set of attributes against the other, implicitly revealing their preferences regarding the different attributes.

The SCE alternative started with McFadden (1974) and was further developed by Louviere and Hensher (1982). More recently, it has become popular in energy and health economics (Johnson and Desvousges, 1997; Johnson et al., 1998; Telser and Zweifel, 2002). In the context of nuclear energy risks, SCE have been found to yield qualitatively and quantitatively plausible results (Schneider and Zweifel, 2004). However, that study neglected the spatial dimension of risk associated both with the operation and waste disposal of nuclear plants.

3.3.2 Experimental Design

In the present context, the 'stated choice' method allows individuals to choose among different types of electricity. During the decision process, the attributes (among them price) of electricity are traded off against each other. Participants in the experiment are asked to pairwise evaluate several different electricity products by indicating their preferred choice. By observing a number of choices, it is possible to approximate an indifference curve in attribute space and therefore estimate

	Description	**Mean**	**Median**	**Std.**
DISTANCE	distance in kilometers from respondents residence to nearest nuclear power plant	45	36	30
PESSIMIST	= 1 if respondent considered a nuclear accident at least ten times more probable than experts	0.59	1	0.49
OPPONENT	= 1 if respondent said to be against nuclear energy even if there was no waste disposal problem	0.21	0	0.41
OWNER	= 1 if respondent owns his or her dwelling	0.35	0	0.48
SEXM	= 1 if respondent is male, = 0 if female	0.52	1	0.5
INCOME	yearly income in CHF. Seven income categories were used in the questionnaire. 44 percent did not reveal their income	47,500	60,000	35,400
INC_MISSG	= 1 if income was missing	0.44	0	0.5

Table 3.2: Sample description of explanatory variables.

how much income (through higher electricity prices) respondents are prepared to give up in return for an increased amount of some other desired attribute.

For a SCE, it is necessary to define the product under consideration (here: electricity) by but a few relevant attributes. In a telephone survey preceding the main survey, 500 Swiss residents were asked to indicate how important they considered several electricity attributes. The following five emerged as the most important: size of area exposed to hazard (DAMAGE), secure and sustainable waste disposal (NOWASTE), reliability (BLACKOUT, low frequency of blackouts), financial compensation of victims in case of an accident (COVERAGE), and average price per kwh (PRICE). Since the study is concerned with insurance against financial risks of a nuclear accident, DAMAGE was defined as billions of CHF at risk rather than area exposed to hazard. The relevant attributes are summarized in Table 3.1.

The questionnaire for the main survey was divided in three parts: warm-up questions, the actual choice experiment, and socioeconomic information. In the first part, data on monthly electricity outlay, attitudes towards nuclear energy, and the importance of choice between different types of electricity was collected. Respondents then had to read a description of the risks of nuclear and hydro power plants (see the appendix for exact wording). Emphasis was put on possible worstcase scenarios and their financial consequences. Respondents were also told that nuclear power plants were already mandated to have liability insurance but that coverage fell far short of possible financial loss in case of a major accident. The government would possibly provide relief by imposing a special tax. Alternatively, mandated insurance coverage could be stepped up to reduce reliance on the tax system.

The second part of the questionnaire consisted of the actual DCE. Respondents were confronted with 14 different choice situations where they had to decide whether they preferred a proposed type of power to the status quo. Note that respondents could always opt out by stating "cannot decide". In the third and last part of the questionnaire, standard socioeconomic data was collected, summarized in Table 3.2. Specifically, DISTANCE from the nearest nuclear power plant was calculated using Zip codes provided by respondents. PESSIMIST=1 obtains if on a visual analog scale, respondents marked their estimated accident probability at least one order of magnitude higher than experts.

Face-to-face interviews were performed in the German-speaking part of Switzerland during September and October 2001 (in the aftermath of 9/11). In total, 391 persons were interviewed: Each respondent evaluated 14 choice scenarios, resulting in 5,474 recorded decisions. After excluding "cannot decide" answers and missing values, a total of 4,613 observations were retained.

3.4 Expected Effects of Distance on Willingness-to-Pay

3.4.1 The Confounding Effect of Locational Choice

Given the possibility of choosing residential location according to attitudes towards nuclear power plants, one would expect respondents to be sorted according to distance to nuclear power plants, with the more skeptical types to be found farther away from the plant. Therefore, people located farther away from the plant might be willing to pay more for additional insurance coverage than those located in the vicinity of the plant. Whether the distance gradient of WTP for risk reduction is positive or negative thus depends on the degree of sorting that took place. For without sorting, living farther away from the plant does serve as self-insurance, reducing the risk of radiation and causing a lowered WTP for coverage of financial loss.

Since the cost of relocating is not trivial, sorting is expected to be less than perfect. Thus, the combined effect of self-insurance and sorting is ambiguous. However, to the extent that estimation succeeds in controlling for attitude and perceived risk, the distance gradient should be negative, i.e. WTP for additional insurance coverage should decrease with distance from plants.

Three indicators for respondents' attitude towards nuclear power plant are used in the estimation, (1) whether they perceive nuclear accidents to be much more likely than experts' best estimates (PESSIMIST), (2) whether they state to be opposed to nuclear energy in principle (OPPONENT), and (3) male sex (SEXM). The last indicator is based on several studies. Hartog et al. (2002) analyze the influence of individual characteristics on risk aversion and find survey evidence that men are less risk-averse than women. Nielsen et al. (2003) present evidence that men have a lower perception of risks than women. SEXM is therefore predicted to display a negative relationship with WTP. Furthermore, to the extent that locational decisions concerning residence more strongly reflect main breadwinners' preferences (who are still predominantly men in Switzerland), interacting SEXM with DISTANCE controls for sorting. Thus, the net effect of DISTANCE on WTP is predicted to be less strongly negative for men than women.

3.4.2 The Difference Between Radiation and Waste Disposal Risks

The risk of nuclear waste disposal differs importantly from that of plant operation. In view of the fact that a national site for nuclear waste disposal has not been designated yet, putting more distance between one's residence and a nuclear plant has little effect. Therefore, DISTANCE should not be a relevant predictor of WTP for solving the problem of nuclear waste (WTP_W). This statement needs to be qualified in the following way. For geological reasons, a future disposal site is unlikely to be near existing plants (which are all located on rivers). Therefore, being located at a great DISTANCE to the plant does not protect individuals from the risk associated with the shipping and disposal of nuclear waste. Yet transports of radioactive waste and spent fuel necessarily originate from plants, from where they will be directed to whatever national disposal site will be chosen (at present, destinations are Le Hague in France and Sellafield in Great Britain). This implies that there is and will be an increased exposure to the risk of nuclear waste in the vicinity of the plant. For most values of DISTANCE, however, the effect of distance on WTP_W is expected to be zero.

3.4.3 Hypotheses to Be Tested

The preceding arguments may be summed up as follows. (1) Marginal WTP for higher liability insurance coverage (MWP_C) may decrease or increase with distance from the nearest nuclear plant, depending on whether the direct effect of distance (risk effect) or the sorting effect prevails. (2) By controlling for respondents' attitude, parts of the pure risk effect can be estimated and is expected to be negative. (3) WTP for solving the waste problem (WTP_W) is predicted not to depend on distance. (4) Ceteris paribus, increasing values of MWP_C and WPT_W are expected with higher income since the marginal utility loss caused by an increase in the price of power should be decreasing in income. Since some 40 percent of respondents refused to indicate their income, restricting the sample to those individuals with information on income has to be avoided. The solution retained is to equate missing values to zero (INCOME = 0) while creating a dummy variable INC_MISSG that takes on the value of one if income information is not available. Interaction terms are limited to OUTLAY; this is sufficient to represent differences in marginal utility of income.

3.5 Econometric Specification

An individual may be confronted with a discrete choice, e.g. whether to buy a certain product or not. Given this choice, individuals maximize their utility with respect to their budget constraints, obtaining certain utility values. These values define an indirect utility function (conditional on the alternative selected) that depends on individuals' characteristics, their incomes, on the particular attributes of the alternative (including price) as well as on various unobservable and therefore random effects (random utility specification, see McFadden, 2001).

In the present stated choice experiment (SCE), respondents were confronted with 14 binary choice situations, involving the status quo and an alternative. The dependent variable y_i equals to one if respondents chose the alternative and zero if they stayed with the status quo. Respondent i's utility of the alternative in choice situation j is denoted by V_{ij}; the one of the status quo, by V_{mj}. Respondents therefore chose the alternative ($y_i = 1$) if $V_{ij} - V_{im} \geq 0$. The utility function to be estimated thus reads

$$\begin{aligned} V_{ij} = \beta_0 + \beta_1 \cdot \text{COVERAGE}_j + \beta_2 \cdot \text{NOWASTE}_j + \beta_3 \cdot \text{BLACKOUT}_j + \beta_4 \cdot \text{DAMAGE}_j \\ + \beta_5 \cdot \text{OUTLAY}_j + \beta_6 \cdot \text{OUTLAY}_j^2 + \beta_7 \cdot \text{DIST}_i \cdot \text{COVERAGE}_j + \beta_8 \cdot \text{DIST}_i \cdot \text{NOWASTE}_j \\ + \beta_9 \cdot \text{DIST}_i \cdot \text{DAMAGE}_j + \beta_{13} \cdot \text{DIST}_i \cdot \text{PESS}_i \cdot \text{COVERAGE}_j + \beta_{34} \cdot \text{INCOME}_i \cdot \text{OUTLAY}_j \\ + \ldots + \epsilon_{ij}, \end{aligned} \quad (3.1)$$

where $\epsilon_{ij} = \mu_i + \nu_{ij}$ has a two-way random specification to account for the fact that the same individual i decides in the several situations j. Note that $\text{INCOME}_i \cdot \text{OUTLAY}_j$ and OUTLAY_j^2 permit marginal utility of income to vary with income (since $\Delta\text{OUTLAY}_j = -\Delta\text{INCOME}_i$). Since only differences $V_{ij} - V_{im}$ are relevant for an individual's decision, regressors such as OUTLAY_j are measured as differences from the status quo, causing socioeconomic variables to drop out of the equation unless interacted with regressors that vary between situations, at least under the random utility specification. The random effects probit model was estimated using maximum likelihood. The estimated utility function, , permits to calculate marginal WTP for the different product attributes, defined as the marginal utility of the attribute divided by the marginal utility of income,

$$MWP_i(\text{COVERAGE}) := \frac{\partial \hat{V}_i / \partial \text{COVERAGE}}{\partial \hat{V}_i / \partial \text{OUTLAY}} \quad (3.2)$$

In the case of WTP for solving the nuclear waste problem, one has

$$WP_i(\text{NOWASTE}) := \frac{\hat{V}_i[\text{NOWASTE} = 1] - \hat{V}_i[\text{NOWASTE} = 0]}{\partial \hat{V}_i / \partial \text{OUTLAY}}. \quad (3.3)$$

MWP values reported are in US$ per year. Using the delta method (see Greene, 2003, p.70), standard errors of MWP are derived using the standard errors of the parameters contained in the estimated utility function.

3.6 Results

Selected estimation results are displayed in Table 3.3 (full results are relegated to the appendix). All coefficients of product attributes (COVERAGE, NOWASTE, BLACKOUT, OUTLAY, OUTLAY2) with the exception of DAMAGE show the expected sign and are highly significant, indicating that respondents were (on average) willing to

	Sign[a)]		**Coefficient**		**Std.Err.**	**z**
COVERAGE	+	β_1	0.009887	***	0.002264	4.37
NOWASTE	+	β_2	0.516774	***	0.168756	3.06
BLACKOUT	-	β_3	-0.362759	***	0.049799	-7.28
DAMAGE	-	β_4	0.001445		0.000939	1.54
OUTLAY	-	β_5	-0.003341	***	0.000303	-11.04
OUTLAY2	+	β_6	1.19E-07	***	0.000000	11.06
DIST·COVERAGE	-	β_7	-1.03E-04	**	0.000042	-2.47
DIST·NOWASTE	0	β_8	-0.001138		0.003068	-0.37
DIST·SEXM·COVERAGE	?	β_{10}	1.54E-04	***	0.000044	3.47
DIST·SEXM·NOWASTE	0	β_{11}	0.005323		0.003347	1.59
DIST·PESS·COVERAGE	+	β_{13}	1.22E-04	**	0.000049	2.49
DIST·PESS·NOWASTE	0	β_{14}	-0.0042		0.003589	-1.17
DIST·OPP·COVERAGE	+	β_{16}	-7.70E-05		0.00006	-1.29
DIST·OPP·NOWASTE	0	β_{17}	0.002981		0.00452	0.66
SEXM·COVERAGE	-	β_{22}	-0.008248	***	0.002285	-3.61
SEXM·NOWASTE	-	β_{23}	-0.427121	**	0.173976	-2.46
PESS·COVERAGE	+	β_{25}	-0.005563	**	0.002417	-2.3
PESS·NOWASTE	?	β_{26}	0.285538		0.181611	1.57
OPP·COVERAGE	+	β_{28}	0.00336		0.003009	1.12
OPP·NOWASTE	+	β_{29}	0.273641		0.224882	1.22
INCOME·OUTLAY	+	β_{34}	7.19E-09	***	0.000000	2.52
Log likelihood = -2,178.2026, N=4,613, 376 respondents						

a) Theoreticall y expected sign.

Table 3.3: Selected estimation results. Dependent variable is the probability of accepting the alternative type of power. Estimation results for the full equation (containing 37 rather than 21 explanatory variables) are given in the appendix.

make tradeoffs among the different attributes. Furthermore, the relative magnitudes of marginal utilities associated with product attributes are intuitively plausible. Note that COVERAGE measures the increase in insurance coverage in percentage points, whereas NOWASTE is an all-or-nothing variable indicating whether or not there are any problems regarding nuclear waste. Multiplying the coefficient of COVERAGE (0.001) by 100 for making it roughly comparable to that of NOWASTE (0.52), one obtains 1.00, suggesting that the two attributes are valued similarly. Moreover, there are first indications that the hypotheses formulated in section 3.4.3 may be confirmed. (1) While DIST·COVERAGE has a negative coefficient, the one of DIST·PESS·COVERAGE is positive, turning the overall effect of DISTANCE around. (2) The effect of DISTANCE interacted with COVERAGE again changes sign when further interacted with OPPONENT. (3) The coefficient of DIST·NOWASTE is insignificant, in contradistinction to that of DIST·COVERAGE. (4) Higher income mitigates the disutility caused by higher outlay on electricity, pointing to diminishing marginal utility of income.

Therefore, respondents are not only concerned about the risks associated with nuclear energy (COVERAGE, NOWASTE) but also about the frequency of power outages (BLACKOUT) and about the costs of electricity (OUTLAY and OUTLAY2), with the positive coefficient of OUTLAY2 pointing to a diminishing marginal disutility of loss of income and hence decreasing marginal utility of income.

3.6.1 The Effect of Attitudinal Variables on WTP

Using eq. (3.2), MWP is evaluated for different values of SEXM and PESSIMIST while keeping the remaining variables at their median values. The results in Table 3.4 reproduce the well-known fact of women being more concerned with the well-being of future generations than men. This was already borne out by the negative coefficient of SEXM·COVERAGE in Table 3.3. Men (SEXM=1) do not value both additional insurance coverage and solving the waste disposal problem as much as women (SEXM=0). WTP values reported in Table 3.4 confirm this finding. Pessimistic women are willing to pay more than twice as much as comparable men for a marginal increase in insurance coverage (1.47 US$/year compared to 0.62 US$/year) and roughly 50 percent more than men for solving the waste disposal problem (182 US$/year compared to 111 US$/year). The relative differential is less for non-pessimistic women w.r.t. coverage, viz. some 85 percent, but the same (100 percent) w.r.t. waste disposal. This may be surprising at first sight.

	Value	s.e.	z
$MWP_C^{*)}$			
pessimistic men	0.6198	0.3660	1.6900
pessimistic women	1.4677	0.4031	3.6400
non-pessimistic men	1.0357	0.3706	2.8000
non-pessimistic women	1.8837	0.4085	4.6100
WTP_W			
pessimistic men	110.6924	27.8928	3.9700
pessimistic women	181.9186	31.7356	5.7300
non-pessimistic men	71.5003	28.5781	2.5000
non-pessimistic women	142.7264	31.0601	4.6000

*) percentage point, e.g. from 1 to 2 percent of maximum loss

Table 3.4: Marginal willingness-to-pay for increased coverage (MWP_C) and for solving the waste disposal problem (WTP_W) evaluated at median distance (35km) in US$ per year.

Respondents who perceive a nuclear accident to be at least ten times more probable than experts (PESSIMIST=1) are expected to be willing to pay more for coverage than non-pessimistic individuals. In fact, PESSIMIST is the one item of the questionnaire permitting respondents who do not want to be openly against nuclear power (OPPONENT=1, chosen by only 21 percent, see Table3.2) to express their skepticism. However, skeptical individuals likely are more concerned about the non-financial risks associated with nuclear power such as the disposal of radiating waste. The results in Table 3.4 thus suggest that PESSIMIST more likely serves as an indicator of individuals' general attitude towards nuclear energy rather than of their perceived accident probability. Therefore, it makes sense that their WTP for additional coverage should be somewhat lower than that of non-pessimistic respondents, while their WTP for the solution of the waste problem is much higher.

3.6.2 The Effect of Distance on WTP

Returning to Table 3.3 for estimation results, one can see that the hypotheses formulated in 3.4.3 are confirmed. In keeping with hypothesis (1), the coefficient of DIST·COVERAGE, reflecting the risk effect, is significantly negative. By way of contrast, the coefficient of DIST·NOWASTE lacks significance, suggesting that WTP for solving the waste problem does not depend on distance from the nearest nuclear plant, as predicted by hypothesis (3).

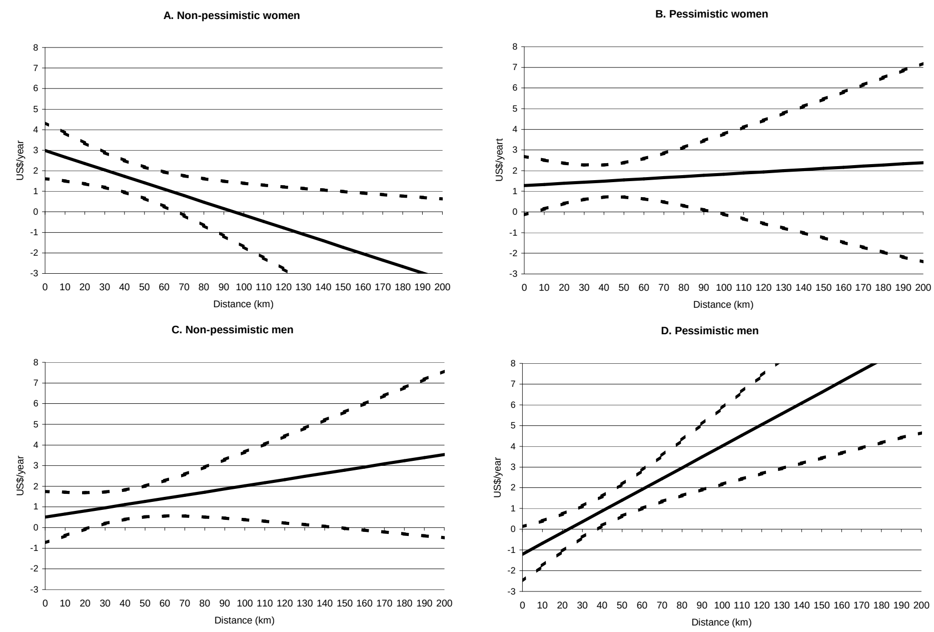


Figure 3.1: Trading off different product attributes.

Prediction (2) states that these rather clear-cut results are due to the fact that the effect of sorting is controlled for by attitudinal variables in the regression. This in turn means that pessimists should exhibit a smaller distance gradient w.r.t. MWP for coverage. And indeed, DIST·PESS·COVERAGE has a positive coefficient (+1.22E-4), which even more than counterbalances the negative one of DIST·COVERAGE (-1.03E-4). The same effect can be observed for men (who are hypothesized to decide about residential location in a family), for DIST·SEXM·COVERAGE also has a countervailing positive coefficient, amounting to 1.54E-4. And while DIST·OPP·COVERAGE has a negative partial effect (-0.77E-4), it is not significant. However, hypothesis (3) also states that sorting should not make a difference w.r.t. WTP for having the waste problem resolved. Indeed, DIST·SEXM·NOWASTE, DIST·PESS·NOWASTE, and DIST·OPP·NOWASTE all fail to attain statistical significance. Finally, prediction (4) states that marginal WTP for safety is an increasing function of income because of decreasing marginal utility of income. In Table 3.3, the negative coefficient of OUTLAY (and the much smaller coefficient of OUTLAY2) indicate that the probability of accepting the alternative to the status quo (and hence utility) decreases because of the reduction in disposable income caused by outlay on electricity. However, the positive coefficient of INCOME·OUTLAY means that this loss of utility weighs less heavily when income is high. This certainly constitutes a sign of decreasing marginal utility derived from income. Moreover, the fact that INC_MISSG · OUTLAY lacks significance (see bottom of the table in the appendix) suggests that the 40 percent of respondents who refused to provide information on their income (see section 3.4.3) do not differ systematically from the others w.r.t. their marginal utility. Thus, hypothesis (4) is confirmed as well.

Since the indicators SEXM, PESSIMIST, and OPPONENT are designed to capture the sorting effect, the WTP of non-pessimistic women not opposed to nuclear energy (SEXM=0, PESSIMIST=0, OPPONENT=0) represents the pure risk effect. By hypothesis (2) of section 3.4.3, their WTP for more comprehensive insurance coverage should be decreasing in distance from plant. Panel A of Figure 3.1 shows that non-pessimistic women who do not oppose nuclear energy exhibit positive MWP for coverage at small distances from plant. However, MWP_C becomes indistinguishable from zero (at the 5 percent significance level) at 65 kilometers, turning negative farther away.

Pessimistic women, by way of contrast, may be claimed to have sorted themselves away from nuclear plants, resulting in a slightly positive distance gradient of MWP (which however itself is significantly positive only between 0 and 85 kilometers, see panel B of Figure 3.1). Among pessimistic men, this positive gradient is far more marked (panel D), in accordance with the view that they are the ones who decide about residential location. Finally, non-pessimistic men (panel C) exhibit a much weaker positive gradient than the pessimists of panel D, combined with positive WTP for coverage between distance of 20 and 135 kilometers from plant. The two groups seem to differ in terms of their sorting w.r.t. distance, in accordance with hypothesis (1) of section 3.4.3.

By way of contrast, the distance gradients of WTP for solving the waste disposal problems are flat in all cases (not shown). Evaluating WTP_W it at median sample values (which includes a remaining life expectancy of some 44 years for women and discounting at 15 percent), one obtains a lifetime WTP_W of \$1,087 for a distance of 35 kilometers. The hundredfold of MWP for insurance coverage also corresponds to the full solution of a problem, this time the one of financial risk associated with operation of the plant. In this case, lifetime WTP at a distance of 35 km amounts to \$1,439, suggesting that (full) financial coverage is valued higher than solving the waste disposal problem, at least by the Swiss population. For residents located at the power plant, lifetime WTP is maximum at \$2,280. It decreases by $24 per km, or $15 per mile, much less than the $200 to 300 per mile reported by Faber (1998) for the United States.

3.7 Conclusion

The objective of this paper is to analyze the effect of distance from nuclear plant on the WTP for a reduction of two types of risk emanating from these plants, using survey evidence. In the case of Switzerland, respondents had ample opportunity to choose their residential location according to their preferences regarding nuclear power. In the case of radioactive risk associated with the operation of a nuclear plant, this causes distance to play an ambiguous role. If spatial sorting of individuals is indeed important, one would expect to find more strongly concerned people residing at a greater distance from plants. This could result in a positive rather than negative distance gradient in their marginal WTP for risk reduction. In the case of nuclear waste disposal, however, distance from plant is predicted to be irrelevant as long as the final disposal site is not decided (as in Switzerland).

In a Stated Choice Experiment, with statistical inference based on the Random Utility Model, the attributes of electric power (degree of coverage by nuclear liability insurance, solution of the waste problem, but also number of blackouts, size of damage, and price of electricity) are found to be valued as hypothesized. More importantly, DISTANCE proves to be a significant predictor of marginal WTP for insurance coverage but not of WTP for having the waste disposal problem solved. Controlling for attitudes towards nuclear energy and nuclear sorting in space, the distance gradient turns out to be significantly negative with regard to marginal WTP for increased insurance coverage. Starting with WTP for full insurance coverage of $2,280 at zero distance from nuclear power plants, WTP decreases by $24 per km [$15 per mile, compared to $200 to 300 according to Faber (1998) for the United States] and eventually falls to zero at a distance of 95 km.

In sum, this research suggests that distance from an environmental disamenity may have unexpected effects on WTP for risk reduction. Data on housing prices, being contaminated by regional supply shift effects, are unlikely to permit discovering the demand effects caused by the sorting in space performed by individuals when choosing their residential location.

3.8 Appendix

Full estimation results. Dependent variable is the probability of accepting the alternative type of power.

		Coefficient	Std.Err.	z
COVERAGE	β_1	0.009887	0.002264	4.37
NOWASTE	β_2	0.516774	0.168756	3.06
BLACKOUT	β_3	-0.362759	0.049799	-7.28
DAMAGE	β_4	0.001445	0.000939	1.54
OUTLAY	β_5	-0.003341	0.000303	-11.04
OUTLAY2	β_6	1.19E-07	0.000000	11.06
DIST·COVERAGE	β_7	-0.000103	0.000042	-2.47
DIST·NOWASTE	β_8	-0.001138	0.003068	-0.37
DIST·DAM	β_9	-0.00005	0.000016	-3.06
DIST·sexm·COVERAGE	β_{10}	0.000154	0.000044	3.47
DIST·SEXM·NOWASTE	β_{11}	0.005323	0.003347	1.59
DIST·SEXM·DAMAGE	β_{12}	0.000038	0.000018	2.1
DIST·PESS·COVERAGE	β_{13}	0.000122	0.000049	2.49
DIST·PESS·NOWASTE	β_{14}	-0.0042	0.003589	-1.17
DIST·PESS·DAMAGE	β_{15}	-0.000029	0.000019	-1.47
DIST·OPP·COVERAGE	β_{16}	-0.000077	0.00006	-1.29
DIST·OPP·NOWASTE	β_{17}	0.002981	0.00452	0.66
DIST·OPP·DAMAGE	β_{18}	0.000092	0.000024	3.89
DIST·OWN·COVERAGE	β_{19}	-0.000055	0.00005	-1.11
DIST·OWN·NOWASTE	β_{20}	0.006622	0.003738	1.77
DIST·OWN·DAMAGEe	β_{21}	0.000031	0.00002	1.51
SEXM·COVERAGE	β_{22}	-0.008248	0.002285	-3.61
SEXM·NOWASTE	β_{23}	-0.427121	0.173976	-2.46
SEXM·DAMAGE	β_{24}	-0.000822	0.00095	-0.86
PESS·COVERAGE	β_{25}	-0.005563	0.002417	-2.3
PESS·NOWASTE	β_{26}	0.285538	0.181611	1.57
PESS·DAMAGE	β_{27}	-0.000379	0.000987	-0.38
OPP·COVERAGE	β_{28}	0.00336	0.003009	1.12
OPP·NOWASTE	β_{29}	0.273641	0.224882	1.22
OPP·DAMAGE	β_{30}	-0.003633	0.001204	-3.02
OWNER·COVERAGE	β_{31}	0.003946	0.002528	1.56
OWNER·NOWASTE	β_{32}	-0.167968	0.19344	-0.87
OWNER·DAMAGE	β_{33}	-0.002436	0.001071	-2.27
INCOME·OUTLAY	β_{34}	7.19E-09	0.000000	2.52
INC_MISSG·OUTLAY	β_{35}	0.000601	0.000328	1.83
const	β_0	0.659389	0.08581	7.68
Log likelihood = -2,178.2026, N=4,613, 376 respondents				

Bibliography

Clark, D. E. and T. Allison (1999). Spent nuclear fuel and residential property values: The influence of proximity, visual clues and public information. *Papers in Regional Science 78*, 703–421.

Davis, L. W. (2005). The effect of health risk on housing values: Evidence from a cancer cluster. *American Economic Review 94*, 1693–1704.

Dionne, G. and L. Eeckhoudt (1985). Self-insurance, self-protection and increased risk aversion. *Economic Letters 17*, 39–42.

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. *The Journal of Political Economy 80*, 623–648.

Faber, S. (1998). Undesirable facilities and property values: A summary of empirical studies. *Ecological Economics 24*, 1–14.

Folland, S. and R. Hough (2000). Externalities of nuclear power plants: Further evidence. *Journal of Regional Science 40*, 735–753.

Gamble, H. B. and R. H. Downing (1982). Effects of nuclear power plants on residential property values. *Journal of Regional Science 22*, 457–478.

Gawande, K. and H. Jenkins-Smith (2001). Nuclear waste transport and residential property values: Estimating the effects of perceived risks. *Journal of Environmental Economics and Management 42*, 207–233.

Greene, W. H. (2003). *Econometric Analysis* (Fifth Edition ed.). Prentice Hall International.

Hartog, J., A. F. i Carbonell, and N. Jonker (2002). Linking measured risk aversion to individual characteristics. *Kyklos 55*, 3–26.

Johnson, R. F. and W. H. Desvousges (1997). Estimating stated preferences with rated-pair data: Environmental, health, and employment effects of energy programs. *Journal of Environmental Economics and Management 34*, 79–99.

Johnson, R. F., M. C. Ruby, and W. H. Desvousges (1998). Using stated preference measurement and health-state classification to estimate the value of health effects of air pollution. *Working Paper T-9807, Triangle Economic Research*.

Kelly, M. and A. E. Kleffner (2003). Optimal loss mitigation and contract design. *The Journal of Risk and Insurance 70*, 53–72.

Louviere, J. J. and D. A. Hensher (1982). On the design and analysis of simulated or allocation experiments in travel choice modelling. *Transportation Research Record 890*, 11–17.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics 3*, 303–28.

McFadden, D. (2001). Economic choices. *American Economic Review 91*, 351–378.

Nelson, J. P. (1981). Three mile island and residential property values: Empirical analysis and policy implications. *Land Economics 57*, 363–372.

Nielsen, J. B., D. Gyrd-Hansen, I. S. Kristiansen, and J. Nexoe (2003). Impact of socio-demographic factors on willingness to pay for the reduction of future health risk. *Journal of Environmental Plannign and Management 46*, 39–47.

Riddel, M., C. Dwyer, and D. W. Shaw (2003). Environmental risk and uncertainty: Insights from yucca mountain. *Journal of Regional Science 43*, 435–457.

Schneider, Y. and P. Zweifel (2004). How much internalization of risk through liability insurance? *Journal of Risk and Uncertainty 29*, 219–240.

Smith, K. V. and W. H. Desvousges (1986). The value of avoiding a lulu: Hazardous waste disposal sites. *The Review of Economics and Statistics 68*, 293–299.

Telser, H. and P. Zweifel (2002). Measuring willingness-to-pay for risk reduction: An application of conjoint analysis. *Health Economics 11*, 129–139.

Switching Costs, Firm Size, and Market Structure

Simon Loertscher and Yves Schneider *

*Loertscher: University of Bern, Economics Department, Schanzeneckstrasse 1, CH-3012 Bern Email: simon.loertscher@vwi.unibe.ch. Schneider: Socioeconomic Institute, University of Zurich, Hottingerstrasse 10, CH-8032 Zurich. Email: yschneider@soi.unizh.ch. We want to thank Michael Breuer, Winand Emonds, Edna Epelbaum, Martin Everts, Sergei Koulayev, Armin Schmutzler, Peter Zweifel, and seminar participants at SSES 2005 in Zurich and VfS 2005 in Bonn for most valuable comments and discussions. The first author acknowledges financial support by the Swiss National Science Foundation via grant PBBE1-103015. Any errors and omissions are ours.

Chapter 4

Switching Costs, Firm Size, and Market Structure

4.1 Introduction

A bus trip from New York City to Boston is a fairly homogenous good. It takes about four hours and twenty minutes and costs US$55 at the Greyhound/Peter Pan desk and US$15 at the Fung-Wah desk.[1] Similarly, a big cup of milk coffee in the Big Cup Café on 8th Avenue in Manhattan costs US$3.60, while the largest cup of café latte in the Starbucks café on the other side of the avenue is sold at US$3.95.[2] Most strikingly perhaps, the airfare for a return flight from Berlin to Cologne-Bonn costs Euro 395 if one flies with Lufthansa and Euro 53 if one travels with German Wings.[3] What is the common feature of these three pricing patterns? First, arguably homogenous goods are sold at, sometimes substantially, different prices. Second, one of the sellers is a large firm that is more or less globally active and known by almost every potential consumer, while the other seller is a small

[1]Prices are as of May 2005. If one buys the tickets online, one pays US$28-35 with Greyhound/Peter Pan and US$15 with Fung-Wah. Greyhound/Peter Pan trips begin in Midtown Manhattan on 42nd street, while Fung-Wah trips start in Chinatown in Manhattan on Canal street. Both trips end at Boston South station.

[2]Both cafés are between 21st and 22nd street. Prices are as of spring 2005.

[3]Sources: www.lufthansa.de and www.germanwings.com. We choose return flights because these are cheaper than one-way tickets for major carriers such as Lufthansa. The price of the German Wings return ticket is the sum of two one-way tickets. The date of booking was July 21, 2005. Lufthansa's airport in Berlin is Tegel, while German Wings flies from and to Berlin Schönefeld. For an outbound flight from Berlin to Cologne-Bonn, we arbitrarily chose July 28 round 8 a.m. For the return flight, we chose August, 1, round 7 p.m. Though the price differences vary as a function of various factors such as date and flexibility, there can be little doubt that German Wings is substantially cheaper than Lufthansa. Moreover, the fact that German Wings is a partner of Lufthansa does not refute that the two carriers set different prices and face different demand functions.

local firm that is most probably only known by customers familiar with the locality. Third, the large firm charges the high price.

The purpose of this paper is to provide a parsimonious model that explains pricing patterns such as these. As the larger firms sell at higher prices, it is clear from the outset that economies of scale *cannot* explain these patterns. What seems to be at work here is a non-convexity in the consumption technology. Potential customers of local firms must first learn about the existence of the local provider. Once they know this, they have to experiment whether the goods and services provided by the local store suit their preferences, and eventually have to learn how to best consume these. If this type of search and experimentation is costly, buying from a new provider involves set-up costs. Thus, these set-up costs are a kind of switching costs.

Of course, the same is true for new customers of global firms, or chain stores, as we call them. The twist, though, is that if customers are mobile and consume repeatedly, they have to incur the set-up cost only once when buying from the chain store, whereas these costs have to be borne each time they buy from another local store. Moving from one location to the other with an exogenous probability, consumers cannot always buy from the same local firm. Consequently, they risk to incur the set-up costs anew when first buying from a local firm. Since chain stores help consumers save switching costs, they may be able to charge higher prices and yet to attract more customers than do local firms.

Put in a nutshell, this is the explanation our paper suggests. So as to back up the intuition just outlined, we develop the following model. Consumers are located in two identical cities and live for two periods. With a given probability, they move from one city to the other after the first period. In each period, they can either buy from a local store or from a chain store, which both supply the same good. Before they buy from a given store for the first time, each consumer incurs a set-up cost. While all consumers value the good in the same way and face the same probability of moving, they are heterogenous with respect to the switching costs. We show that in the unique equilibrium both types of stores are active. The chain store charges a higher price and attracts more consumers than do local stores. Low switching cost consumers buy from the local stores and high switching costs consumers buy from the chain store. Moreover, the relative profitability of the chain store increases as consumers become more mobile.

The remainder of the paper is structured as follows. The next section relates the paper to the existing literature. Section 4.3 introduces the model. Section 4.4 analyzes the benchmark case with two local monopolies, while section 4.5 derives the unique equilibrium for the market structure with two local stores competing with a chain. Section 4.6 then shows that the market structure with a local store in each city and a chain store active in both cities is the unique stable market structure if there is a small, positive entry cost. Section 4.7 concludes.

4.2 Related Literature

To the best of our knowledge, the idea that larger firms may gain more customers while charging higher prices than smaller firms merely because of consumers' switching costs has not been fully recognized in the previous formal literature. For example, Stahl (1982) notes that a merger of local stores to a chain store "appears exclusively connected to the input side of the retailing activity, that is, to the exhaustion of economies of scale in purchasing and distributing inputs."

Switching costs as understood in this paper are a short-cut to search and experimentation costs à la Nelson (1970), where consumers have to search and experiment so as to find their most preferred good. Insofar as our model does not allow for dynamic price competition, it is in some contrast to a part of the switching cost literature. For example, Klemperer (1987, 1995)'s major concern is with the dynamic aspects of price competition when consumers are locked in with their supplier due to switching costs, so that sellers are tempted to use 'bargains followed by ripoffs'- pricing schemes (Farrell and Klemperer, 2004). However, our approach is perfectly in line with von Weizsäcker (1984), whom we follow by assuming that firms do not set different prices over time.

Two papers that deal with search costs but are not concerned with switching costs are Stahl (1982) and Wolinsky (1983). Stahl illustrates how a model of demand externalities creates a similar agglomeration effect. Wolinsky presents a model where imperfect information creates the need to search for a suitable buy, leading firms to cluster at one location in order to reduce search costs.

Baye and Morgan (2001) provide a model with equilibrium dispersion of publicized prices, which arise because some consumers decide not to bear the cost required to become informed about prices.[4] Insofar as in our model high switching cost consumers prefer paying higher prices to bearing the switching cost, this is

[4] For empirical evidence, see Baye et al. (2004).

very similar to our model. In the model of Baye and Morgan, though, the high and low priced firms are not determined ex ante because the price dispersion stems from a mixed strategy equilibrium. Consequently, in their model firm size does not matter.

Because chain stores are physically differentiated from local stores in that they are active in more locations than local stores, the paper also relates to the product differentiation literature initiated by Hotelling (1929). Janssen et al. (2003) study competition between two firms with multiple outlets (chain stores) on the Salop (1979) circle, where firms sell differentiated products to heterogenous consumers. In their model, outlets from the same chain are homogenous but outlets across chains are heterogenous. Whereas Janssen et al. (2003) are concerned with location and pricing decisions of two chains, we are interested in the effect of homogeneity of outlets from the same chain on consumer choice if alternatively they can buy from heterogenous single outlet firms.

Aside from explaining the above mentioned price patterns, our model also provides a simple explanation for the remarkable asymmetry in firms size as observed, e.g., in the retail and hotel industries because our model predicts the local stores' market share is at most one third in equilibrium.[5] For alternative explanations for such asymmetries, see, e.g., Besanko and Doraszelski (2004), Athey and Schmutzler (2001), Bagwell et al. (1997) and Hausman and Leibtag (2004).

4.3 The Model

There are two cities E (East) and W (West), each hosting one unit of risk neutral consumers. Firms sell a homogenous product and each consumer is assumed to bear exogenously given switching costs $s \in [0, \sigma]$ prior to buying the good for the first time in any given type of store. The timing of events is as illustrated in Figure 4.1. After firms choose their prices at date zero, each consumer observes the prices in his home city. Consumers then decide at $t = 1$ from which firm to buy the good. At the intermediate stage, each consumer moves to the other city with an exogenously given probability $\alpha \in (0, 1)$. Throughout it is assumed that consumers and firms know this probability but that ex ante neither firms nor consumers know

[5] For retailing, see, e.g., Bagwell and Ramey (1994), Bagwell et al. (1997), Dinlersoz (2004) or www.stores.org. According to the last source the sales of Wal-Mart, the largest retailer in the U.S., were approximately four times as large as those of the second ranked Home Depot in 2003. For the hotel industry, Michael and Moore (1995) report that 39 percent of all sales are accounted for by franchise chains.

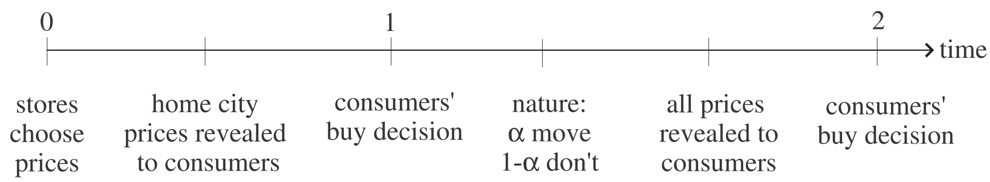


Figure 4.1: Time line.

whether a particular consumer will move between period one and two. Consumers who move to the other city then learn the prices prevalent in this city.[6] At date two, regardless of whether the consumer still is in the same city or not, he decides again from which firm to buy the good or whether not to buy at all. For simplicity, there is no discounting of future payoffs.

Though for the purpose of a consistent illustration, we stick to the literal two-city interpretation, the basic framework also applies to many other situations. For example, if consumers commute within a metropolitan area, customers of retailers will face problems that are very similar to those in the two-city interpretation. On the other hand, it is also clear that only the literal interpretation is appropriate for hotel chains.[7]

4.3.1 Interpretation of Switching Costs

The switching costs considered in this paper have the interpretation of a fixed cost of consumption. This is easily understood if one considers two different supermarkets, each of whom sells a set of products (or brands) that at most partially overlaps with those sold by its competitor. Assume also that each consumer can find his optimal consumption bundle in either one of the two supermarkets, but that finding or putting together this consumption bundle involves a fixed cost $s > 0$, which may be due to time spent searching for the products or to the (opportunity) cost of

[6]For simplicity, we assume that all consumers learn all prices in period two. But this is clearly without any loss since the period two decisions of consumers who do not move are not affected by the price of the local store in the other city.

[7]An additional or alternative interpretation is that α is the probability of a preference shock for related goods, say, cosmetics. A chain store or brand like, e.g., NIVEA, can then insure consumers against the cost of switching by offering several cosmetic products.

experimenting with different products.[8] If a consumer has invested s for one of the two supermarkets, he will no longer be indifferent between the two supermarkets though he would have been indifferent between the two ex ante. Thus, the fix cost s is equivalent to a switching cost.

A similar reasoning applies in the case of hotels, although for hotels, search costs for customers do typically not accrue when searching within a given establishment but when searching across different hotels in a given city. So as to minimize search costs, a consumer who has found a suitable hotel that is part of a chain in one city may want to go to a hotel belonging to the same chain when staying in another city.

Viewing switching costs in this way also motivates the informational assumption of the model. If consumers do not know what kind of stores to expect in a yet unfamiliar city, they are quite probably also uncertain about the prices prevailing in this city. Therefore, consumers only learn all prices in the other city after moving to this city. However, if a chain store is present in their home city as well as in the other city, consumers know exactly what prices to expect at the chain store in the other city. A consumer deciding whether to buy at the local or chain store at $t = 1$ thus knows the local store's price in his home city and the chain store's price charged in both cities. He is, however, uncertain about the price of the local store in the other city.

4.3.2 Consumers

There is a continuum of consumers with heterogenous switching costs. Consumers' switching costs s are uniformly distributed on $[0, \sigma]$, so that the density is $f(s) = \frac{1}{\sigma}$ for $0 \leq s \leq \sigma$ and zero otherwise. The probability $\alpha \in (0, 1)$ of moving to the other city in period two is independent of type s. Consumers decide at $t = 1$ and $t = 2$ whether to buy one unit of the good, thereby generating gross utility $u > \sigma$ or not to buy, in which case they get zero utility. A consumer who buys twice from the same store at price p has net utility of $(u - p - s) + (u - p)$, while a consumer who buys from two different stores at prices p' and p'' gets a net utility of $(u - p' - s) + (u - p'' - s)$.

[8] Note that though consumers are modelled as homogenous with respect to utility generated by consumption, this is without loss since the only thing that matters is that absent switching costs, each consumer is indifferent between two different sellers if they set the same price.

4.3.3 Firms

All firms have constant unit costs of production, which are normalized to zero. This simplifying assumption allows to disentangle the effects of consumer mobility and switching costs from the effects of increasing returns to scale. We assume also that firms are committed to charge the same prices in both periods. There are several possible and plausible justifications for this assumption. First, period length may simply be too short to make changing prices worthwhile. For example, if consumers commute and shop at different locations in a metropolitan area on a daily basis, then changing prices from day to day will probably not be optimal for retailers.[9] Second, though this is not part of the present paper, one can imagine a dynamic game where the number of newcomers in every period is sufficiently large, so that the bargain-and-rip-off strategy of low initial and high second period prices does not pay if, as seems realistic, new and old customers cannot be distinguished (see also von Weizsäcker, 1984).[10] Third, for industries where chains are important the assumption of uniform prices over time seems to be more in accord with casual empirical observations than bargain-and-rip-off pricing. Finally, uniform prices make the analysis much more tractable. Though we have no definite results for the alternative with time varying prices, we do not believe that the assumption of uniform prices is in any way crucial for our main findings that chain stores are profitable because they help mobile consumers to economize switching costs.

Firms are also restricted to charge the same price in all locations. This assumption is obviously of no consequences for local stores. It is, however, restrictive for chain stores. A chain could choose a low price in one city and a high price in the other city in order to implement some kind of bargain-and-rip-off strategy. In the present model, however, a bargain-and-rip-off strategy as experienced by consumers, say, in W is also a rip-off-and-bargain strategy when viewed from the perspective of consumers in E. For the case of perfect information, no pure strategy equilibrium where the chain stores charge different prices in the two cities exists. The only pure strategy equilibrium with perfect information is in symmetric prices, but it exists only for a subset of α's.[11]

[9]Clearly, this argument applies much less for hotel chains because of the arguably greater time length that elapses between purchases.

[10]Note that old customers are very unlikely to reveal their type if as a "reward" for this they have to pay higher prices. However, if firms are patient enough, they may play alternating bargain-and-rip-off strategies in equilibrium; see Farrell and Klemperer (2004).

[11]More precisely, the pure strategy equilibrium exists only for $\alpha < 0.35$. The proof is available

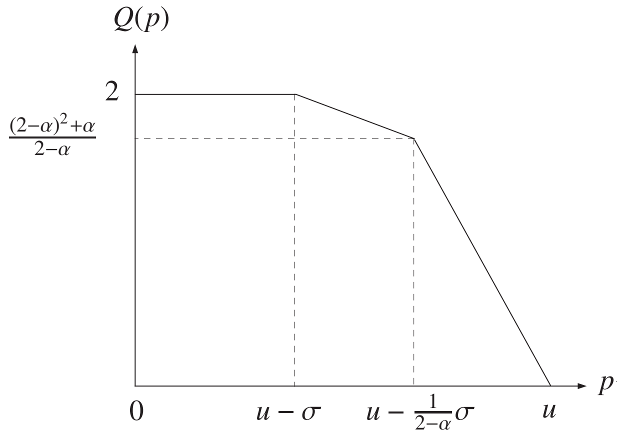


Figure 4.2: Demand faced by a local monopolist.

4.4 Local Monopolies

To set the stage for analyzing the role played by a chain store, the benchmark case of a local monopolist in each city is considered first. When there are only local stores, they cannot help consumers save switching costs. Because a consumer's decision to buy from a given store will only depend on this store's price, each store acts independently of the other one.

Consider the local store in $k \in \{E, W\}$. Throughout, we use $-k$ to denote the city other than k. A consumer s in k at $t = 1$ will choose to shop at this store if his expected net utility exceeds his switching costs, i.e., if $(2 - \alpha)(u - p_k^l) - s \geq 0$. If the same consumer moves to the other city in $t = 2$, he will choose the local store in $-k$, whenever $u - p_{-k}^l \geq s$. An analogous argument applies for consumer s in $-k$. A generic local store charging price p is thus confronted with a demand function consisting of three segments. If the local stores price is very low (i.e., lower than $u - \sigma$), all consumers shop and total demand is 2, consisting of the $2 - \alpha$ consumers from the original city and the fraction α who moves over from the other city.

If price is increased, then consumers who move over shop if and only if $u - p - s > 0 \Leftrightarrow s < u - p$, while still all consumers from the original city shop, i.e. $2 - \alpha$. This amounts to a demand of $(2 - \alpha) + \alpha(u - p)/\sigma$, where $2 - \alpha$ is overall demand from the home city and $\alpha \frac{u-p}{\sigma}$ is the mass of consumers who move and who consume.

upon request.

If price is increased further, then also some consumers from the original city prefer not to shop at all. Consumers with $(2-\alpha)(u-p)-s<0$ do not shop at all. Total demand then amounts to $(2-\alpha)(2-\alpha)(u-p)/\sigma+\alpha(u-p)/\sigma$. In sum, the local store faces demand $Q(p)$ with

$$Q(p)=\begin{cases} 2 & p\leq u-\sigma \\ 2-\alpha+\alpha\frac{u-p}{\sigma} & u-\sigma<p\leq u-\frac{1}{2-\alpha}\sigma \\ \left[(2-\alpha)^2+\alpha\right]\frac{u-p}{\sigma} & u-\frac{1}{2-\alpha}\sigma<p\leq u \end{cases} \tag{4.1}$$

Figure 4.2 provides an illustration. Optimal price is obtained by piecewise maximizing $pQ(p)$.

First note that optimal price will never be lower than $u-\sigma$. Otherwise the local store could increase its price without losing any customers. Consider next the second segment of demand which applies for prices $u-\sigma<p\leq u-\frac{1}{2-\alpha}\sigma$. It is easy to see that the price elasticity of demand (defined negatively) is always greater than minus one if $\frac{u}{\sigma}<\frac{2(2-\alpha)+\alpha^2}{\alpha(2-\alpha)}$.[12] In this case, the local store always prefers a higher price, thus driving price up to the upper bound of this segment, yielding

$$\hat{p}=u-\frac{1}{2-\alpha}\sigma$$

as optimal price and $Q(\hat{p})=\frac{(2-\alpha)^2+\alpha}{2-\alpha}$ as quantity demanded.

If the elasticity of demand is always smaller than minus one,[13] i.e., if $\frac{u}{\sigma}<\frac{2+\alpha}{\alpha}$, then price will be lowered until the lower bound for this segment, $u-\sigma$, is reached. For values of $\frac{u}{\sigma}$ in between these two threshold values, optimal price is given by the first order condition from maximizing profit, yielding

$$p^*=\frac{2-\alpha}{\alpha}\frac{1}{2}\sigma+\frac{1}{2}u.$$

For the third segment of $Q(p)$, the elasticity of demand is smaller than minus one if $u-\frac{1}{2-\alpha}\sigma>\frac{1}{2}u\Leftrightarrow\frac{u}{\sigma}>\frac{2}{2-\alpha}$, in which case the optimal price is as low as possible, i.e., is equal to the lower bound of the segment. Otherwise, the optimal price is given by the first order condition on this segment, yielding

$$p^*=\frac{1}{2}u.$$

[12]The elasticity is $-\frac{\alpha}{\sigma}\frac{p}{2-\alpha+\alpha\frac{u-p}{\sigma}}$, which is bigger than -1 if and only if $u>2p-\frac{\sigma}{\alpha}(2-\alpha)$. Since p is at most $u-\frac{1}{2-\alpha}\sigma$, the right-hand side is not greater than $2u-(\frac{2}{2-\alpha}+(2-\alpha))\sigma$. Re-arranging and simplifying yields the condition in the text.

[13]From the previous footnote, $-\frac{\alpha}{\sigma}\frac{p}{2-\alpha+\alpha\frac{u-p}{\sigma}}<-1\Leftrightarrow u<2p-\frac{\sigma}{\alpha}(2-\alpha)$. Since p is at least $u-\sigma$, the right-hand side is larger than $2(u-\sigma)-\frac{\sigma}{\alpha}(2-\alpha)$, whence the condition in the text is obtained after some re-arranging.

Summarizing, the optimal price p^* is given by

$$p^* = \begin{cases} u-\sigma & \frac{2+\alpha}{\alpha} < \frac{u}{\sigma} \\ \frac{2-\alpha}{\alpha}\frac{1}{2}\sigma + \frac{1}{2}u & \frac{2(2-\alpha)+\alpha^2}{\alpha(2-\alpha)} < \frac{u}{\sigma} \leq \frac{2+\alpha}{\alpha} \\ u - \frac{1}{2-\alpha}\sigma & \frac{2}{2-\alpha} < \frac{u}{\sigma} \leq \frac{2(2-\alpha)+\alpha^2}{\alpha(2-\alpha)} \\ \frac{1}{2}u & 1 \leq \frac{u}{\sigma} \leq \frac{2}{2-\alpha} \end{cases}. \tag{4.2}$$

Welfare Given the zero-one-nature of consumption, maximum welfare with two local monopolies is achieved when all consumers buy the good in both periods. The price then only serves a distributional function, shifting rents from consumers to firms. Note, however, that the local monopolists choose a price p^* sufficiently low to induce all consumers choose to buy the good in both periods only if $\frac{u}{\sigma} > \frac{2+\alpha}{\alpha}$. In all other cases, the local monopoly creates a welfare loss. If two local stores were allowed in each city, then standard Bertrand competition would drive prices down to zero and restore the welfare optimum. But this comes at the price of zero profits for the firms involved and is thus not feasible under costly entry. In Section 4.6 below, we discuss market structure with costly entry in more detail.

4.5 Two Local Stores Compete with a Chain Store

The previous section analyzed equilibrium when a local monopolist serves consumers in each city. The model is now extended by introducing a chain store that operates an outlet in each city and competes with local stores. The game is outlined in Figure 4.3. The advantage of patronizing the chain store instead of local stores is that consumers can economize switching costs: Even if they move to the other city, they can visit the chain store in the new city without incurring additional set-up or switching costs if they have visited it in period one.

Let p_k^l denote the price of the local store in city $k \in \{E, W\}$ and p^c the chain store's price. Remember that consumers in k observe the price of the local store in $-k$ only at $t = 2$ but not at $t = 1$. Denote by Ep_{-k}^l consumers' expected price of the local store in $-k$ from the perspective of consumers living in k at $t = 1$.

The game has one proper subgame, starting with consumers' $t = 2$ buy-decision. In $t = 2$, a consumer s who lived in city k at $t = 1$ has to decide which store to patronize. At this point in time there is no uncertainty with respect to the relevant prices. He will thus patronize the store offering him the lowest price net of switching costs. Clearly, net prices depend on whether he had to move to the other city

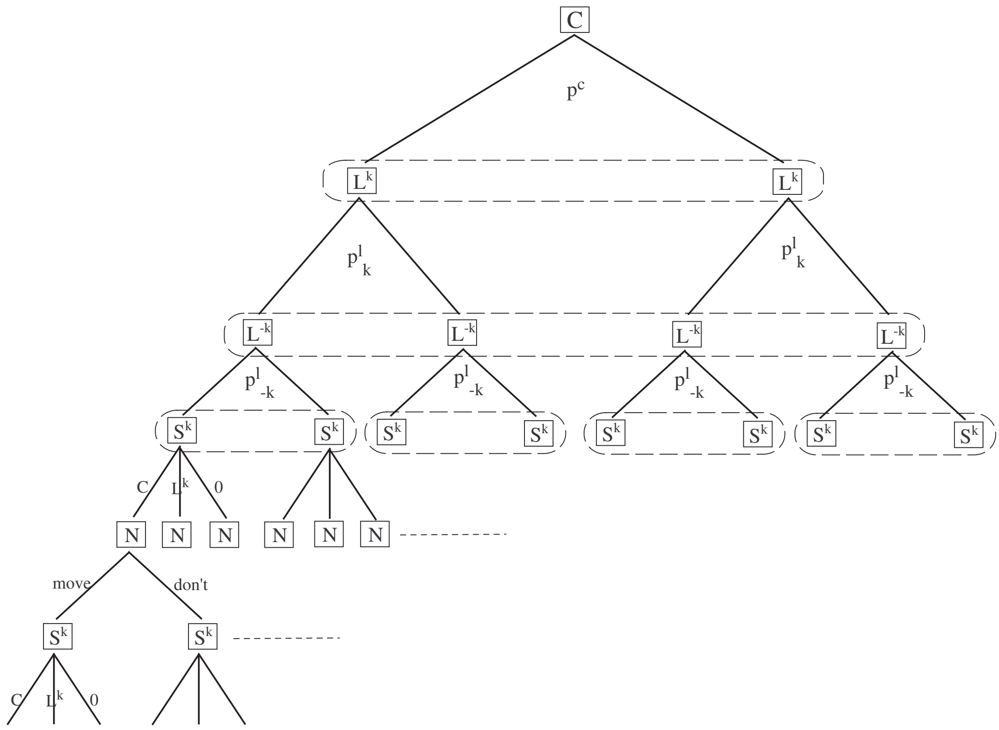


Figure 4.3: Game with two local stores and a chain from the perspective of city k consumers. C stands for the chain, L^k denotes the local store in city k, S^k stands for a representative consumer in k and N is for nature.

(city $-k$) and whether he chose the local or the chain store at $t = 1$. Figure 4.4 summarizes consumer s's optimal decision at $t = 2$.

Consider next what prices firms charge in equilibrium. If all firms charge the same price, all consumers will choose to patronize the chain because it economizes on expected switching costs, thus leaving local stores with zero profits. The next lemma shows that the chain store charges a higher price than local store's in equilibrium.

Lemma 1. *In any subgame perfect pure strategy equilibrium* $0 < p_k^l < p^c < u$ *for* $k \in \{E, W\}$.

Proof. We first show that $0 < p_k^l < p^c$. Suppose that $p_k^l \geq p^c$. Then nobody in k chooses the local store in $t = 1$. In $t = 2$ new consumers arrive, who either chose the chain or local store in $-k$ at $t = 1$. Those who chose the local store in $-k$ will choose the chain in k since it is cheaper. The same reasoning applies for the chain store customers from $-k$. The consumers who were already in k in $t = 1$ all chose the chain store in $t = 1$ and will do so again in $t = 2$. In summary, with $p_k^l \geq p^c$ (assuming that everybody visits the chain in case of a tie) the local store in k will have no customers at all. The only situation where this could be part of an equilibrium is when prices are such that $p_k^l \geq p^c = 0$ because in this case (and only in this case) the local store is indifferent between having customers and having none. We now show that $p_k^l \geq p^c = 0$ cannot be an equilibrium. To see this, note that $s > 0$ for a positive measure of consumers. Consequently, the chain can make positive profits by setting a sufficiently small but positive price p^c, so that it attracts a positive measure of consumers. By setting a price somewhat smaller than p^c but still strictly positive, the local store attracts those consumers with very low switching costs, so that it realizes positive profits. Hence, $p_k^l \geq p^c$ cannot be.

We next show that $p^c < u$. Consider first price setting by the chain. Suppose, contrary to the statement in the lemma, that $p^c \geq u$. In this case, no consumer will patronize the chain in $t = 2$. Given that consumers do not choose the chain in $t = 2$, they will not choose the chain in $t = 1$ either. By setting its price above u the chain thus makes zero profits. If a local store sets $p^l > 0$, the chain can make positive profit for sure by lowering its price just below $\min\{p^l, u\}$.

□

Lemma 1 shows that both firms will charge positive prices which are smaller than maximal utility u. Charging a price of zero is not optimal because with a

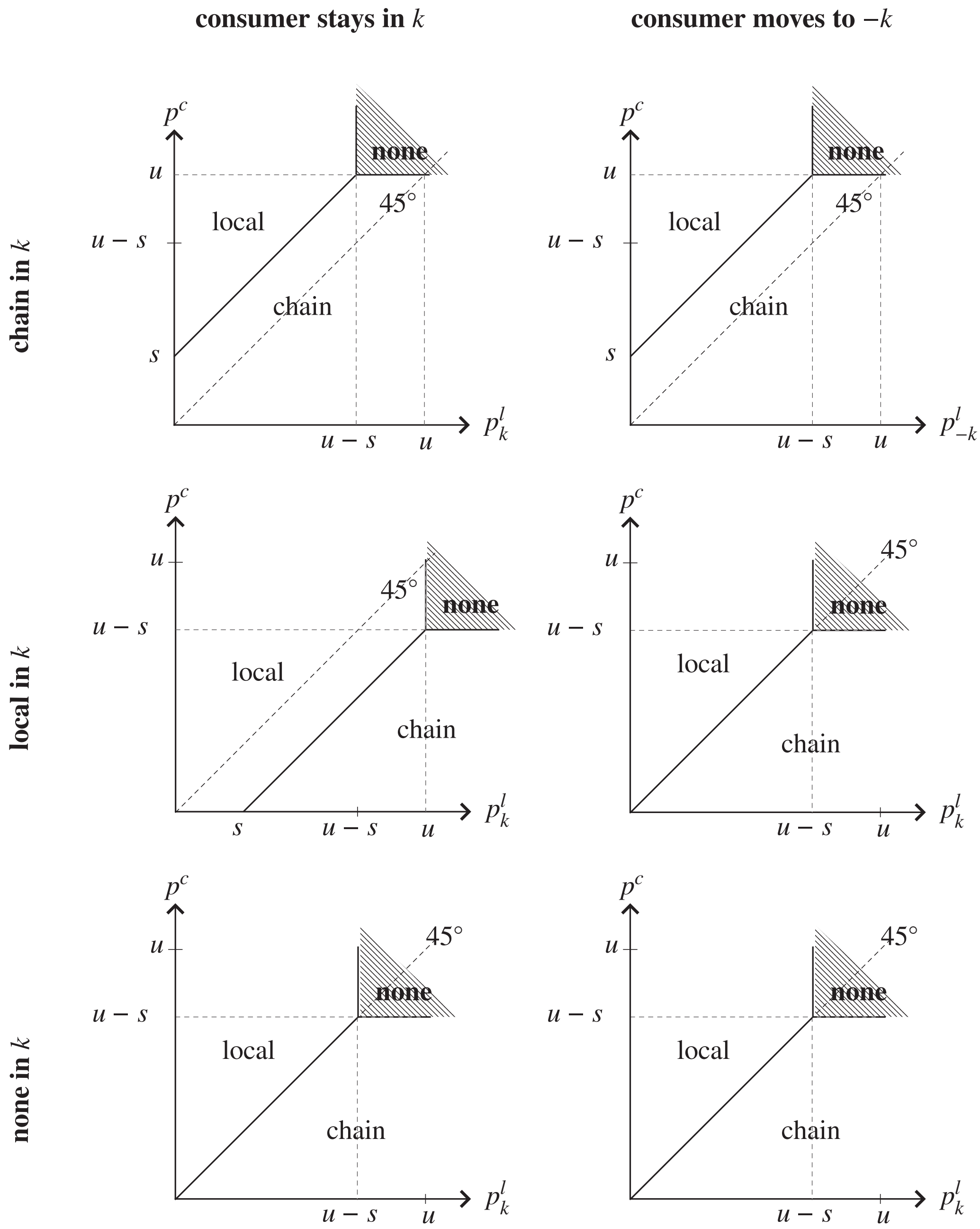


Figure 4.4: Consumer choice in period $t = 2$.

positive mobility parameter α the two firms do no longer sell identical products and can thus attract different consumers: Low switching cost consumers prefer the local store while high switching cost consumers prefer the chain. Setting the price above u is neither optimal because in this case, the respective store has no customers at all.

With Lemma 1 restricting firms' optimal price setting behavior, consumers optimal strategies can be narrowed down to three alternatives: (1) always choose the chain. (2) always choose the local. (3) choose the local in $t = 1$ and when not moved also in $t = 2$. If moved, do not shop at all. The following analysis proves this claim.

Lemma 2. *Whoever buys from the chain buys from it in both periods, whether he moves or not.*

Proof. Denote by $V(x, y, z)$ the expected utility of a consumer who plays the strategy (x, y, z), meaning "buy from x in $t = 1$, from y in $t = 2$ if not moved and from z if moved" with $x, y, z \in \{0, c, l\}$, where 0 stands for not buying at all, c for buying from the chain and l for buying from the local. The strategy of the proof is to show that any strategy that contains at least one c and at most two c's is dominated by a strategy that does either contain no c or by the strategy (c, c, c).

First, it is straightforward to check that (c, c, c) dominates any strategy that contains one or two c's and 0 elsewhere because by Lemma 1, $p^c < u$.

Second, one can show that buying from the chain only in $t = 2$ and from at least one local in any other city or period is dominated by a strategy that does not contain any c. The proof is immediate because one can replace any c that appears in the strategy by an l: The switching cost is borne in either case, but the local's price is smaller.

Third, consider strategies where the chain is chosen in $t = 1$, but some local is chosen in $t = 2$. The basic procedure of the proof is again the same: Replace any c by an l. Complications arise only when establishing that $V(c, l, c) < \max\{V(c, c, c), V(l, l, l)\}$. To see that this indeed holds, consider a consumer who is initially in k and notice that

$$\begin{aligned} V(c,c,c) = 2(u - p^c) - s &> 2u - (1+\alpha)p^c - (1-\alpha)p_k^l - (2-\alpha)s = V(c,l,c) \\ &\Leftrightarrow \\ s &> p^c - p_k^l. \end{aligned}$$

On the other hand, $V(c, l, c) =$

$$\begin{aligned} 2u - (1+\alpha)p^c - (1-\alpha)p_k^l - (2-\alpha)s &> 2u - (2-\alpha)p_k^l - \alpha E p_{-k}^l - (1+\alpha)s = V(l,l,l) \\ &\Leftrightarrow \\ (2\alpha - 1)s &> p^c - p_k^l + \alpha(p^c - E p_{-k}^l). \end{aligned}$$

Since the summands on the right-hand side are positive by Lemma 1, the condition requires s to be smaller than something negative for $\alpha < \frac{1}{2}$, which cannot be. For $\alpha > \frac{1}{2}$, the condition reads

$$s > \frac{1}{2\alpha - 1}(p^c - p_k^l) + \frac{\alpha}{2\alpha - 1}(p^c - E p_{-k}^l).$$

A necessary condition for this condition to be satisfied is $s > (p^c - p_k^l)$. But if $s > (p^c - p_k^l)$ holds, then $V(c, c, c) > V(c, l, c)$ holds. Thus, either $V(c, c, c) > V(c, l, c)$ or $V(l, l, l) > V(c, l, c)$. This completes the proof. □

Now, suppose that the consumer chose the local store in $t = 1$ (second row in Figure 4.4). Since the chain is always more expensive than local stores (Lemma 1), the chain will never be chosen in $t = 2$. Together with Lemma 2 this immediately leads to the following corollary.

Corollary 1. *Consumers do not change type of stores from $t = 1$ to $t = 2$.*

Proof. Local store customers do not switch to the chain store because $p_k^l < p^c$ is given by Lemma 1. According to Lemma 2, chain store customers never switch to local stores. □

Due to Lemma 1, there will always be some consumers in each city choosing the local store in $t = 1$ in any equilibrium. If some consumer s_0 chooses to patronize the local store in $t = 1$, then so will any consumer with $s \in [0, s_0]$. Since according to Corollary 1, consumers do not switch the type of store between $t = 1$ and $t = 2$ and since according to Lemma 2, all consumers who choose the chain in $t = 1$ will do so again in $t = 2$, there remain three relevant strategies for consumers in k:

- always patronize local stores, (l, l, l), with payoff
  $V^k_{(l,l,l)}(s) := (2 - \alpha)(u - p_k^l) - s + \alpha(u - E p_{-k}^l - s)$,

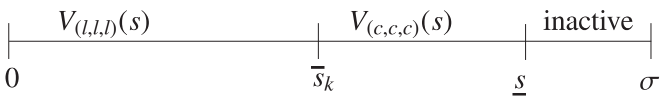


Figure 4.5: Partition of the set of consumers in k provided condition (4.3) does not hold.

- patronize local store in k and patronize no store in $-k$ if moved, $(l, l, 0)$, with payoff:
  $V^k_{(l,l,0)}(s) := (2 - \alpha)(u - p^l_k) - s$ and
- always patronize the chain store, (c, c, c), with payoff:
  $V^k_{(c,c,c)}(s) := (2 - \alpha)(u - p^c) - s + \alpha(u - p^c) = 2(u - p^c) - s.$

Note that if the strategy $(l, l, 0)$ is the preferred strategy for consumer s, then it must be true that $V^k_{(l,l,0)}(s) > V^k_{(c,c,c)}(s)$, i.e.,

$$(2 - \alpha)(p^c - p^l_k) > \alpha(u - p^c). \tag{4.3}$$

Since this condition is independent of s, no consumer at all will choose the chain store in k if it holds. Note that condition (4.3) can only hold in equilibrium, if $p^l_k \neq p^l_{-k}$. Otherwise, condition (4.3) holds for both cities and the chain has no customers at all. This can only be an equilibrium if $p^l_k = p^l_{-k} = 0$, contradicting $p^l_k \neq p^l_{-k}$.

Suppose that condition (4.3) does not hold. In this case the strategy to shop at the local store as long as not moved $(l, l, 0)$ is dominated by always choosing the chain store (c, c, c). Consumers are thus divided into three groups. Low switching cost consumers with $s \leq \overline{s}_k$ always choose local stores, where

$$\overline{s}_k := \frac{2 - \alpha}{\alpha}(p^c - p^l_k) + (p^c - Ep^l_{-k}). \tag{4.4}$$

Medium switching cost consumers with $s \in (\overline{s}_k, \underline{s}]$ always choose chain stores, where

$$\underline{s} := \min\{2(u - p^c), \sigma\}. \tag{4.5}$$

High switching cost consumers with

$$s \in [\underline{s}, \sigma]$$

do not shop at all; see Figure 4.5. The min-operator in (4.5) is necessary because the support of s is $[0, \sigma]$. Notice that the set of high switching cost consumers who do not shop can be empty.

In deriving the demand functions below, it is assumed that all consumers shop (i.e., that this set is empty), implying $\underline{s} = \sigma$. Proposition 1 shows that this is indeed the case in equilibrium.

Given some prices $p_k^c \leq p^c$, and $Ep_k^l < p^c$ for both k, the local store in k thus faces the demand function

$$\begin{aligned} Q_k^l &:= (2-\alpha)\frac{1}{\sigma}\left[\frac{2-\alpha}{\alpha}(p^c - p_k^l) + (p^c - Ep_{-k}^l)\right] \\ &+\alpha\frac{1}{\sigma}\left[\frac{2-\alpha}{\alpha}(p^c - p_{-k}^l) + (p^c - Ep_k^l)\right]. \end{aligned} \tag{4.6}$$

Maximizing $Q_k^l(p_k^l)p_k^l$ with respect to p_k^l for both k yields the first order condition for the local store in k

$$0 = 4p^c - 2\alpha(2-\alpha)Ep_{-k}^l - \alpha^2 Ep_k^l - 2(2-\alpha)^2 p_k^l \quad \text{with} \quad k = E, W. \tag{4.7}$$

A local store's best response function is

$$p_k^{l*}\left(Ep_k^l, Ep_{-k}^l\right) = \frac{4p^c - 2\alpha(2-\alpha)Ep_{-k}^l - \alpha^2 Ep_k^l}{2(2-\alpha)^2}. \tag{4.8}$$

The chain store faces demand

$$Q^c(p^c) := \left(2 - Q_k^l\right) + \left(2 - Q_{-k}^l\right) \tag{4.9}$$

and maximizes $Q^c(p^c)p^c$ with respect to p^c. Its first order condition is

$$0 = -8p^c + 2\sigma\alpha + (2-\alpha)p_k^l + \alpha Ep_k^l + (2-\alpha)p_{-k}^l + \alpha Ep_{-k}^l \tag{4.10}$$

If a subgame perfect equilibrium (SPE) in pure strategies exists, then consumers' expectation about the local store's price in the other city must be correct, i.e., equilibrium prices must be a solution to

$$Ep_k^l = p_k^l \quad \text{for} \quad k = E, W. \tag{4.11}$$

The three first order conditions (4.7) for $k = E, W$ and (4.10) and the two expectation consistency conditions (4.11) constitute a linear system of five equations in $p^c, p_k^l, p_{-k}^l, Ep_k^l$ and Ep_{-k}^l. The following proposition states that such an equilibrium exists and that it is unique.

Proposition 1. *There exists a unique SPE in pure strategies. Equilibrium prices are*

$$p^{l*} := \frac{\alpha}{(2-\alpha)^2+2}\sigma \qquad p^{c*} := \frac{\alpha[(2-\alpha)^2+4]}{4[(2-\alpha)^2+2]}\sigma$$

with $p_k^{l*} = p_{-k}^{l*} = p^{l*}$*. Equilibrium quantities and profits are, respectively,*

$$Q^{l*} := \frac{(2-\alpha)^2}{(2-\alpha)^2+2} \qquad Q^{c*} := \frac{2[(2-\alpha)^2+4]}{(2-\alpha)^2+2}$$

$$\Pi^{l*} := \frac{\alpha(2-\alpha)^2}{[(2-\alpha)^2+2]^2}\sigma \qquad \Pi^{c*} := \frac{\alpha[(2-\alpha)^2+4]^2}{2[(2-\alpha)^2+2]^2}\sigma,$$

where $Q_k^{l*} = Q_{-k}^{l*} = Q^{l*}$ *and* $\Pi_k^{l*} = \Pi_{-k}^{l*} = \Pi^{l*}$.

Proof. **Existence** The system of five linear equations (4.7), (4.10) and (4.11) with $k = E, W$ has a unique solution, which is given by the prices in the proposition. At these prices, all consumers shop in both cities. We thus have $\underline{s}^* = \sigma$. Since the chain store's profit function was derived under the assumption that $\underline{s} \equiv \min\{2(u - p^c), \sigma\} = \sigma$, it is necessary to verify whether the chain has an incentive to deviate, adopting a high price such that $\underline{s} < \sigma$. However, this cannot occur because when deriving the prices, too many consumers were assumed to buy from the chain store if the assumption $\underline{s} = \sigma$ does not hold. That is, we imposed a too favorable demand facing the chain store. Consequently, if under this assumption the chain does not choose a price sufficiently high to induce $\underline{s}^* < \sigma$, then it will a fortiori not choose such a high price when demand is smaller.

Next, it needs to be verified whether the chain store has an incentive to deviate from equilibrium to $p^c = p^{l*}$ in order to attract all consumers, leaving the local store with zero demand. If the chain store chooses price p^{l*}, it attracts Q^{l*} additional customers in each city. The additional revenue thereby generated is $p^{l*}Q^{l*}$ per city. However, in each city the chain store loses the revenue $(p^{c*} - p^{l*})(2 - Q^{l*})$ on the customers it would have attracted even without the deviation. Deviation to p^{l*} is therefore profitable if only if

$$\Delta\Pi := Q^{l*}p^{l*} - \left(2 - Q^{l*}\right)(p^{c*} - p^{l*}) > 0.$$

Note that $p^{c*} - p^{l*} = \frac{(2-\alpha)^2}{4} \geq 1$ for all α, $2 - Q^{l*} = \frac{1}{2}Q^{c*}$ and $\frac{1}{2}Q^{c*} > Q^{l*}$. Therefore,

$$\Delta\Pi = p^{l*}\left[Q^{l*} - \frac{1}{2}Q^{c*}\frac{(2-\alpha)^2}{4}\right] \leq p^{l*}\left[Q^{l*} - \frac{1}{2}Q^{c*}\right] < 0.$$

Hence, it is not profitable for the chain store to deviate to p^{l*}, or to any lower price.

Alternatively, a local store could deviate to a lower price in order to push the chain store out of the market completely. Fix the chain store's and the other local store's prices at p^{c*} and p^{l*} respectively. The local store in k could then set its price so low as to make condition (4.3) hold. To this end it must choose price $p^l_k \leq p^D$, where p^D is such that condition (4.3) holds with equality, i.e.,

$$\begin{aligned}(2-\alpha)(p^{c*}-p^D) &= \alpha(u-p^{c*}) \\ &\Leftrightarrow \\ p^D &:= \frac{2p^{c*}-\alpha u}{2-\alpha}.\end{aligned}$$

But

$$2p^{c*} = \frac{(2-\alpha)^2+4}{2(2-\alpha)^2+4}\alpha\sigma < \alpha\sigma < \alpha u,$$

where the last inequality holds by assumption. Hence,

$$p^D < 0$$

follows, proving that the deviation does not pay for a local store. This completes the proof that the strategy profile stated in the proposition constitutes a SPE.

Uniqueness In deriving the above equilibrium prices, two crucial assumptions on the prevailing demand structure were made: (i) all consumers shop, i.e., $\underline{s}^* = \sigma$ and (ii) that condition (4.3) does not hold.

(i) Suppose that $\underline{s} \equiv \min\{2(u-p^c), \sigma\} < \sigma$. That is, the set of high switching cost consumers who do not shop at all is non-empty. Notice that this does not affect the local stores' profit functions. Consequently, their first order conditions are still given by (4.7). With $\underline{s} = 2(u-p^c)$, the chain store's profit function is

$$\Pi^c(p^c) = \left(2\frac{2(u-p^c)}{\sigma} - Q^l_k\right)p^c + \left(2\frac{2(u-p^c)}{\sigma} - Q^l_{-k}\right),$$

yielding the first order condition

$$0 = 4\alpha u + 2p^l_{-k} + 2p^l_k - 8\alpha p^c + \alpha E p^l_{-k} - \alpha p^l_k + \alpha E p^l_k - \alpha p^l_{-k} - 8p^c \tag{4.12}$$

In addition, conditions (4.11) have to be satisfied. Candidate equilibrium prices are given as solution to equations (4.7), (4.12) and (4.11). These prices are

$$\begin{aligned}\tilde{p}^l &:= \frac{2\alpha u}{6+\alpha[(2-\alpha)^2+\alpha]} \\ \tilde{p}^c &:= \frac{(2-\alpha)^2+4}{4}\tilde{p}^l.\end{aligned}$$

These prices, however, imply

$$\underline{s} = 2u - \frac{\alpha(\alpha^2 - 4\alpha + 8)}{\alpha^3 - 3\alpha^2 + 4\alpha + 6}u > u > \sigma,$$

where the first inequality follows because the fraction is less than one for all α and the second inequality holds by assumption. Thus, there is no equilibrium with $\underline{s} < \sigma$.

(ii) Suppose that condition (4.3) holds. That is, assume that the chain attracts no customers in k. If the chain neither attracts consumers in $-k$, then lowering its price is until it attracts some generates positive revenue. So, assume that the chain attracts some consumers in $-k$. Then condition (4.3) must not hold in $-k$, implying

$$p^l_{-k} > p^l_k. \tag{4.13}$$

The chain store's demand is then given by some consumers in $-k$ and by those of them who move from $-k$ to k. Total demand is thus

$$Q^c(p^c) = \frac{2}{\sigma}\left[\underline{s} - \overline{s}_{-k}\right] = \frac{2}{\sigma}\left[\underline{s} - \frac{2-\alpha}{\alpha}(p^c - p^l_{-k}) - (p^c - Ep^l_k)\right].$$

For the case $\underline{s} = \sigma$, the first order conditions from maximizing $p^c Q^c$ yields the chain store's reaction function

$$p^{c*}(p^l_{-k}, Ep^l_k) = \frac{1}{4}\left[\alpha\sigma + (2-\alpha)p^l_{-k} + \alpha Ep^l_k\right].$$

Now condition (4.3) for city k requires that $2p^c > \alpha u + (2-\alpha)p^l_k$. Using $Ep^l_k = p^l_k$ and inserting $p^{c*}(p^l_{-k}, Ep^l_k)$ for p^c, condition (4.3) reads

$$\begin{aligned} \frac{1}{4}\left[\alpha\sigma + (2-\alpha)p^l_{-k} + \alpha p^l_k\right] &> \frac{1}{2}\left[\alpha u + (2-\alpha)p^l_{-k}\right] \\ &\Leftrightarrow \\ \alpha p^l_k > \alpha(\underbrace{2u - \sigma}_{>u}) + (2-\alpha)p^l_{-k} &> \alpha u \end{aligned}$$

where the last inequality is due to $u > \sigma$. Condition (4.3) in k thus requires $p^l_k > u$. But this cannot be an equilibrium since the local store in k has no consumers in this case.

For $\underline{s} = 2(u - p^c) < \sigma$, the chain store's profit function is

$$\Pi^c(p^c) = \frac{2}{\sigma}\left[2(u - p^c) - \frac{2-\alpha}{\alpha}(p^c - p^l_{-k}) - (p^c - Ep^l_k)\right]p^c.$$

Thus, its reaction function is

$$p^{c*}\left(p^l_{-k}, Ep^l_k\right) = \frac{1}{4(1+\alpha)}\left[2\alpha u + (2-\alpha)p^l_{-k} + \alpha E p^l_k\right].$$

Proceeding as before, one gets

$$\begin{aligned}
\frac{1}{4(1+\alpha)}\left[2\alpha u + (2-\alpha)p^l_{-k} + \alpha E p^l_k\right] &> \frac{1}{2}\left[\alpha u + (2-\alpha)p^l_{-k}\right] \\
&\Rightarrow \\
\frac{1}{4}\left[2\alpha u + (2-\alpha)p^l_{-k} + \alpha E p^l_k\right] &> \frac{1}{2}\left[\alpha u + (2-\alpha)p^l_{-k}\right] \\
&\Leftrightarrow \\
\alpha p^l_k &> (2-\alpha)p^l_{-k} \\
&\Leftrightarrow \\
p^l_k &> p^l_{-k}
\end{aligned}$$

for condition (4.3). But this contradicts (4.13). There exists thus no equilibrium with condition (4.3) holding. Hence, the equilibrium is unique. □

Discussion *Market Shares and Profits.* According to Proposition 1, the chain store's profits in each city (which are equal to half of the chain store's total profit Π^{c*}) are larger than the profits of a local store. Prices and profits of both local stores and the chain store increase in α, but Π^{c*} increases faster in α than Π^{l*}. Note also that $Q^{c*}(\alpha)$ is strictly increasing in α. It equals $8/3$ for $\alpha = 0$ and is equal to $10/3$ for $\alpha = 1$. Because equilibrium demand aggregated over both periods and both cities is four, the chain store's market coverage increase from $2/3$ to $5/6$ as α increases from zero to one.

Predicted Price Differences. The model predicts also that local stores charge lower prices than chain stores. To see this, notice that

$$p^{c*} = \frac{(2-\alpha)^2 + 4}{4} p^{l*} > p^{l*}$$

for all α. At first sight, this may seem at odds with empirical facts if one thinks of, say, the retail industry.[14] However, this indicates only that switching costs are not the only driving factor in the retail industry, where increasing returns and market power on the input side may be at least as important. On the other hand, there are other industries where observed pricing patterns are hard to understand without the factors that our model emphasizes. As mentioned at the very beginning, a

[14] See, e.g., Hausman and Leibtag (2004).

local provider of bus trips from New York City to Boston is substantially cheaper than the large chain. Similarly, the regional airline German Wings offers flights that are cheaper by orders of magnitude than those of Lufthansa. Starbucks, the largest coffee house chain, is not exactly known for providing cheap coffee, though the price differences here are certainly less striking than those for bus trips or airfares. Casual empiricism in the hotel industry also suggests that large chains are by no means cheaper than local hotels offering the same quality. More importantly, though, there is also some systematic evidence that is in line with the price pattern predicted by our model from the banking industry. Ishii (2004) estimates the effect of ATM surcharges on retail banking industry structure and welfare.[15] Surcharges for withdrawing cash from banks other than the one at which a customer has his or her deposit account impose a cost of switching banks to the consumer. Ishii finds that consumers prefer banks with larger ATM networks, arguably because of lower expected surcharge payments. She finds that banks with larger ATM networks pay lower interest rates on deposits, which corresponds to charging a higher price in our model.

Public Prices of Local Stores. The assumption that local stores' prices are only known locally has some consequences that are worth a brief discussion. Consider the local store in k. Differentiating its best response function (4.8) with respect to consumers' expectations Ep_k^l yields

$$\frac{\partial p_k^{l*}(Ep_k^l)}{\partial Ep_k^l} = -\frac{1}{2}\left(\frac{\alpha}{2-\alpha}\right)^2 < 0.$$

That is, the lower the expected price, the higher the optimal price of the local store in k. The reason for this is straightforward. A lower expected price implies a larger demand, and the larger demand in turn induces the local store to set a higher price. However, since in equilibrium consumers cannot be fooled, $Ep_k^l = p_k^l$ must hold, implying that a high price and low expected price are not consistent.

This behavior is reminiscent of the well known problem of the durable goods seller uncovered by Coase (1972), in which consumers' (correct) expectations of lower future prices reduce demand in the presence, as a consequence of which price in the presence is reduced as well. A durable goods seller who could commit not to lower its price in the future would make a larger profit. Very similarly, the local stores in our model could gain if they could credibly communicate their

[15] See also Knittel and Stango (2005).

prices in both cities in period one, thereby committing themselves not to "cheat" on consumers.[16]

Having said that, we should emphasize that the assumption that local stores' prices are not completely public information is not only more realistic than assuming that they are known in both cities, but it is also without consequences for the qualitative predictions of our model. If we assumed instead that the prices of local stores are known in both cities in period one, the chain store would still set a higher price and make a larger profit than local stores.

4.6 Industry Equilibrium with Costly Entry

So far, we took market structures as given. In section 4.4, we analyzed the market structure with local monopolies, and in the previous section we analyzed the interplay of a chain store that competes with a local store in every city. An interesting question is whether one of these configurations is stable in the sense that all firms that are active make non-negative profits and that no additional firms have incentives to enter the market.

As has already been seen, though local monopolies make positive profits, the market structure with a local monopoly in each city is not stable because a chain can profitably enter. So as to show that the market structure of section 4.5 is stable, we thus have to show that no additional local store and no additional chain has an incentive to enter if this market structure prevails.

Lemma 3. *If there are two or more stores of the same type (local or chain) in a city, at least two of them charge a price of zero, and all stores of the this type make zero profits.*

[16] Assuming that only the local store in k communicates its price in both cities, Ep_k^l has to be replaced by p_k^l in the demand functions of the two local stores and of the chain. Everything else remains the same. Equilibrium prices then are:

$$\hat{p}_k^l = \frac{1}{2}\sigma\alpha\frac{-12\alpha + 5\alpha^2 + 8}{48 - 104\alpha + 93\alpha^2 - 40\alpha^3 + 8\alpha^4},$$

$$\hat{p}_{-k}^l = \hat{E}p_{-k}^l = \sigma\alpha\frac{4 - 6\alpha + 3\alpha^2}{48 - 104\alpha + 93\alpha^2 - 40\alpha^3 + 8\alpha^4},$$

and

$$\hat{p}^c = \sigma\alpha\frac{8 - 16\alpha + 13\alpha^2 - 5\alpha^3 + \alpha^4}{48 - 104\alpha + 93\alpha^2 - 40\alpha^3 + 8\alpha^4}.$$

It can be checked that $\hat{p}_k^l > p_k^{l*}$ holds and that equilibrium profit is larger when price is known in both cities.

Proof. Assume to the contrary that some firm makes positive profits. The only way that this can happen is that it charges a positive price. But given that this firm serves customers at a positive price, another firm of the same type will have an incentive to slightly undercut this price and get all the customers from this firm. Clearly, this race to the bottom will only stop if one of the firms charges a price equal to zero. So that a firm that charges a price of zero has no incentive to raise its price, it must be the case that another firm sets a price of zero as well. This proves the claim about equilibrium prices. As to profits, note first that all firms that charge a price of zero trivially make zero profits. Second, any firm that charges a higher price will have no customers and consequently will make zero profits, too. □

Lemma 3 implies that the market structure with one local store in each city and one chain serving both cities is the unique market structure if entry into the industry is associated with some positive costs. Starting with no firms at all, either a local or a chain store can profitably enter the market. If a local store enters, no other local store will enter the same city, since profits would be zero. Another local store will only enter in the other city. If there is a local store in each city, a chain can still enter profitably but, due to Lemma 3, not more than one chain will enter. Thus, we have:

Proposition 2. *With small but positive entry costs, the unique stable market structure consists of a chain store with an outlet in both cities and a local store in each city.*

Regarding welfare, this market structure only achieves second best. Since $u > \sigma$ and since there are no production costs, it is optimal that all consumers consume the good in both periods. As Proposition 1 showed, all consumers shop in equilibrium in both periods. However, some of them shop at local stores and are thus confronted with expected switching costs of $(1 + \alpha)s$.

With two competing chains, prices are zero and all consumers shop in both periods. But now expected switching costs are only s for all consumers. Therefore, first best would be achieved by two competing chains. But as just argued, this is not a stable market structure. The only stable market structure with one chain and a local shop in each city, thus generates higher welfare than do two local monopolists, but does not attain first best welfare.

4.7 Conclusions

We study a two city model where mobile consumers face costs of switching sellers. Since consumers change the city with an exogenous probability, they can reduce expected switching costs by shopping at a chain store rather than at a local store. If consumers differ with respect to switching costs, firm size serves as a means of product differentiation, where local stores serve low switching cost consumers and chain stores serve high switching cost consumers.

This model provides three key insights. First, the market structure with a local store and a chain store in each city is the unique stable market structure if there is a small, positive cost of entry. That is, local stores coexist in equilibrium with the chain store. Second, the chain store charges a higher price than local stores. Third, as consumers become more mobile, the market share of the chain store increases, and so do profits and prices of all stores. Moreover, the chain store becomes more profitable relative to local stores as mobility increases.

Bibliography

Athey, S. and A. Schmutzler (2001). Investment and market dominance. *RAND Journal of Economics 32*(1), 1–26.

Bagwell, K. and G. Ramey (1994). Coordination economics, advertising, and search behavior in retail markets. *American Economic Review 84*(3), 498–517.

Bagwell, K., G. Ramey, and D. F. Spulber (1997). Dynamic retail price and investment competition. *RAND Journal of Economics 28*(2), 207–27.

Baye, M. R. and J. Morgan (2001). Information gatekeepers on the internet and the competitiveness of homogenous product markets. *American Economic Review 91*(3), 454–74.

Baye, M. R., J. Morgan, and P. Scholten (2004). Price dispersion in the small and in the large: Evidence from an internet comparison site. *Journal of Industrial Economics LII*(4), 463–496.

Besanko, D. and U. Doraszelski (2004). Capacity dynamics and endogenous asymmetries in firm size. *RAND Journal of Economics 35*(1), 23–49.

Coase, R. H. (1972). Durability and monopoly. *Journal of Law and Economics 15*(1), 143–149.

Dinlersoz, E. M. (2004). Firm organization and the structure of retail markets. *Journal of Economics & Management Strategy 13*(2), 207–40.

Farrell, J. and P. Klemperer (2004). Coordination and lock-in: Competition with switching costs and network effects.

Hausman, J. and E. Leibtag (2004). Consumer benefits from increased competition in shopping outlets. *Working paper*.

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

Ishii, J. (2004). Interconnection pricing and compatibility in network industries: ATM networks in the banking industry. Working Paper.

Janssen, M. C., V. A. Karamychev, and P. van Reeven (2003). Multi-store competition: Market segmentation or interlacing? Tinbergen Institute Discussion Paper No. 03-033/1.

Klemperer, P. (1987). The competitiveness of markets with switching costs. *RAND Journal of Economics 18*(1), 138–150.

Klemperer, P. (1995). Competition when consumers have switching costs. *Review of Economic Studies 62*, 515–39.

Knittel, C. and V. Stango (2005). Compatibility with indirect network effects: Evidence from ATMs. *NBER Working Paper Nr. 10774*.

Michael, S. C. and H. J. Moore (1995). Returns to franchising. *Journal of Corporate Finance 2*, 133–155.

Nelson, P. (1970). Information and consumer behavior. *The Journal of Political Economy 78*, 311–329.

Salop, S. (1979). Monopolistic competition with outside goods. *Bell Journal of Eonomics 10*, 141–56.

Stahl, K. (1982). Location and spatial pricing theory with nonconvex transportation cost schedules. *Bell Journal of Economics 13*(2), 575–82.

von Weizsäcker, C. C. (1984). The cost of substitution. *Econometrica 52*(5), 1085–1116.

Wolinsky, A. (1983). Retail trade concentration due to consumers' imperfect information. *Bell Journal of Economics 14*(1), 275–282.

Can Conglomerates Ease Credit Constraints?

Yves Schneider [*]

[*]I gratefully acknowledge Philippe Aghion, Karolin Becker, Michael Breuer, Matthias Dewatripont, Boris Krey, Armin Schmutzler and Peter Zweifel for most valuable comments and discussion. Any errors are mine.

Chapter 5

Can Conglomerates Ease Credit Constraints?

5.1 Introduction

In contrast to the Modigliani-Miller-Theorem, it is widely acknowledged in the literature that a firm's financial structure is relevant for its profitability. Departing from a world of perfect contracts where all information is verifiable, the corporate finance literature studies the incentive effects of a firm's financial structure. [1]

Financing constraints arising from asymmetric or non-verifiable information also shed light on the issue of defining the boundaries of the firm. Does owner-provided investment strengthen or undermine managers' incentives to act in the interest of the investor? As was pointed out by Zingales (2000), a satisfactory answer to this question must be motivated by the costs and benefits of allocating funds through a firm relative to having arms-length market transactions. A theory of the firm is inevitably necessary for analyzing the financing decision in general and the relative efficiency of internal or external financing in particular. The internal capital market hypothesis states that "internal markets of diversified firms enable them to fund profitable projects that, because of information asymmetries and agency costs, the external capital market would not be able to finance."(Shin and Stulz, 1998, p.531).

This tradeoff between internal and external finance has received considerable attention during the last decade (Stein (1997), Gertner et al. (1994), Inderst and Müller (2003), Brusco and Panunzi (2005)). Still, most of these papers do not fully address an important issue. They assume that headquarters have some superior ca-

[1] A recent survey is given by Hart (2001). See also Harris and Raviv (1991).

pabilities compared to an external investor in dealing with agency problems. Due to the incompleteness of contracts, external investors cannot reap the full value of cash flows produced by the firm, which are (at least in part) not verifiable. However, headquarters somehow manage to get their hands on these nonverifiable cash flows.

For example, in Stein (1997)'s winner picking model, headquarters can improve a conglomerate's financing possibilities by reallocating funds between projects of different profitability. Headquarters are assumed to be able to monitor the projects it oversees and it is endowed with authority to reallocate resources. This authority also allows headquarters to appropriate a fraction of its project manager's private benefits. The model I present below aims at explaining where this power to appropriate private benefits comes from. It is argued that headquarters informational advantage comes with a contractual advantage. Although formal contracting possibilities are the same for outside investors and headquarters, the latter can resort to informal contracts.

Similarly, Inderst and Müller (2003) compare centralized (via headquarters) to decentralized borrowing (each project is directly financed by outside funds) under the premise that cash flows are not verifiable. Conglomerates can improve upon stand alone projects through cash flow pooling: Excess cash flow from one project can be used to buy continuation of another, liquidity constrained, project. However, in analyzing centralized financing they implicitly assume that cash flows are verifiable to headquarters since headquarters do not face any difficulties in appropriating their project managers cash flows (as opposed to outside investors). The model presented below explains that this "increased verifiability" of cash flows for headquarters may stem from its possibility to use an informal employment contract with its project managers.

This article provides a rationale for how headquarters can mitigate a conglomerate's internal agency problem. Due to their informational advantage, headquarters can use informal contracts which are not available to the outside investor (e.g. a bank). Whether this advantage translates into increased investor's expected return depends on the characteristics of the conglomerate's projects. The model derives conditions under which conglomerates are able to finance projects that would not be viable if financed independently on external capital markets. The model compares two cases. Case 1 ("independent financing") is where two identical projects are financed independently using a formal financial contract between entrepreneurs

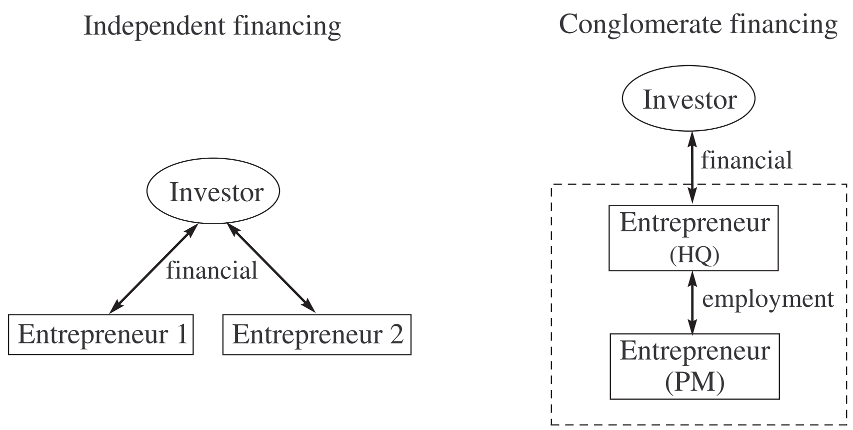


Figure 5.1: Independent project financing vs. conglomerate financing.

and investor. Case 2 ("conglomerate financing") is where two entrepreneurs use an informal contract between themselves and a formal financial contract with the investor (see Figure 5.1).

Empirical studies find that conglomerates trade at a discount compared to an equivalent portfolio of stand-alone firms. Berger and Ofek (1995) estimate a discount of up to 15 percent.[2] Lang and Stulz (1994) document a significant diversification discount throughout the 1980s as well.

Graham et al. (2002) criticize these studies for neglecting potential selection problems. There might be systematic differences between divisions of conglomerates and stand-alone firms that serve as the benchmark. Indeed, Hubbard and Palia (1999) find evidence suggesting that the highest bidder returns in conglomerate mergers during the 1960s were realized by financially unconstrained buyers acquiring financially constrained target firms. This tends to confirm the internal capital market view of conglomerates. Moreover, Fluck and Lynch (1999) note that if conglomerates enable the financing of marginally profitable projects which otherwise would not be viable, conglomerates will as a consequence be less profitable than stand-alone firms, causing them to be traded at a discount.

The model presented below supports this hypothesis. Projects which are financed optimally through a conglomerate are predicted to deliver smaller expected

[2]They analyze a sample of roughly 3,500 firms contained in the Compustat industry segment database which covers US and Canadian firms.

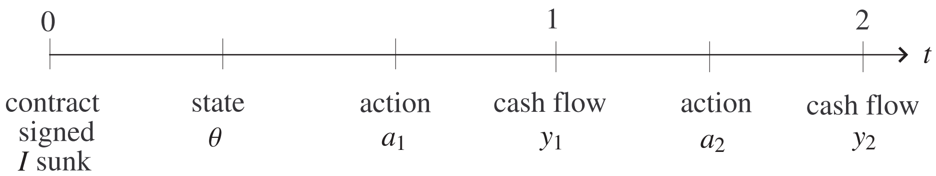


Figure 5.2: Timing.

returns to the investor than projects which are optimally financed as independent entities.

Section 5.2 presents the basic model characterizing the relationship between an entrepreneur and an outside investor (i.e. the case of independent financing). Section 5.3 takes this model as a building block in order to analyze conglomerate financing. After deriving optimality conditions, results are compared to those pertaining to independent financing. Section 5.4 concludes.

5.2 Basic Model

The model is based on Aghion and Bolton (1992). A risk-neutral entrepreneur owns a project which produces cash flows at two dates, being either 1 (successful) or 0 (unsuccessful). Since the entrepreneur himself has no wealth, he requires funds from a risk–neutral investor to finance the project.

Figure 5.2 illustrates the timing. After initial investment I is sunk at $t = 0$, the project can either turn out to be in a good (θ_g) or in a bad state (θ_b). The ex ante probability of the good state is p. After the state of the project realized, a choice of action a_1 is required. There are three option available: "continue" (C), "reorganize" (R) or "liquidate" (L). The impact of this action on expected cash flows depends on the state of the project. In the good state no special action is required (C is optimal), and the project can be continued as before. In this case it delivers high cash flow with probability $\pi_g \equiv E[y_t|\theta = \theta_g, C]$. However, if the project is in the bad state, reorganizing the project increases the probability of success compared to continuing as before (R is optimal). By choosing to reorganize, the project is successful with probability $\pi_R \equiv E[y_t|\theta = \theta_b, R]$ and by choosing to continue, the project is successful with probability $\pi_b \equiv E[y_t|\theta = \theta_b, C] < \pi_R$. Liquidation,

$t = 1$ **action** (a_1)	$\rightarrow$	**actions available at** $t = 2$ (a_2)
continue (C)	$\rightarrow$	continue (C), reorganize (R), liquidate (L)
reorganize (R)	$\rightarrow$	continue reorganization (R), liquidate (L)
liquidate (L)	$\rightarrow$	none

Table 5.1: Summary of available actions for $t = 1$ and $t = 2$.

finally, refers to either putting the project's assets to its most valuable alternative use or replacing the entrepreneur. The expected cash flow from liquidating is independent of the state of the project and is assumed to be worse than reorganizing. Therefore, liquidation might only be valuable as a strategic threat. For future reference, the expected per period cash flows $E[y|\theta, a]$ for all different combinations of project state (θ) and action chosen (a) are summarized below.

$E[y|\theta, a] =$

Project state	action		
	$a = C$	$a = R$	$a = L$
$\theta = \theta_g$	π_g	π_b	π_0
$\theta = \theta_b$	π_b	π_R	π_0

After first period cash flow (y_1) realized, a second action choice (a_2) leading to a second cash flow realization (y_2) is required. In principle, the set of available options is the same as in the first period: $a_2 \in \{C, R, L\}$. However, if liquidation was chosen before ($a_1 = L$) then the project is terminated at the end of period 1 and, as a consequence, no action choice is required in period two. Similarly, if reorganization was chosen in period one already ($a_1 = R$), it is not possible to reverse it in period two. In this case options are limited to either continue reorganizing ($a_2 = R$) or to liquidate the project ($a_2 = L$). Finally, if continuation was chosen in period one ($a_1 = C$), all three options are available in period two ($a_2 \in \{C, R, L\}$). Table 5.1 summarizes available actions at $t = 1$ and $t = 2$. The project terminates after second period cash flow was realized.

Besides allowing for two consecutive phases of action choice and cash flow realization instead of one, the model presented here differs from Aghion and Bolton (1992) in two important ways. First, there is no verifiable signal on the projects profitability except for realized cash flow. Second, the signal in Aghion and Bolton is observed before action choice and does therefore contain no information on which action was chosen. Here, however, first period cash flow which also serves as a signal on the project's state is observed after action chosen and thus contains information about the choice of action.

Expected returns are structured as follows (recall that expected returns are 1 or 0):

$$\pi_g \geq \pi_R > \frac{1}{2} > \pi_b. \tag{5.1}$$

The investor who provides investment I only cares about his share in total monetary returns $y_t(\theta, a)$. The entrepreneur who came up with the project-proposal and is running it not only cares about his share in monetary returns but also about private benefits $B(a_t)$. These private benefits depend on the action chosen but not on the state of the project. It is assumed that $B(C) \equiv B$ and $B(R) = B(L) = 0$. The entrepreneur thus derives private benefits B from continuing the project regardless of the state θ.

Although the entrepreneur forgoes private benefits by reorganizing a bad project, the additional cash flows realized by reorganizing more than compensate this loss:

$$\pi_R > \pi_b + B. \tag{5.2}$$

If the entrepreneur is the full claimant on cash flows, he therefore prefers to reorganize b ad projects. However, if his share of cash flow is sufficiently los, he prefers to continue bad projects and thereby to collect his private benefits.

Private Benefits from Continuation. There are several plausible reasons for the existence of private benefits as postulated here. For example, an entrepreneur interacting with his employees on a daily basis does not like to confront them with unpleasant decisions like lay offs or restructuring. Moreover, an entrepreneur might adopt his employees' (or other stake holders') preferences, generating a preference for continuation (see Hart and Holmstrom, 2002). An outside party, like banks or consultants, is not locked into this 'corporate culture' and shy away from unpleasant decisions. As a consequence, it does not enjoy any private benefits from continuation.

Bertrand and Mullainathan (2003) empirically estimate the effect of relaxation of corporate governance as measured by the introduction of anti-takeover laws on managerial behavior. They conclude that behavior resembles preferences for a quiet live. After anti–takeover laws were passed, less plants were closed down but new plants were also less likely to be built. Also do employees' wages increase after the introduction of such laws. This evidence supports the notion of private benefits from continuation (and contradicts the notion of empire building, which presumes that managers prefer to actively build new plants or divisions).

Private benefits from project continuation create a potential conflict of interest between entrepreneur and outside investor. If the entrepreneur had sufficient wealth to fund his project out of his own pocket, no problem would arise because in this case he is the full claimant of cash flows. If, however, he has to approach an investor for funds, the conflict of interest provoked by the presence of private benefits generates an agency problem. In order to induce the entrepreneur to reorganize bad projects, the investor must grant him some minimal share of the project's cash flow, reducing the investor's expected return. This reduction might be sufficient to make it unprofitable for the investor to finance the project. Before analyzing this agency problem in detail, the first best contract is derived as a benchmark.

5.2.1 First Best and Laissez-faire Contracts

The first best contract is associated with the optimal contract chosen, if it is possible to include all available and relevant variables in a verifiable contract.[3] Such a contract needs to specify which actions are to be implemented in which state of the project. A contract s is thus given by $s = \{a(\theta), t\}$, where t refers to a monetary transfer from the investor to the entrepreneur. The optimal contract s^* maximizes total expected surplus

$$E\left[y_1(\theta, a_1) + B(a_1) + y_2(\theta, a_2) + B(a_2)\right]. \tag{5.3}$$

Due to (5.2), this expression is maximized for $a_1^*(\theta_g) = a_2^*(\theta_g) = C$ and $a_1^*(\theta_b) = a_2^*(\theta_b) = R$.

Clearly, the contract implemented depends on the bargaining power of the investor and the entrepreneur. If the investor can make a take-it-or-leave-it offer to the entrepreneur, no transfer will be included in the contract, i.e $t = 0$. If, on the other hand, the entrepreneur can make a take-it-or-leave-it offer to the investor, the contract would specify the largest possible transfer to the entrepreneur that still guarantees an expected return of I to the investor, i.e. $t = E(y_1|s^*) + E(y_2|s^*) - I = 2E(y|s^*) - I$. Because cash flows are identically and independently distributed in both periods, expected cash flows are the same in both periods, that is $E(y|s^*) = p\pi_g + (1-p)\pi_R$. To simplify exposition, the remainder of the paper assumes that all bargaining power is vested with the investor.

[3] First best can also be associated with the counterfactual case where the entrepreneur has unlimited own funds and does not need to contract with an external investor at all.

Assumption 1. *The investor has all bargaining power, i.e. she is always in the position to make a take-it-or-leave-it offer to the entrepreneur.*

With Assumption 1, the first best contract has $t^* = 0$, that is no transfers at all to the entrepreneur. The entrepreneur thus receives expected private benefits $2pB$ but no share in cash flows. Investor's expected return net of investment is $2E(y|s^*) - I$. All projects with a positive net present value (NPV) are thus carried out and no profitable project is credit constrained.

Note that from the perspective of efficiency, all positive NPV projects should be financed. Whether this is achieved with or without any monetary transfers between entrepreneur and investor is irrelevant. Given that the project is implemented and run until $t = 2$, any transfer between entrepreneur and investor has only distributional consequences and does not affect efficiency.

Since state (θ), actions (a_t) and private benefits (B) are not verifiable, moral hazard issues arise in the contracting relationship. The standard solution would be to devise an incentive contract specifying transfers from and to the investor conditional on output (cash flow). In the present context of risk neutral players, such an incentive contract would guarantee that all positive NPV projects are implemented. However, such a contract generally is not compatible with limited liability on the part of the entrepreneur. Since a central aspect of the present paper is to discuss financing possibilities for wealth-constrained entrepreneurs, limited liability is naturally given here. Entrepreneurs' wealth (or utility) is bounded from below by zero. Although all players are risk neutral, limited liability imposes costs for project financing and makes it impossible to fund all positive NPV projects.

As an alternative to an incentive contract, the investor could let the entrepreneur choose his preferred action. The investor could simply provide initial investment and then claim all future cash flows. In this case the entrepreneur will never reorganize bad projects, resulting in an expected return of $2\left[p\pi_g + (1-p)\pi_b\right]$ for the investor. In order for the financing problem to be interesting at all, it is assumed that the expected return from such a "laissez-faire"-contract is smaller than investment costs I, yielding a negative NPV for the project. A "laissez-faire"-contract will therefore never be chosen by the investor.

Before discussing the optimal contract (Section 5.2.3), the informational environment in which investor and entrepreneur interact needs to be clarified.

5.2.2 Entrepreneur's Role and Informational Assumptions

Given the model outlined so far, the most obvious solution for a financially constrained entrepreneur is to sell his project to a financially unconstrained investor. If the investor is as good as the entrepreneur in running the project, all agency problems are resolved by this transaction. An investor is willing to pay a maximum of $2E(y) - I$, that is the full expected net present value of the project.[4]

If all actions taken by the entrepreneur can as well be taken by the investor, the entrepreneur is not necessary for running his project. The value of the entrepreneur is in this case reduced to the value of his business idea. Once the idea is sold to the investor, the latter can run the project herself. This is not very plausible and even if it were, it constitutes a case of minor interest. The issue of project financing becomes important if the entrepreneur's human capital is inalienable to some degree, causing the investor to depend on the entrepreneur.

Assumption 2. *The entrepreneur is necessary for running his project. Without the entrepreneur, the investor's only option is to liquidate the project. Liquidating yields an expected cash flow of π_0, where $\pi_R > \pi_0 \geq \pi_b$ and $2\pi_0 < I$.*

So far it was implicitly assumed that investor and entrepreneur share the same information (both observe θ). Intuitively, one would suppose that only the entrepreneur knows exactly what is going on in his market or industry and that an outside investor has much less detailed knowledge of the entrepreneur's industry.[5] Detailed knowledge requires playing an active role in the market and communicating with employees, customers and suppliers while constantly analyzing competitors' behavior and decisions. This knowledge cannot be acquired by monitoring an industry, like some specialized consulting firm would do, or by inspecting the firm's books. Clearly, specialized consulting firms are able to learn more about an industry than an outside investor like a bank, but there is still some valuable knowledge which can only be acquired by actively participating in this industry.

For the remainder of the paper it is assumed that an outside investor has less information than does an entrepreneur who is actively involved in his industry. This is in line with Shin and Stulz (1998, p.531), who emphasize the importance of asymmetric information in explaining the potential advantage of internal capital

[4] This option is also feasible in Aghion and Bolton (1992) although it is not discussed there. A Pareto-efficient allocation is always achievable by selling the project to the investor.

[5] See also Stein (1997, p.117).

markets. Moreover, actions, state of the project and private benefits are assumed to be non-verifiable.

Assumption 3. *Cash flows y_t, are verifiable but actions a_t, project state θ and private benefits B, are not. In addition, state θ is only observed by the entrepreneur but not by an outside investor.*

The presence of private benefits from continuing the project might actually be related to having superior knowledge. If superior knowledge is only obtainable through investment in project- or firm-specific human capital, individuals having access to such knowledge would naturally derive private benefits from it. These benefits would however be lost in the event of reorganization, which would cause the entrepreneur to lose his project-specific human capital.

5.2.3 Optimal Contract

A financial contract between entrepreneur and investor specifies cash flow shares $w_t(y_t)$, depending on cash flow realizations y_t. In addition, control over the project, that is the right to choose actions a_1 and a_2, could be allocated to one of the two contracting parties. Due to Assumption 3, control over action a_1 has to be allocated unconditionally to the entrepreneur. Remember that the first verifiable information y_1 is observed after the choice of action a_1 only. If the investor had control over the project, she could only choose to liquidate the project at $t = 1$ (continuation and reorganization are only available to the entrepreneur). This is not in her interest because this guarantees a negative NPV for the project. However, action a_2 can be allocated conditionally on realized first-period cash flow y_1. Denote by $\beta \in [0, 1]$ the probability that control is allocated to the entrepreneur at $t = 1$. The most general contract is thus given by $s = (w_1, \beta(y_1), w_2(y_1))$, with control initially allocated to the entrepreneur. After first-period cash flow was realized at $t = 1$, revenues are shared according to $1 \geq w_1 \geq 0$. Hence, control is allocated to the entrepreneur with probability $\beta(y_1)$, and revenues at $t = 2$ are shared according to $w_2(y_1)$.

Allocating control to the investor at $t = 1$ contingent on first-period cash flow (y_1) serves to increase the entrepreneur's incentive to reorganize. Zero first-period cash flow increases the investor's belief that the project is in the bad state θ_b and that the entrepreneur did not reorganize the bad project ($a_1 \neq R$). In order to induce the entrepreneur to reorganize bad projects and thereby increase the probability of positive first-period cash flow, the investor could threaten to liquidate the project given that first-period cash flow is zero. If the project is liquidated, the entrepreneur does not receive any cash flow at $t = 2$ and he loses private benefits B in case he would choose to continue. The investor, on the other hand, would secure herself an expected return of π_0. However, the following lemma shows that such a liquidation threat by the investor is not credible.

Lemma 4. *It is in the investor's interest to allocate control unconditionally to the entrepreneur:* $\beta(y_1) = 1$ *for all* y_1.

Proof. First note that this paper focusses on pure strategies. The entrepreneur will therefore either choose to reorganize a bad project or to continue a bad project. He is not allowed to mix.

Further note that because $E(y_2|y_1 = 0, \theta = \theta_b, R) > \pi_0$, the investor will only use liquidation as a threat in order to induce the entrepreneur to reorganize. Now, suppose that the entrepreneur beliefs that the investor will liquidate a project returning zero cash flow. Suppose further that this threat is effective, that is the entrepreneur will indeed prefer to reorganize bad projects to continuing them. But in this case, the investor "knows" that the entrepreneur will reorganize bad projects. Even if zero cash flow realizes, the investor knows that the entrepreneur would have chosen to reorganize if the project was bad. The investor has no (ex post) incentive to actually liquidate the project because this guarantees lower expected return than continuing. Liquidating is thus not sequentially rational.

□

Given that in equilibrium such a threat effectively deters the entrepreneur from continuing a bad project, the investor will not liquidate the project even if first-period cash flow was zero because $E(y_2|y_1 = 0, \theta = \theta_b, R) > \pi_R > \pi_0$. Liquidating the project is thus not sequentially rational.

Incentives similarly to these induced by a liquidation-threat can be provided by choosing the entrepreneur's second-period share of cash flow, $w_2(y_1)$, to be zero in case first-period cash flow was zero. The entrepreneur does not receive any second-period cash flow if first-period cash flow was zero (as with liquidation) but he is able to retain his private benefits should he choose to continue a bad project. As Proposition 3 below shows, this is the investor's preferred contract for a large and plausible range of model parameters.

Note that because action choices are assumed to be non-verifiable ex ante *and* ex post, message games are not feasible even with symmetric information between investor and entrepreneur.[6]

Expected first-period cash flow is $E(y_1)$. Expected second-period cash flow is $E(y_2) = E(y_1)$. The investor can distinguish two conditional expectations for second-period cash flow.[7]The first is expected second-period cash flow given that first-period cash flow was equal to one (which occurs with probability $Pr(y_1 = 1) \equiv E(y_1)$). The second is expected second-period cash flow given that first-period cash flow was equal to zero (which occurs with probability $Pr(y_1 = 0) \equiv 1 - E(y_1)$). In the former case expected second-period cash flow is $E[y_2|y_1 = 1]$, and in the latter case it is $E[y_2|y_1 = 0]$. One therefore has

$$Pr(y_1 = 1)E[y_2|y_1 = 1] = p\pi_g^2 + (1-p)\pi_R^2 \tag{5.4}$$

$$Pr(y_1 = 0)E[y_2|y_1 = 0] = p(1-\pi_g)\pi_g + (1-p)(1-\pi_R)\pi_R. \tag{5.5}$$

Investor's expected return $V(\cdot)$ is then given by

$$V(w_1, w_2(1), w_2(0)) := [1 - w_1]E(y_1) + [1 - w_2(0)]Pr(y_1 = 0)E(y_2|y_1 = 0) + [1 - w_2(1)]Pr(y_1 = 1)E(y_2|y_1 = 1). \tag{5.6}$$

The entrepreneur's incentive constraint must guarantee that the he will reorganize bad projects. If the project is in the good state (θ_g), no incentive issues arise since continuation is preferred by entrepreneur and investor. If, however, the project is in the bad state (θ_b), the project will produce a positive cash flow ($y_1 = 1$) with probability π_R if the project is reorganized but only with probability π_b if continued.

First-period cash flow provides valuable information for the investor who does not observe θ by permitting her to update her prior belief concerning the probability

[6]If actions were verifiable ex post and entrepreneur and investor shared the same information, the following message game could be envisaged. After observing the state of nature, both players announce the state of nature to a third party. If the announced states coincide, prespecified action $a(\hat{\theta})$ with $a(\hat{\theta}_g) = C$ and $a(\hat{\theta}_b) = R$ is implemented and transfer $t(\hat{\theta})$ is made from investor to entrepreneur, with $t(\theta_g) = z > 0$ and $t(\theta_b) = 2B + z$, and z slightly larger than $2\pi_b$. If the announced states do not coincide, the project is liquidated and the proceeds from liquidation are allocated to the third party. Abstracting from renegotiation and enforceability concerns, investor's expected return from this contract is $2E(y) - (1-p)2B - 2\pi_b$. Depending on the size of π_b, this return might be too small for the investor to recover her investment costs I. In this case, the project cannot be financed using message contracts. However, if actions are not verifiable ex post, message games are not feasible.

[7]Note that $E(y_2) = Pr(y_1 = 0)E[y_2|y_1 = 0] + Pr(y_1 = 1)E[y_2|y_1 = 1] = [1 - E(y_1)]E[y_2|y_1 = 0] + E(y_1)E[y_2|y_1 = 1]$.

that the project is in the good state. On the other hand, first-period cash flow does not provide any additional information to the entrepreneur because he directly observes the project's state and thus has full information. Therefore, even if first-period cash flow was positive in the bad state (which occurs with probability π_R or π_b, depending on the action a_1 chosen), expected second-period cash flow is still π_R or π_b. The entrepreneur will thus reorganize the project if

$$w_1\pi_R + w_2(1)\pi_R\pi_R + w_2(0)(1-\pi_R)\pi_R \geq w_1\pi_b + w_2(1)\pi_b\pi_b + w_2(0)(1-\pi_b)\pi_b + 2B \quad (5.7)$$

Rewriting the objective function V defined in (5.6) and incentive constraint IC given in (5.7), the investor's problem is given by[8]

$$\max_{w_1, w_2(0), w_2(1)} \quad (5.8)$$

$$2E(y) - w_1E(y) - w_2(0)E(y)E[y_2|y_1 = 0] - w_2(1)E(y)E[y_1|y_2 = 1]$$

subject to

$$w_1 + w_2(1)(\pi_R + \pi_b) + w_2(0)(1 - \pi_R - \pi_b) \geq 2B/(\pi_R - \pi_b). \quad (5.9)$$

Objective function as well as incentive constraint are linear in entrepreneurs' shares of cash flow $w_t(h)$. Whether incentives are most effectively provided through w_1, $w_2(1)$ or $w_2(0)$ depends on the coefficients $\gamma(w) := -IC_w/V_w$ where IC_w is the partial derivative of the incentive constraint with respect to w and V_w the partial derivative of the objective function with respect to the same w. $\gamma(w)$ expresses the efficiency of providing incentives through share w. The larger the effect of w on incentives (IC_w) and the smaller the associated loss in investor's expected return (V_w), the higher is the efficiency coefficient $\gamma(w)$. Incentives are thus preferably provided through the share w with the highest $\gamma(w)$.

A simple equity contract would have the same share of cash flow w in both periods, regardless of the level of first-period cash flow. Proposition 3 below shows that such a standard equity contract will generally not be optimal. The optimal contract includes second-period shares that are contingent on cash flows.

Proposition 3. *An uninformed investor financing a single project will optimally choose the following contract.*

[8] Because $E(y_1)E[y_2|y_1 = 1] + [1 - E(y_1)]E[y_2|y_1 = 0] = E(y_1)$ and $E(y) \equiv E(y_1) = E(y_2)$, the objective function can be written $2E(y) - w_1E(y) - w_{21}E(y)E[y_1|y_2 = 1] - w_{20}E(y)E[y_2|y_1 = 0]$.

(i) If $\pi_R + \pi_b > E(y_2|y_1 = 1)$ *then*

$$w_2(1)^* = \frac{2B}{\pi_R^2 - \pi_b^2} \qquad \textit{and} \qquad w_1^* = w_2^*(0) = 0.$$

(ii) If $\pi_R + \pi_b < E(y_2|y_1 = 1)$ *then*

$$w_2(0)^* = \frac{2B}{\pi_R - \pi_b} \qquad \textit{and} \qquad w_1^* = w_2^*(1) = 0.$$

(iii) If $\pi_R + \pi_b = E(y_2|y_1 = 1)$ *then*

$$w_1^* = w_2^*(0) = w_2^*(1) = \frac{B}{\pi_R - \pi_b} \qquad \textit{(Simple equity contract).}$$

Proof. The proof has three steps. It shows (i) that from $\pi_R + \pi_b > E(y_2|y = 1)$ it follows that $\gamma(w_2(1)) > \gamma(w_1) > \gamma(w_2(0))$ and incentives are thus best provided by $w_2(1)$. Next it is shown that (ii) from $\pi_R + \pi_b < E(y_2|y = 1)$ it follows that $\gamma(w_2(1)) < \gamma(w_1) < \gamma(w_2(0))$ and incentives are thus best provided by $w_2(0)$. Finally, with (iii) $\pi_R + \pi_b = E(y_2|y = 1)$ all γ's are identical in which case the investor is indifferent between the three available cash flow shares.

(i) Note that if $\pi_R + \pi_b > 1$, $\gamma(w_2(0)) < 0$ and $\gamma(w_2(1)) > \gamma(w_1)$. Incentives are thus best provided through $w_2(1)$. Remains the case where $\pi_R + \pi_b \leq 1$.

With $1 \geq \pi_R + \pi_b > E(y_2|y_1 = 1) > E(y)$, it follows that

$$\frac{1 - \pi_R - \pi_b}{[1 - E(y)]E(y_2|y_1 = 0)} < \frac{1 - E(y_2|y_1 = 1)}{[1 - E(y)]E(y_2|y_1 = 0)}. \tag{5.10}$$

Since $E(y_2) = E(y)E(y_2|y_1 = 1) + [1 - E(y)]E(y_2|y_1 = 0)$ and $E(y_2) = E(y)$ it follows that

$$\frac{1 - E(y_2|y_1 = 1)}{[1 - E(y)]E(y_2|y_1 = 0)} = \frac{1}{E(y)}. \tag{5.11}$$

Furthermore, because of $\pi_R + \pi_b > E(y_2|y_1 = 1)$, we have

$$\frac{1}{E(y)} < \frac{1}{E(y)} \frac{\pi_R + \pi_b}{E(y_2|y_1 = 1)}. \tag{5.12}$$

Combining (5.10), (5.11) and (5.12) yields

$$\frac{\pi_R + \pi_b}{E(y)E(y_2|y_1 = 1)} > \frac{1}{E(y_2)} > \frac{1 - \pi_R - \pi_b}{[1 - E(y)]E(y_2|y_1 = 0)} \tag{5.13}$$

which is equal to $\gamma(w_2(1)) > \gamma(w_1) > \gamma(w_2(0))$, implying that incentives are more effectively provided through $w_2(1)$. Setting $w_1 = w_2(0) = 0$ in the incentive constraint (5.9) and solving for $w_2(1)$ yields $w_2(1)^* = 2B/(\pi_R^2 - \pi_b^2)$.

(ii) With $\pi_R + \pi_b < E(y_2|y_1 = 1)$ the inequalities in (5.10) and (5.12) are inverted, leading to $\gamma(w_2(1)) < \gamma(w_1) < \gamma(w_2(0))$. In this case, incentives are best provided by $w_2(0)$. Setting $w_1 = w_2(1) = 0$ in the incentive constraint (5.9) and solving for $w_2(0)$ yields $w_2(0)^* = 2B/(\pi_R - \pi_b)$.

(iii) Clearly, with $\pi_R + \pi_b = E(y_2|y_1 = 1)$ the inequalities in (5.10) and (5.12) turn into equalities. Since in this case the investor is indifferent between the available cash flow shares, one particular solution is to use all three shares equally. Setting $w \equiv w_1 = w_2(0) = w_2(1)$ and solving (5.9) for w yields $w = B/(\pi_R - \pi_b)$.

Substituting optimal shares into the objective function of (5.8), investor's maximum expected return becomes

$$V^* = E(y_1) + E(y_2) - \frac{2b}{\pi_R^2 - \pi_b^2} E(y_2) E(y_1|y_2 = 1)$$

$$= 2\left[1 - \frac{BE[y_2|y_1 = 1]}{(\pi_R^2 - \pi_b^2)}\right] E(y_2). \quad (5.14)$$

□

Proposition 3 shows under which conditions which shares w_1 and $w_2(y_1)$ are most effective in providing incentives to the entrepreneur. If the share $w_2(1)$ is used then the entrepreneur's gain in expected cash flow from reorganizing bad projects is $\pi_R^2 - \pi_b^2 = (\pi_R + \pi_b)(\pi_R - \pi_b)$. Investor's "marginal costs" of using $w_2(1)$ is $E(y_2|y_1 = 1)$, resulting in per unit costs of providing additional incentives to the entrepreneur of $\frac{\pi_R^2 - \pi_b^2}{E(y_2|y_1=1)}$. The respective costs of providing additional incentives to the entrepreneur by using the share w_1 are $\frac{\pi_R - \pi_b}{E(y_1)}$. Condition (i) of Proposition 3 states that if the latter cost is lower than the former then the share $w_2(1)$ is to be used. The remaining two conditions in the proposition state under which conditions the other two shares are optimal.

Empirical studies for the venture capital (VC) industry find financial arrangements that are similar to the optimal contract in Proposition 3. For instance Kaplan and Strömberg (2003, 2001), argue that cash flow rights are often conditional on observable measures of financial performance. This corresponds with the optimal contract derived above, where claims to second-period cash flow are only granted to the entrepreneur if first-period cash flow was high.

Empirical findings also suggest that control over the project shifts conditional on observable performance measures. The optimal contract in the present model, however, does not allow for contingent control allocation because of a (strong) informational assumption of the model. The investor or venture capitalist lacks the necessary information from deriving a benefit from controlling the project. This contrasts with findings by Sahlman (1990), where venture capitalists are reported to take an active part in operational decisions and do in fact have control rights.

Although the model presented here captures one prominent feature of venture capital finance (namely the contingent allocation of cash flow rights), a richer model which relaxes the informational assumption is necessary to allow for contingent allocation of control to be a valuable governance instrument. As argued above, assuming symmetric information would not be satisfactory, since in this special case the entrepreneur's value is reduced to the value of his project idea. A promising extension would be to allow external investors to possess information that is *complementary* to that of the entrepreneur.

5.3 Conglomerate Financing

This section analyzes the possibility of two entrepreneurs financing their projects jointly within a conglomerate. The external investor finances the conglomerate. This implies joint finance of projects and joint observation of cash flows. Cash flows of individual projects are no longer verifiable to the outside investor.

Assumption 4. *For two projects pursued within a conglomerate, only (joint) conglomerate cash flow is verifiable. The cash flows of individual projects are no longer verifiable.*

Nonverifiability of individual cash flows inside a conglomerate can be motivated by the fact that headquarters are able transfer cash flows from one project to the other without leaving a verifiable trace. Accounts only state conglomerate cash flow rather than project-specific cash flows.

Since according to Assumption 4 an investor can only contract on the conglomerate's joint cash flow, the entrepreneurs' incentives to reorganize bad projects are reduced. Each entrepreneur will partially free ride on the other's success. It is evident that incentives are more focussed if the investor finances each project separately. Therefore, pooling cash flows per se makes conglomerate financing inferior to independent project financing. This contrasts the result from Inderst and Müller (2003), where conglomerate's advantage is explained precisely with the pooling of cash flows. In their model it is assumed that cash flows are not verifiable at all. In order to provide entrepreneurs with an incentive to repay their loans, investors threaten not to finance continuation of projects in case entrepreneurs default. Investors therefore inefficiently liquidate the project with positive probability. Cash flow pooling reduces this probability since headquarters can use a successful project's cash flow to secure the continuation of an unsuccessful project. In the present paper, however, joint cash flow is verifiable, and the only restriction for the investor is to provide entrepreneurs with sufficient incentives to reorganize bad projects. Because cash flow realizations depend on entrepreneurs' actions, incentives are provided via cash flow shares. Since projects are assumed to be independent[9] of each other, cash flow pooling leads to a loss of information with regard to the behavior of each entrepreneur. It is no longer possible for an investor to distinguish each project's cash flow within total cash flow.

The advantage of a conglomerate must therefore stem from another source. This paper claims that the advantage of a conglomerate is the result of making use of informal contracts between the two entrepreneurs. Due to Assumption 3, they both observe each other's state of the project θ. While this information as well as choice of action is not verifiable, they are free to codify it in an informal contract. For the sake of clarity, such a contract will be called an "employment contract" in contrast to the formal contract between investor and entrepreneur, which will be called a "financial contract".[10]

As illustrated in Figure 5.1, two entrepreneurs, called headquarters (HQ) and project manager (PM), form a conglomerate to jointly finance their projects. HQ and investor use a formal contract to secure investment funds for both projects while HQ and PM use an informal contract to guarantee the desired behavior on part of the PM.

Before analyzing this informal employment contract in detail, the structure of the investor's problem when financing a conglomerate is studied.

[9] The states of the projects are assumed to be uncorrelated although they are in the same industry. This is somewhat counterintuitive, but simplifies the analysis considerably. See also the discussion in Stein (1997, B.1 and B.2).

[10] The terms "employment" and "financial" serve only to distinguish the two contracts. Both contracts are incentive contracts and there is no reason why a priori one should be viewed as an employment contract while the other as a financial contract. As Dybvig and Zender (1991) show, the incentives provided by a financial contract can equally well be provided by an employment/incentive contract, making the distinction between these two contracts artificial.

5.3.1 Structure of the Investor's Problem

In order to study the contracting problem between investor and HQ in isolation, it is assumed for the moment that the PM will always reorganize bad projects. The contract between HQ and PM and its implication for the outside investor is discussed in section 5.3.2 below.

Conglomerate cash flow at the end of period t is denoted by $Y_t = y_t^1 + y_t^2$, where $y_t^i \in \{0, 1\}$ denotes cash flow of project i at date t. Note that compared to financing each project independently, contracting possibilities are richer now for the investor, since conglomerate cash flow can attain three levels, $Y_t \in \{0, 1, 2\}$. As with single project financing, the contract specifies HQ's shares of cash flow w_1 and w_2. As before, these shares can be specified contingent on previous and current cash flow levels: $w_1(Y_1)$ and $w_2(Y_1, Y_2)$. Let $Pr(Y_1 = h)$ denote the probability that first-period cash flow is equal to h and let $Pr(Y_2 = h'|Y_1 = h)$ denote the conditional probability that second-period cash flow is equal to h' given that first-period cash flow was equal to h. The investor's objective function now is[11]

$$\max_{w_1(h), w_2(h,h')} \sum_{h=1}^{2} [1 - w_1(h)] Pr(Y_1 = h) h + \sum_{h=0}^{2} \sum_{h'=1}^{2} [1 - w_2(h, h')] Pr(Y_1 = h) Pr(Y_2 = h'|Y_1 = h) h'. \quad (5.15)$$

This objective function is more flexible than the objective function in the context of single project financing because a conglomerate's cash flow level is determined by the sum of cash flows from independent projects. Note that each of these two projects can be in two different states, namely θ_g and θ_b. This amounts to four possible states for the conglomerate, (θ_g, θ_g), (θ_g, θ_b), (θ_b, θ_g) and (θ_b, θ_b).

From the perspective of HQ's incentives, only those two states of the conglomerate are relevant in which his own project is in the bad state: (θ_b, θ_g) and (θ_b, θ_b). This gives rise to two incentive constraints, one for each of these states. The next Lemma allows to simplify these incentive constraints. It shows that it is never optimal to provide incentives through first-period shares $w_1^*(0)$ and $w_1^*(1)$.

Lemma 5. *The optimal contract has* $w_1(0)^* = w_1(1)^* = 0$.

[11] Note: $Pr(Y_1 = h \text{ and } Y_2 = h') = Pr(Y_1 = h) Pr(Y_2 = h'|Y_1 = h)$. Furthermore, $E(Y_2) = \sum_h \sum_{h'} Pr(Y_1 = h \text{ and } Y_2 = h')h' = \sum_h \sum_{h'} Pr(Y_1 = h) Pr(Y_2 = h'|Y_1 = h)h'$.
.

Proof. Consider the effect of first-period share of cash flow on HQ's incentives to reorganize. Suppose the conglomerate is in state (θ_b, θ_b) and that the PM reorganizes his project.

If HQ reorganizes his project, first-period conglomerate cash flow is zero with probability $(1 - \pi_R)^2$, one with probability $2\pi_R(1 - \pi_R)$ and two with probability π_R^2. If he continues his project, conglomerate cash flow is zero with probability $(1-\pi_R)(1-\pi_b)$, one with probability $\pi_R(1-\pi_b)+(1-\pi_R)\pi_b$ and two with probability $\pi_R\pi_b$. HQ thus reorganizes his project, if

$$- (\pi_R - \pi_b)(1 - \pi_R)w_1(0) \cdot 0 + (\pi_R - \pi_b)\overbrace{(1 - 2\pi_R)}^{<0}w_1(1) \cdot 1 \\ + (\pi_R - \pi_b)\pi_R w_1(2) \cdot 2 \geq 2B.$$

Since positive values for $w_1(0)$ and $w_1(1)$ reduce incentives, it is optimal to choose $w_1(0)^* = w_1(1)^* = 0$. The same result holds, if the PM's project is in the good state. □

Next, consider HQ's incentive constraint for state (θ_b, θ_b). Lemma 5 limits the relevant first-period shares to $w_1(2)$. This share applies only if first-period cash flow was equal to 2. Since π_R is the probability that each single project produces a cash flow of one, first-period conglomerate cash flow is equal to 2 with probability π_R^2.

Second-period shares are more complicated than first-period shares because they are allowed to depend on first-period cash flow. Conditioning second-period shares on first-period cash flow might improve HQ's incentives because high first-period cash flow is only possible if both projects produce high cash flow. For the same reason it might be preferable to grant claims to second-period cash flow to the HQ only if second-period cash flow is high.

With probability $(1 - \pi_R)^2$, first-period cash flow is zero, with probability $2\pi_R(1 - \pi_R)$ it is one, and with probability π_R^2 it is equal to two. For each level of first-period cash flow, second-period cash flow can either be 0, 1 or 2. Neglecting the share for zero second-period cash flow leads to two relevant second-period cash flow levels for each of the three possible first-period cash flow levels. In sum, HQ's expected return from reorganizing his bad project is

$$
\begin{aligned}
&\pi_R^2 w_1(2)\cdot 2\\
&+(1-\pi_R)^2\left[2(1-\pi_R)\pi_R w_2(0,1)\cdot 1+\pi_R^2 w_2(0,2)\cdot 2\right]\\
&+2\pi_R(1-\pi_R)\left[2(1-\pi_R)\pi_R w_2(1,1)\cdot 1+\pi_R^2 w_2(1,2)\cdot 2\right]\\
&+\pi_R^2\left[2(1-\pi_R)\pi_R w_2(2,1)\cdot 1+\pi_R^2 w_2(2,2)\cdot 2\right].
\end{aligned}
\tag{5.16}
$$

Analogously, if HQ does not reorganize his own project his expected return is

$$
\begin{aligned}
&\pi_b\pi_R w_1(2)\cdot 2\\
&+(1-\pi_R)(1-\pi_b)\left[((1-\pi_R)\pi_b+\pi_R(1-\pi_b))\,w_2(0,1)\cdot 1+\pi_R\pi_b w_2(0,2)\cdot 2\right]\\
&+\left[\pi_R(1-\pi_b)+(1-\pi_R)\pi_b\right]\left[((1-\pi_R)\pi_b+\pi_R(1-\pi_b))\,w_2(1,1)\cdot 1+\pi_R\pi_b w_2(1,2)\cdot 2\right]\\
&+\pi_R\pi_b\left[((1-\pi_R)\pi_b+\pi_R(1-\pi_b))\,w_2(2,1)\cdot 1+\pi_R\pi_b w_2(2,2)\cdot 2\right]\\
&+2B.
\end{aligned}
\tag{5.17}
$$

As with independent financing, HQ reaps private benefits B in each period if it continues its own project. HQ will therefore reorganize, if (5.16) is larger than (5.17), i.e. if

$$
\begin{aligned}
&2\pi_R\cdot w_1(2)\\
&+(1-\pi_R)\left[\pi_R(\pi_R+\pi_b-2)+(1-\pi_R)(1-\pi_R-\pi_b)\right]\cdot w_2(0,1)\\
&+2(1-\pi_R)\pi_R(1-\pi_R-\pi_b)\cdot w_2(0,2)\\
&+\{\pi_R\left[\pi_R(\pi_R+\pi_b-2)+(1-\pi_R)(1-\pi_R-\pi_b)\right]\\
&+(1-\pi_R)\left[\pi_R(1-\pi_R-\pi_b)+(1-\pi_R)(\pi_R+\pi_b)\right]\}\cdot w_2(1,1)\\
&+2\pi_R\left[\pi_R(1-\pi_R-\pi_b)+(1-\pi_R)(\pi_R+\pi_b)\right]\cdot w_2(1,2)\\
&+\pi_R\left[\pi_R(1-\pi_R-\pi_b)+(1-\pi_R)(\pi_R+\pi_b)\right]\cdot w_2(2,1)\\
&+2\pi_R^2(\pi_R+\pi_b)\cdot w_2(2,2)\\
&\geq 2B/(\pi_R-\pi_b).
\end{aligned}
\tag{5.18}
$$

The incentive constraint for state (θ_b, θ_g) is

$$\begin{aligned}
&2\pi_g \cdot w_1(2) \\
&+(1-\pi_g)\left[\pi_g(\pi_R+\pi_b-2)+(1-\pi_g)(1-\pi_R-\pi_b)\right]\cdot w_2(0,1) \\
&+2(1-\pi_g)\pi_g(1-\pi_R-\pi_b)\cdot w_2(0,2) \\
&+2\Big\{(1-\pi_g)\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right] \\
&\quad+\pi_g\left[\pi_g(\pi_R+\pi_b-2)+(1-\pi_g)(1-\pi_R-\pi_b)\right]\Big\}\cdot w_2(1,1) \\
&+2\pi_g\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]\cdot w_2(1,2) \\
&+\pi_g\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]\cdot w_2(2,1) \\
&+2\pi_g^2(\pi_R+\pi_b)\cdot w_2(2,2) \\
&\geq 2B/(\pi_R-\pi_b).
\end{aligned} \tag{5.19}$$

See the appendix for details.

Fortunately it is possible to simplify these incentive constraints. The next lemma shows that it is never optimal to use the shares $w_2(0,1)$ and $w_2(1,1)$. The reason is that the factors in (5.18) and (5.19) for these shares are negative, leading to a reduction of HQ's reorganization-incentives.

Lemma 6. $w_2^*(0,1) = w_2^*(1,1) = 0.$

Proof. It is first shown that $w_2^*(0,1) = 0$ is optimal. The expression in square brackets of eq. (5.18) in the factor for $w_2(0,1)$ can be rewritten as

$$\pi_R(\pi_R+\pi_b-2)+(1-\pi_R)(1-\pi_R-\pi_b) = (1-2\pi_R)(1-\pi_R-\pi_b)-\pi_R. \tag{5.20}$$

Since $1-2\pi_R < 0$ and $|1-\pi_R-\pi_b| < \pi_R$, this expression is always negative. In state (θ_b, θ_g), expression (5.20) changes to $(1-2\pi_g)(1-\pi_R-\pi_b)-\pi_g$, which is also negative. If the investor set $w_2^*(0,1)$ positive, entrepreneur's incentives are reduced. It is therefore optimal to set $w_2^*(0,1) = 0$. Note that negative shares are excluded by limited liability on behalf of the entrepreneur.

It is shown next that the factor in front of $w_2(1,1)$ in (5.18) is also negative. This factor can be rewritten as

$$\begin{aligned}
&\pi_R\left[1-3\pi_R-(1-2\pi_R)(\pi_R+\pi_b)\right] \\
&+(1-\pi_R)\left[\pi_R+(1-2\pi_R)(\pi_R+\pi_b)\right] \\
=&(1-2\pi_R)2\pi_R+(1-2\pi_R)^2(\pi_R+\pi_b) \\
\leq&(1-2\pi_R)2\pi_R+(1-2\pi_R)^2 2\pi_R,
\end{aligned} \tag{5.21}$$

where the last inequality holds because π_b is smaller than π_R. Since $-1 < 1-2\pi_R < 0$, this expression is negative. This proves that the factor of $w_2(1,1)$ is negative, providing negative incentives to the entrepreneur. In state (θ_b, θ_g), the last line of (5.21) is $(1-2\pi_g)2\pi_g + (1-2\pi_g)^2 2\pi_g$. This expression is also negative since $-1 < 1-2\pi_g < 0$. The investor therefore optimally chooses $w_2(1,1)^* = 0$. □

Although Lemma 6 reduces the investor's set of relevant shares, there are still many possibilities left. In order to decide which share provides incentives most effectively, their efficiency coefficient γ are calculated. Let $\gamma_w^s \equiv -IC_w^s/V_w$, where $s = b$ refers to the state (θ_b, θ_b), while $s = g$ refers to the state (θ_b, θ_g). The partial derivative of the incentive constraint with respect to the share w is denoted by IC_w and the partial derivative of the investor's objective function with respect to the same share w is denoted by V_w. Using (5.15) and (5.18), the relevant efficiency parameters γ_w^θ are

$$\begin{aligned}
\gamma_1^b &= \frac{2\pi_R}{2Pr(Y_1=2)} \\
\gamma_{0,2}^b &= \frac{2\pi_R(1-\pi_R)(1-\pi_R-\pi_b)}{2Pr(Y_1=0)Pr(Y_2=2|Y_1=0)} \\
\gamma_{1,2}^b &= \frac{2\pi_R\left[\pi_R(1-\pi_R-\pi_b)+(1-\pi_R)(\pi_R+\pi_b)\right]}{2Pr(Y_1=1)Pr(Y_2=2|Y_1=1)} \\
\gamma_{2,1}^b &= \frac{\pi_R\left[\pi_R(1-\pi_R-\pi_b)+(1-\pi_R)(\pi_R+\pi_b)\right]}{Pr(Y_1=2)Pr(Y_2=1|Y_1=2)} \\
\gamma_{2,2}^b &= \frac{2\pi_R^2(\pi_R+\pi_b)}{2Pr(Y_1=2)Pr(Y_2=2|Y_1=2)}
\end{aligned} \tag{5.22}$$

from the incentive constraint for state (θ_b, θ_b), and

$$\begin{aligned}
\gamma_1^g &= \frac{2\pi_g}{2Pr(Y_1=2)} \\
\gamma_{0,2}^g &= \frac{2\pi_g(1-\pi_g)(1-\pi_R-\pi_b)}{2Pr(Y_1=0)Pr(Y_2=2|Y_1=0)} \\
\gamma_{1,2}^g &= \frac{2\pi_g\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]}{2Pr(Y_1=1)Pr(Y_2=2|Y_1=1)} \\
\gamma_{2,1}^g &= \frac{\pi_g\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]}{Pr(Y_1=2)Pr(Y_2=1|Y_1=2)} \\
\gamma_{2,2}^g &= \frac{2\pi_g^2(\pi_R+\pi_b)}{2Pr(Y_1=2)Pr(Y_2=2|Y_1=2)}
\end{aligned} \tag{5.23}$$

from the incentive constraint for state (θ_b, θ_g).

Lemma 8 in the appendix shows that $Pr(Y_1 = 1)Pr(Y_2 = 2|Y_1 = 1) = Pr(Y_1 = 2)Pr(Y_2 = 1|Y_1 = 2)$. Therefore, $\gamma^b_{1,2} = \gamma^b_{2,1}$ and $\gamma^g_{1,2} = \gamma^g_{2,1}$, and attention can be limited to γ^s_1, $\gamma^s_{0,2}$, $\gamma^s_{1,2}$ and $\gamma^s_{2,2}$ with $s = g, b$.

For any given w, only the lower of the two values of γ is relevant, since a lower γ is associated with lower efficiency and since only the incentive constraint associated with the lower efficiency of any given share w is binding. By inspection of (5.22) and (5.23) and remembering that $\pi_g > \pi_R > \frac{1}{2}$, the set of relevant γ's can be narrowed down to

$$\begin{aligned}
\gamma_1 &:= \gamma^b_1 = \frac{\pi_g}{Pr(Y_1 = 2)} \\
\gamma_{0,2} &:= \gamma^g_{0,2} = \frac{\pi_g(1-\pi_g)(1-\pi_R-\pi_b)}{Pr(Y_1 = 0)Pr(Y_2 = 2|Y_1 = 0)} \\
\gamma_{1,2} &:= \gamma^g_{1,2} = \frac{\pi_g\left[\pi_g(1-\pi_R-\pi_b) + (1-\pi_g)(\pi_R+\pi_b)\right]}{Pr(Y_1 = 1)Pr(Y_2 = 2|Y_1 = 1)} \\
\gamma_{2,2} &:= \gamma^b_{2,2} = \frac{\pi_g^2(\pi_R+\pi_b)}{Pr(Y_1 = 2)Pr(Y_2 = 2|Y_1 = 2)}.
\end{aligned} \tag{5.24}$$

The investor will choose the share with the highest efficiency parameter γ. This choice depends on the parameters B, π_g, π_R, π_b and p. No single value of γ is maximum for all conceivable parameter configurations. Note that B only affects the level of the incentive constraints and does not influence the relative attractiveness of the different cash flow shares. If B were zero, the incentive constraints are trivially satisfied, and there are no incentive problems associated with the outside financing of the conglomerate. All positive NPV projects are financed. Increasing B results in higher cash flow shares for the HQ and hence to incentive problems. If B is increased above a certain threshold value, these problems become too serious, causing the project to become nonviable. The following proposition summarizes the previous analysis and specifies the contract an investor will optimally choose to provide HQ with incentives to reorganize his own project.

Proposition 4. *An outside investor uses the following optimal contract to provide reorganization incentives to the HQ of a conglomerate:*

$$\begin{aligned}
w_1^*(2) &= \frac{B}{\pi_R(\pi_R-\pi_b)} && \text{if } \gamma_1 = \gamma_{max} \\
w_2^*(0,2) &= \frac{B}{(1-\pi_g)(1-\pi_R-\pi_b)(\pi_R-\pi_b)} && \text{if } \gamma_{0,2} = \gamma_{max} \\
w_2^*(1,2) &= \frac{B}{\pi_g\left[(1-2\pi_g)(\pi_R+\pi_b)+\pi_g\right](\pi_R-\pi_b)} && \text{if } \gamma_{1,2} = \gamma_{max} \\
w_2^*(2,2) &= \frac{B}{\pi_g^2(\pi_R+\pi_b)(\pi_R-\pi_b)} && \text{if } \gamma_{2,2} = \gamma_{max},
\end{aligned}$$

where $\gamma_{max} := \max\{\gamma_1, \gamma_{0,2}, \gamma_{1,2}, \gamma_{2,2}\}$. *Her expected return is*

$$V^{c*} = 4E(y) - R^*(p, \pi_b, \pi_R, \pi_g, B),$$

where

$$R^*(p, \pi_b, \pi_R, \pi_g, B) = 2 \cdot \begin{cases} w_1^*(2)Pr(Y_1 = 2) & \text{if } \gamma_1 = \gamma_{max} \\ w_2^*(0,2)Pr(Y_1 = 0)Pr(Y_2 = 2|Y_1 = 2) & \text{if } \gamma_{0,2} = \gamma_{max} \\ w_2^*(1,2)Pr(Y_1 = 1)Pr(Y_2 = 2|Y_1 = 2) & \text{if } \gamma_{1,2} = \gamma_{max} \\ w_2^*(2,2)Pr(Y_2 = 2)Pr(Y_2 = 2|Y_1 = 2) & \text{if } \gamma_{2,2} = \gamma_{max}, \end{cases}$$

denotes HQ's expected share of cash flow.

Proof. Incentives are most effectively provided through the share w with the highest γ.

If γ_1 is maximum, first-period share $w_1(2)$ is most efficient. In this case (5.24) shows that the incentive constraint for state (θ_b, θ_b) is binding. From (5.18), we have $w_1^*(2) = \frac{B}{\pi_R(\pi_R - \pi_b)}$.

If $\gamma_{0,2}$ is maximum, second-period share $w_2(0,2)$ is most efficient. In this case (5.24) shows that the incentive constraint for state (θ_b, θ_g) is binding. From (5.19), one has $w_2^*(0,2) = \frac{B}{(1-\pi_g)(1-\pi_R-\pi_b)(\pi_R-\pi_b)}$.

If $\gamma_{1,2}$ is maximum, second-period share $w_2(1,2)$ is most efficient. In this case (5.24) shows that the incentive constraint for state (θ_b, θ_g) is binding. From (5.19), it follows that $w_2^*(0,2) = \frac{B}{\pi_g\left[(1-2\pi_g)(\pi_R+\pi_b)+\pi_g\right](\pi_R-\pi_b)}$.

Finally, if $\gamma_{2,2}$ is highest, second-period share $w_2(2,2)$ is most efficient. In this case (5.24) shows that the incentive constraint for state (θ_b, θ_b) is binding. From (5.18), $w_2^*(2,2) = \frac{B}{\pi_g^2(\pi_R+\pi_b)(\pi_R-\pi_b)}$ follows.

The investor's maximum expected return is calculated by plugging the optimal cash flow share into the investor's objective function (5.15). Note that $\sum_{h=0}^{2}\sum_{h'=0}^{2} Pr(Y_1 = h)Pr(Y_2 = h'|Y_1 = h)h' = \sum_{h=0}^{2} Pr(Y_1 = h)h = 2E(y)$. Maximum expected return is thus equal to $4E(y)$ minus entrepreneur's expected share of cash flow $R^*(p, \pi_b, \pi_R, \pi_g, B)$ as defined in the proposition.

This proofs the proposition. □

So far only HQ's incentives with regard to its own project were considered. Proposition 4 takes it for granted that the PM reorganizes his own project. In order to isolate HQ's incentives, it was assumed that the PM behaves optimally. A conglomerate, however, needs to take its internal agency problem between HQ and PM serious, making sure that the PM has an incentive to reorganize his project.

However, the financial contract between investor and HQ studied so far does not provide any such incentive to the PM.

The next section analyzes the relationship between HQ and PM. As it turns out, the above financial contract between HQ and investor remains valid, but the investor will have to provide some extra liquidity in excess of $2I$.

5.3.2 Informal Employment Contract Between HQ and PM

As argued above, conglomerate financing is disadvantageous compared to independent project financing in providing incentives because financial contracts can only be conditioned on total conglomerate cash flow. For conglomerate financing to be an attractive alternative at all, HQ and PM must take advantage of their superior knowledge. Since an informed party observes the state of the project as well as action chosen, an *informal* contract conditioning on these two variables is possible as long as this contract is self-enforcing.

Such an employment contract has the HQ "employing" the PM and commanding him to run the project. In order to induce PM to reorganize a bad project, HQ can threaten to fire him if he does not behave. If fired, the PM is separated from his project and receives nothing. HQ on the other hand may be able to liquidate the project.

An employment contract defines a transfer $z(\theta, a)$ from HQ to PM depending on the PM's state of the project and choice of action. Because the game ends at $t = 2$, there will be no transfers at $t = 2$. All transfers to the PM must thus be made at $t = 1$. Analogous to Assumption 1, it is assumed that the HQ has all the bargaining power vis-à-vis the PM and that the investor has all the bargaining power vis-à-vis the HQ. Since PM needs only be compensated for his loss of private benefits in state θ_b, HQ will choose $z(\theta_g, C) = z(\theta_g, R) = z(\theta_b, C) = 0$ and $z(\theta_b, R) = 2B$.

The conglomerate itself being liquidity-constrained, it must secure itself funds in addition to investment costs $2I$ in order to be able to make payment $2B$ to the PM. Otherwise, HQ's share of cash flow at $t = 1$ might not be sufficient to pay PM. More importantly, if HQ had to finance the transfer $z(\theta_b, R)$ out of its share of cash flow, then the conglomerate will be less profitable than two independent projects. The reason is that mitigating HQ's liquidity problem by increasing its cash flow shares nullifies its informational advantage because claims to cash flow cannot be made contingent on the states realized or action chosen.

Because the optimal financial contract defined in Proposition 4 does not allow for positive cash flow shares in period one, the funds needed to compensate the PM must be made available to the HQ by the investor. The minimally required compensation is $2B$. The investor therefore needs to finance the conglomerate with $2I+2B$ to guarantee investment ($2I$) and excess liquidity for the case the PM has to be paid the wage $2B$ at $t = 1$. It is reasonable to assume that such wage payments are verifiable. After all, conglomerate cash flow was assumed to be verifiable as well. If HQ does not pay this wage, the investor will therefore reclaim $2B$ at $t = 1$ from the HQ.

For the HQ to have an incentive to pay the wage to the PM at $t = 1$, the return from making this transfer must be larger than the return from liquidating the project. If the PM reorganized his bad project, then by paying him $2B$ at $t = 1$, HQ secures itself a higher expected conglomerate cash flow for $t = 2$. Would HQ not make the payment, the PM quits and the project must be liquidated, yielding lower expected cash flow.

Remember that the optimal contract in Proposition 4 guarantees the HQ cash flow shares contingent on first-period cash flow being equal to 2. If $Y_1 < 2$ HQ knows that it will receive no cash flow at all in $t = 2$ and is as a consequence indifferent between paying the PM and not paying him. This difficulty is circumvented by granting an arbitrarily small but positive share of cash flow to the HQ for any level of first-period cash flow realization. HQ will therefore pay the wage if the PM reorganizes his bad project. Note that HQ cannot divert excess liquidity to itself in case it does not pay the wage to the PM since it has to repay all excess liquidity to the investor at $t = 1$.

On the other hand, HQ will fire the PM and liquidate the project, if he does not reorganize a bad project. If he did not liquidate the project, it would be successful with probability π_b, whereas liquidating yields a success probability of $\pi_0 \geq \pi_b$. HQ will therefore fire the PM if he does not reorganize his bad project.

Finally, the PM weakly prefers to quit at $t = 1$ if he does not receive wage $2B$ after having reorganized a bad project. Whether he stays or quits, the PM receives nothing either way. Similarly he weakly prefers to stay after having received $2B$ for reorganizing a bad project. He receives nothing at $t = 2$ whether he stays or quits.

There is some ambiguity with respect to the fraction of the liquidity grant $2B$ repaid to the investor. The contract forces HQ to pay the $2B$ either to his PM or to the investor. As argued above, HQ will pay the $2B$ to his PM in case the PM reorganized a bad project. But what happens if the PM's project was in the good state? In this case, HQ is indifferent between paying the PM or paying the investor. If HQ's expected second-period share of profit at $t = 1$ is positive, then the PM could try to extract the $2B$ from HQ. The PM can threaten to quit[12], in which case HQ has to liquidate the project wherefrom he earns a fraction of π_0. Since this is smaller than π_R while HQ is indifferent between paying and not paying $2B$ to the PM, he will cave in to PM's threat and pay him $2B$.

Furthermore, it seems reasonable to expect HQ to make the payment $2B$ to his PM rather than to the outside investor in case of indifference. After all, HQ has much closer ties to his employees than to the investor causing him to weigh PM's preferences more strongly than those of an outsider. [13]

Total investment for financing the conglomerate therefore amounts to $2I + 2B$. With independent project financing, it is only $2I$; on the other hand the investor has to provide higher shares of cash flow to the single entrepreneur.

The investor's expected return from the contract with the conglomerate is

$$4E(y) - 2B - R^*\left(p, \pi_b, \pi_R, \pi_g, b\right), \tag{5.25}$$

where $R^*\left(p, \pi_b, \pi_R, \pi_g, B\right)$ denotes HQ's expected cash flow share, as derived in Proposition 4.

5.3.3 Conglomerate Financing Compared to Independent Financing

As Proposition 4 demonstrates, no clear-cut result for the investor's problem is available without limiting the model's parameter space. In order to illustrate the results, private benefits B are fixed at 0.2. Furthermore, attention is limited to four different sets of parameters:

$$\begin{aligned} (a) &:= \left\{p = 0.5, \pi_R = 0.6, \pi_g = 0.9\right\} \\ (b) &:= \left\{p = 0.5, \pi_R = 0.8, \pi_g = 0.9\right\} \\ (c) &:= \left\{p = 0.25, \pi_R = 0.6, \pi_g = 0.9\right\} \\ (d) &:= \left\{p = 0.25, \pi_R = 0.8, \pi_g = 0.9\right\}. \end{aligned} \tag{5.26}$$

This restriction allows to analyze the effect of a change in the a priori probability of the project being in the good state (p) as well as the effect of an increase of the

[12]The PM's threat is weakly credible, since absent any wage transfer he earns nothing from either staying or quitting.

[13]See Hart and Holmstrom (2002), where a similar assumption is made.

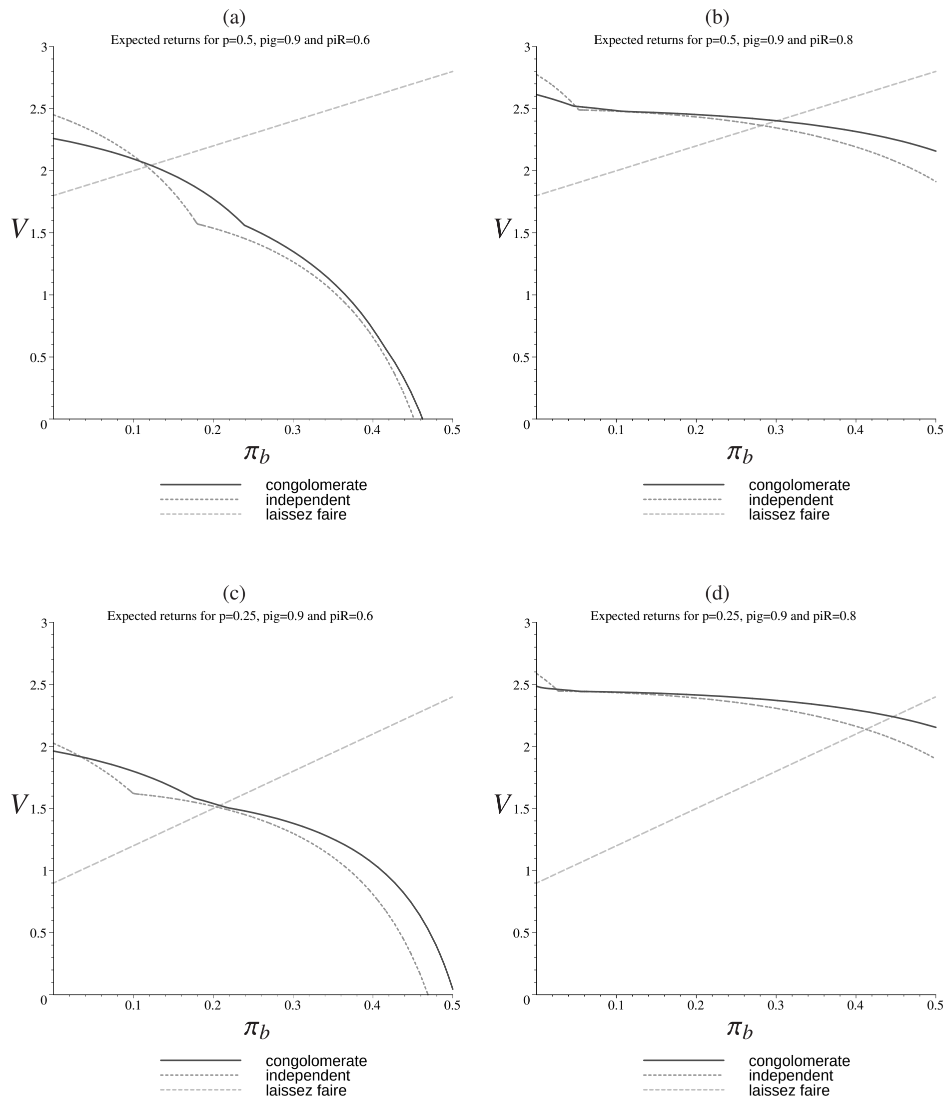


Figure 5.3: Comparison of expected returns V between independent financing and conglomerate financing for four different sets of parameters p, π_g (pig) and π_R (piR), see (5.26). Private benefits are fixed at $B = 0.2$.

difference $\pi_g - \pi_R$. For each set of parameters, expected returns from independent financing is compared to expected returns from conglomerate financing at varying levels of π_b.[14] The results are illustrated in Figure 5.3. The kinks in the investor's expected return functions are due to the piecewise nature of the optimal contracts as derived in Propositions 3 and 4.

Standard Moral Hazard. With the a priori probability of the project being successful $p = 0$, the contracting parties are certain to find themselves in the bad state. The contracting problem then reduces to a standard moral hazard problem with a risk neutral agent "protected" by limited liability. Financing both projects jointly via a conglomerate relaxes this limited liability constraint. Even if HQ's own project produces zero cash flow, HQ can still be punished by taking him away any cash flow generated by the other project.[15] With single project financing this is not possible because the entrepreneur must only worry about losing the cash flow generated by his own project. A decrease of p therefore boosts the advantage of the conglomerate ceteris paribus.

Probability of Success Identical in Both Project States ($\pi_R = \pi_g$). Note that incentive constraints are identical for states of the conglomerate (θ_b, θ_b) and (θ_b, θ_g) if the probability of success is identical in each project-state, i.e. if $\pi_R = \pi_g$. Moreover, the next lemma shows that conglomerate financing is always preferred to independent financing if $\pi_R = \pi_g$.

Lemma 7. *If $\pi_R = \pi_g$, the uninformed investor prefers conglomerate financing to independent financing if $\pi_b > 0$. Expected returns from conglomerate financing is given by*

$$V^c = 4\pi_g - \left[1 + \frac{\pi_g^2}{\pi_g^2 - \pi_b^2}\right] 2B.$$

Proof. With $\pi_R = \pi_g$, the optimal contract for financing a conglomerate is unambiguously given by $w_2^*(2,2) = b/\pi_g^2(\pi_g^2 - \pi_b^2)$, see Proposition 4. All other shares

[14] The γ's can be interpreted as functions of π_b Note that γ_1 is independent of π_b, $\gamma_{2,2}$ is increasing in π_b while $\gamma_{0,2}$ and $\gamma_{1,2}$ are both decreasing in π_b.

[15] This effect is reminiscent of Laux (2001), where the advantage of financing multiple projects jointly stems from a relaxed limited liability constraint.

are equal to zero. Investor's expected returns are

$$
\begin{aligned}
V^c &= 4E(y_1) - w_2^*(2,2)E(y_1)^2 E(y_2|y_1 = 1)^2 2 - 2B \\
&= 4\pi_g - \frac{\pi_g^4 + \pi_g^2(\pi_g^2 - \pi_b^2)}{\pi_g^2(\pi_g^2 - \pi_b^2)} 2B.
\end{aligned}
\tag{5.27}
$$

Note that $E(y) = E(y_2|y_1 = 1) = \pi_g$. If the projects are financed independently, the optimal contract requires $w_2^*(1) = 2B/(\pi_g^2 - \pi_b^2)$ and zero for all other shares. Investor's expected returns are

$$
\begin{aligned}
2 \cdot V^s &= 2 \cdot \left[2E(y_1) - w_2^*(1)E(y_1)E(y_2|y_1 = 1)\right] \\
&= 4\pi_g - \frac{2\pi_g^2}{\pi_g^2 - \pi_b^2} 2B.
\end{aligned}
\tag{5.28}
$$

Expected returns from financing a conglomerate are greater than expected return from independent financing, if

$$
\begin{aligned}
\frac{\pi_g^4 + \pi_g^2(\pi_g^2 - \pi_b^2)}{\pi_g^2(\pi_g^2 - \pi_b^2)} &< \frac{2\pi_g^2}{\pi_g^2 - \pi_b^2} \\
\Leftrightarrow 2\pi_g^2 - \pi_b^2 &< 2\pi_g^2,
\end{aligned}
\tag{5.29}
$$

which is always satisfied as long as $\pi_b > 0$. This proves Lemma 7. □

Lemma 7 makes clear that for independent financing to be preferable to conglomerate financing at all, the success probability for project-state θ_b must be strictly smaller than the success probability for project-state θ_g.

The intuition is the following. With $\pi_R = \pi_g$, the probability of success does not depend on the realized state. The only difference between the two states is that the entrepreneur forgoes some private benefits in the bad state and must be compensated accordingly. Even if the bad state is realized, the project is as profitable as in the good state. In contrast, $\pi_R < \pi_g$ represents a different environment. If the bad state is realized, the project's probability of success is lower. This case might represent a project which is reasonably profitable if the business environment develops as assumed in the project's business plan but where it is ambiguous what happens if the world develops differently. The entrepreneur still looses his private benefits in the bad state, but in addition the project is now less profitable. In such an environment, low cash flow is a less informative signal for the behavior of both the single entrepreneur and HQ than if $\pi_R = \pi_g$. Since $w_2^*(2,2)$ only rewards HQ if the project was very successful (highest cash flow realized in both periods), it is generally not optimal to use $w_2(2,2)$ if $\pi_R < \pi_g$.

But if incentives are no longer provided by $w_2(2,2)$ then HQ can profit from the PM's project being successful. Even if his own project delivers zero cash flow, PM's project might still deliver positive cash flow. This reduces HQ's incentive to reorganize. Such a "cross-subsidy" is not available with independent project financing, thus boosting its advantage relative to conglomerate financing.

Increasing π_b. Remember that π_b is a bad project's expected return if it is not reorganized. The higher π_b, the smaller the loss associated with a non-performing entrepreneur. However, in this case the investor must offer a higher expected share of cash flow to the entrepreneur in order to induce him to reorganize. In case of a conglomerate, the informal employment contract between HQ and PM causes the cost of inducing the PM to reorganize to be independent of π_b. Rather it is constant equal to $2B$. An increase in π_b therefore eventually favors conglomerates.

Note also that an increase in π_b increases the attractiveness of a laissez-faire contract (see section 5.2.1) which does not provide incentives to reorganize at all. For a sufficiently high value of π_b, such a contract dominates both conglomerate and independent project financing. This is illustrated by the linearly increasing line in Figure 5.3.

In sum, the smaller the difference $\pi_g - \pi_R$ and the smaller the a priori probability of the good project-state (p), the more attractive is conglomerate financing. A third important finding is that projects optimally financed independently exhibit higher expected return than projects optimally financed as conglomerates. This prediction is consistent with empirical findings which suggest that conglomerates trade at a discount (Berger and Ofek, 1995; Lang and Stulz, 1994). However, the reason for the discount is not that conglomerates are less efficient than independently financed projects. Rather the reason is selectivity. This is in line with arguments put forth by Fluck and Lynch (1999) and Graham et al. (2002) that an important reason for the observed conglomerate discount might be due to endogeneity.

5.4 Conclusion

This paper presents a model based on Aghion and Bolton (1992), where pennyless entrepreneurs need to contract for funds with investors to start their projects. A project lasts two periods and requires different action choices in different states. Entrepreneurs' private benefits associated with a specific action (continuation of the project) generates a costly moral hazard problem. The optimal contract specifies claims to cash flow for the entrepreneur devised to guarantee the that investor's preferred action is chosen. Since the investor can no longer appropriate all cash flow, not all projects with positive net present value will be financed. Entrepreneurs are thus credit-constrained.

These credit constraints can be mitigated through "conglomeration". Conglomeration refers to two entrepreneurs financing their projects jointly. One acting as the conglomerate's headquarters, the other as the project manager. By treating entrepreneurs inside and outside a conglomerate symmetric in terms of capabilities, no superior abilities are assumed for headquarters. Both projects require a founding entrepreneur, and entrepreneurs enjoy private benefits from continuing rather than reorganizing their projects. The sole potential advantage of a conglomerate stems from the superior information of entrepreneurs with regard to the quality of their projects.

In contrast to existing literature this paper shows how and under which conditions the superior information of headquarters increases the investor's expected return. The difficulty is that although headquarters has more information, it is not verifiable to outside parties. Hence it is not guaranteed that this informational advantage is of any use to headquarters. The solution proposed is that informal contracting is available inside the conglomerate whereas an external investor has to rely on formal contracting alone.

Since headquarters and product manager share the same symmetric information, they can contract on an informal basis and need not rely on verifiable contracts. Due to the financial contract entered with the investor, headquarters has an interest in project managers behaving in a cash flow maximizing manner, using an informal employment contract to induce project managers to behave. The advantage of forming a conglomerate thus is that it enables the investor to benefit from headquarters superior information about the quality of the project run by the project manager.

The disadvantage of forming a conglomerate is that the investor can no longer contract on the cash flows of individual projects. While total cash flow is observable and verifiable, headquarters can transfer funds internally at his discretion, making individual cash flows no longer verifiable to third parties (namely, courts). This results in a loss of information to the investor causing the cost of providing incentives to increase.

The tradeoff between benefits from headquarters' superior information and information loss due to cash flow pooling permits to formulate predictions concerning the types of projects that preferably are financed independently and which are better financed as a conglomerate. For projects characterized by a high a prior probability of the good state and by low expected return in case a bad project is not reorganized, independent financing is preferable. In other words, if the agency problem is not too strong, independent financing is superior to conglomerate financing.

Furthermore, it turns out that projects that are preferably financed independently deliver higher expected returns to the investor than projects which are preferably financed as conglomerates. This supports the hypothesis put forth in the literature (see, e.g., Fluck and Lynch, 1999; Graham et al., 2002), that the observed conglomerate discount is due to selection by investors rather than to conglomerates being inherently less efficient than stand-alone firms.

Appendix

Lemma 8. *Some properties of joint probabilities:*
(i):

$$\begin{aligned}
Pr(Y_2 = 0|Y_1 = 1) &= Pr(y_2 = 0|y_1 = 1)Pr(y_2 = 0|y_1 = 0)\\
Pr(Y_2 = 0|Y_1 = 2) &= Pr(y_2 = 0|y_1 = 1)^2\\
Pr(Y_2 = 1|Y_1 = 1) &= Pr(y_1 = 1|y_2 = 0)Pr(y_2 = 0|y_1 = 1)\\
&\quad + Pr(y_2 = 0|y_1 = 0)Pr(y_2 = 1|y_1 = 1)\\
Pr(Y_2 = 1|Y_1 = 2) &= 2Pr(y_2 = 1|y_1 = 1)Pr(y_2 = 0|y_1 = 1)\\
Pr(Y_2 = 2|Y_1 = 0) &= Pr(y_2 = 1|y_1 = 0)^2\\
Pr(Y_2 = 2|Y_1 = 1) &= Pr(y_2 = 1|y_1 = 1)Pr(y_2 = 1|y_1 = 0)\\
Pr(Y_2 = 2|Y_1 = 2) &= Pr(y_2 = 1|y_1 = 1)^2.
\end{aligned}$$

(ii):

$$Pr(y_1 = 1)Pr(y_2 = 0|y_1 = 1) = Pr(y_1 = 0)Pr(y_2 = 1|y_1 = 0). \tag{5.30}$$

Proof. Proof of Part (i): Remember that $Y_t = y_t^1 + y_t^2$ and that y^1 and y^2 are independent. Then

$$\begin{aligned}
Pr(Y_2 = 2|Y_1 = 1) &= Pr(Y_2 = 2|(y_1^1 = 1 \text{ and } y_1^2 = 0) \text{ or } (y_1^1 = 0 \text{ and } y_1^2 = 1))\\
&= \tfrac{1}{2}Pr(Y_2 = 2|(y_1^1 = 1 \text{ and } y_1^2 = 0)) + \tfrac{1}{2}Pr(Y_2 = 2|y_1^1 = 0 \text{ and } y_1^2 = 1)\\
&= Pr(y_2 = 1|y_1 = 1)Pr(y_2 = 1|y_1 = 0).
\end{aligned}$$

Analogously, $Pr(Y_2 = 0|Y_1 = 1) = Pr(y_2 = 0|y_1 = 1)Pr(y_2 = 0|y_1 = 0)$. Furthermore,

$$\begin{aligned}
Pr(Y_2 = 1|Y_1 = 2) &= Pr(Y_2 = 1|(y_1^1 = 1 \text{ and } y_1^2 = 1)\\
&= Pr(y_2^1 = 1|y_1^1 = 1)Pr(y_2^2 = 0|y_1^2 = 1)\\
&+ Pr(y_2^1 = 0|y_1^1 = 1)Pr(y_2^2 = 1|y_1^2 = 1)\\
&= 2Pr(y_2 = 1|y_1 = 1)Pr(y_2 = 0|y_1 = 1)
\end{aligned}$$

and

$$\begin{aligned}
Pr(Y_2 = 2|Y_1 = 2) &= Pr(Y_2 = 2|y_1^1 = 1 \text{ and } y_1^2 = 1)\\
&= Pr(y_2^1 = 1|y_1^1 = 1 \text{ and } y_1^2 = 1) \cdot Pr(y_2^2 = 1|y_1^1 = 1 \text{ and } y_1^2 = 1)\\
&= Pr(y_2^1 = 1|y_1^1 = 1) \cdot Pr(y_2^2 = 1|y_1^2 = 1)\\
&= Pr(y_2 = 1|y_1 = 1)^2.
\end{aligned}$$

Analogously,

$$Pr(Y_2 = 2|Y_1 = 0) = Pr(y_2 = 1|y_1 = 0)^2 \quad \text{and } Pr(Y_2 = 0|Y_1 = 0) \quad = Pr(y_2 = 0|y_1 = 0)^2.$$

Finally,

$$\begin{aligned} Pr(Y_2 = 1|Y_1 = 1) &= Pr(Y_2 = 1|(y_1^1 = 1 \text{ and } y_1^2 = 0) \text{ or } (y_1^1 = 0 \text{ and } y_1^2 = 1)) \\ &= \tfrac{1}{2}Pr(Y_1 = 2|(y_1^1 = 1 \text{ and } y_1^2 = 0)) + \tfrac{1}{2}Pr(Y_2 = 1|y_1^1 = 0 \text{ and } y_1^2 = 1) \\ &= Pr(y_2 = 1|y_1 = 1)Pr(y_2 = 0|y_1 = 0) \\ &\quad + Pr(y_2 = 0|y_1 = 1)Pr(y_2 = 1|y_1 = 0). \end{aligned}$$

This proof part (i).

For part (ii), note that

$$\begin{aligned} &Pr(y_1 = 1)Pr(y_2 = 0|y_1 = 1) \\ &\quad = Pr(y_1 = 1)\left(\frac{p\pi_g}{Pr(y_1 = 1)}(1 - \pi_g) + \frac{(1 - p)\pi_R}{Pr(y_1 = 1)}(1 - \pi_R)\right) \\ &\quad = p\pi_g(1 - \pi_g) + (1 - p)\pi_R(1 - \pi_R). \end{aligned}$$

and that

$$\begin{aligned} &Pr(y_1 = 0)Pr(y_2 = 1|y_1 = 0) \\ &\quad = Pr(y_1 = 0)\left(\frac{p(1 - \pi_g)}{Pr(y_1 = 0)}\pi_g + \frac{(1 - p)(1 - \pi_R)}{Pr(y_1 = 0)}\pi_R\right) \\ &\quad = p\pi_g(1 - \pi_g) + (1 - p)\pi_R(1 - \pi_R), \end{aligned}$$

which proofs the claim in the lemma. □

Conglomerate IC for state (θ_b, θ_g) HQ's return from reorganizing in state (θ_b, θ_g), assuming the PM reorganizes his project, is:

$$\begin{aligned} &\pi_g\pi_R w_1(2) \cdot 2 \\ &+(1 - \pi_g)(1 - \pi_R)\left\{[\pi_g(1 - \pi_R) + \pi_R(1 - \pi_g)]w_2(0, 1) \cdot 1 + \pi_R\pi_g w_2(0, 2) \cdot 2\right\} \\ &+[\pi_g(1 - \pi_R) + (1 - \pi_g)\pi_R]\left\{[\pi_g(1 - \pi_R) + \pi_R(1 - \pi_g)]w_2(1, 1) \cdot 1 + \pi_R\pi_g w_2(1, 2) \cdot 2\right\} \\ &+\pi_R\pi_g\left\{[\pi_g(1 - \pi_R) + \pi_R(1 - \pi_g)]w_2(2, 1) \cdot 1 + \pi_R\pi_g w_2(2, 2) \cdot 2\right\} \end{aligned} \tag{5.31}$$

If HQ does not reorganize, his return is

$$
\begin{aligned}
&\pi_g \pi_b w_1(2) \cdot 2 \\
&+(1-\pi_g)(1-\pi_b)\left\{[\pi_g(1-\pi_b)+\pi_b(1-\pi_g)]w_2(0,1)\cdot 1+\pi_b\pi_g w_2(0,2)\cdot 2\right\} \\
&+[\pi_g(1-\pi_b)+(1-\pi_g)\pi_b]\left\{[\pi_g(1-\pi_b)+\pi_b(1-\pi_g)]w_2(1,1)\cdot 1+\pi_b\pi_g w_2(1,2)\cdot 2\right\} \\
&+\pi_g\pi_b\left\{[\pi_g(1-\pi_b)+\pi_b(1-\pi_g)]w_2(2,1)\cdot 1+\pi_b\pi_g w_2(2,2)\cdot 2\right\} \\
&+2B
\end{aligned}
\tag{5.32}
$$

He will therefore reorganize, if (5.31) is larger than (5.32), i.e. if

$$
\begin{aligned}
&2\pi_g \cdot w_1(2) \\
&+(1-\pi_g)\left[\pi_g(\pi_R+\pi_b-2)+(1-\pi_g)(1-\pi_R-\pi_b)\right]\cdot w_2(0,1) \\
&+2(1-\pi_g)\pi_g(1-\pi_R-\pi_b)\cdot w_2(0,2) \\
&+\left\{\pi_g\left[\pi_g(\pi_R+\pi_b-2)+(1-\pi_g)(1-\pi_R-\pi_b)\right]\right. \\
&\left.+(1-\pi_g)\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]\right\}\cdot w_2(1,1) \\
&+2\pi_g\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]\cdot w_2(1,2) \\
&+\pi_g\left[\pi_g(1-\pi_R-\pi_b)+(1-\pi_g)(\pi_R+\pi_b)\right]\cdot w_2(2,1) \\
&+2\pi_g^2(\pi_R+\pi_b)\cdot w_2(2,2) \\
&\geq 2B/(\pi_R-\pi_b)
\end{aligned}
\tag{5.33}
$$

Bibliography

Aghion, P. and P. Bolton (1992). An incomplete contracts approach to financial contracting. *The Review of Economics Studies 59*, 473–494.

Berger, P. and E. Ofek (1995). Diversification's effect on firm value. *Journal of Financial Economics 37*, 39–65.

Bertrand, M. and S. Mullainathan (2003). Enjyoing the quiet life? corporate governance and managerial preferences. *Journal of Political Economy 111*(5), 1043–1075.

Brusco, S. and F. Panunzi (2005). Reallocation of corporate resources and managerial incentives in internal capital markets. *European Economic Review 49*(3), 659–681.

Dybvig, P. H. and J. F. Zender (1991). Capital structure and dividend irrelevance with asymmetric information. *The Review of Financial Studies 4*, 201–219.

Fluck, Z. and A. W. Lynch (1999). Why do firms merge and then divest? a theory of financial synergy. *Journal of Business 72*, 319–346.

Gertner, R. H., D. S. Scharfstein, and J. C. Stein (1994). Internal versus external capital markets. *Quarterly Journal of Economics 109*, 1211–1230.

Graham, J. R., M. L. Lemmon, and J. G. Wolf (2002). Does corporate diversifiaction destroy value? *The Journal of Finance 57*, 695–720.

Harris, M. and A. Raviv (1991). The theory of capital structure. *The Journal of Finance 46*, 297–355.

Hart, O. (2001). Financial contracting. *Journal of Economic Literature 39*(4), 1079–1100.

Hart, O. and B. Holmstrom (2002). Vision and firm scope. Working paper.

Hubbard, G. and D. Palia (1999). A reexamination of the conglomerate merger wave in the 1960s: An internal capital markets view. *Journal of Finance 54*, 1131–1152.

Inderst, R. and H. M. Müller (2003). Internal versus external financing: An optimal contracting approach. *The Journal of Finance 58*, 1033–1062.

Kaplan, S. N. and P. Strömberg (2001). Venture capitalsits as principals: Contracting, screening and monitoring. *American Economic Review 91*(2), 45–50. Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association.

Kaplan, S. N. and P. Strömberg (2003). Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *Review of Economic Studies 70*, 281–315.

Lang, L. and R. Stulz (1994). Tobin's *q*, corporate diversification and firm performance. *The Journal of Political Economy 102*, 1248–1280.

Laux, C. (2001). Limited-liability and incentive contracting with multiple projects. *Rand Journal of Economics 32*, 514–526.

Sahlman, W. A. (1990). The structure and governance of venture-capital organizations. *Journal of Financial Economics 27*, 473–521.

Shin, H.-H. and R. M. Stulz (1998). Are internal capital markets efficient? *The Quarterly Journal of Economics 113*, 531–552.

Stein, J. C. (1997). Internal capital markets and the competition for corporate ressources. *The Journal of Finance 52*, 111–133.

Zingales, L. (2000). In search of new foundations. *The Journal of Finance 55*, 1623–1653.

Chapter 6

Conclusion

The conclusion is confined to a brief discussion of possible extensions to the models presented in chapters 4 and 5.

The third essay on switching costs provides interesting testable predictions. First, prices of chain stores are predicted to be higher than prices of comparable local stores. The hotel industry is a good candidate to empirically test this prediction. Constructing a dataset of hotels in a given city region (e.g. Manhattan, New York) and regressing room rates on characteristics such as quality, location and an indicator for chain stores would provide an empirical test for the alleged price pattern. Second, as mobility increases the model predicts increasing market share for chain stores relative to local stores. If increased demand is met by an increased number of outlets, the model predicts that more chain store outlets relative to local stores are observed along axes of high mobility. Here, the coffee-shop industry (Starbucks) would provide a potentially fruitful field for research.

An important drawback of the conglomerate model in chapter 5 is that control over actions is always allocated to the entrepreneur. Since the investor cannot run or reorganize projects herself, her only option available is to liquidate them. Therefore, she prefers the entrepreneur to be in charge of the project and to devise an optimal incentive contract, such that bad projects are always reorganized by the entrepreneur. This contrasts with reported evidence from the venture capital industry (see, e.g., Sahlman, 1990; Kaplan and Strömberg, 2001), which states that allocation of control rights contingent on performance is an important governance instrument. An interesting extension of the model presumably leading to contingent control allocation would be to allow for different types of entrepreneurs. If only a fraction of the population of entrepreneurs is competent and thus capable of reorganizing bad projects, control rights are relevant. It is conjectured that in such a world control would be allocated to the investor if first period cash flow was low. The investor would then replace the existing entrepreneur by a new one or liquidate the project.

Bibliography

Kaplan, S. N. and P. Strömberg (2001). Venture capitalsits as principals: Contracting, screening and monitoring. *American Economic Review 91*(2), 45–50. Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association.

Sahlman, W. A. (1990). The structure and governance of venture-capital organizations. *Journal of Financial Economics 27*, 473–521.

Lebenslauf

Yves Schneider, geboren am 16. Juli 1974 in Biel wuchs in Lyss auf wo er die Primar- und Sekundarschule besuchte. Nach dem Gymnasium in Biel und dem Studium der Wirtschaftwissenschaften in Bern begann er 2000 das Doktorat an der Wirtschaftswissenschaftlichen Fakultät der Universität Zürich. Während dieser Zeit war er auch Verwaltungsratsmitglied im elterlichen Grosshandelsbetrieb, in welchem er zudem während knapp einem Jahr für diverse Projekte verantwortlich war.