Essays in Information Economics

Manuscript committee: Prof. dr. Janneke Plantenga
Prof. dr. Jeroen Hinloopen
Prof. dr. Jeroen de Jong
Prof. dr. Robert Dur
Prof. dr. Marco Haan

ISBN 978-94-91870-25-5
Tjalling C. Koopmans Dissertation Series
USE 040
Printed by Ridderprint, Ridderkerk

Essays in Information Economics

Essays in de Informatie Economie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus,
prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
woensdag 31 mei 2017 des middags te 12.45 uur

door

Najmeh Rezaei Khavas
geboren op 2 september 1976 te Tehran, Iran

Promotor: Prof. dr. S. Rosenkranz
Copromotor: Dr. K.J.M. De Jaegher

Acknowledgments

I would like to sincerely thank my supervisors Prof. Dr. Stephanie Rosenkranz and Dr. Kris De Jaegher for giving me the opportunity to start my academic career at USE. I greatly appreciate all the time both of you invested in discussing and revising my thesis. Stephanie, your encouragements were sometimes all that kept me going. Thank you very much for your genuine care of others and your consideration of their differences. Kris, your maddening attention to details drove me to become a stronger writer. I cannot thank you enough for everything you taught me, from analytical thinking to rigorous academic writing.

During my PhD, I spent two years at UCLA as a visiting graduate researcher. I would like to warmly thank my host professor Prof. Marek Pycia for providing me with this valuable opportunity.

I would also like to thank the members of my reading committee, Prof. dr. Janneke Plantenga, Prof. dr. Jeroen Hinloopen, Prof. dr. Jeroen de Jong, Prof. dr. Robert Dur, and Prof. dr. Marco Haan for devoting their time to reading and providing feedback on my thesis. I am specially grateful to Prof. dr. Marco Haan for his constructive comments on my work.

Bahar (Najmeh) Rezaei
Santa Monica, April 2017

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The theory of asymmetric information deals with the study of decisions in transactions in which one party has more or better information than the other. Expecting the other party to have better information can lead to unfavorable changes in behavior from the perspective of efficiency. This area of research was developed in the 1970s and pioneered with: the market for lemons, Akerlof (1970); job market signaling, Spence (1973); and the theory of screening, Stiglitz (1975). Akerlof, Spence and Stiglitz shared a Nobel Memorial Prize in Economic Sciences in 2001 for their contribution in analyzing markets with asymmetric information. Since then, a rich literature on this subject has been developed, and it has been among the most active areas of research in microeconomic theory.[1] One remarkable achievement of the economics of asymmetric information is that many different economic phenomena can be understood using the same theoretical tools. Two major incentive problems caused by asymmetric information are commonly referred to in the literature as *hidden information* and *hidden action*. Problems of hidden information are also referred to as *adverse selection* while problems of hidden actions are also referred to as *moral hazard* (Bolton and Dewatripont, 2005, p.15).

Asymmetric information theory has diverse applications in economics and finance. One of the most common applications appears in the context of principal-agent problems. The principal-agent problem, first introduced by Berle Jr. and Means (1930), is a framework for analyzing problems in which one party (the principal) hires another party (the agent) to perform a certain task. However, a conflict of interest between the principal and the agent exists, and whenever an information asymmetry exists, complete honesty is not optimal for one player. For example, in a labor relationship, the owner of a firm (the principal) wants her employee (the agent) to work hard, but the employee wants to minimize his effort. Parties involved in this relationship are constrained either by asymmetric information or by their inability to monitor each other's actions. Well-known instances of applications of asymmetric information theory include insurance markets, Rotschild and Stiglitz (1976) and Cohen (2005); price discrimination, Spence (1977) and Armstrong and Vickers (2001); credit rationing, Stiglitz and Weiss (1981); and the effects of asymmetric information on bargaining, Samuelson (1984) and Ausubel et al. (2002).

[1] See Riley (2001) for a survey of developments in the economics of information.

Information asymmetries have been traditionally studied under the framework of the *expected utility theory* (Von Neumann and Morgenstern, 1944). In the principal-agent model, for example, it is assumed that decision makers compare their expected utilities of different possible decisions when making risky or uncertain decisions. Therefore, incentives are designed in such a way that the agent is willing to act on behalf of the principal. Experiments, however, have shown that the preferences of economic agents do not always follow the axioms of expected utility theory (see, for example, Kahneman et al., 1991).

More recently, assumptions of *prospect theory* (Kahneman and Tversky, 1979; 1992)—i.e., loss aversion and reference-dependent preferences—have been utilized to explain these inconsistencies. Prospect theory states that agents have reference-dependent preferences and evaluate their possible outcomes in risky decisions with respect to a *reference point* rather than evaluating outcomes in absolute terms. Outcomes below this reference point are seen as losses while outcomes above this reference point as gains. This theory also says that losses hurt more, proportionally, than same-sized gains give pleasure (loss aversion).[2] One example of incorporating prospect-theoretic preferences into a principal-agent framework is presented by Herweg et al. (2010), who explain why the optimal contract often takes a bonus form with two possible wage levels.

Although prospect theory, as a descriptive model of decision making under risk, provides a strong framework to study asymmetric information problems, employing the traditional expected utility theory framework for newly developed models continues to be relevant. When an application of asymmetric information theory is completely new, it is reasonable to first examine the rational approach of the expected utility theory as a benchmark. Although the following chapters of this dissertation are all concerned with hidden information situations, individual chapters employ different theoretical frameworks. Chapters 2 and 3 investigate two classical theoretical mechanisms in information economics—i.e., screening and signaling—by employing the behavioral framework (prospect theory), while Chapter 4 uses the rational framework (expected utility theory) to explore an application of information economics, that is, microcredit lending, which is a relatively new concept.

This study is written as a collection of separately readable papers, and for this reason, there is some overlap between Chapters 2 and 3. Each paper will provide its own detailed introduction and literature review. In what follows, we briefly explain what is discussed in each chapter, what motivates our studies, and how our main findings relate to findings of other relevant research.

1.1 Theoretical Background

One of the most challenging questions in the literature that incorporates unsophisticated agents, that is, agents with reference-dependent preferences, into standard economic models is the way in which a reference point is determined. This question is an inherently

[2]Prospect theory also includes the reflection effect and probability weighting, but we do not apply these concepts in this dissertation, as the model becomes too complicated, and loss aversion is really the key factor in prospect theory.

difficult one to answer. Since the emergence of prospect theory, researchers have proposed many different determinants of the reference point such as the effect of framing (Tversky and Kahneman, 1981), the status-quo (Samuelson and Zeckhauser, 1988; Kahneman et al., 1991), aspirations (Lopes and Oden, 1999), expectations (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991), and so forth. More recent empirical and experimental studies have emphasized the key role of expectations in the formation of reference points (see, for example, Mas, 2006; Pope and Schweitzer, 2011; Crawford and Meng, 2011; Abeler et al., 2011; Gill and Prowse, 2012).

A significant theory in clarifying how people form reference points and evaluate gains and losses is Kőszegi and Rabin's (2006; 2007; 2009) theory of rational expectations. The authors assume that reference points correspond to agent's probabilistic beliefs (anticipations) about the outcome of their choices that are formed in the recent past and are always correct. In this theory, since expectations determine the reference point, as well as what is expected is a lottery, the reference point itself should be stochastic.

In Chapters 2 and 3, we employ Kőszegi and Rabin's rational expectations theory to include expectation-based loss aversion in our models of monopoly pricing and labor market signaling. We assume that the total utility of consumption consists of a standard *intrinsic utility* and a *gain-loss utility*. Intrinsic utility is the utility derived directly from consumption and depends on consumers' types while gain-loss utility is an aggregation of gains and losses that are experienced with respect to the reference point—i.e., rational expectations. We also assume that consumers always choose to be in a *personal equilibrium*, in which beliefs and behavior are jointly determined by the requirement that behavior given past beliefs must be consistent with those beliefs (Kőszegi, 2010). Therefore, we may have multiple personal equilibria for a different set of expectations. If multiple personal equilibria do exist, Kőszegi and Rabin (2006; 2007) argue that senders can ex-ante compare them and choose the one that provides the highest payoff, namely the *preferred personal equilibrium*.

Our study in Chapter 4 belongs to a literature different from the literature of Chapters 2 and 3. In this chapter, we assume rational agents with standard preferences to analyze repeated microcredit lending to groups of borrowers under *joint liability*. Jointly liable members of a group are mutually in charge to repay the total group loan. In other words, repaying members have to jointly repay for all defaulting members of their group in addition to their own loan repayment. Group lending contracts under joint liability are considered one of the most important factors in the relative success of microcredit lending programs. Such contracts are believed to improve the lending outcome by reducing market imperfections created by asymmetric information (for a review, see Ghatak, 1999; Banerjee, 2013).

1.2 Outline of the Dissertation

The economics and marketing literature indicate that reference dependence and loss aversion affects consumer behavior in the market. For example, Genesove and Mayer (2001) observe that in the Boston housing market, sellers set higher prices if they have experienced a loss relative to their purchase price. For evidence from financial and labor markets, see Odean (1998) and Beweley (1998). Literature also confirms that firms are aware of the effect of reference dependence and loss aversion on consumer behavior (Blinder et al., 1998, Marketing News 1985). Inspired by this literature, we investigate the effect of reference dependence and loss aversion on two classical models of price discrimination and job market signaling in Chapters 2 and 3, respectively. More specifically, we explore the strategic pricing behavior of a profit-maximizing monopolist facing loss-averse consumers with reference-dependent preferences in Chapter 2 and the strategic signaling behavior of loss-averse job market candidates with reference-dependent preferences facing profit-maximizing employers in Chapter 3.

The classical literature on price discrimination and labor market signaling, which follows the expected utility theory and assumes rational agents, only finds separating equilibria by assuming that utility functions of the high-type and low-type agents satisfy the single-crossing property. This property, in terms of the price discrimination model presented in Chapter 2, requires the higher type consumer to have a higher marginal utility of consumption, while in terms of the job market signaling model of Chapter 3, the single-crossing property demands the higher type worker to have a lower marginal cost of acquiring a higher education level. In Chapters 2 and 3, we discuss that employing the assumptions of prospect theory and considering agents—who are loss averse and have reference-dependent preferences—in both price discrimination and labor-market signaling mechanisms, help separating equilibria to exist even when the single-crossing property is violated. In other words, loss aversion and reference-dependent preferences facilitate self-selection.

Our result, however, is not in line with the results of the current literature that similarly follows prospect theory and considers loss-averse agents with reference-dependent preferences. The literature, contrary to our findings, argues that reference-dependent preferences lead to pooling equilibria or bunching equilibria (see, for example, Heidhues and Kőszegi, 2004; 2008 and Hahn et al., 2014). The reason for this divergence is that we assume that people form their expectations and therefore their reference points *after* they have learned about their types; while in the literature, it is commonly assumed that people form their expectations *before* learning their types. As these latter agents expect something before knowing if they can afford it, it is reasonable for the other party in the contract to offer them what they expect. Although in some situations it is reasonable to assume that people do not know their types when they form their expectations,[3] in other contexts it may be more appropriate to assume that people actually learn their types first and then form their expectations accordingly. For example, a low-income student who considers buying a used car knows that he cannot afford a car that has a high maintenance cost. Therefore, prior to going to a dealer, he sets an expectation of buying a low-maintenance cost car, even though a luxury car offers more convenience and safety options.

[3] Examples of these situations are discussed in Heidhues and Kőszegi (2004; 2008) and Hahn et al. (2014).

In particular, in Chapter 2, we study the optimal pricing strategy of a monopolist who faces consumers that have heterogeneous private tastes and reference-dependent preferences, and are subject to loss aversion. Asymmetric information exists, and the monopolist does not observe consumers' valuations. Consider an airline, which deals with passengers, who can be either rich or poor. Because the airline is not able to categorize its passengers, it designs two kinds of tickets: a business class ticket and an economy class ticket. The business class ticket is more expensive but has higher quality; the economy class ticket is cheaper but has lower quality. The tickets are designed in a way that rich passengers self-select the business class ticket and poor passengers self-select the economy class ticket. The model we employ is based on the classical model of monopoly pricing under asymmetric information, Mussa and Rosen (1978), which we extend by assuming that consumers are loss averse and have reference-dependent preferences. In Chapter 2, we answer the following research question:

Research question 1: *Is it easier or more difficult for the monopolist to make consumers self-select (i.e., to choose the separating equilibrium) when agents have expectation-based loss aversion?*

Assuming that the monopolist can make consumers expect to buy the desired variety of the good, we prove that for consumers with expectation-based loss aversion, menu pricing is possible for a wider range of parameters. We argue that loss-averse consumers with different expectations self-select their types naturally. Intuitively, a loss-averse rich consumer strongly avoids low quality. Therefore, if he considers buying the low-quality ticket, he focuses more on the loss in quality than on the gain of paying a lower price. Subsequently, such a consumer has a stronger incentive not to shift to economy class ticket. Similarly, a loss-averse poor consumer has a stronger incentive not to shift to business class ticket. Because, if he considers buying the business class ticket, he focuses more on the loss of paying a higher price than on the gain of acquiring a higher quality. Consequently, the incentive compatibility constraints are easier to satisfy for the monopolist when consumers are loss averse.

In Chapter 3, as our leading example, we consider an employer who faces two types of workers. One type is highly productive and the other is not very productive, but the employer cannot distinguish between them. The workers invest in obtaining some level of education in order to signal their high productivity to the employer. The model we analyze is built on Spence's (1973) model of job market signaling with the added feature that senders are subject to loss aversion and have reference-dependent preferences. We want to answer the following question in Chapter 3:

Research question 2: *Is it more feasible or less feasible to achieve credible signaling (i.e., separating equilibrium) when senders have expectation-based loss aversion?*

We argue in this chapter that loss-averse senders are naturally more inclined to signal credibly when each sender type has a different expectation. Intuitively, high-productivity workers acquire a high education because they want to avoid the loss in wages that would follow if they chose to acquire a low education. In addition, low-productivity workers acquire a low education because they want to avoid the loss in the form of incurring higher educational costs if they chose to acquire a high education. As is well known in the literature, the single-crossing property, saying that high types have a lower marginal cost of signaling is a necessary condition for the existence of separating equilibria in signaling games. In this chapter, we show that with expectation-based loss aversion, the separating equilibrium can arise even when the single-crossing property is not satisfied.

In our models in Chapters 2 and 3, we may have multiple personal equilibria for different sets of expectations. Multiple personal equilibria are treated differently in these two chapters. In Chapter 2, we assume that the monopolist is able to frame consumers' reference points strategically. In other words, she can manipulate consumers' expectations in a way that they prefer to choose the personal equilibrium that is best for the monopolist. Therefore, in the case of multiple personal equilibria, the separating equilibrium can still be chosen by consumers. In Chapter 3, however, we employ Kőszegi and Rabin's (2006; 2007) theory of the preferred personal equilibrium. Thus, even when multiple personal equilibria exist, the separating equilibrium may still be selected as the preferred personal equilibrium by senders. We do not use the idea of strategic framing in the signaling model of Chapter 3 because in this game, unlike the screening game of Chapter 2, the informed player (the worker) moves first, and the uninformed player (the employer) does not obtain a chance to manipulate his reference point.

The microcredit lending literature has a question about the efficacy of group lending as opposed to individual lending. Chapter 4 addresses the question of which type of contract (mechanism) offers greater welfare and feasibility. Extant literature has asserted that group lending performs worse than individual lending without strong social sanctions; furthermore, the literature pays little attention to the effect of group size. This chapter fills this gap by considering an arbitrary group size and proposing a joint liability contract that can do better than an individual lending contract.

In Chapter 4, as previously mentioned, no loss aversion or reference-dependent preferences are assumed. The model in this chapter extends existing models on group lending by endogenizing the number of jointly liable borrowers. For this extension it is reasonable to first assume standard preferences and use the rational approach as a starting point. In this chapter, we analyze an infinitely repeated microcredit lending game, in which a benevolent lender decides to lend to a group of borrowers under a joint liability contract. Each borrower invests his loan on a risky project. Although the lender cannot observe the outcome of each project, members of the group can. Therefore, if borrowers default on their loan repayments strategically, the lender cannot realize it. We investigate the following question in Chapter 4:

Research question 3: *How large should the group of borrowers be so that each borrower's benefit is maximized while the lender breaks even?*

Intuitively, under a joint liability contract, being a member of a large group of borrowers has both benefits and costs. On the one hand, it can improve the chance of assured repayment for a defaulting member. On the other hand, there is also a higher threat of repayment if the other members of the group default. While for risky projects the insurance provided by larger groups becomes more attractive, it is less necessary for low-risk projects. In this chapter, we argue that larger group sizes can improve the lending outcome when the chance of project success is not that high, and we characterize the optimal group size given the chance of project success and borrowers' discount factor.

The existing literature on microcredit lending mostly concludes that when borrowers are unable to impose strong social sanctions on each other, the lender will be better off choosing a contract with individual liability over joint liability (Besley and Coate, 1995; Armendáriz de Aghion, 1999). In Chapter 4, we show that this is not necessarily true when group members play grim-trigger against each other—i.e., when the repaying group members deprive any strategically defaulting member of the group from a future loan at least for a T period of time. In particular, we discuss that although joint liability lending is feasible under a smaller parameter setting, has a positive effect on borrowers' welfare and the repayment rate compared with individual lending. Furthermore, we show that joint liability, when feasible, outperforms individual liability in terms of the maximum loan that can be offered to borrowers.

Further in this chapter, we examine how the severity of punishment employed against strategically defaulting borrowers can influence the group size. When the length of the punishment phase is reduced from infinitely many periods to T periods, it is clearly less costly for borrowers to default strategically. We prove that, when the punishment phase is not that long, we need a larger group to guarantee the total repayment. In this chapter, we also discuss that larger group sizes can additionally improve the lending outcome when projects are correlated. We assume that when projects are correlated, borrowers have higher incentives to help each other. Therefore, members of a larger group can positively contribute to each others' success.

Chapter 2

Screening Loss-Averse Consumers

2.1 Introduction

Recent researches in behavioral industrial organization literature suggests that a profit-maximizing firm facing loss-averse consumers with reference-dependent preferences and a heterogeneous willingness to pay can increase profits by offering a single product type rather than multiple product types and screening consumers (see, for example, Heidhues and Kőszegi, 2004; 2008; Herweg and Mierendorff, 2013; Hahn et al., 2014). As we show in this paper, the reason that the behavioral literature concludes that loss aversion limits the extent of price discrimination is due to the assumption that consumers form reference points before learning their types. If we instead assume that consumers form their reference points after learning their types, loss aversion could actually make price discrimination more feasible, and more desirable to a profit-maximizing firm , and it could enable easier screening. The fact that in the real world price discrimination is still limited may be explained by firms' costs of maintaining a large price list and/or consumers' difficulty in processing an extensive price list (e.g., Dixit, 1977; Spence, 1980).

The basic mechanism at play in our model may be understood through the following example. Consider an airline that faces poor and rich passengers, but is not able to tell them apart. The airline attempts to design a business class ticket and an economy class ticket in which the latter is cheaper and has lower quality, in such a way that rich passengers self-select the business class ticket and poor passengers self-select the economy class ticket. Is it easier or more difficult for the airline to do this when the passengers have reference-dependent preferences and are subject to loss aversion (Kőszegi and Rabin, 2006; 2007)? The following intuition suggests that reference-dependent preferences make self-selection easier. Let it be the case that rich consumers expect to buy business class tickets, and poor consumers expect to buy economy class tickets, and let this also determine their respective reference points. When considering whether to buy the economy class ticket instead of the business class ticket, the loss-averse rich consumer focuses more on the loss in quality than on the gain of paying a lower price, and is thereby less inclined to switch to economy class. In the same way, when considering to buy the business class ticket instead of the economy class ticket, the loss-averse poor consumer focuses more on the loss caused by paying a higher price than on the gain in quality, and will be less inclined to upgrade to business class.

We introduce reference-dependent preferences and loss aversion to the classic monopoly pricing model under asymmetric information (Mussa and Rosen, 1978). We assume that consumers have expectation-based reference-dependent preferences in the sense of Kőszegi and Rabin (2006). Our motivation for this assumption is that the important role of expectations in the formation of reference points is emphasized by several recent studies (see, for example, Pope and Schweitzer, 2011; Crawford and Meng, 2011; Abeler et al., 2011). We follow Kőszegi and Rabin's (2006) model of reference-dependent preferences, where consumer's total utility consists of a standard *intrinsic utility*, which depends on the consumers' type, and of a *gain-loss utility*. Gains and losses are experienced with respect to the reference point determined by the consumers' rational expectations. It is assumed that consumers are always in a *personal equilibrium (PE)* in which they select their optimal actions while their expectations are taken into account.

Our analysis shows two ways in which the fact that consumers have reference-dependent preferences may facilitate menu pricing compared to the case of standard preferences. First, with reference-dependent preferences, menu pricing becomes possible even if the consumer intrinsic utility do not satisfy the *single-crossing property (SCP)*. Rich consumers who consider buying low quality consider this as a loss, and for this reason attach a higher overall marginal utility (including the gain-loss utility part of their utility) to quality. Poor consumers who consider buying high quality focus on the loss of having to pay a higher price, and therefore to the poor consumer, the marginal utility of quality is small relative to the marginal utility of money.

Second, it is possible that when facing consumers without reference-dependent preferences, the monopolist prefers offering a single high price and thus prefers selling only to high-valuation consumers rather than menu pricing, whereas when facing consumers with reference-dependent preferences, the monopolist prefers menu pricing, when the portion of high-valuation consumers is not very high among consumers. In general, the self-selection necessary for menu pricing to work succeeds only if the monopolist offers high quality at a discount. The monopolist may shy away from menu pricing if the discount needed to make menu pricing work is too large. However, a rich consumer with reference-dependent preferences is less inclined to choose the low-quality variant of the monopolist's product given his aversion to quality loss. For this reason, the monopolist does not need to offer such a high discount and is more likely to prefer menu pricing.

Furthermore, we discuss in this paper that loss aversion and reference-dependent preferences can generate multiple PEs. For example, for the same price and quality, both cases in which the consumer buys the good because he expects to buy it, and he does not buy the good because he does not expect to buy it may be PEs. In such a case, although some additional PEs coexist with the separating PE, consumers may still choose to play the separating equilibrium if the monopolist is able to frame consumers' expectations and therefore their reference points strategically.[1] That is, by using strategic framing, the monopolist can encourage consumers to play the PE that is most preferred by the monopolist—i.e., the separating equilibrium.

Section 2 discusses the related literature. Section 3 introduces our model of menu pricing in which consumers have reference-dependent preferences. Section 4 treats the case of symmetric information between the monopolist and consumers as a benchmark, followed by the case of asymmetric information in Section 5. Section 6 explores whether multiple PEs exist, and in such a case, what would be the preferred PE. We end with a conclusion in Section 6.

2.2 Literature Review

Several recent studies relate to our paper. In general, this paper is part of a strand of research that investigates firm behavior when profit-maximizing firms face naive consumers (for an overview, see Ellison, 2006; Armstrong and Huck, 2010; Spiegler, 2011). Within this literature, our paper is part of a growing body of research that integrates consumers with prospect theoretic preferences or reference-dependent preferences into standard economic models (for an overview, see Kőszegi, 2014; Kim and Lee, 2014). The difference between this literature and our paper is that we find an argument for price variation with reference-dependent preferences, whereas the literature on the contrary shows how reference-dependence can lead to price stickiness or less price variation (see, for example, Heidhues and Kőszegi, 2004; 2008). Within the behavioral industrial organization literature, we are also related to studies that discuss if and how firms can influence consumers' choices via framing (for example, see Spiegler, 2011; Piccione and Spiegler, 2012).

Closer to our paper, Hahn et al. (2014) study the menu pricing model with reference-dependent consumers. An essential difference of Hahn et al. (2014) with our paper is that the authors assume that consumers form their expectations, and therefore also form their reference points, *before* they learn their own valuation of the good. Intuitively, prices are then sticky around this average expectation, so that it becomes more likely that the monopolist does not prefer menu pricing but rather offers a single price. In a setting similar to Hahn et al., Herweg and Mierendorff (2013) demonstrate that consumer loss aversion can explain the prevalence of flat-rate contracts relative to measured tariffs, although they may not minimize consumers' expected billing amounts. In their model, consumers' reference points, which are only reference points with respect to the price, are set *before* they accept the contract and learn about their consumption levels. Therefore, loss-averse consumers with uncertain future demands prefer a flat-rate contract to insure themselves against fluctuations in their billing amounts. In the same way, in their model of product differentiation, Heidhues and Kőszegi (2008) assume that consumers form expectations *before* they learn their types. Thus, as heterogeneous consumers maintain the same reference price, prices

[1] In the literature, it is believed that firms can manipulate consumers' reference points through their marketing practices. See, for examples, Puto, (1987); Karle (2013) ; Spiegler (2014); and Greenzy et al., (2014).

are sticky around this single reference price, and only a single price may be maintained in equilibrium. Our model instead is based on the assumption that, for example, rich and poor consumers form their reference points *after* they have learned if they are rich or poor, leading us to a result that is diametrically opposed to Hahn et al., Herweg and Mierendorff, or Heidhues and Kőszegi.

The paper that is closest to ours may be Carbajal and Ely (2016), who study price discrimination for loss-averse consumers who have state-contingent reference points. Similar to us, they consider consumers reference points that are formed *after* consumers learn their types (i.e., or "ex-post consistent reference points"). A key difference between Carbajal and Ely (2016) and our paper is that in the former, a reference point is only a reference quality and not a reference price. They find that loss aversion increases the optimal quality level of all consumers compared to a loss-neutral case, and propose that quality allocation for high-type consumers may become inefficiently distorted upwards. Intuitively, when consumers are loss averse only in the quality direction, a higher reference point increases the consumers' willingness to pay, and thus loss aversion increases the quality level. In our model, however, the reference point is both a reference quality and a reference price. Thus, for the poor loss-averse consumer, loss in price matters more than gain in quality, while for the rich consumer, loss in quality matters more than gain in price. Therefore, it is easier to convince different types of consumers not to mimic the other type and thus easier to achieve separating equilibrium.

Using framing for screening, the study by Salant and Siegel (2013) is also related to our paper. Similarly, in their model, a seller can employ a frame and manipulate buyers' preferences. However, the authors assume that the effect of framing is temporary, and that the agent can reconsider his decision after the effect of framing declines. Consequently, contrary to our study, they prove that framing can only occasionally increase profit.

2.3 Model

In stage 1, Nature decides on the type of the consumer (he). The consumer is of type θ_1 with probability $(1-q)$, and of type θ_2 with probability q. The consumer learns his type, but the monopolist (she) does not. In stage 2, the monopolist offers the consumer a menu of possible price-quality combinations that allow the consumer to buy a single unit of quality s at price p. As only two types exist, the monopolist sets at most two price-quality combinations, which are denoted as (s_1, p_1) and (s_2, p_2), where $s_2 \geq s_1$. The price-quality combinations may be identical or different. When they are different, the monopolist will design them such that consumer θ_1 chooses (s_1, p_1), and consumer θ_2 chooses (s_2, p_2). As the consumer is better off as the quality increases (see below), it only makes sense for the monopolist to offer two price-quality combinations if $p_2 \geq p_1$, as the consumer otherwise opts for one of the price-quality combinations independent of his type.

In stage 3, the consumer either chooses a price-quality combination from the menu or does not buy the good at all. In stage 4, if the consumer chooses one of the price-quality combinations from the menu, the monopolist provides the good with the quality and the price as agreed upon, and the players obtain their payoffs. The payoff of the monopolist equals $p - C(s)$, where $C(s)$ is the monopolist's cost function of quality. The cost function is increasing and convex—i.e., $C'(s) > 0$ (including $C'(0) > 0$) and $C''(s) > 0$. The consumer's utility contains an intrinsic utility part, and a gain-loss utility part, where his total utility is additively separable in his intrinsic utility and his gain-loss utility. The intrinsic utility of any choice that the consumer makes equals $U(\theta, s) - p$, where $U(\theta, s)$ is the consumer's utility of obtaining quality s when he is of type θ. We assume throughout that function $U(\theta, s)$ is strictly increasing and concave—i.e., $U_s(\theta, s) > 0$, $U_{ss}(\theta, s) < 0$ with $U(\theta, 0) = 0$, and $U_s(\theta, 0) = \infty$. $U(\theta, 0) = 0$ means that for a consumer who does not pick any of the price-quality combinations, his intrinsic utility is zero. Further, we assume throughout that for any s,

$$(A1): \qquad U(\theta_2, s) > U(\theta_1, s),$$

meaning that type θ_2 receives a higher utility from quality than type θ_1. This gives the monopolist a motive for attempting to price discriminate, and also implies also $U_\theta(\theta, s) > 0$.

Finally, we formulate the standard SCP, where we immediately note that we do not systematically assume this property: For any $s_j > s_k$,

$$(SCP): \qquad U(\theta_2, s_j) - U(\theta_2, s_k) > U(\theta_1, s_j) - U(\theta_1, s_k),$$

which implies that $U_{s\theta}(\theta, s) > 0$.

The consumer's gain-loss utility is determined by the reference price and the reference quality that he forms. Following Kőszegi and Rabin (2006), we assume that the consumer's reference price and quality are determined each time by the price-quality combination that he expects to choose. When the consumer expects to buy at price p_r but instead buys it at a lower price p, he experiences a gain $(p_r - p)$, which is added to his total utility. When the price at which he buys is higher than p_r, he experiences a loss $\lambda(p - p_r)$, which is subtracted from his total utility. The parameter λ, with $\lambda \geq 1$, reflects the consumer's degree of loss aversion. In the same way, when the consumer expects to buy a quality of s_r but instead buys higher quality s, he experiences a gain $(s - s_r)$, which is added to his intrinsic utility. If instead he buys a lower quality, he experiences a loss $\lambda(s_r - s)$, which is subtracted from his intrinsic utility. Thus, for instance, a consumer of type θ_2 who expects to choose high quality s_2 and pay high price p_2 for it, but instead ends up with lower quality s_1 and lower price p_1, gets overall utility

$$U(\theta_2, s_1) - p_1 - \lambda(s_2 - s_1) + (p_2 - p_1).$$

Following Kőszegi and Rabin (2006), we adopt a rational expectations approach, in which the consumer's expectations about what he will choose are also fulfilled. The resulting choice is then referred to as a PE. In other words, a PE is a situation in which the consumer chooses his optimal product given his reference point, while his reference point is set by his expected optimal choice of product. This raises the possibility of multiple PEs. For one and the same price-quality combination offered by the monopolist, there may be two PEs, a PE in which the consumer prefers to buy the good because he expects to buy it, and a PE in which the consumer prefers not to buy the good because he expects not to buy it. For two price-quality combinations offered by the monopolist, it is possible that a PE exists in which the consumer always chooses one price-quality combination, and a PE in which each type of consumer chooses a different price-quality combination. In this case, Kőszegi and Rabin assume that the consumer can ex-ante compare the different PEs, and picks out the PE best for him, which is then referred to as the preferred personal equilibrium (PPE).

An alternative way of choosing between multiple PEs that we use in our analysis is framing. We assume that the monopolist has the power to influence the consumer's expectations and thereby his reference point, to her advantage.

2.4 Symmetric Information

As a benchmark, we first look at a case in which not only the consumer, but also the monopolist learns the consumer's type. We then have a simple model of third-degree price discrimination, wherein the monopolist can offer a different price-quality combination to each type, with an additional feature that the consumer has reference-dependent preferences. The monopolist maximizes her expected profit with respect to the participation constraint of each consumer type, wherein the assumption is made that the monopolist can ensure that the consumer expects to buy the good, thus

$$\begin{aligned} \underset{p_1,p_2,s_1,s_2}{maximize} \quad & (1-q)\left[p_1 - C(s_1)\right] + q\left[p_2 - C(s_2)\right] && (2.1)\\ s.t. \quad & U(\theta_1, s_1) - p_1 \geq -\lambda U(\theta_1, s_1) + p_1 && (2.2)\\ & U(\theta_2, s_2) - p_2 \geq -\lambda U(\theta_2, s_2) + p_2 && (2.3) \end{aligned}$$

The left-hand side of the participation constraint takes a standard form; that is, as previously mentioned, if the consumer chooses the equilibrium price-quality combination, he achieves his reference price and quality, and does not experience gains or losses. If he chooses not to buy the good, given our assumption $U(\theta, 0) = 0$, the consumer obtains intrinsic utility zero. However, the consumer also experiences a loss of not obtaining the quality he expected, and a gain of not having to pay a price.

As shown in Proposition 2.1, as long as loss aversion exists, reference-dependent preferences result in a higher quality and a corresponding higher price in both price-quality combinations. Intuitively, when a loss-averse consumer who has a reference point that he buys the good, is considering not to buy the good, he focuses more on the loss of not receiving the good than on the gain of not having to pay any price, whereby the consumer's willingness to pay increases. Further, since self-selection is not a problem, loss aversion distorts the two qualities in the same manner; hence, while both qualities go up, there is no relative distortion. The higher the loss aversion, the more weight the consumer puts on the quality, and the less weight he puts on the price, thus the higher the quality in both price-quality combinations. Denote $s_i^{FB,R}$ and $p_i^{FB,R}$ ($s_i^{FB,NR}$ and $p_i^{FB,NR}$) as respectively the optimal quality and price offered to the consumer of type i when symmetric information exists and when the consumer has reference-dependent preferences (no reference-dependent preferences). FB stands for first-best, and superscripts R and NR refer to reference-dependent preferences and no reference-dependent preferences, respectively.

Proposition 2.1. *The following statements hold under symmetric information.*

1. It is always better to offer two price-quality combinations rather than one, regardless of whether consumers have reference-dependent preferences or standard preferences.

2. $s_i^{FB,R} > s_i^{FB,NR}$, $p_i^{FB,R} > p_i^{FB,NR}$.

3. $\frac{\partial s_i^{FB,R}}{\partial \lambda} > 0$, *and* $\frac{\partial p_i^{FB,R}}{\partial \lambda} > 0$.

Proof. Note first that the price-quality combination offered to one type of consumer does not affect the price-quality combination offered to another type of consumer. If we let $(\gamma_1, \gamma_2) \geq 0$ be the multipliers on constraints (2.2), (2.3), respectively, the Kuhn-Tucker conditions for this problem can then be written as:

$$1 - q - 2\gamma_1 = 0 \tag{2.4}$$

$$q - 2\gamma_2 = 0 \tag{2.5}$$

$$-(1-q)C'(s_1) + \gamma_1(1+\lambda)U_s(\theta_1, s_1) \leq 0 \tag{2.6}$$

$$-qC'(s_2) + \gamma_2(1+\lambda)U_s(\theta_2, s_2) \leq 0, \tag{2.7}$$

along with the complementary slackness conditions for constraints (2.2) and (2.3).

Solving (2.5) and (2.6) we get $\gamma_1 = \frac{1-q}{2}$ and $\gamma_2 = \frac{q}{2}$, which both are strictly positive so that the participation constraints (2.2) and (2.3) are binding. Substituting these values of γ_1, γ_2 into (2.6) and (2.7) we obtain

$$C'(s_1^{FB,R}) = \frac{1+\lambda}{2} U_s(\theta_1, s_1^{FB,R}) \tag{2.8}$$

$$C'(s_2^{FB,R}) = \frac{1+\lambda}{2} U_s(\theta_2, s_2^{FB,R}). \tag{2.9}$$

(2.8) and (2.9) characterize the optimal values $s_1^{FB,R}$ and $s_2^{FB,R}$, respectively. It is clear now that for $\lambda > 1$, $s_i^{FB,R} > s_i^{FB,NR}$ for $i = 1, 2$.

The optimal values for $p_1^{FB,R}$ and $p_2^{FB,R}$ are then determined from (2.2) and (2.3), which we have seen hold with equality at the optimal solution

$$p_1 = \frac{1+\lambda}{2} U(\theta_1, s_1^{FB,R}) \tag{2.10}$$

$$p_2 = \frac{1+\lambda}{2} U(\theta_2, s_2^{FB,R}). \tag{2.11}$$

It follows that, if $\lambda > 1$, as s_1 and s_2 increase compared to the case without reference-dependent preferences, p_1 and p_2 also increase—i.e., $p_i^{FB,R} > p_i^{FB,NR}$. Further from (2.8) and (2.9), we have $\frac{\partial s_i^{FB,R}}{\partial \lambda} > 0$, and from (2.10) and (2.11), we have $\frac{\partial p_i^{FB,R}}{\partial \lambda} > 0$. □

Multiple personal equilibria and preferred personal equilibria: The assumption that the monopolist can always ensure that the consumer expects to buy the good, and accordingly forms his reference point, is not a trivial one. A PE in which the consumer does not buy because he expects not to buy exists if:

$$0 \geq U(\theta_1, s_1) - \lambda p_1 \tag{2.12}$$

$$0 \geq U(\theta_2, s_2) - \lambda p_2, \tag{2.13}$$

which are always valid for large enough λ. Thus, for a large enough degree of loss aversion, there are always multiple PEs. In this case, the PE obtained in Proposition 2.1 is a PPE if additionally:

$$U(\theta_1, s_1) - p_1 \geq 0 \tag{2.14}$$

$$U(\theta_2, s_2) - p_2 \geq 0. \tag{2.15}$$

Thus, for $\lambda > 1$, the true constraints that the monopolist faces are constraints (2.14) and (2.15). In this case, the equilibrium is indistinguishable from the one without reference-dependent preferences, and both the consumer and the monopolist are exactly equally well off with and without reference-dependent preferences.

2.5 Asymmetric Information

We now look at the case of asymmetric information. We first look at the optimal prices and qualities in three cases, namely, where the monopolist offers a different quality level to both consumers (Proposition 2.2), where the monopolist only offers high quality, which is not bought by low-valuation consumers (Proposition 2.3), and where the monopolist offers the same low-quality level to both consumer types (Proposition 2.4). The next step is to determine which of these outcomes is best for the monopolist (Proposition 2.5 and 2.6).

The first case is where the firm sets two price-quality combinations and adapts prices and qualities such that each consumer type self-selects the right combination. The maximization problem is now the same as in (2.1) to (2.3), but we now additionally have two incentive compatibility constraints:

$$U(\theta_1, s_1) - p_1 \geq U(\theta_1, s_2) - p_2 + [U(\theta_1, s_2) - U(\theta_1, s_1)] - \lambda(p_2 - p_1) \tag{2.16}$$

$$U(\theta_2, s_2) - p_2 \geq U(\theta_2, s_1) - p_1 - \lambda\, [U(\theta_2, s_2) - U(\theta_2, s_1)] + (p_2 - p_1). \tag{2.17}$$

Denote $s_i^{SB,R}$ and $p_i^{SB,R}$ ($s_i^{SB,NR}$ and $p_i^{SB,NR}$) for $i = 1, 2$ the optimal price-quality combinations under asymmetric information with reference-dependent preferences (without reference-dependent preferences). *SB* stands for second-best, and *R* and *NR* refer to reference-dependent preferences and no reference-dependent preferences, respectively.

Proposition 2.2, besides giving a sufficient condition for self-selection when there is loss aversion, compares the second-best without reference-dependent preferences to the second-best with reference-dependent preferences. The first part of this proposition suggests that with reference-dependent preferences, self-selection is possible under a condition that is less strict than the SCP, and applies even if high- and low-valuation consumers have the same marginal utility of quality. Therefore, it is more plausible to have possible menu pricing if consumers have reference-dependent preferences.

The second part of Proposition 2.2 compares the case of consumers having reference-dependent preferences with the case of consumers having standard preferences, when there is asymmetry of information. As is shown, the effect of reference-dependent preferences on both menus is ambiguous except for the quality of the expensive option that is higher when there is reference-dependency and loss aversion, and this effect is increasing in the degree of loss aversion. Intuitively, the loss-averse high-valuation consumer that cares about a gain in higher quality more than a loss in higher price, is offered a higher quality in order to convince him to self-select.

The last part of Proposition 2.2 presents a comparison between the case of asymmetric information and symmetric information when consumers have reference-dependent preferences. As is shown, for the high-valuation consumer, the quality is not changed in the second-best, but the price is lowered. For the low-valuation consumer, the quality and the price are reduced in the second-best when the SCP is valid. However, when the SCP is strictly violated, both the quality and the price for the low-valuation consumer are increased in the second-best. The intuition can be explained as follows.

In the second-best, the firm offers a discount to the high-valuation consumer in order to stop him from pretending to be the low-valuation consumer. The firm can keep the discount limited by manipulating the quality offered to the low-valuation consumer in the following manner. When the SCP is valid, the firm can reduce the quality for the low-valuation consumer. In this case, because the high-valuation consumer cares more about quality reductions than the low-valuation consumer, he will stick to the expensive option. When the SCP is strictly violated, the firm can increase the quality for the cheaper option, and thereby also increase its price. In this case, because the high-valuation consumer cares less about quality increases than the low-valuation consumer, pretending to have a low-valuation will be less attractive to the high-valuation consumer. When the SCP is weakly violated, if the quality for the cheap option is changed the price changes accordingly, in such a way that choosing the cheap option remains exactly equally attractive to the high-valuation consumer. Thus, in this case, there is no reason to distort the quality of the cheap option.

Proposition 2.2. *Consider the case of asymmetric information.*

1. For any $s_2 > s_1$, if

$$\frac{1+\lambda}{2}\left[U(\theta_2,s_2)-U(\theta_2,s_1)\right] > \frac{2}{1+\lambda}\left[U(\theta_1,s_2)-U(\theta_1,s_1)\right],$$

then the monopolist can make the loss-averse consumer self-select contracts (s_1,p_1) and (s_2,p_2).

2. Comparing asymmetric information with and without reference-dependent preferences:

a) The relations between $s_1^{SB,R}$ and $s_1^{SB,NR}$ and between $p_1^{SB,R}$ and $p_1^{SB,NR}$ are ambiguous.

b) $s_2^{SB,R} > s_2^{SB,NR}$. The relation between $p_2^{SB,R}$ and $p_2^{SB,NR}$ is ambiguous.

c) $\frac{\partial s_2^{SB,R}}{\partial \lambda} > 0$.

3. Comparing asymmetric information with symmetric information when consumers have reference-dependent preferences:

a) For $U_{s\theta} > 0$, we have $s_1^{SB,R} < s_1^{FB,R}$, $p_1^{SB,R} < p_1^{FB,R}$.

b) For $U_{s\theta} = 0$, we have $s_1^{SB,R} = s_1^{FB,R}$, $p_1^{SB,R} = p_1^{FB,R}$.

c) For $U_{s\theta} < 0$, we have $s_1^{SB,R} > s_1^{FB,R}$, $p_1^{SB,R} > p_1^{FB,R}$.

d) Regardless of whether the SCP is satisfied, we have $s_2^{SB,R} = s_2^{FB,R}$, and $p_2^{SB,R} < p_2^{FB,R}$.

Proof. This proof is presented in 6 steps.
Step 1. We first show that participation constraint (2.3) is slack. Using (2.17), and (A1), it follows that

$$U(\theta_2,s_2)-p_2 \geq U(\theta_1,s_1)-p_1-\lambda\left[U(\theta_2,s_2)-U(\theta_2,s_1)\right]+(p_2-p_1). \tag{2.18}$$

Given participation constraint (2.2), and given that (2.18) is valid, it is then certainly true that

$$U(\theta_2,s_2)-p_2 \geq -\lambda U(\theta_1,s_1)+p_1-\lambda\left[U(\theta_2,s_2)-U(\theta_2,s_1)\right]+(p_2-p_1),$$

or more simplified

$$U(\theta_2,s_2)-p_2 \geq -\lambda U(\theta_2,s_2)+p_2-\lambda\left[U(\theta_1,s_1)-U(\theta_2,s_1)\right].$$

By $(A1)$, $[U(\theta_2,s_1)-U(\theta_1,s_1)] > 0$, it follows that (2.3) is slack. Therefore, we need only consider (2.2), (2.16) and (2.17). Letting γ_1, ϕ_1, and ϕ_2 be the multipliers on constraints (2.2), (2.16), and (2.17), respectively, the Kuhn-Tucker conditions for this problem can be written as

$$(1-q)-2\gamma_1-\phi_1(1+\lambda)+2\phi_2=0 \tag{2.19}$$
$$q+\phi_1(1+\lambda)-2\phi_2=0 \tag{2.20}$$
$$-(1-q)C'(s_1)+\gamma_1(1+\lambda)U_s(\theta_1,s_1)+2\phi_1U_s(\theta_1,s_1)-\phi_2(1+\lambda)U_s(\theta_2,s_1)\leq 0 \tag{2.21}$$
$$-qC'(s_2)-2\phi_1U_s(\theta_1,s_2)+\phi_2(1+\lambda)U_s(\theta_2,s_2)\leq 0, \tag{2.22}$$

along with the complementary slackness conditions for (2.2), (2.16), and (2.17).
Step 2. By (2.20), if $\phi_2 = 0$, then $\phi_1 < 0$, which is not possible. Thus, $\phi_2 > 0$ and (2.17) is binding.
Step 3. (2.16) and (2.17) can be rewritten as

$$p_2 - p_1 \geq \frac{2}{1+\lambda}[U(\theta_1, s_2) - U(\theta_1, s_1)] \tag{2.23}$$

$$p_2 - p_1 \leq \frac{1+\lambda}{2}[U(\theta_2, s_2) - U(\theta_2, s_1)]. \tag{2.24}$$

If

$$\frac{1+\lambda}{2}[U(\theta_2, s_2) - U(\theta_2, s_1)] \leq \frac{2}{1+\lambda}[U(\theta_1, s_2) - U(\theta_1, s_1)],$$

then the monopolist is not able to achieve self-selection, and the two proposed price-quality combinations are necessarily identical. If

$$\frac{1+\lambda}{2}[U(\theta_2, s_2) - U(\theta_2, s_1)] > \frac{2}{1+\lambda}[U(\theta_1, s_2) - U(\theta_1, s_1)],$$

it follows that either (2.16) or (2.17) is slack. Since in Step 2 we showed that (2.17) is binding, it follows that (2.16) is slack, so that $\phi_1 = 0$. Note that $[U(\theta_2, s_2) - U(\theta_2, s_1)] > [U(\theta_1, s_2) - U(\theta_1, s_1)]$ is not a necessary condition for $\phi_1 = 0$. If $[U(\theta_2, s_2) - U(\theta_2, s_1)] \leq [U(\theta_1, s_2) - U(\theta_1, s_1)]$, the monopolist can still achieve self-selection if the degree of loss aversion is sufficiently large ($\lambda \gg 1$).
Step 4. By adding (2.19) and (2.20), we obtain that $\gamma_1 = \frac{1}{2}$. It follows that (2.2) is binding.
Step 5. Given that $\phi_1 = 0$, $\gamma_1 = \frac{1}{2}$, it follows from (2.19) that $\phi_2 = \frac{q}{2}$. Plugging these values into (2.21) and (2.22), and evaluating at respectively $s_1 = 0$ and $s_2 = 0$, we obtain

$$-(1-q)\underbrace{C'(0)}_{>0} + \frac{1+\lambda}{2}\underbrace{U_s(\theta_1, 0)}_{=+\infty} - q\frac{1+\lambda}{2}\underbrace{U_s(\theta_2, 0)}_{=+\infty} \tag{2.25}$$

$$-q\underbrace{C'(0)}_{>0} + q\frac{1+\lambda}{2}\underbrace{U_s(\theta_2, 0)}_{=+\infty}. \tag{2.26}$$

It follows that both constraints are positive for zero levels of quality, so that neither quality is optimally at zero.
Step 6. Given that the optimal qualities are non-zero, for the qualities under asymmetric information and reference-dependent preferences, from (2.21) and (2.22) we have

$$-(1-q)C'(s_1^{SB,R}) + \frac{1+\lambda}{2}U_s(\theta_1, s_1^{SB,R}) - q\frac{1+\lambda}{2}U_s(\theta_2, s_1^{SB,R}) = 0$$

$$-qC'(s_2^{SB,R}) + q\frac{1+\lambda}{2}U_s(\theta_2, s_2^{SB,R}) = 0.$$

Solving these equations, we will have

$$C'(s_1^{SB,R}) = \frac{1+\lambda}{2}U_s(\theta_1, s_1^{SB,R}) - \frac{q(1+\lambda)}{2(1-q)}\left[U_s(\theta_2, s_1^{SB,R}) - U_s(\theta_1, s_1^{SB,R})\right] \quad (2.27)$$

$$C'(s_2^{SB,R}) = \frac{1+\lambda}{2}U_s(\theta_2, s_2^{SB,R}). \quad (2.28)$$

The following answers two questions:

First, what is the effect of reference-dependent preferences on s_1 and s_2 in the second best? In (2.27), we know that $\frac{1+\lambda}{2} > 1$ increases the $C'(s_1^{SB,R})$ and consequently increases s_1. We also know that $U_s(\theta_1, s_1^{SB,R})$ and $U_s(\theta_2, s_1^{SB,R})$ are both decreasing functions in s_1, but it is not clear how $\left[U_s(\theta_2, s_1^{SB,R}) - U_s(\theta_1, s_1^{SB,R})\right]$ is affected by the changes of s_1. Therefore, the effect of reference-dependent preferences on s_1 is ambiguous. From (2.28), it is clear that $s_2^{SB,R} > s_2^{SB,NR}$.

Second, what is the effect of asymmetric information on s_1 and s_2 when agents have reference-dependent preferences? Given (2.27) and (2.8), it is clear that $s_1^{SB,R} < s_1^{FB,R}$ if and only if

$$U_s(\theta_2, s_1^{SB,R}) - U_s(\theta_1, s_1^{SB,R}) > 0$$

(i.e., the standard SCP). $s_1^{SB,R} = s_1^{FB,R}$ if and only if

$$U_s(\theta_2, s_1^{SB,R}) - U_s(\theta_1, s_1^{SB,R}) = 0,$$

and $s_1^{SB,R} > s_1^{FB,R}$ if and only if

$$U_s(\theta_2, s_1^{SB,R}) - U_s(\theta_1, s_1^{SB,R}) < 0.$$

As (2.28) is identical to (2.9), it follows that $s_2^{SB,R} = s_2^{FB,R}$.

We now look at the optimal prices. Given that (2.2) and (2.17) are binding, we have

$$p_1 = \frac{1+\lambda}{2}U(\theta_1, s_1) \quad (2.29)$$

$$p_2 = \frac{1+\lambda}{2}U(\theta_2, s_2) - \frac{1+\lambda}{2}\left[U(\theta_2, s_1) - U(\theta_1, s_1)\right]. \quad (2.30)$$

The following answers two questions:

First, what is the effect of reference-dependent preferences on p_1 and p_2? From (2.30), given that the relation between $s_1^{SB,R}$ and $s_1^{SB,NR}$ is ambiguous, it follows that the relation between $p_1^{SB,R}$ and $p_1^{SB,NR}$ is also ambiguous. In (2.30), we know that $s_2^{SB,R} > s_2^{SB,NR}$, and $U(\theta_2, s_2)$ increases. We also know that $\frac{1+\lambda}{2} > 1$, ceteris paribus, increases p_2. However, given that the relation between $s_1^{SB,R}$ and $s_1^{SB,NR}$ is ambiguous, it follows that the effect of reference-dependent preferences on $[U(\theta_2, s_1) - U(\theta_1, s_1)]$ and, as a result on p_2, is also ambiguous.

Second, what is the effect of asymmetric information on p_1 and p_2? From (2.10) and (2.30). If $U_{s\theta} > 0$, $s_1^{SB,R} < s_1^{FB,R}$ and then $p_1^{SB,R} < p_1^{FB,R}$. If $U_{s\theta} = 0$, $s_1^{SB,R} = s_1^{FB,R}$ and then $p_1^{SB,R} = p_1^{FB,R}$. If $U_{s\theta} < 0$, $s_1^{SB,R} > s_1^{FB,R}$ and then $p_1^{SB,R} > p_1^{FB,R}$. Knowing that $s_2^{SB,R} = s_2^{FB,R}$, by comparing (2.11) and (2.30), it is clear that $p_2^{SB,R} < p_2^{FB,R}$. □

Note that it is not necessarily optimal to sell a different level of quality to both consumers. It could be better to sell the cheap menu to everyone and to not distort its quality; intuitively this happens when a majority of consumers are of the low-valuation type. On the other hand, it might be better to sell only the expensive menu and not to offer a discount, accepting the fact that the low-valuation consumers refuse to buy in this case; intuitively, this happens when the majority of consumers are of high-valuation type.

We next consider the monopolist's optimal price and quality when only offering high quality and a high price to the high-valuation consumers, and where the low-valuation consumers do not buy the good. Denote $s^{SB,R}$ and $p^{SB,R}$ ($s^{SB,NR}$ and $p^{SB,NR}$) the optimal price-quality combination under asymmetric information with reference-dependent preferences (without reference-dependent preferences). *SB* stands for second-best, and *R* and *NR* refer to reference-dependent preferences and no reference-dependent preferences respectively.

In Proposition 2.3, two types of comparisons are offered: first, a comparison between the first-best with and without reference-dependent preferences; second, a comparison between the first-best and the second-best when consumers have reference-dependent preferences. In this scenario, the monopolist sets the same price and quality as offered to the high-valuation consumer in the first-best. When consumers have standard preferences, the monopolist could offer a lower price-quality combination.

Proposition 2.3. *Consider the case of asymmetric information. Consider a single price-quality combination (s, p) that is bought only by high-valuation consumers. Then,*

1. $s^{SB,R} = s_2^{FB,R}$, $p^{SB,R} = p_2^{FB,R}$.
2. $s^{FB,NR} < s^{FB,R}$, $p^{FB,NR} < p^{FB,R}$.

Proof. The monopolist maximizes:

$$\begin{aligned} \underset{p,s}{maximize} \quad & q\,[p - C(s)] \\ s.t. \quad U(\theta_1, s) - \lambda p \quad & \leq 0 && (2.31)\\ U(\theta_2, s) - p \quad & \geq -\lambda U(\theta_2, s) + p. && (2.32) \end{aligned}$$

The constraints can be rewritten as:

$$\begin{aligned} p &\geq \frac{1}{\lambda} U(\theta_1, s) \\ p &\leq \frac{1+\lambda}{2} U(\theta_2, s). \end{aligned}$$

Note that by (A1), $\frac{1}{\lambda}U(\theta_1,s) < \frac{1+\lambda}{2}U(\theta_2,s)$. As the monopolist wants the price to be as high as possible for any given quality, it follows that $p = \frac{1+\lambda}{2}U(\theta_2,s)$, so that the monopolist sets $\frac{1+\lambda}{2}U_s(\theta_2,s) = C'(s)$. Therefore, the monopolist offers the same quality at the same price in terms of the high-valuation type.

Using the second result of Proposition 2.1, $s_i^{FB,R} > s_i^{FB,NR}$, $p_i^{FB,R} > p_i^{FB,NR}$, the monopolist could set a lower price and lower quality menu to be offered to high-valuation consumer, when high-valuation consumers have standard preferences instead of reference-dependent preferences. □

We finally consider the monopolist's optimal price and quality when offering the same low quality to both types of consumers. Proposition 2.4 offers two types of comparisons: first, a comparison between the second-best with and without reference-dependent preferences; second, a comparison between the first-best and the second-best when consumers have reference-dependent preferences.

The first part of Proposition 2.4 shows that when consumers have standard preferences, the price and quality offered should be lower than the case in which they have reference-dependent preferences. The second part of Proposition 2.4 suggests that when consumers have reference-dependent preferences, the monopolist sets the same price and quality in the second-best that would be offered to the low-valuation consumer in the first-best.

Proposition 2.4. *Consider the case of asymmetric information. Consider a single quality-price combination (s,p) that is bought by all consumers. Then,*

1. $s^{SB,R} = s_1^{FB,R}$, $p^{SB,R} = p_1^{FB,R}$.
2. $s^{SB,NR} < s^{SB,R}$, $p^{SB,NR} < p^{SB,R}$.

Proof. The monopolist maximization problem is:

$$\begin{aligned} \underset{p,s}{maximize} \quad & p - C(s) \\ s.t. \quad U(\theta_1,s) - p &\geq -\lambda U(\theta_1,s) + p \qquad (2.33) \\ U(\theta_2,s) - p &\geq -\lambda U(\theta_2,s) + p. \qquad (2.34) \end{aligned}$$

The above constraints can be rewritten as

$$p \leq \frac{1+\lambda}{2}U(\theta_1,s)$$
$$p \leq \frac{1+\lambda}{2}U(\theta_2,s).$$

Given (A1), the binding constraint is the first one. It follows that the monopolist ensures that the first constraint is met with equality and maximizes

$$\underset{s}{maximize} \frac{1+\lambda}{2}U(\theta_1,s) - C(s).$$

This leads to the low price and low quality offered in the first-best.

Recalling the second result of Proposition 2.1, $s_i^{FB,R} > s_i^{FB,NR}$, $p_i^{FB,R} > p_i^{FB,NR}$, the monopolist could set an even lower price, a lower quality menu to be offered to everyone, when consumers have standard preferences instead of reference-dependent preferences. $\square$

We are now ready to determine the conditions under which the monopolist prefers to offer two price-quality combinations rather than one. In Proposition 2.5, we limit ourselves to comparing the case of menu pricing to a situation of selling only to high-valuation, loss-averse consumers. We show that, with loss aversion, when few high-valuation consumers exist in the market, it is more likely that the monopolist prefers menu pricing. It may even be that menu pricing is the only desirable pricing if consumers are highly loss averse.

Proposition 2.5. *When consumers have reference-dependent preferences:*

1. There exists q^* such that for any $q < q^*$, the monopolist prefers menu pricing to only selling to the high-valuation consumers.

2. If $U_{s\theta} = 0$, then $\frac{\partial q^*}{\partial \lambda} > 0$.

Proof. By Proposition 2.2, given that high quality is the same as in the first-best, the monopolist's expected profit with menu pricing when agents have reference-dependent preferences is equal to

$$(1-q)\left[\frac{1+\lambda}{2}U(\theta_1, s_1^{SB,R}) - C(s_1^{SB,R})\right] +$$
$$q\left[\frac{1+\lambda}{2}U(\theta_2, s_2^{FB,R}) - \frac{1+\lambda}{2}\left[U(\theta_2, s_1^{SB,R}) - U(\theta_1, s_1^{SB,R})\right] - C(s_2^{FB,R})\right]. \quad (2.35)$$

By Proposition 2.4, given that quality is the same as offered to the high-valuation consumer in the first-best, the monopolist's expected profit when selling only to high-valuation consumers who have reference-dependent preferences is equal to

$$q\left[\frac{1+\lambda}{2}U(\theta_2, s_2^{FB,R}) - C(s_2^{FB,R})\right]. \quad (2.36)$$

Using (2.35) and (2.36), it can be calculated that the monopolist prefers menu pricing if and only if (2.35) > (2.36), that is

$$q < \frac{\frac{1+\lambda}{2}U(\theta_1, s_1^{SB,R}) - C(s_1^{SB,R})}{\frac{1+\lambda}{2}U(\theta_2, s_1^{SB,R}) - C(s_1^{SB,R})} = q^*.$$

To see the effect of changes in the level of loss aversion on the monopolist's decision, we now look at the derivative of q^* to λ,

$$\frac{\partial q^*}{\partial \lambda} = \frac{\left(\frac{1+\lambda}{2}\right)^2 \times \frac{ds_1^{SB,R}}{d\lambda}\left[U_s(\theta_1, s_1^{SB,R})U(\theta_2, s_1^{SB,R}) - U_s(\theta_2, s_1^{SB,R})U(\theta_1, s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}U(\theta_2, s_1^{SB,R}) - C(s_1^{SB,R})\right]^2}$$
$$+ \frac{\frac{1}{2}C(s_1^{SB,R})\left[U(\theta_2, s_1^{SB,R}) - U(\theta_1, s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}U(\theta_2, s_1^{SB,R}) - C(s_1^{SB,R})\right]^2}$$
$$+ \frac{\left(\frac{1+\lambda}{2}\right) \times \frac{ds_1^{SB,R}}{d\lambda}C(s_1^{SB,R})\left[U_s(\theta_2, s_1^{SB,R}) - U_s(\theta_1, s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}U(\theta_2, s_1^{SB,R}) - C(s_1^{SB,R})\right]^2}$$
$$- \frac{\left(\frac{1+\lambda}{2}\right) \times \frac{ds_1^{SB,R}}{d\lambda}C'(s_1^{SB,R})\left[U(\theta_2, s_1^{SB,R}) - U(\theta_1, s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}U(\theta_2, s_1^{SB,R}) - C(s_1^{SB,R})\right]^2}.$$

When $U_{s\theta} = 0$, given (2.27), the above derivative can be simplified to

$$\frac{\partial q^*}{\partial \lambda} = \frac{C(s_1^{SB,R})\left[U(\theta_2, s_1^{SB,R}) - U(\theta_1, s_1^{SB,R})\right]}{2\left[\frac{1+\lambda}{2}U(\theta_2, s_1^{SB,R}) - C(s_1^{SB,R})\right]^2},$$

that is always strictly positive. □

Proposition 2.6 complements Proposition 2.5 by comparing the case of menu pricing to the case of selling low quality and low price to all consumers. We show that if consumers are loss averse, it is more likely that the monopolist prefers menu pricing to selling the same low quality and low price to all consumers, when there are enough high-valuation consumers in the market. When consumers have an equal marginal utility of quality, regardless of the degree of their loss aversion, the monopolist prefers menu pricing

Proposition 2.6. *When consumers have reference-dependent preferences:*

1. There exists q^{**}, such that for any $q > q^{**}$, the monopolist prefers menu pricing to offering the same price and quality to all consumers.

2. If $U_{s\theta} = 0$, then $q^{**} = 0$ and $\frac{\partial q^{**}}{\partial \lambda} = 0$.

Proof. By Proposition 2.3, given that price and quality are the same as offered to the low-valuation consumer in the first-best, the monopolist's expected profits when offering one menu to all consumers is equal to

$$\left[\frac{1+\lambda}{2}U(\theta_1, s_1^{FB,R}) - C(s_1^{FB,R})\right]. \tag{2.37}$$

Therefore, the monopolist prefers menu pricing, if and only if (2.35) > (2.37), that is,

$$q > \frac{\frac{1+\lambda}{2}\left[U(\theta_1, s_1^{FB,R}) - U(\theta_1, s_1^{SB,R})\right] - \left[C(s_1^{FB,R}) - C(s_1^{SB,R})\right]}{\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]} = q^{**}.$$

From Proposition 2.2, we know that if $U_{s\theta} = 0$, then $s_1^{FB,R} = s_1^{SB,R}$. In this case, $q^{**} = 0$. Thus, when $U_{s\theta} = 0$, for any $q > 0$, the monopolist is more likely to choose menu pricing over selling the same price and quality to all consumers.

To see the effect of loss aversion on q^{**}, we look at its derivative with respect to λ (please see $\frac{\partial q^{**}}{\partial \lambda}$ in Appendix). According to Proposition 2.2, if $U_{s\theta} = 0$, then $s_1^{FB,R} = s_1^{SB,R}$. Consequently $\frac{\partial q^{**}}{\partial \lambda} = 0$. □

As seen, in Propositions 2.5 and 2.6, we were only able to come up with solid results when $U_{s\theta} = 0$, and so $s_1^{SB,R} = s_1^{FB,R}$. When $U_{s\theta} \neq 0$, distortions in $s_1^{SB,R}$ relative to $s_1^{FB,R}$ that are created partly by loss aversion makes it difficult to find unambiguous results. However, from (2.27), it is clear that if $U_{s\theta}$ is close to zero, distortions in s_1 are very small and close to zero. Therefore, the results of Propositions 2.5 and 2.6 holds for $U_{s\theta}$ just above and just below zero.

2.6 Multiple Personal Equilibria

In this section, in the case of asymmetric information, we will check whether there are other PEs, when multiple PEs exist, what is the PPE, and whether our analysis in the previous sections is maintained even when considering the possibility that multiple PEs exist.

The separating PE that we have described in Proposition 2.2 is achieved under the assumption that consumers expect to buy the good, and moreover, that rich consumers expect to receive high quality and poor consumers expect to receive low quality. If consumers may also expect not to buy the good, or expect to buy the other variety of the good, there may be other PEs. A poor consumer who expects to buy high quality, picks the high quality if and only if

$$U(\theta_1, s_2) - p_2 \geq U(\theta_1, s_1) - p_1 - \lambda\left[U(\theta_1, s_2) - U(\theta_1, s_1)\right] + (p_2 - p_1), \tag{2.38}$$

and a rich consumer who expects to buy low quality, picks the low quality if and only if

$$U(\theta_2,s_1)-p_1 \geq U(\theta_2,s_2)-p_2+[U(\theta_2,s_2)-U(\theta_2,s_1)]-\lambda(p_2-p_1). \tag{2.39}$$

(2.38) and (2.39) can be summarized as:

$$\frac{2}{1+\lambda}[U(\theta_2,s_2)-U(\theta_2,s_1)] \leq p_2-p_1 \leq \frac{1+\lambda}{2}[U(\theta_1,s_2)-U(\theta_1,s_1)]. \tag{2.40}$$

From (2.29) and (2.30), we know that p_2-p_1 in (2.40) can be replaced by

$$\frac{1+\lambda}{2}[U(\theta_2,s_2)-U(\theta_2,s_1)].$$

If (2.16) and (2.17) coexist with (2.40), multiple PEs exist. When the SCP is violated, for very large degrees of loss aversion, (2.40) is always valid, and thus multiple PEs always exist. For small and moderate degrees of loss aversion, if (2.40) is still valid, (2.16) and (2.17) will be slack. When neither (2.38) nor (2.39) are valid, the separating PE described by (2.16) and (2.17) coexists with the completely reversed separating PE in which the rich consumer picks the low-quality variety and the poor consumer picks the high-quality variety. In this case, the two pooling equilibria are, at the same time, also PEs. However, it may be that only one of the (2.38) or (2.39) is not valid. In this case, the separating PE described by (2.16) and (2.17) coexists with a pooling PE.

Similarly, the PE described in Proposition 2.4 is achieved under the assumption that the high-type consumer expects to buy while the low-type consumer expects not to buy. This PE can be seen as a separating PE, as only the high-type consumer buys the good and thus different types of consumers decide differently. For the same quality-price offer, multiple PEs may exist if consumers expect differently—i.e., the low-type consumer expects to buy and the high-type consumer expects not to buy. A poor consumer who expects to buy will buy if and only if

$$U(\theta_1,s)-p \geq -\lambda U(\theta_1,s)+p, \tag{2.41}$$

and a rich consumer who does not expect to buy, will not buy if and only if

$$0 \geq U(\theta_2,s)-\lambda p. \tag{2.42}$$

These conditions can be summarized as:

$$\frac{1}{\lambda}U(\theta_2,s) \leq p \leq \frac{1+\lambda}{2}U(\theta_1,s). \tag{2.43}$$

If (2.31) and (2.32) coexist with (2.43), multiple PEs exist. For large-enough degrees of loss aversion, (2.43) is always valid, and so multiple PEs will exist. For small and average degrees of loss aversion, (2.43) makes (2.31) and (2.32) slack. When both (2.41) and (2.42) are violated, the separating PE described by (2.31) and (2.32) coexists with a completely reverse PE in which the low-type consumer buys the good and the high-type consumer opts out. When only (2.41) is violated, the separating PE described by (2.41) and (2.42) coexists with the pooling PE where neither type buys. Finally, when only (2.42) is violated, the separating PE described by (2.41) and (2.42) coexists with the pooling PE that either type buys.

The pooling PE described in Proposition 2.4 was calculated based on the assumption that both high-valuation and low-valuation consumers expect to buy the good. Thus, multiple PEs may exist when one or both types expect differently. A poor consumer who expects not to buy, will not buy if and only if

$$0 \geq U(\theta_1, s) - \lambda p, \tag{2.44}$$

and a rich consumer who does not expect to buy, will not buy if and only if

$$0 \geq U(\theta_2, s) - \lambda p. \tag{2.45}$$

(2.44) and (2.45) can be rewritten as

$$\begin{aligned} p &\geq \frac{1}{\lambda} U(\theta_1, s) \\ p &\geq \frac{1}{\lambda} U(\theta_2, s). \end{aligned}$$

Given (A1), the first constraint is slack. Therefore, multiple PEs exist if (2.33) and (2.34) coexist with (2.45), that is,

$$p \leq \frac{1+\lambda}{2} U(\theta_1, s).$$

For large enough λ, (2.45) is always valid and multiple PEs exist. As (2.44) is slack, if (2.45) holds, a pooling PE exists where neither type buys.

In the case, multiple PEs exist, Kőszegi and Rabin (2006) assume that the consumer is able to choose his PPE. When will the separating PE described by (2.16) and (2.17) be the PPE? The condition for the rich consumer's PPE to be that he chooses the high-quality variety, and for the poor consumer that he chooses the low-quality variety, leads to incentive constraints that are indistinguishable from the self-selection constraints of consumers that do not have reference-dependent preferences. Along with the condition that the consumers also should prefer a PE in which they buy a variety of the good to a PE in which they do not, one obtains a model that is indistinguishable from the one without reference-dependent preferences. However, once you allow for the PPE concept, the analysis in the previous sections cannot be maintained. For example, for the optimal separating prices given by (2.30) and (2.30), when the high- and low-quality options are both available as PEs for the high-valuation consumer, he will not pick the high-quality option as the PPE.

An alternative theory that we could use to select equilibria, when multiple PEs exist, is strategic framing. Framing may take place if the monopolist is able to influence consumers' expectations and thereby their reference points. According to Armantier and Boly (2012), " An agent's reference point may be manipulated through framing, a phenomenon the reference-dependent models recently proposed by (e.g.) Kőszegi and Rabin (2006; 2007) cannot explain at this point".

The idea behind framing is that the decision problem can be viewed from different perspectives. For example, consider a monopolist that can advertise her product through different magazines that are used by different types of consumer. Thus, the monopolist has the chance to advertise differently for different types of consumer. She can highlight low price in the magazine read mainly by poor consumers and high quality in the magazine read mainly by rich consumers. When consumers are susceptible to framing, they do not necessarily play their PPE, but rather the PE that is most preferred by the monopolist. We assume that the monopolist can ensure that a consumer forms expectations such that he expects to buy the menu designed for his type when there are multiple PEs available.

2.7 Conclusion

In this paper, we explore optimal menu pricing by a profit-maximizing monopolist who faces two types of consumers, a high type (rich) and a low type (poor). Consumers are loss averse and have reference-dependent preferences in the sense of Kőszegi and Rabin (2006); that is, consumers compare their consumption outcomes to their reference points that are set by their rational expectations about their consumption outcomes. Consumers' types affect their expectations of the consumption outcome—i.e., their reference points. We examine whether it becomes more or less difficult for the monopolist to make consumers self-select when consumers have reference-dependent preferences and are loss averse. Our paper follows the line of research pioneered by Heidhues and Kőszegi (2008), Hahn et al. (2014), and Herweg and Mierendorff (2013) among others, in studying asymmetric information situations with behavioral biased agents. This paper explains price discrimination in that when consumers are loss averse, price discrimination is possible for a wider range of parameters.

A comparison between our results with the results of the above-mentioned literature shows that the timing of reference-point formation is an influential factor in determining the optimal pricing behavior of firms. We show that when consumers' reference points are formed after they learn their types, and thus each consumer's reference point is dependent upon his type, consumers are more prone to self-select their types, and thus menu pricing is facilitated. Type-dependent reference points are also studied by Carbajal and Ely (2016).

In particular, we show that with reference-dependent preferences, price discrimination becomes easier, as the single-crossing property does not need to be satisfied in order to run price discrimination; whereas the single-crossing property requires that richer consumers have a higher marginal utility of quality. With reference-dependent preferences, consumers self-select their types naturally once rich consumers expect to buy high quality at a high price, and poor consumers expect to buy low quality at a low price.

Further, we prove that the reference quality and price expected by consumers has an effect on the optimal menu. This effect is caused by loss aversion under symmetric and asymmetric information. With reference-dependent preferences, the monopolist prefers menu pricing to offering a single high price only to rich consumers, and she also prefers menu pricing to offering a low quality and low price to all consumer when both types have the same marginal utility of quality. When consumers do not have the same marginal utility, the story may be different. In general, self-selection, which is necessary for menu pricing to work, succeeds only if the monopolist offers high quality at a discount. The monopolist may avoid menu pricing if the discount needed to make it work is too large. However, the monopolist does not need to offer such a high discount, since a rich consumer with reference-dependent preferences is less inclined to choose the low-quality variant of the monopolist's product given his high aversion to quality loss. Therefore, the monopolist is more likely to prefer menu pricing.

In our model, multiple personal equilibria may exist, because the consumer has reference-dependent preferences and the same quality, price combination can support several expectation sets. Put differently, for the same price and quality it may be a personal equilibrium that the consumer buys the good because he expects to buy it, and it may also be a personal equilibrium that he does not buy it because he does not expect to buy it. If there are multiple personal equilibria, we assume that the monopolist can ensure that the consumer forms expectations such that he expects to buy a particular variety of the good. This may take place if the monopolist is able to influence consumer expectations and thereby consumer reference point.

The model that we consider in this study is simple and abstracts from many aspects that would make it more realistic. For example, the degree of loss aversion may not be the same for all consumers. Different consumers do not comprise just two types, and may put different weights on the gain-loss utility. As first step, we did not start out with these general assumptions, which call for more complex mathematical challenges, to be able to focus merely on the effect of loss aversion on the outcome of optimal menu pricing.

An interesting direction for extending our paper is analyzing how consumer loss aversion and reference-dependent preferences affect the pricing behavior of a multiple-product monopolist. As suggested by Heidhues and Kőszegi (2004), if consumers purchase some joint products at the same time, a firm might decide to correlate prices across goods. We leave these topics for future research.

2.8 Appendix

$$\frac{\partial q^{**}}{\partial \lambda} = \frac{-\frac{1}{2}\left[U(\theta_1, s_1^{FB,R}) - U(\theta_1, s_1^{SB,R})\right]\left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\right]^2}$$

$$+ \frac{\frac{1}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right]\left[C(s_1^{FB,R}) - C(s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\right]^2}$$

$$+ \frac{-\frac{1+\lambda}{2} \times \frac{ds_1^{SB,R}}{d\lambda}\left[U_s(\theta_1, s_1^{FB,R}) - U_s(\theta_1, s_1^{SB,R})\right]\left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\right]^2}$$

$$+ \frac{\frac{1+\lambda}{2} \times \frac{ds_1^{SB,R}}{d\lambda}\left[U_s(\theta_2, s_2^{FB,R}) - U_s(\theta_2, s_1^{SB,R})\right]\left[C(s_1^{FB,R}) - C(s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\right]^2}$$

$$+ \frac{-\frac{1+\lambda}{2} \times \frac{ds_1^{SB,R}}{d\lambda}\left[C'(s_1^{FB,R}) - C'(s_1^{SB,R})\right]\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\right]^2}$$

$$+ \frac{\frac{1+\lambda}{2} \times \frac{ds_1^{SB,R}}{d\lambda}\left[C'(s_2^{FB,R}) - C'(s_1^{SB,R})\right]\left[U(\theta_1, s_1^{FB,R}) - U(\theta_1, s_1^{SB,R})\right]}{\left[\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\right]^2}$$

$$+ \frac{\left(\frac{1+\lambda}{2}\right)^2 \times \frac{ds_1^{SB,R}}{d\lambda}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right]\left[U_s(\theta_1, s_1^{FB,R}) - U_s(\theta_1, s_1^{SB,R})\right]}{\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]}$$

$$+ \frac{-\left(\frac{1+\lambda}{2}\right)^2 \times \frac{ds_1^{SB,R}}{d\lambda}\left[U(\theta_1, s_1^{FB,R}) - U(\theta_1, s_1^{SB,R})\right]\left[U_s(\theta_2, s_2^{FB,R}) - U_s(\theta_2, s_1^{SB,R})\right]}{\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]}$$

$$
+ \frac{\left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]\left[C_s(s_1^{FB,R}) - C_s(s_1^{SB,R})\right]\frac{ds_1^{SB,R}}{d\lambda}}{\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]}
$$

$$
+ \frac{-\left[C(s_1^{FB,R}) - C(s_1^{SB,R})\right]\left[C_s(s_2^{FB,R}) - C_s(s_1^{SB,R})\right]\frac{ds_1^{SB,R}}{d\lambda}}{\frac{1+\lambda}{2}\left[U(\theta_2, s_2^{FB,R}) - U(\theta_2, s_1^{SB,R})\right] - \left[C(s_2^{FB,R}) - C(s_1^{SB,R})\right]}.
$$

Chapter 3

Signaling Game with Loss-Averse Senders

3.1 Introduction

Examples of communication through costly signaling among individuals with divergent goals or interests are numerous in nature and human society (Bradbury, 1998; Riley, 2001). This subject has attracted a lot of attention from researchers of differenet fields, specially in the areas of economics, finance, and biology (Kreps and Wilson, 1982; Leland and Pyle, 1977; Grafen, 1990). Models of costly signaling suggest that proper signal costs can encourage honest communication or credible signaling by making high-quality signals hard to fake. If sending a false signal is costly enough, sending it becomes counterproductive.[1]

Most of the classical literature on costly signaling games consider senders to be rational and achieve credible signaling (i.e., separating equilibrium) by assuming the *single-crossing property (SCP)* (see Riley, 1979, and Cai et al., 2007).[2] The SCP requires that higher quality senders have either lower marginal costs of signaling or higher marginal benefits.This paper aims at explaining how reference-dependent preferences and loss aversion may additionally push senders to signal credibly. We use job market signaling (Spence, 1973) as our leading example throughout the paper.

To gain an intuition, consider the following example of strategic signaling. An employer faces two types of workers, one with high productivity and one with low productivity, but she cannot tell them apart. The workers invest in obtaining some degree of education in order to signal their high productivity to the employer. The question we want to investigate is whether it is more or less feasible to achieve credible signaling when senders have reference-dependent preferences and are subject to loss aversion in the sense of Kőszegi

[1] Mas-Colell et al. (1995, p.450) defines signaling in the labor market context as situations where high-ability workers may have actions they can take to distinguish themselves from their low-ability counterparts. Spence (1973) formalized this idea and analyzed its implications. In his model, education serves as a pre-contractual signal of productivity. Informal versions of signaling can be found in sociological literature before Spence (see Berg 1970).

[2] When considering violation of the SCP, classical literature only finds pooling or bunching equilibria (see, for examples, Bernheim, 1994; Bagwell and Bernheim, 1996; Glazer and Konrad, 1996).

and Rabin (2006; 2007). This study suggests that senders with reference-dependent preferences are naturally more inclined to signal credibly. The intuition behind this thinking can be explained as follows. Assume that high-type workers expect to acquire a high level of education, and low-type workers expect to acquire a low level of education, assume also that these expectations determine their reference points, respectively. If the loss-averse, high-type worker considers acquiring a low education level instead of a high education level, he will focuse more on the loss in wages than on the gain in education costs, and is therefore less inclined to switch to acquiring a low education level. Similarly, if the loss-averse, low-type worker considers acquiring a high education level instead of a low education level, he will focuse more on the loss of the education cost than on the gain in salary, and will be less inclined to upgrade to a high education level.

We build a simple version of Spence's (1973) model of job market signaling,[3] with the added feature that senders are subject to loss aversion and have reference-dependent preferences. We follow Kőszegi and Rabin's (2006) model of reference-dependent preferences, in which senders' total utility is composed of two additive terms: *intrinsic utility*,[4] and *gain-loss utility*. Gains and losses are experienced with respect to the reference point determined by the senders' rational expectations. It is assumed that senders are always in a *personal equilibrium (PE)*, in which their optimal action is chosen and their expectations are fulfilled. In Kőszegi and Rabin's model, there may be multiple PEs, wherein for the same wage offered, it may both be a PE that the worker obtains higher education because he expects to obtain it, and that he does not obtain it because he does not expect to obtain it. If there are multiple PEs, senders can ex ante compare the different PEs, and pick out the PE best for them, which is referred to as the *preferred personal equilibrium (PPE)*.

The first key result of our study suggests that if senders are loss averse and have reference-dependent preferences, the separating equilibrium may exist even when the SCP is violated. In terms of job market signaling, we argue that the fact that workers have reference-dependent preferences may facilitate achieving credible signaling, compared to the standard model in which workers have standard preferences, even when the SCP is violated.

The second result of this study shows that if expectation-based loss aversion and reference-dependent preferences give rise to multiple PEs, that is, the separating PE coexists with some additional PEs, the separating PE could still be selected as the PPE by workers. At least we will never have the case in which "the pooling PE is always the PPE", as long as the senders are highly loss averse.

The following section discusses the related literature briefly. Section 3.3 introduces our signaling model. Section 3.4 discusses all equilibria and looks at the possibility for multiple personal equilibria to arise and derives conditions under which separating is preferred by senders. Section 3.5 concludes.

---

[3]Our model could be also built on Zahavi (1975), which is a counterpart of Spence (1973) in a biology context.

[4]Utility derived directly from consumption that is dependent on the consumers' types.

3.2 Literature Review

This study belongs to the line of research that adopts the assumptions of prospect theory (Kahneman and Tversky, 1979) and incorporates agents with prospect-theoretic preferences or reference-dependent preferences into standard industrial organization models (for an overview, see Kőszegi, 2014, and Kim and Lee, 2014). We depart from this literature by finding separating equilibrium when agents are loss averse and have reference-dependent preferences, whereas the literature finds pooling equilibria (see, for example, Heidhues and Kőszegi, 2004; 2008).

Among the studies of the effect of loss aversion and reference-dependent preferences on market outcome, Hahn et al. (2014), Heidhues and Kőszegi (2008), Herweg and Mierendorff (2013), and Carbajal and Ely (2016) are related to our paper, although they are not within the signaling theory literature. These authors adopt Kőszegi and Rabin's (2006; 2007) theory of rational expectations as a reference point to study firm's pricing behavior.

Hahn et al. (2014) and Heidhues and Kőszegi (2008) predict price stickiness[5] when consumers have reference-dependent preferences. In a setting similar to Hahn et al., Herweg and Mierendorff (2013) demonstrate that consumer loss aversion can explain the prevalence of flat-rate contracts. The reason that these authors find pooling equilibria is because they assume that the reference point is formed *before* the senders learn their types. This causes stickiness around this reference point even though it is a stochastic point. We assume instead that senders form their reference points *after* they have discovered whether they are of high or low type, and we find separating equilibria. Intuitively, when the reference point is set ex ante, there is stickiness around it, making separation more difficult. When the reference point is set ex post, there is stickiness for each type around the menu chosen, thus making separation easier.[6] Carbajal and Ely (2016), in their monopoly price discrimination model, consider different ways of reference point formations and argue that the specifications of the optimal menu depends on the formation of the reference point. They show that depending on the reference plan, there may be different answers for the optimal price discrimination problem.

3.3 Model

Assume that there are at least two employers and some workers of two different types interacting in a job market. In stage 1, Nature decides on the type of the worker. The worker is of type θ_1 with probability $1-q$, or of type θ_2 with probability q. The worker learns his type, but the employer does not. In stage two, the worker, based on his knowledge of his type, chooses to send either e_1 or e_2 as a signal, where $e_1 < e_2$.

[5]The resistance of prices to changes in cost or demand.

[6]Baker et al. (2016) and Shapiro and Zhuang (2015) should be mentioned as further relevant studies. These authors study dividend policy in which investors are loss averse to dividend cuts and their reference points are established by previous dividends. They show that the likelihood of dividend initiation and dividend size is influenced not only by firms' characteristics, but also by investors' preferences. However, they assume that the receivers have reference-dependent preferences rather than the sender in our model.

In stage three, the employer observes this signal—but not the type of the worker—and offers a wage to the worker. For simplicity, we assume that the employer can choose only one of two possible wages, w_1 and w_2, which respectively equal the productivity of the θ_1 and θ_2—i.e., $w_i(e_i) = \theta_i$. In a separating equilibrium, competition between the employers forces them to pay wages that equal the worker's expected productivity. This is enforced when more than one employer exists in the market (see Mas-Colell et al., 1995, p.437). In a pooling equilibrium, we assume that the employer pays whichever of the two wages that is closest to the expected productivity.

In stage four, workers accept or reject the offered wage; both players, worker and employer, receive payoffs depending on the worker's type, the signal chosen by the worker, and the wage schedule offered by the employer. The payoff of the employer will be a function of the productivity received minus the wage paid $u_E(\theta - w)$. The payoff of the workers, as they are loss averse, consists of two components: intrinsic utility and gain-loss utility. The intrinsic utility is directly obtained from the wage received minus the cost incurred by acquiring education $w - c(e,\theta)$. Following the standard model of job market signaling (Spence, 1973), we assume that the cost of zero education is zero for everyone, $c\,(0,\theta_i) = 0$, a higher effort costs more for all types, $c_e(e,\theta_i) > 0$, and the signal is less costly for the sender in the higher state than for the sender in the lower state, $c_\theta(e,\theta) < 0$. However, contrary to Spence, we do not systematically assume the SCP that can be formulated as follows: For any $e_2 > e_1$,

$$(SCP): \qquad U(\theta_2, s_j) - U(\theta_2, s_k) > U(\theta_1, s_j) - U(\theta_1, s_k),$$

meaning that the marginal cost of education is smaller for the higher types.

The gain-loss utility of each worker is determined by reference wage w_r and reference level of education e_r that he forms. Following Kőszegi and Rabin (2006), we assume that the worker's reference wage and education level are determined each time by the wage- and education-level combination that he expects to choose. When the worker expects to receive wage w_r but instead receives the higher wage w, he experiences a gain $\eta\,(w - w_r)$, which is added to his intrinsic utility. When his received wage is lower than w_r, he experiences a loss $\eta\lambda\,(w_r - w)$, which is subtracted from his intrinsic utility. The parameter λ, with $\lambda \geq 1$, reflects the worker's degree of loss aversion, and the parameter η, with $\eta \geq 0$, reflects the weight that he puts on the gain-loss utility part of his total utility function (if $\eta = 1$ means that the worker cares about intrinsic utility and gain-loss utility equally). In the same way, when the worker expects to obtain education e_r at cost $c\,(e_r,\theta)$, but instead obtains lower education e at lower cost $c\,(e,\theta)$, he experiences a gain $\eta\,[c\,(e_r,\theta) - c\,(e,\theta)]$, which is added to his intrinsic utility. If instead he obtains a higher level of education, he experiences a loss due to higher costs of higher education $\eta\lambda\,[c\,(e,\theta) - c\,(e_r,\theta)]$, which is subtracted from his intrinsic utility. For example, the overall utility of a worker of type θ_2 who expected to acquire education e_2 and obtain wage w_2, but ends up acquiring $e_1 < e_2$

and obtaining $w_1 < w_2$ equals

$$u_W(\theta_2) = w_1 - c(e_1, \theta_2) + \eta\left[c(e_2, \theta_2) - c(e_1, \theta_2)\right] - \eta\lambda(w_2 - w_1).$$

We thus obtain a simple model of signaling with an added feature that workers have reference-dependent preferences. We employ Kőszegi and Rabin's (2006) notion of a personal equilibrium (PE) in which each sender chooses his optimal action given his reference point, while his reference point is decided by the optimal action that he expects to take. Therefore, a PE exists when the optimal choice of a sender given his reference point, and his reference point given his expected choice, are consistent. In our model, as in Kőszegi and Rabin's, there may be multiple PEs. Because of loss aversion, for exactly the same wage and education level, it may be a PE for a sender to acquire an education level because he expects to, or to not acquire an education level because he does not expect to. If there are multiple PEs, Kőszegi and Rabin assume that the only equilibrium that can ever be played is the preferred personal equilibrium (PPE), which is the PE that leaves the sender best off. It is implicitly assumed that the sender is able to anticipate his utility of having each specific expectation, and he can choose his reference point in a way that maximizes his utility. Therefore, he chooses the PE that makes him better off, namely the PPE.

3.4 Equilibrium Characterization

Let us first briefly look at the case with standard preferences. Considering that in any equilibrium, wages are equal to productivity of workers, the incentive constraints of workers, in order to choose the separating equilibrium, can be written as follows:

$$\theta_2 - \theta_1 \leq c(e_2, \theta_1) - c(e_1, \theta_1) \tag{3.1}$$

$$\theta_2 - \theta_1 \geq c(e_2, \theta_2) - c(e_1, \theta_2). \tag{3.2}$$

A necessary condition for satisfaction of 3.1 and 3.2 and consequently for the existence of a separating equilibrium, is that the SCP must be satisfied. If the SCP is valid, separating equilibria exist if and only if the incentive constraints 3.1 and 3.2 are satisfied; i.e.,

$$c(e_2, \theta_2) - c(e_1, \theta_2) \leq \theta_2 - \theta_1 \leq c(e_2, \theta_1) - c(e_1, \theta_1).$$

We continue this section by looking at a case in which workers have reference-dependent preferences. We consider both a situation where the SCP is valid, and where it is violated. Below, Proposition 3.1 shows that if workers are highly loss averse, separating equilibria may exist even when the SCP is violated. Intuitively, if the high-productivity worker expects a high wage, he experiences a loss when receiving the low wage rather than the high wage, thus increasing his willingness to invest in a higher level of education. In turn, if the low-productivity worker expects a lower cost of education, he experiences a loss when paying a high cost of education, thus increasing his willingness to put up with a lower wage. Figure 3.1 illustrates the situation discussed in Proposition 3.1.

Proposition 3.1. *The following statements hold for the two-state signaling model.*

1. If the SCP is valid, then

a) when there is no loss aversion, separating equilibrium exists if and only if

$$c(e_2,\theta_2)-c(e_1,\theta_2)\leq\theta_2-\theta_1\leq c(e_2,\theta_1)-c(e_1,\theta_1);$$

b) when there is loss aversion, separating equilibrium exists if and only if

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)]\leq\theta_2-\theta_1\leq\frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)].$$

2. If the SCP is violated, then

a) no separating equilibrium exists, when the degree of loss aversion is low;

b) separating equilibrium still exist when the degree of loss aversion is sufficiently high.

Proof. 1. Assume the SCP is valid.

a) As we discussed in the beginning of this section, formulas (3.1) and (3.2) provide necessary and sufficient conditions for the existence of separating equilibrium when no loss aversion exists. The incentive constraints for workers to choose the separating equilibrium can be written as follows:

$$\begin{aligned}w_1-c(e_1,\theta_1)\geq& w_2-c(e_2,\theta_1)+\eta(w_2-w_1)-\eta\lambda\left[c(e_2,\theta_1)-c(e_1,\theta_1)\right]\\ w_2-c(e_2,\theta_2)\geq& w_1-c(e_1,\theta_2)+\eta\left[c(e_2,\theta_2)-c(e_1,\theta_2)\right]-\eta\lambda(w_2-w_1),\end{aligned}$$

as $w_i=\theta_i$ in equilibrium, the above conditions can be rewritten as

$$\theta_2-\theta_1\leq\frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)] \tag{3.3}$$

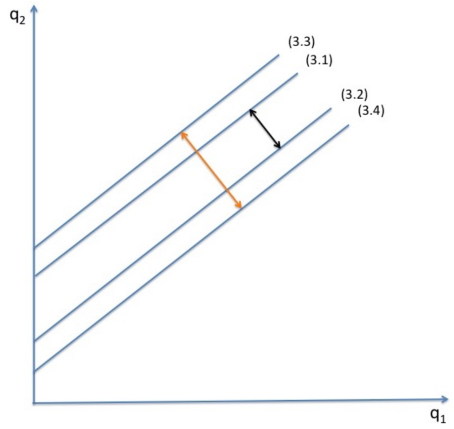
$$\theta_2-\theta_1\geq\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)]. \tag{3.4}$$

b) A necessary condition for satisfaction of (3.3) and (3.4) and consequently for the existence of the separating equilibria is:

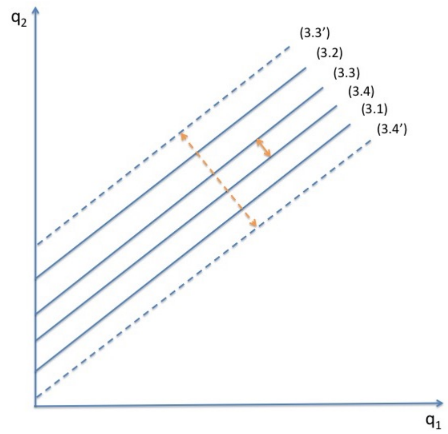
$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)]\leq\frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)],$$

which is equal to the SCP for $\lambda=1$. A separating equilibria exist if and only if the incentive constraints (3.3) and (3.4) are satisfied; i.e.,

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)]\leq\theta_2-\theta_1\leq\frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)],$$



(a) SCP is valid, $\lambda > 1$.



(b) SCP is violated, λ is large enough to guarantee the existence of separating equilibria.

Figure 3.1: Relative positions of the incentive constraints when the SCP is valid and when it is violated.

which may hold even when the SCP is violated if λ is sufficiently large.

2. Assume the SCP is violated.

a) If loss aversion is low (i.e., λ is close to 1), formulas (3.3) and (3.4) are equivalent to (3.1) and (3.2), and the situation is similar to standard signaling. In this case, separating equilibria do not exist at all when the SCP is violated.

b) The higher the degree of loss aversion, the larger the feasible area that allows for the existence of separating equilibria that were defined by (3.3) and (3.4). □

Figure 3.1 shows the relative positions of the incentive constraints (3.1), (3.2), (3.3), and (3.4) when the SCP is valid and when it is violated. Orange arrow: the feasibility area for the existence of separating equilibria when loss aversion exists. Black arrow: the feasibility area for the existence of separating equilibria without loss aversion. Dashed orange arrow: feasibility area for the existence of separating equilibria when the degree of loss aversion is sufficiently large. As we see in Figure 3.1a, when the SCP is valid, (3.3) and (3.4) define a larger area than (3.1) and (3.2); that is, loss aversion leads to a larger feasibility area for the existence of separating equilibria. In Figure 3.1b, when the SCP is violated, (3.2) lies above (3.1); that is, without loss aversion there are no θ_1 and θ_2 that satisfy the incentive constraints, and thus no separating equilibrium. While with loss aversion, however, there may be still some θ_1 and θ_2 that satisfy the incentive constraints and result in a separating equilibrium if λ is sufficiently large. The larger the degree of loss aversion, the larger the area defined by (3.3) and (3.4). In Figure 3.1, this is shown by dashed lines. Thus, in general, we should say that loss aversion increases the scope for separating equilibria no matter whether the SCP is valid or violated.

Because multiple PEs are a possibility, in the reference-dependent model, it does not suffice to check only whether a separating PE exists; it may be that for the same wage scheme, a pooling equilibrium is also a PE. We recall our assumption that only θ_1 or θ_2 could be offered as a wage in a pooling equilibrium. In the following propositions, we treat two cases that differ in whether θ_1 or θ_2 is offered in a pooling equilibrium.

First, in Proposition 3.2 and 3.3, we look at a case in which the majority of workers are of the high type and thus θ_2 would be offered in any pooling equilibrium. Proposition 3.2 derives conditions for the coexistence of the separating PE and the pooling PE by checking whether separating wages support pooling. In addition, Proposition 3.3 will discuss the situation in which the separating PE will be the PPE.

Proposition 3.2. *Assume pooling wage θ_2 is offered in any pooling equilibrium.*

1. When the SCP is valid, the pooling PE (e_2, θ_2) coexists with the separating PE for the same wage schedule if and only if

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)] \leq \theta_2-\theta_1 \leq \frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)],$$

and when the condition below is satisfied only the separating PE is played

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)] \leq \theta_2-\theta_1 \leq \frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)].$$

2. When the SCP is violated and a separating PE exists, the pooling PE (e_2, θ_2) coexists with the separating PE for the same wage schedule everywhere.

Proof. Assume that two wages θ_1 and θ_2 are offered, and incentive constraints 3.3 and 3.4 are satisfied, and thus a separating PE exists. Assume that the majority of workers are of the high type, and so θ_2 is always offered in a pooling equilibrium, and the low-productivity worker expects a high wage and a high education. A worker of low-productivity type would choose to do pooling if he is better off choosing a high wage and a high education instead of a low wage and a low education; that is, the constraint below must be true for him,

$$\theta_2 - c(e_2,\theta_1) \geq \theta_1 - c(e_1,\theta_1) + \eta\,[c(e_2,\theta_1)-c(e_1,\theta_1)] - \eta\lambda\,(\theta_2-\theta_1),$$

which can be simplified to

$$\theta_2-\theta_1 \geq \frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)]. \tag{3.5}$$

Multiple PEs exist if and only if the separating incentive constraints (3.3) and (3.4) hold simultaneously with (3.5).

Constraint (3.5) holds for strong enough loss aversion. Hence, for strong enough loss aversion, whether the low-productivity worker expects a high wage, a high education or a low wage, a low education, his expectations can lead to a PE. Thus, if the degree of loss aversion is sufficiently large, there exists a high wage, a high education pooling PE on top of the separating PE, and both equilibria are supported by the same wage schedule.

Technically, as far as the area defined by (3.3) and (3.4) is not empty, it overlaps with the area defined by (3.5) because the right-hand side (RHS) of (3.5) is always smaller than the RHS of (3.3). However, the RHS of (3.5) can be larger or smaller than the RHS of 3.4, depending on the SCP being valid or violated:

1. If the SCP is valid, the RHS of (3.5) is larger than the RHS of (3.4), and multiple PEs exist as far as (3.3) and (3.5) are satisfied; that is,

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)] \leq \theta_2-\theta_1 \leq \frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)].$$

Note that only the separating equilibrium is played everywhere between (3.4) and (3.5); that is,

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)] \leq \theta_2-\theta_1 \leq \frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)],$$

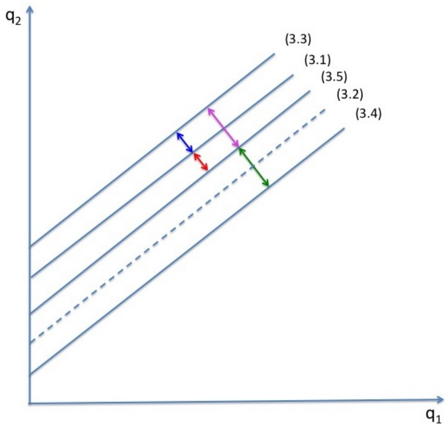
(see Figures 3.2a and 3.2b for intuition).

2. If the SCP is violated, the RHS of (3.5) is smaller than the RHS of (3.4), and multiple PEs exist as far as (3.3) and (3.4) are satisfied, or as far as separating equilibria exist (see Figures 3.2c and 3.2d for intuition). □

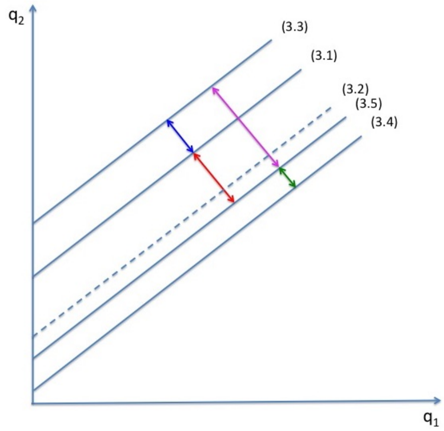
Consider the relative position of (3.5) with respect to positions of incentive constraints (3.3) and (3.4) in Figure 3.2. If the SCP holds, in the area between (3.5) and (3.4), only the separating PE is played while multiple PEs exist in the area between (3.3) and (3.5). If the SCP is violated, multiple PEs exist in the entire area between (3.3) and (3.4).

Even though loss aversion increases the scope for separating equilibriua, it is not clear whether separating equilibria can be also the PPE, when separating equilibria coexist with the pooling PE. Proposition 3.3 provides conditions under which the separating PE will be the PPE, when θ_2 is offered as a pooling wage and multiple PEs exist. The first result of Proposition 3.3 shows that, when the SCP is valid—i.e., the marginal cost of education is lower for the high type than for the low type—the low-type worker will choose the separating PE and obtain a low education as far as the difference between high and low productivity is smaller than the difference between the cost of high and low education. Moreover, the higher the degree of loss aversion, the more likely the low-type worker chooses separating and obtains low education. In other words, when there is loss aversion, the cost of obtaining higher education is more painful to the low-type worker, which makes him more likely to choose a lower education and a lower wage.

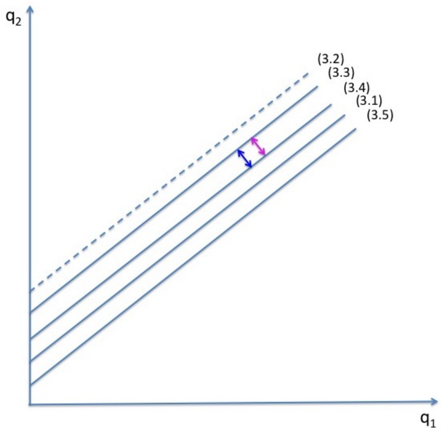
The second result of Proposition 3.3 suggests that when the SCP is violated—i.e., the high-type worker has at least the same marginal cost of education as the low-type worker—the high-type worker may doubt obtaining high education. In this case, he will choose separating and obtaining a high education if his degree of loss aversion is high enough, and thus he cares enough about the loss of obtaining a lower rather than a higher wage that he would receive in obtaining a low education.



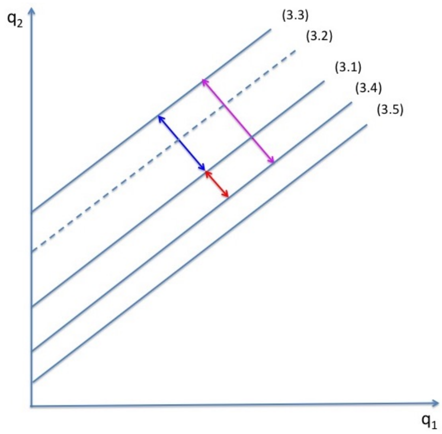
(a) SCP valid, λ small.



(b) SCP valid, λ large.



(c) SCP violated, λ small (but large enough to guarantee the existence of separating equilibria).



(d) SCP violated, λ large.

Figure 3.2: When the pooling PE of high education, high wage coexists with the separating PE for the same wage schedule.

Proposition 3.3. *The following statements hold if the separating PE coexists with the pooling PE (e_2, θ_2) for the same wage schedule.*

1. When the SCP is valid, the separating PE is the PPE if and only if

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)] \leq \theta_2-\theta_1 \leq c(e_2,\theta_1)-c(e_1,\theta_1),$$

which is non-empty for any $\lambda > 1$; otherwise, the pooling PE is the PPE.

2. When the SCP is violated, the separating PE is the PPE if and only if

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)] \leq \theta_2-\theta_1 \leq c(e_2,\theta_1)-c(e_1,\theta_1),$$

which can be non-empty only for large enough λ; otherwise, the pooling PE is the PPE.

Proof. In case the pooling PE (e_2, θ_2) coexists with the separating PE, for the high-productivity worker, pooling and separating strategies are the same. As workers are able to anticipate their expected utilities, the low-productivity worker prefers separating expectations to pooling expectations if the anticipated outcome of a separating strategy is higher than the anticipated outcome of a pooling strategy:

$$\theta_1 - c(e_1,\theta_1) > \theta_2 - c(e_2,\theta_1),$$

which is the same as (3.1). Hence, a separating equilibrium can be the PPE for some $\theta_2 - \theta_1$ that satisfy both (3.1) and (3.5) as far as incentive constraints (3.3) and (3.4) hold.

Note that the RHS of (3.1) is always larger than the RHS of (3.5); and the RHS of (3.3) is always larger than the RHS of (3.1). As is shown in Figure 3.2, (3.1) lies always above (3.5), and (3.3) lies always above (3.1). However, the relative position of (3.5) and (3.4) depends on the SCP being valid or violated:

1. When the SCP is valid, the RHS of (3.5) is always larger than the RHS of (3.4). Thus, for all θ_1 and θ_2 satisfying (3.1) and (3.5) or simply for all θ_1 and θ_2 satisfying

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_1)-c(e_1,\theta_1)] \leq \theta_2-\theta_1 \leq c(e_2,\theta_1)-c(e_1,\theta_1),$$

all the (3.3), (3.4), (3.5), and (3.1) are satisfied, and the separating PE is the PPE (see Figures 3.2a and 3.2b). For all θ_1 and θ_2 that do not satisfy (3.1) but satisfy (3.3), the pooling PE is the PPE.

2. When the SCP is violated, the RHS of (3.5) is always smaller than the RHS of (3.4). In this case, separating PE is the PPE for all $\theta_2 - \theta_1$ limited by (3.1) and (3.4). Let us see if this interval can be empty.

Depending on the magnitude of λ, the RHS of (3.4) can be either smaller than the RHS of (3.1) or larger than it:

a) If the RHS of (3.4) is smaller than the RHS of (3.1), then

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)] < c(e_2,\theta_1)-c(e_1,\theta_1),$$

which can hold only for large enough λ. In this case, for all $\theta_2-\theta_1$ limited by (3.1) and (3.4)or simply for all $\theta_2-\theta_1$ satisfying

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)] \leq \theta_2-\theta_1 \leq c(e_2,\theta_1)-c(e_1,\theta_1),$$

the separating PE is preferred to the pooling PE (e_2,θ_2) (see Figures 3.2c and 3.2d). For all other θ_1 and θ_2 that bring up multiple PEs, the pooling PE (e_2,θ_2) is preferred to the separating PE (see Figure 3.2d).

b) If the RHS of (3.4) is larger than the RHS of (3.1), then the area defined by (3.1) and (3.5) has no overlap with the area defined by (3.3) and (3.4) (see Figure 3.2c). As a result, the pooling PE (e_2,θ_2) is preferred to the separating PE. Note that this can happen only for smal enough λ (λ should still be large enough in order to have any separating equilibria). □

Figure 3.2 visualizes the situations discussed in Propositions 3.2 and 3.3: the case that the pooling PE (e_2,θ_2) coexists with the separating PE for the same wage schedule. Purple, green, red, and blue arrows respectively specify where multiple PEs exist, only the separating PE exists, the separating PE is the PPE, and the pooling PE is the PPE. In Figure 3.2, we can see the relative positions of (3.1), (3.2), (3.3), (3.4), and (3.5) in which (3.1) appears both as the incentive constraint of the low-type consumer when he has standard preferences, and a necessary condition for the separating PE to be the PPE. If the SCP holds, in the area between (3.5) and (3.4), only the separating PE is played; multiple PEs exist in the area between (3.3) and (3.5) from which the area between (3.1) and (3.5) allows the separating PE to be the PPE. If the SCP is violated, multiple PEs exist in the entire area between (3.3) and (3.4) from which the area between (3.1) and (3.4), which is non-empty only for larger λ, allows the separating PE to be the PPE.

Constraints (3.1) and (3.2) are also contained in Figure 3.2 (specified by dashed lines), so that we can compare the scope for separating equilibria provided by (3.1) and (3.2) in the case of standard preferences and provided by (3.3) and (3.4) in case of reference-dependent preferences. As is shown, reference-dependent preferences allow for the separating PE to be played for a larger set of parameters, θ_1 and θ_2.

Broadly speaking, the conditions suggested in Proposition 3.3 for the separating PE to be the PPE are more difficult to satisfy than what Proposition 3.1 suggests for the existence of the separating PE. Therefore, there may be situations where the separating PE exists but it is not the PPE. In terms of Figure 3.2, to the extent that (3.4) lies below (3.1), for the entire area between (3.3) and (3.1), the pooling PE is the PPE. However, as long as the degree of loss aversion is sufficiently large, the case in which “the pooling PE is always the PPE" will not occur. In Figure 3.2, the case where "the pooling PE is always the PPE" happens only in 3.2d when the SCP is violated, and the degree of loss aversion is small.

In Propositions 3.2 and 3.3, we analyzed the case, in which the majority of workers are of the high type. Analogously, in the alternative case, in which the majority of workers are of low type and θ_1 will be offered in any pooling equilibrium, it can be proved that when the SCP is valid, the pooling PE (e_1, θ_1) coexists with the separating PE for the same wage schedule if and only if

$$\frac{1+\eta}{1+\eta\lambda}[c(e_2,\theta_2)-c(e_1,\theta_2)] \leq \theta_2-\theta_1 \leq \frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_2)-c(e_1,\theta_2)],$$

and only the separating PE is played if

$$\frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_2)-c(e_1,\theta_2)] \leq \theta_2-\theta_1 \leq \frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)].$$

When the SCP is violated and separating PE exists, the pooling PE (e_1, θ_1) coexists with the separating PE for the same wage schedule everywhere.

In the same way, if the separating PE coexists with the pooling PE (e_1, θ_1) for the same wage schedule, the following can be proved. When the SCP is valid, the separating PE is the PPE if and only if

$$c(e_2,\theta_2)-c(e_1,\theta_2) \leq \theta_2-\theta_1 \leq \frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_2)-c(e_1,\theta_2)],$$

which is non-empty for any $\lambda > 1$; otherwise, the pooling PE is the PPE. When the SCP is violated, the separating PE is the PPE if and only if

$$c(e_2,\theta_2)-c(e_1,\theta_2) \leq \theta_2-\theta_1 \leq \frac{1+\eta\lambda}{1+\eta}[c(e_2,\theta_1)-c(e_1,\theta_1)],$$

which can be non-empty only for large enough λ; otherwise, the pooling PE is the PPE.

3.5 Conclusion

This paper aims at analyzing the strategic signaling behavior of senders who have reference-dependent preferences and are loss averse in the sense of Kőszegi and Rabin (2006), that is, they compare their outcomes with their rational expectations regarding their outcomes. Senders' choice of education levels and wages as well as their expectations of the outcomes of their choices, are affected by their types. The main question we try to answer in this study is whether it is easier or more difficult to achieve credible signaling when senders have reference-dependent preferences and are subject to loss aversion. Our paper follows the line of research pioneered by Heidhues and Kőszegi (2008), Hahn et al. (2014), and Herweg and Mierendorff (2013) among others, in studying asymmetric information situations with agents who are subject to expectation-based loss aversion.

We depart from the abovementioned literature by assuming that agents learn their types first, and then form their reference points based upon those types. This assumption causes our results to be diametrically opposed to the literature results, which assumes that people form their reference points without knowing their types, and argues in favor of price-stickiness (i.e., pooling equilibrium). A comparison between our paper and the literature suggests that the timing of reference-point formation is an influential factor in determining the signaling choice of the workers. Type-dependent reference points are also employed by Carbajal and Ely (2016).

Particularly, we show that loss-averse senders with heterogeneous expectations are naturally inclined to self-select. High-productivity workers acquire a higher education level, because they strongly dislike the loss in wages that they suffer if they acquire a low education. Low-productivity workers acquire a lower education level because they dislike the loss in the form of higher educational costs that they face if they acquire a high education level. We proved that separating equilibria exist for a larger variety of productivity differences when the sender has reference-dependent preferences and his utility function contains gain-loss utility. Moreover, with a sufficiently high degree of loss aversion, separating equilibria may exist even if the single-crossing property is violated.

In our model, multiple personal equilibria may arise because the sender has reference-dependent preferences, and the same wage scheme can support several expectation sets. Even though the sender's reference-dependent preferences create an additional equilibrium, it is just one of the personal equilibria and not necessarily the preferred personal equilibrium. The notion of preferred personal equilibrium is supported by the assumption that senders are able to anticipate the outcomes of their expectations, and that they always pick the expectation that creates a higher level of outcome. In case of multiple personal equilibria, we prove that the separating personal equilibrium can be the PPE if the degree of loss aversion is high enough, even when the single-crossing property is violated. Although there may be situations where the separating personal equilibrium exists, but it is not preferred to the pooling personal equilibrium, as long as loss aversion is sufficiently high; that is, the case in which "the pooling personal equilibrium is always the preferred personal equilibrium" will not occur.

The model that we consider in this study is simple and restricted to many assumptions. As a first step, this has allowed us to focus solely on the effect of loss aversion and reference-dependent preferences. An obvious restriction of our model is that consumers are limited to be only two types while, in the real world, infinitely types of senders exist. Considering several types of senders or possibly a continuum of senders' types will enrich the model. Apart from natural extensions by relaxing the assumptions, an interesting direction for extending our paper as a behavioral economic paper, is to examine our results experimentally and test how realistic they are.

Chapter 4

Optimal Group Size in Microcredit Contracts

4.1 Introduction

Small-scale businesses are considered a major source of employment, especially in developing countries where such businesses employ more than half of the economically active population (De Mel et al., 2008). One of the most efficient tools for the development of microenterprises is access to a source of finance (Berge et al., 2014). However, these enterprises are usually excluded from conventional financial services, either because they are unable to offer formal guarantees or are located too far from the financial networks (Prior and Argandoña, 2009). In Peru in 2008, for example, there were 3,080,000 microenterprises that employed 76% of the country's labor force and accounted for 42% of the country's GDP, but the enterprises were never regarded as relevant by the formal banking system (Chu, 2015).

Microcredits may equip financial institutions with the right tools to rehearse their specific social responsibilities of 'integrating people in the active population and combating the cause of social and financial exclusion' (Lacalle-Calderón and Rico-Garrido, 2006) and reducing poverty (Littlefield et al., 2003) in developing countries. Microcredits are small loans with little or no collateral offered to microenterpreneurs who are usually excluded from conventional financial services. This type of lending has had some success in lending to the poor and has been publicly seen as a recent improvement in financial institutions to support development, although it has also been claimed that microcredit only reaches moderately poor people and not the poorest of the poor (Scully, 2004; Marr, 2003). Although microcredit lending has had a long history,[1] it received substantial attention after the successful experiment of Grameen bank in Bangladesh in the mid-1970s conducted by Dr. Muhammad Yunus, who received the 2006 Nobel Peace Prize for his efforts in poverty reduction. The great success of the Grameen bank's innovative business model has led to

[1] Credit cooperatives were active in the nineteenth century in Germany (Guinnane, 2001).

a rapid reproduction of similar types of lending all over the world.[2] Beside being of interest to socially responsible investors, charitable institutions, and governments, microcredit has also attracted enormous academic attention as a potential key to economic and social returns (for a survey, see Milana and Ashta, 2012; Banerjee, 2013).

In the present paper, we study the following simple model. Consider a benevolent lender (she) who wants to provide loans to a group of n borrowers (he) with *joint liability (JL)*, that is, members accept to be jointly liable for each member's loan. The loans will be invested in n projects that can be either disjointed or correlated. We assume that the realized output of each project is known to the group members but unknown to the lender. This is a plausible assumption when groups of borrowers are self-selected, and possess better information about each other than the lender. Moreover, group members have the means for observing each other's output.

In this study, we try to answer three questions. First, what kind of contracts should be offered to borrowers so that the lender could receive her money back while borrowers' lifetime profit is maximized? More specifically, what penalty function should be employed against defaulting borrowers so that they are encouraged to repay, but not punished too harshly? Should such borrowers be excluded from lending forever or should they get another chance? In this paper, we reconsider the optimal design of uncollateralized lending contracts with joint liability. We compare joint liability contracts, which deprive the strategically defaulting member from future loans forever, with joint liability contracts, which deprive them for only T periods and let them rejoin the group afterwards. We examine positive and negative effects of the less severe punishment on the lending outcome.

The second question we are exploring is how large should group size n be in order to maximize the borrowers' benefit while leaving the lender to break even? On the one hand, a larger group size can have a positive effect on the repayment rate, as having more people liable for repaying defaulted payments assures a higher rate of repayment. On the other hand, a larger group size can be a threat toward members who repay their loans successfully, since they should pledge repayment for all their defaulting peers, and it may happen that everybody else in the group is defaulting on his repayment. We propose a method to find the optimal group size that enables larger loans to borrowers under JL.

The third question that leads our study is how project correlation can affect the lending outcome. A natural assumption to make is that the chance of a project's success is an increasing function of the group size, as jointly liable group members are likely to help each other succeed. However, the marginal desirability of forming larger groups should decrease, as large groups must handle higher tensions.

We model the lending situation described above as an infinitely repeated game. Inspired by Tedeschi (2006), we define two phases in our model: a *lending phase* in which the game starts and continues until the group defaults on repayment, and a *punishment phase*

[2]Famous examples of successful microfinance institutions include: Equitas Microfinance in India (Narayanan and Rangan, 2010), Mibanco in Peru (Chu and Gustavo, 2009), Tibet Poverty Alleviation Fund (TPAF) in the Tibet Autonomous Region (Stuart et al., 2008), Kashf Foundation in Pakistan (Mahmood and Hui, 2007), Accion International, a U.S. organization that is active in Latin America (Quelch and Laidler, 2003).

in which no new loans are extended to the borrowers. We stay aligned with Bhole and Ogden's (2010) simple group-lending model that is designed for groups of two borrowers and extend their results to groups of n borrowers. In our model, borrowers receive contracts individually that determine the amount of loan L and repayment R. Borrowers invest their loans on n projects that are identical in terms of mean return and chance of success, however they can be disjoint or correlated. Projects will be either successful with high return or unsuccessful with low return. After the outcomes of projects are realized, each borrower decides to repay his loan or not. Only those with a successful project can repay. We make a difference between *strategic default* and *non-strategic default*. Borrowers default strategically when they refuse to repay while they had high outcome from their projects; and they default non-strategically when they do not repay because they had low or no outcome from their project due to bad luck. The lender cannot identify strategic default, but members of the group can. Under joint liability, if someone defaults non-strategically, his group members will repay his loan. The lender deprives a defaulting group from loans in the future in order to decrease incentives for strategic default, however, non-strategic defaults that are a result of bad luck are unavoidable and repayment is not completely insured. We assume that group members play a grim-trigger strategy in the strategic game between themselves. As such if someone defaults strategically, other members of the group will not repay his loan as well as their own loans, and no further loans are granted to the group members. Later, we relax the grim-trigger assumption and define a *flexible joint liability (FJL),* in which successful members of a group repay the loan of someone who defaults strategically, but punish the person by excluding him from receiving loans for the next T periods.

The results of our study offer three sets of comparisons, joint liability vs. *individual liability (IL)* lending, unlimited length vs. limited length punishment phase, and disjoint vs. correlated projects. The first comparison shows that both the JL and the FJL lending contracts are feasible under a smaller set of parameter settings than the IL lending contract, but have higher performances than the IL contract in terms of borrowers' welfare and the repayment rate. Both the JL and the FJL lending contracts can also outperform the IL lending contract in terms of maximum loans that can be offered to borrowers if the group size is not too large. The intuition behind this result is that groups can offer higher repayment insurance than individuals, thus is it less risky to offer them larger loans. However, if the group size becomes too large, the risk of project failure and more defaulting members also goes up. This risk is even higher when borrowers invest in riskier projects. We calculate the optimal group size given the maximum and minimum returns of projects, the discount factor of borrowers for future loans, and the chance of project's success.

The second comparison proves that the length of the punishment phase does not have a significant effect on borrowers' repayment amount and their welfare, and that both borrowers' repayment amount and their welfare stays the same under the JL and the FJL lending contract. However, the maximum loan that can be offered to borrowers when the punishment is less severe—i.e., under the FJL lending contract—should be lower compared to the JL lending contract. More specifically, the maximum loan that can be offered to borrowers under the FJL lending contract is increasing in the length of the punishment phase and is highest when the length of the punishment phase goes to infinity, that is, the maximum loan that can be offered to borrowers is highest under the JL lending contract, which is not surprising given what we know from studies on repeated games of cooperation. We also show that, in the FJL lending contract, when the length of the punishment phase is not too long, larger groups should be formed. Intuitively, when the punishment is not as severe, members are more prone to default strategically, therefore, a larger group is needed to increase the repayment insurance.

Finally, the third comparison provides evidence that when any positive correlation exists between project returns, we could increase the feasibility of joint liability contracts by enlarging the borrowers group.

The following section reviews the related literature. Section 4.3 describes our model. Sections 4.4 and 4.5 characterize the optimal JL and FJL contracts. These sections also present comparisons between the JL and the IL contracts, between the JL and the FJL contracts, and between the FJL and the IL contracts and discusses conditions under which a joint liability contract is the optimal decision. Section 4.6 looks at the effect of project correlation on joint liability contracts. Appendix presents proofs of lemmas.

4.2 Related Literature

In general, our study is in line with recent and growing theoretical literature in microfinance that explores conditions under which group lending can enhance lending results and alleviate informational asymmetry and enforcement problems that are specifically present in developing countries (see, for example, Ghatak, 2000; Armendáriz de Aghion and Gollier, 2000). Within this literature, our study is related to research that is concerned with the optimal design of contracts that should be offered to borrowers as well as punishment that should be used against defaulting members (see e.g. Bhole and Ogden, 2010; Tedeschi, 2006). The literature mostly concludes that when borrowers are unable to impose strong social sanctions on each other, the lender would be better off to offer the IL contracts rather than the JL contract (see, for example, Besley and Coate, 1995; Armendáriz de Aghion, 1999). In this paper, we assume that borrowers of one group play a grim-trigger strategy against each other, and we discuss that JL lending can actually have a positive effect on borrowers' welfare and repayment amounts compared to IL lending, even in the absence of social sanctions, although the JL contracts are feasible under a smaller parameter set. We also argue that the JL contracts, when feasible, outperform the IL contracts in terms of the maximum loans that can be offered. Furthermore, we discuss that less severe punishment does not change a borrower's lifetime benefits and repayment amounts, thus the borrower's problem stays unchanged whether the joint liability contract comes with a severe punishment (i.e., JL) or with a more flexible punishment (i.e., FJL).

Literature investigating group lending with joint liability has paid little attention to group size as a potentially influential factor in the relative success of group lending. Theoretical studies have mostly analyzed group lending models of two borrowers, while experimental and empirical studies have suggested the importance of group size (Abbink et al., 2006; Galak et al., 2011). We argue that group size is an important factor in increasing borrowers' welfare and the repayment rate in microcredit lending, and we determine the optimal group size endogenously.

The existing literature that considers group size as an important factor have arrived at different conclusions. On the one hand, arguments exist in favor of larger group sizes, such as provided by Conning (2004) and Ahlin (2015), who also suggest that group size cannot become too large. Conning (2004) specifies that it becomes increasingly costly to contain free-riding as a group size increases. Ahlin (2015) shows that the presence of local borrower information is necessary for large groups to have any impact. Our assumption that borrowers play a grim-trigger strategy against each other helps us to treat free-riding inside a group of borrowers, as well as replace the local borrower-information effect. On the other hand, there are arguments in the literature that favor a smaller group size. For example, Bourjade and Schindele (2012) prove that if group members have social ties, a rational lender should choose a group of limited size. They explain that a trade-off exists between raising profits through increased group size and providing incentives for borrowers with less social ties. Similar to our findings, Baland et al. (2013) show that the optimal group size depends on project characteristics. We demonstrate that for risky projects, although the repayment insurance provided by a larger group is higher, the risk of being in charge of the group is also higher. Therefore, riskier projects should be handled in relatively larger groups, but the group size cannot become too large.

Literature on group lending with joint liability has generally assumed that borrowers run independent projects (see, for example, Armendáriz de Aghion, 1999). Studies exist claiming that the JL contracts are less feasible when project outcomes are positively correlated. For example, Ghatak (2000) argues that when projects are likely to succeed or fail simultaneously, the joint liability part of the contract is practiced less often. This is because if one project fails, other projects fail as well, and the group cannot make the repayment. We argue, in contrast, that joint liability can cause the formation of project externalities. As jointly liable groups are more likely to contribute to each other's success, lending under joint liability is more feasible when correlation exists between project outcomes. In contrast, Katzur and Lensink (2012) argue that positive correlation of project returns may improve the efficiency of group lending contracts. They stay aligned with Ghatak (2000), who reasoned that if projects are positively correlated, the probability of paying the joint liability payment decreases. However, in the model by Katzur and Lensink (2012), borrowers accept to pay higher interest rates in exchange for joint liability that adjusts the negative effect of project correlation. Apart from various modeling differences, our work is distinguished by modeling the project correlation through the assumption that group members' social interactions positively influence their chance of project success while Katzur and Lensink (2012) integrate the statistical correlation of project returns in their model.

4.3 Model

Consider an infinitely repeated lending game with a benevolent lender and n borrowers playing a grim-trigger strategy against each other. We consider a two-phase model in which the lender and borrowers start in a lending phase, if one loan is successfully repaid by the group, another loan is provided in the next period. In any period, if borrowers default, the lender and borrowers engage in a punishment phase, in which no new loans are extended to borrowers. Each period of our game has three steps:

$s = 0$ Each borrower receives an individual contract (L, R) that specifies the amount of loan L and the repayment R ($R \geq L + \varepsilon$, where ε is the interest rate).

$s = 1$ Each borrower invests L on his project that will be either successful with chance of $\alpha \in [0, 1]$ and yield high return Y^H, or not be successful with chance of $1 - \alpha$, and yield low return Y^L, such that $0 \leq Y^L < Y^H$. Project returns are public information to all members of the group of borrowers.

$s = 2$ Borrowers decide simultaneously to repay their share R or not, and the lender announces the remaining payment. In this case, if i members default on their loan, the other $n - i$ members will be asked to additionally pay an amount $\frac{iR}{n-i}$ to the lender for their defaulting peers. If the total repayment is equal to nR or more, the group receives future financing. Otherwise, the entire group will be excluded from financing next period by the lender.

These three stages will be played repeatedly until the lender realizes that borrowers are not entitled for financing next period, and in each period of not receiving a loan, borrowers' utility will be zero. In the first step, we assume that projects do not differ in their riskiness (i.e., α is the same for all borrowers). We also assume that it is not known to the lender whether projects have high or low returns, and that each borrower always invests in the same project.

Two types of defaults are possible: *strategic default,* in which the borrower does not repay although he had high outcome Y^H, and *nonstrategic default*, in which the borrower cannot repay because he had low outcome Y^L resulting from bad luck. We assume that the lender is unable to observe whether a borrower's default is strategic or nonstrategic. However, borrowers are able to observe strategic defaults of their peers costlessly.

Borrowers play a repeated game among themselves. They begin by cooperating and continue repaying if they can as long as all others do repay or default non-strategically. In the first step, we assume that borrowers play a grim-trigger strategy against each other, meaning that if at some point some members default strategically on their repayments, other members stop repaying for themselves, as well as for the defaulting players' shares. They will therefore not receive a loan in the next period. In the second step, we relax the grim-trigger assumption by assuming that if some members default strategically at some point, other group members will still repay their own loans as well as the defaulting members' share, but they punish the strategically defaulting members by excluding them from the lending game for T periods.

The benevolent lender strives to maximize the payoff of each borrower contingent on the following conditions. First, each borrower must be willing to accept a loan (the repayment amount must be affordable for him). Second, each borrower must have the correct incentive to repay for himself and his share for each defaulting peer when he is able to pay (in the worst case, that all other members default, he must be still willing to repay for the entire group). Third, the lender must break even, meaning that she must maintain a sustainable lending operation over the entire loan portfolio by charging the appropriate repayment. The exact terms of the maximization problem depend on whether we are examining individual lending or group lending and will be discussed in the following sections.

4.4 Joint Liability (JL) Contracts

In this section, we formalize the model assuming that if some borrowers default strategically, their group members punish them by playing a grim-trigger strategy, meaning that successful members refuse to pay for strategically defaulting members, as well as for themselves. In this case, they will not receive a loan next round. The probability that the total loan is paid and lending can continue to the next round is

$$\alpha^n \binom{n}{n} + \alpha^{n-1}(1-\alpha)\binom{n}{n-1} + ... + \alpha^1 (1-\alpha)^{n-1}\binom{n}{1} = [1-(1-\alpha)^n].$$

The expected repayment for a borrower who plays the repayment strategy in each period t can be stated as follows:

$$\sum_{i=0}^{n-1}\binom{n-1}{i}\alpha^{n-i}(1-\alpha)^i\left(R+\frac{i}{n-i}R\right) = [1-(1-\alpha)^n]R.$$

Thus, the expected lifetime utility of a borrower who plays a repayment strategy at any period onwards in which he gets financing is determined as

$$V_{JL}^R = \mathbb{E}(Y) - [1-(1-\alpha)^n]R + [1-(1-\alpha)^n]\,\delta V_{JL}^R,$$

where $0 \leq \delta \leq 1$ is the borrower's discount factor that determines his valuation of tomorrow's utility of financing. The expected lifetime utility of a borrower can be rewritten as

$$V_{JL}^R = \frac{\mathbb{E}(Y) - R[1-(1-\alpha)^n]}{1-\delta[1-(1-\alpha)^n]}. \tag{4.1}$$

Now the lender's optimization problem for any $0 \leq Y^L < Y^H$, $0 \leq \alpha \leq 1$ and $0 \leq \delta \leq 1$ can be stated as $\underset{L,R}{maximize} V_{JL}^R$ subject to:

1. The stipulated repayment amount for a successful borrower cannot exceed his output (even in the worst case, i.e., when everyone else's projects has failed), that is, it must be affordable,
$$nR \leq Y^H. \tag{4.2}$$

2. "Each borrower is repaying when his project is successful" is a subgame perfect equilibrium, if the one-shot deviation constraint is satisfied. Thus, for a successful borrower, the payoff of strategically defaulting cannot be larger than the payoff of repaying and being refinanced,
$$Y^H \leq Y^H - nR + \delta V_{JL}^R.$$
The above condition can be simplified to
$$nR \leq \delta V_{JL}^R, \tag{4.3}$$
which also implies $R < \delta V_{JL}^R$, that is, it is also guaranted that a successful borrower pays repayment R when all his partners are successful.

3. The lender must be able to sustain the lending game over periods and at least break even. Thus, the expected repayment amount of the group has to be at least as large as $n(L+\varepsilon)$, that is,
$$\alpha^n \binom{n}{n} nR + \alpha^{n-1}(1-\alpha)\binom{n}{n-1} nR + \ldots + \alpha^1 (1-\alpha)^{n-1}\binom{n}{1} nR \geq n(L+\varepsilon),$$
that can be simplified to
$$R \geq \frac{L+\varepsilon}{[1-(1-\alpha)^n]}. \tag{4.4}$$

If there are some (L,R) that satisfy (4.2), (4.3), and (4.4), then the JL contract will be feasible, and constraints (4.2), (4.3), and (4.4) will define its feasibility region. Note that individual lending can be considered as a special type of group lending with $n = 1$.

Proposition 4.1. *There exist* $\tilde{\delta}_{JL}(n,\alpha,Y^H,Y^L)$, $\tilde{L}_{JL}(n,\alpha,Y^H)$, *and* $\hat{L}_{JL}(n,\alpha,\delta,Y^H,Y^L)$ *such that:*

a) if $\delta \geq \tilde{\delta}_{JL}(n,\alpha,Y^H,Y^L)$, *then the JL contract is feasible if and only if* $L \leq \tilde{L}_{JL}(n,\alpha,Y^H)$;

b) if $\delta \leq \tilde{\delta}_{JL}(n,\alpha,Y^H,Y^L)$, *then the JL contract is feasible if and only if* $L \leq \hat{L}_{JL}(n,\alpha,\delta,Y^H,Y^L)$.

Moreover, whenever the JL contract is feasible: for any $\alpha \neq 0$, *the lender demands an optimal repayment* $R_{JL} = \frac{L+\varepsilon}{[1-(1-\alpha)^n]}$ *from each borrower, and the expected lifetime utility for each borrower will amount to* $V_{JL}^R = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\delta[1-(1-\alpha)^n]}$.

Proof. In the case of JL lending, the optimal contract (L,R) is a solution to the following problem:

$$\underset{L,R}{maximize} \quad V_{JL}^{R} = \frac{\mathbb{E}(Y) - R\left[1-(1-\alpha)^{n}\right]}{1-\delta\left[1-(1-\alpha)^{n}\right]} \tag{4.5}$$

$$s.t. \quad nR \leq \delta V_{JL}^{R} \tag{4.6}$$

$$nR \leq Y^{H} \tag{4.7}$$

$$R \geq \frac{L+\varepsilon}{1-(1-\alpha)^{n}}. \tag{4.8}$$

As V_{JL}^{R} is decreasing in R, the lender would like to set R as low as possible. Constraint (4.8) gives the minimum R required for breaking even, $R_{JL} = \frac{L+\varepsilon}{\left[1-(1-\alpha)^{n}\right]}$, for any $\alpha \neq 0$. As long as R is limited to the upper limits expressed in constraints (4.6) and (4.7), lending is feasible, otherwise, it is not feasible. Replacing $R_{JL} = \frac{L+\varepsilon}{\left[1-(1-\alpha)^{n}\right]}$ in the objective function, (4.5), and constraints (4.6) and (4.7), we will have

$$\begin{aligned}
\underset{L,R}{maximize} \quad V_{JL}^{R} &= \frac{\mathbb{E}(Y) - (L+\varepsilon)}{1-\delta\left[1-(1-\alpha)^{n}\right]} \\
s.t. \quad L &\leq \frac{\delta\mathbb{E}(Y)\left[1-(1-\alpha)^{n}\right]}{n-\delta(n-1)\left[1-(1-\alpha)^{n}\right]} - \varepsilon \equiv \hat{L}_{JL}\left(n,\alpha,\delta,Y^{H},Y^{L}\right) \\
L &\leq \frac{\left[1-(1-\alpha)^{n}\right]Y^{H}}{n} - \varepsilon \equiv \tilde{L}_{JL}\left(n,\alpha,Y^{H}\right).
\end{aligned}$$

Any feasible solution for the above problem must satisfy both constraints. Therefore, we must have $L \leq \min\left\{\hat{L}_{JL}, \tilde{L}_{JL}\right\}$. There are two cases: it is either the case that $\hat{L}_{JL} \leq \tilde{L}_{JL}$ or the case that $\tilde{L}_{JL} \leq \hat{L}_{JL}$.

i) $\hat{L}_{JL} \leq \tilde{L}_{JL}$ if and only if

$$\delta \leq \frac{Y^{H}}{\mathbb{E}(Y) + \frac{n-1}{n}\left[1-(1-\alpha)^{n}\right]Y^{H}} \equiv \tilde{\delta}_{JL}\left(n,\alpha,Y^{H},Y^{L}\right);$$

ii) $\tilde{L}_{JL} \leq \hat{L}_{JL}$ if and only if $\delta \geq \tilde{\delta}_{JL}$.

Thus for $\delta \leq \tilde{\delta}_{JL}$, the JL contract is feasible for any $L \leq \hat{L}_{JL}$, and for $\delta \geq \tilde{\delta}_{JL}$, the JL contract is feasible for any $L \leq \tilde{L}_{JL}$. □

According to Proposition 4.1, the JL contracts can be feasible if and only if $L \leq F_{JL}\left(n,\alpha,\delta,Y^{H},Y^{L}\right)$, where

$$F_{JL}\left(n,\alpha,\delta,Y^{H},Y^{L}\right) = \min\left\{\hat{L}_{JL},\tilde{L}_{JL}\right\} = \begin{cases} \hat{L}_{JL} & if \quad \delta \leq \tilde{\delta}_{JL} \\ \tilde{L}_{JL} & if \quad \delta \geq \tilde{\delta}_{JL} \end{cases}. \tag{4.9}$$

We will call (4.9) the *feasibility function* from now on.

Proposition 4.1 shows that the maximum loan offered to borrowers under the JL contract, when such contracts are feasible, depends on the borrowers' discount factor—i.e. δ—as both determinants of the maximum loan—i.e. $\tilde{L}_{JL}$ and $\hat{L}_{JL}$—are dependent on δ. From this proposition, it can also be inferred that optimal repayment R_{JL} is decreasing in group size n. Therefore, a larger group can be charged less, which in turn increases the borrowers welfare, V_{JL}^{R}. Corollary 1 is a direct result from Proposition 4.1 and follows from substituting $n = 1$.

Corollary 4.1. *Individual lending is feasible if and only if $L \leq \alpha\delta\mathbb{E}(Y)$. For any $\alpha \neq 0$, the lender demands optimally repayment $R_{IL} = \frac{L+\varepsilon}{\alpha}$. The borrower's expected lifetime utility will be $V_{IL}^{R} = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\alpha\delta}$.*

We continue this section determining how the feasibility of the JL contract is affected by the group size as well as determining the optimal group size that results in the maximum feasibility. To describe the properties of the feasibility function with respect to changes of n, in Lemma 4.1 we take a closer look at the changes of $\hat{L}$, $\tilde{L}$ with respect to changes of n when other parameters $\left(\alpha, \delta, Y^{H}, Y^{L}\right)$ are given.

Lemma 4.1. *Assume $\hat{L}_{JL}\left(n, \alpha, \delta, Y^{H}, Y^{L}\right)$ and $\tilde{L}_{JL}\left(n, \alpha, Y^{H}\right)$ are functions defined in Proposition 4.1.*

1) There exists a $\hat{\delta}_{JL}(n, \alpha)$ such that:

i. $\hat{L}_{JL}\left(n, \alpha, \delta, Y^{H}, Y^{L}\right)$ is strictly decreasing in n if $0 < \delta < \hat{\delta}_{JL}(n, \alpha)$.

ii. $\hat{L}_{JL}\left(n, \alpha, \delta, Y^{H}, Y^{L}\right)$ is strictly increasing in n if $\hat{\delta}_{JL}(n, \alpha) < \delta < \tilde{\delta}_{JL}\left(n, \alpha, Y^{H}, Y^{L}\right)$.

2) $\tilde{L}_{JL}\left(n, \alpha, Y^{H}\right)$ is strictly decreasing in n.

Lemma 4.1 shows that $\hat{L}_{JL}$, and consequently the feasibility function, can be increasing n for any $\delta \in \left(\hat{\delta}_{JL}, \tilde{\delta}_{JL}\right)$, only if the interval $\left(\hat{\delta}_{JL}, \tilde{\delta}_{JL}\right)$ is non-empty. As we see in this lemma, for any δ that does not belong to the interval $\left(\hat{\delta}_{JL}, \tilde{\delta}_{JL}\right)$, both $\hat{L}_{JL}$ and $\tilde{L}_{JL}$ are strictly decreasing. In Lemma 4.2, we look at the changes of $\hat{\delta}_{JL}$ and $\tilde{\delta}_{JL}$ when n changes, to see if there are chances that $\left(\hat{\delta}_{JL}, \tilde{\delta}_{JL}\right)$ is non-empty for some n.

Lemma 4.2. *Assume $\tilde{\delta}_{JL}\left(n, \alpha, Y^{H}, Y^{L}\right)$ and $\hat{\delta}_{JL}(n, \alpha)$ are functions defined in Proposition 4.1 and Lemma 4.1.*

1) $\hat{\delta}_{JL}(n, \alpha)$ is strictly increasing in both n and α.

2) For any given n and α, $-\infty < \hat{\delta}_{JL}(n, \alpha) < 1$.

3) $\tilde{\delta}_{JL}\left(n, \alpha, Y^{H}, Y^{L}\right)$ is strictly decreasing in both n and α.

4) For any given n and α, $\frac{n}{2n-1} < \tilde{\delta}_{JL}\left(n, \alpha, Y^{H}, Y^{L}\right) < \frac{Y^{H}}{Y^{L}}$.

A direct result of Lemma 4.2 is that for very small α, regardless of the magnitude of n, we always have $\hat{\delta}_{JL} < \tilde{\delta}_{JL}$, and for very large α and for any $n > 1$, we always have $\hat{\delta}_{JL} > \tilde{\delta}_{JL}$. Thus, the interval $\left(\hat{\delta}_{JL}, \tilde{\delta}_{JL}\right)$ might actually be non-empty for some n. This lemma shows also that the interval $\left(\hat{\delta}_{JL}, \tilde{\delta}_{JL}\right)$ becomes tighter by the increase of n or α and it becomes wider by the decrease of n or α. Therefore, to keep the interval $\left(\hat{\delta}_{JL},\tilde{\delta}_{JL}\right)$ non-empty, the larger the α, the smaller the n must be chosen in order to offset the contraction of the interval resulting from a large α. In general, $n=2$ always provides the widest interval for any given α. Below, in Proposition 4.2, we prove that regardless of the magnitude of n, for a small enough α, the interval $\left(\hat{\delta}_{JL},\tilde{\delta}_{JL}\right)$ is never empty. However, if α becomes too small, then the project may not be considered for financing.

Proposition 4.2. *Assume $\tilde{\delta}_{JL}\left(n,\alpha,Y^H,Y^L\right)$ and $\hat{\delta}_{JL}(n,\alpha)$ are functions defined in Proposition 4.1 and Lemma 4.1.*

1) There exists $\bar{\alpha}$, such that for any $\alpha<\bar{\alpha}$, feasibility of the JL contract is increasing in $n\in\left[2,N_{\alpha,\delta}\right]$ only if $\hat{\delta}_{JL}(n,\alpha)<\delta<\tilde{\delta}_{JL}\left(n,\alpha,Y^H,Y^L\right)$, where

$$N_{\alpha,\delta}=\min\left\{\lfloor\hat{\delta}_{JL}^{-1}(\alpha,\delta)\rfloor,\lfloor\tilde{\delta}_{JL}^{-1}\left(\alpha,\delta,Y^H,Y^L\right)\rfloor\right\}. \tag{4.10}$$

For any $\alpha>\bar{\alpha}$, maximum feasibility happens at $n=2$.

2) For very large n, feasibility of the JL contract will be decreasing.

Proof. According to Lemma 4.1, for given α and δ, if there are some n such that $\hat{\delta}_{JL}<\delta<\tilde{\delta}_{JL}$, then the feasibility function $F_{JL}\left(n,\alpha,\delta,Y^H,Y^L\right)$ will be strictly increasing in n. $F_{JL}\left(n,\alpha,\delta,Y^H,Y^L\right)$ will be strictly decreasing in n in every other case.

1) It can be simply verified that

i. $\lim_{\alpha\to 0}\hat{\delta}_{JL}(n,\alpha)=-\infty$ and $\lim_{\alpha\to 0}\tilde{\delta}_{JL}(n,\alpha)=\infty$;

ii. $\lim_{\alpha\to 1}\hat{\delta}_{JL}(n,\alpha)=1$ and $\lim_{\alpha\to 1}\tilde{\delta}_{JL}(n,\alpha)=\dfrac{n}{2n-1}$.

That is, for very small α, we always have $\hat{\delta}_{JL}(n,\alpha)<\tilde{\delta}_{JL}(n,\alpha)$, and for very large α, we always have $\hat{\delta}_{JL}(n,\alpha)>\tilde{\delta}_{JL}(n,\alpha)$. We also know from Lemma 4.2 that $\hat{\delta}_{JL}$ and $\tilde{\delta}_{JL}$ are monotonic. Therefore, we should conclude that $\hat{\delta}_{JL}$ and $\tilde{\delta}_{JL}$ coincide only once at some critical $\bar{\alpha}\neq 0$. Thus, for any given n, if $\alpha\in(0,\bar{\alpha})$, then $\hat{\delta}_{JL}<\tilde{\delta}_{JL}$. Consequently, for $\alpha\in(0,\bar{\alpha})$, the feasibility function will be increasing in n and will be maximized at

$$N_{\alpha,\delta}=\max\left\{n|\quad\hat{\delta}_{JL}<\delta<\tilde{\delta}_{JL}\right\}$$

Our algorithm to find the maximum n when α and δ are given would be as follows. Start with $n=2$ and increase n one by one until one of the $\hat{\delta}_{JL}$ or $\tilde{\delta}_{JL}$ equals δ so that n cannot be increased further. If $\hat{\delta}_{JL}=\delta$, then $N_{\alpha,\delta}=\lfloor\hat{\delta}_{JL}^{-1}\rfloor$ (note that we are only interested in n that is a natural number) and if $\tilde{\delta}_{JL}=\delta$, then $N_{\alpha,\delta}=\lfloor\tilde{\delta}_{JL}^{-1}\rfloor$ (see Figure 4.1 for intuition). Thus $N_{\alpha,\delta}$ can be rewritten as $N_{\alpha,\delta}=\min\left\{\lfloor\hat{\delta}_{JL}^{-1}\rfloor,\lfloor\tilde{\delta}_{JL}^{-1}\rfloor\right\}$.

A direct result from the previous discussion is that for $\alpha>\bar{\alpha}$, feasibility will be decreasing in n and so maximum feasibility is given at $n=2$.

2) If n grows very large, feasibility of the JL contract will be decreasing in n: $\lim_{n\to\infty}\hat{\delta}_{JL}=1$ and $\lim_{n\to\infty}\tilde{\delta}_{JL}=\frac{Y^H}{\mathbb{E}(Y)+Y^H}\leq 1$. Thus, with the exception of $\alpha\to 0$ and $Y^L=0$ simultaneously, for very large n, $\left(\hat{\delta}_{JL},\tilde{\delta}_{JL}\right)$ will be empty. $\square$

Table 4.1: The Range of Feasible Group Sizes for Some Given α

	$n=2$	$n=3$	$n=4$	$n=5$	$\cdots$	$n=10$	$\cdots$	$n=50$
$\alpha<0.718$	✓							
$\alpha<0.568$	✓	✓						
$\alpha<0.477$	✓	✓	✓					
$\alpha<0.415$	✓	✓	✓	✓				
$\alpha<0.269$	✓	✓	✓	✓		✓		
$\alpha<0.082$	✓	✓	✓	✓		✓		✓

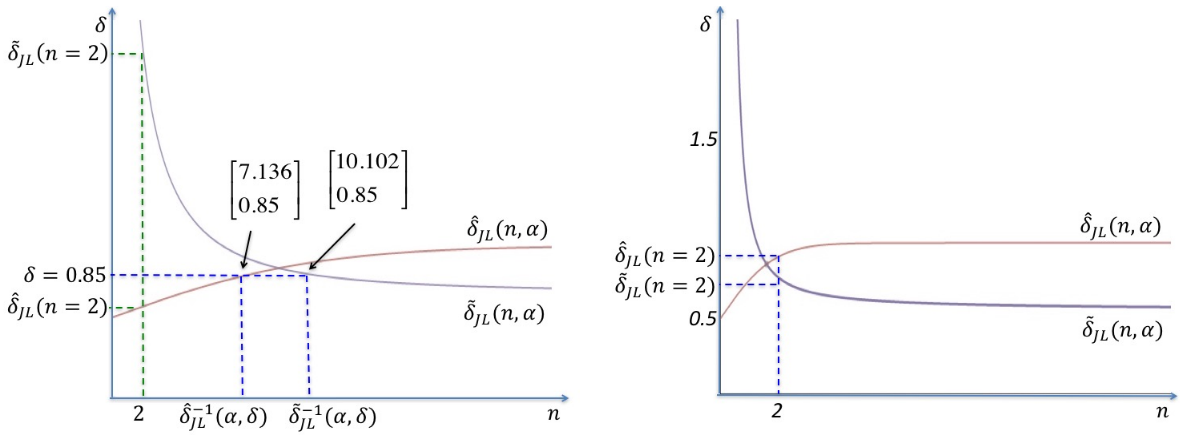
Proposition 4.2 is telling us that if the chance of project success is low, larger loans could be given to larger groups. While for projects with a higher chance of success, the maximum loan that can be given to a group of two members is higher than for any group with $n>2$.

Using the results of Proposition 4.2, it could simply be verified that $\bar{\alpha}$ is approximately 0.5 if $Y_L=0$. For $\alpha<0.718$, the inequality $\hat{\delta}_{JL}<\tilde{\delta}_{JL}$ holds at least for $n=2$. We can see in Table 4.1 that the largest α for which the inequality $\hat{\delta}_{JL}<\tilde{\delta}_{JL}$ holds for more than one n is $\alpha<0.568$. Thus, feasibility can increase in n only if $\alpha<0.568$ and if the given δ is such that $\hat{\delta}_{JL}<\delta<\tilde{\delta}_{JL}$. Note that when $Y^L\neq 0$, α becomes a slightly smaller. For example when Y^L is as large as $\frac{Y^H}{2}$, α is approximately 0.458.

Proposition 4.2 also suggests that the group size cannot grow too large. Intuitively, group size has two countervailing effects. On the one hand, a larger group can provide stronger repayment insurance and is able to handle riskier projects and repay successfully. On the other hand, a large group can be a threat towards the feasibility of group lending. The threat comes from the fact that each successful member is in charge of all defaulting peers and if everybody else fails, he must repay the entire loan of the group. This becomes very difficult if the group is too large. Proposition 4.2 presents a characterization of the optimal size, as a function of all other parameters.

Figures 4.1 and 4.2 can help to derive some intuition about Proposition 4.2. For simplicity, it is assumed that $Y^H=1$ and $Y^L=0$. In Figure 4.1, we depict $\tilde{\delta}$ and $\hat{\delta}$ with respect to n when the chance of project success is small—e.g., $\alpha=0.3$—and when it is large—e.g., $\alpha=0.8$—respectively. As shown in Figure 4.1a, for $\alpha=0.3$, there are some n for which the given $\delta=0.85$ belongs to the interval $\left(\hat{\delta}_{JL},\tilde{\delta}_{JL}\right)$, and the largest of such n lies at the $\min\left\{\hat{\delta}^{-1},\tilde{\delta}^{-1}\right\}$, that is, at $\lfloor\hat{\delta}_{JL}^{-1}\rfloor=\lfloor 7.136\rfloor=7$. Note that if the given δ is very close to 1, the largest n then lies at $\lfloor\tilde{\delta}_{JL}^{-1}\rfloor$. However, in Figure 4.1b, when α is large, for any $n>1$, the interval $\left(\hat{\delta}_{JL},\tilde{\delta}_{JL}\right)$ is empty.

Figure 4.2 depicts $\hat{L}$ and $\tilde{L}$ with respect to n when the chance of project success is small—e.g., $\alpha=0.3$—and when it is large—e.g., $\alpha=0.8$—respectively. As shown in Figure 4.2a, when α is small, $\hat{L}$ defines the boundary for the maximum feasible loan, and it reaches its maximum at $n=7$. Therefore, $n=7$ is the group size that maximizes the feasibility of the JL contract. However, in 4.2b, when α is large, $\tilde{L}$ defines the boundary



(a) It is assumed that $\alpha = 0.3$. (b) It is assumed that $\alpha = 0.8$.

Figure 4.1: Larger group sizes are possible as long as the discount factor belongs to $\left(\hat{\delta},\tilde{\delta}\right)$.

for the maximum feasible loan, which is decreasing in n. Therefore, $n = 2$ is the optimum group size that maximizes the feasibility of the JL contract.

Up to this point, we have discussed that feasibility of the JL contract can be increasing in group size when the chance of project success is small. However, we do not yet know if the JL contract can perform more efficiently than the IL contract. Are there circumstances under which the JL contract outperforms the IL contract? Proposition 4.3 proves formally that the lender can charge borrowers less under the JL contract compared with the IL contract while still breaking even, which may be because "no repayment" happens less often under the JL contract than under the IL contract. In turn, a smaller amount of repayment leads to a higher level of welfare for the borrowers.

Proposition 4.3. *The following statements hold when both the IL and the JL contracts are feasible.*

1) Borrowers' repayment amount is lower and his welfare is higher under the JL contract than the IL contract.

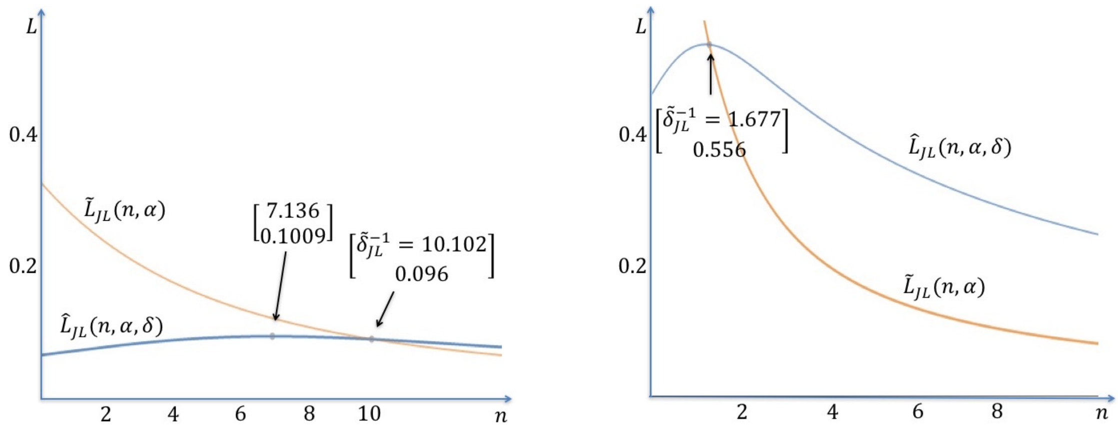
2) There exists $\underline{\alpha}$, such that for $\alpha < \underline{\alpha}$, the JL contract is feasible for a larger loan than the IL contract only if at least for some n,

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} < \delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}; \tag{4.11}$$

for any $\alpha > \underline{\alpha}$, the IL contract is feasible for a larger loan than the JL contract.

4) For very large n, the IL contract can always offer a larger maximum loan than the JL contract.

Proof. 1) When both the IL and the JL contracts are feasible, then each borrower is supposed to repay $R_{JL} = \frac{L+\varepsilon}{[1-(1-\alpha)^n]}$ under the JL, and $R_{IL} = \frac{L+\varepsilon}{\alpha}$ under the IL contract. Clearly



(a) It is assumed that $\delta = 0.85$ and $\alpha = 0.3$. (b) It is assumed that $\delta = 0.85$ and $\alpha = 0.8$.

Figure 4.2: The maximum loan increases in group size as long as the discount factor belongs to $\left(\hat{\delta}, \tilde{\delta}\right)$.

a borrower pays less under the JL contract. Each borrower's expected lifetime utility under the JL contract is $V_{JL}^{R} = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\delta[1-(1-\alpha)^n]}$, and under the IL contract is $V_{IL}^{R} = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\alpha\delta}$. Obviously, a borrower's expected lifetime utility is higher under the JL contract.

2) Is feasibility easier to achieve through the JL contract or the IL contract? There are two cases:

i. If $\delta < \tilde{\delta}_{JL}$, then the JL contract is feasible for all $L \leq \hat{L}_{JL} = \alpha\delta\mathbb{E}(Y) \times \frac{[1-(1-\alpha)^n]}{\alpha[n-\delta(n-1)[1-(1-\alpha)^n]]} - \varepsilon$, and the IL contract is feasible for all $L \leq \alpha\delta\mathbb{E}(Y) - \varepsilon$. Thus, the JL contract can offer a larger range of loans to borrowers than the IL contract if and only if

$$\frac{[1-(1-\alpha)^n]}{\alpha[n-\delta(n-1)[1-(1-\alpha)^n]]} > 1,$$

that can be rewritten as

$$\delta > \frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]}.$$

Otherwise, the IL contract can offer a larger range of loans to borrowers than the JL contract. Since we already have an upper bound for δ, it must be true that

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} < \delta < \tilde{\delta}_{JL}.$$

ii. If $\delta > \tilde{\delta}_{JL}$, then the JL contract is feasible for all $L \leq \tilde{L}_{JL} = \frac{Y^H[1-(1-\alpha)^n]}{n} - \varepsilon$, and again the IL contract is feasible if $L \leq \alpha\delta\mathbb{E}(Y) - \varepsilon$. So the JL contract can offer a larger range of loans to each member than the IL contract if and only if

$$\delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}.$$

Since we already have a lower bound for δ, it must be true that

$$\tilde{\delta}_{JL} < \delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}.$$

Comparing the results of parts (a) and (b), the JL contract is feasible for a larger amount of loans than the IL contract if and only if

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} < \delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}. \tag{4.12}$$

What are the circumstances for which (4.12) holds? A necessary condition to satisfy (4.12) is

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}. \tag{4.13}$$

Both the right-hand side (RHS) and the left-hand side (LHS) of (4.13) are strictly monotonic in $\alpha \in (0,1)$. If α is very small, (4.13) always holds:

$$\lim_{\alpha\to 0}\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} = \frac{1}{2}$$

$$\lim_{\alpha\to 0}\frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)} = \frac{Y^H}{Y^L}.$$

If α is very large, it is easy to see that (4.13) never holds:

$$\lim_{\alpha\to 1}\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} = 1$$

$$\lim_{\alpha\to 1}\frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)} = \frac{1}{n}.$$

Therefore, we should conclude that there exists a critical $\underline{\alpha}$ such that for any $\alpha < \underline{\alpha}$, (4.12) holds, and thus the JL contract performs better than the IL contract in terms of larger feasible maximum loans. For any $\alpha > \underline{\alpha}$, (4.12) never holds, and thus the IL contract performs better than the JL contract in terms of larger feasible maximum loans.

5) For very large n, (4.12) never holds:

$$\lim_{n\to\infty}\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]} = \lim_{n\to\infty}\frac{\alpha n - 1}{\alpha(n-1)} = 1$$

$$\lim_{n\to\infty}\frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)} = 0.$$

□

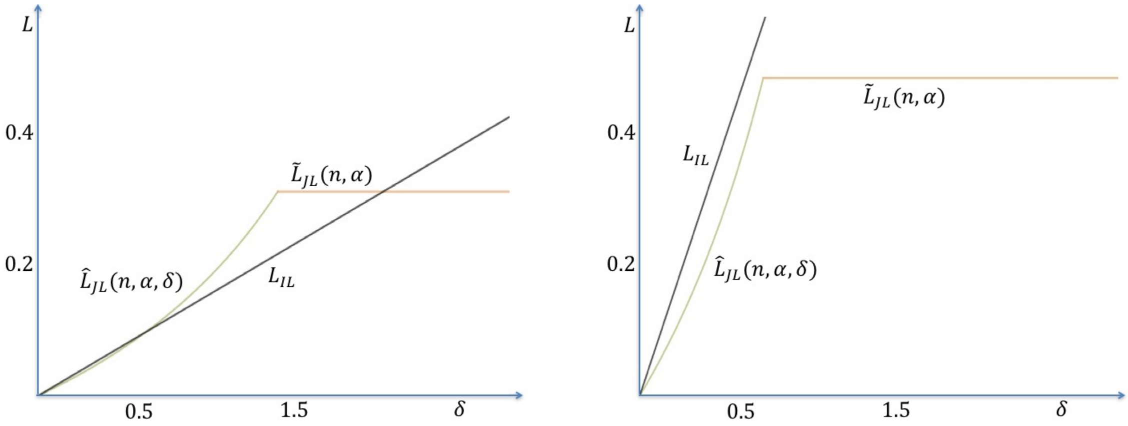
Table 4.2: The Range of Group Sizes for Which the JL Contract Outperforms the IL Contract for Some Given α

	$n=2$	$n=3$	$n=4$	$n=5$	$\cdots$	$n=10$	$\cdots$	$n=50$
$\alpha < 0.764$	✓							
$\alpha < 0.634$	✓	✓						
$\alpha < 0.552$	✓	✓	✓					
$\alpha < 0.495$	✓	✓	✓	✓				
$\alpha < 0.349$	✓	✓	✓	✓		✓		
$\alpha < 0.150$	✓	✓	✓	✓		✓		✓

Proposition 4.3 shows that the JL contract has a positive effect on borrowers' welfare and the repayment rate compared to the IL contract. The proposition suggests a necessary condition, which we will call the *relative feasibility condition*, under which the JL contract can offer a larger maximum loan than the IL contract, and as previously discussed in Proposition 4.2, the size of the loan can be increasing in the group size. Proposition 4.3 also discusses the circumstances necessary for the relative feasibility condition to be satisfied. This proposition proves that if the chance of project success is high (higher than 75%), the maximum loan that can be given to a borrower under the IL contract is higher than under the JL contract. While for projects with a lower chance of success (lower than 75%), larger loans could be given only under the JL contract. As expected, the group size cannot become too large. The JL contracts with groups of more than 10 members can offer larger loans than the IL contract only when projects have only a small chances of success (lower than 35%); otherwise the IL contract can offer a larger loan than the JL contract with groups that have too many members.

If $Y_L = 0$, for $\alpha < 0.764$, (4.11) can be satisfied for $n = 2$. Thus, for $\alpha < 0.764$ and any δ that satisfies (4.11), the JL contract does better than the IL contract at least for groups of $n = 2$. As we can see in Table 4.2, for any other smaller α, more than one group size exists for which the JL contract does better than the IL contract. Therefore, $\alpha = 0.764$ is the maximum α for which the JL contract outperforms the IL contract if the given δ is such that (4.11) holds. Thus, for any $\alpha > 0.764$, the lender should offer an IL contract. Note that if $Y^L \neq 0$, then α must be a slightly smaller. As an instance if $Y^L = \frac{Y^H}{2}$, then we should have $\alpha < 0.697$.

Figure 4.3 illustrates the situation discussed in Proposition 4.3. For simplicity, it is assumed that $Y^H = 1$ and $Y^L = 0$. As shown in 4.3a, when the chance of project success is small enough so that the relative feasibility condition is satisfied (e.g. $\alpha = 0.4$), for any $0 < \delta < 1$, $\hat{L}$ defines the boundary for the maximum feasible loan under the JL contract which is always equal to or larger than the maximum feasible loan under the IL contract. However, when the chance of project success is large (e.g., $\alpha = 0.9$), the maximum feasible loan under the IL contract is strictly higher than maximum feasible loan under the JL contract.



(a) It is assumed that $\alpha = 0.4$.

(b) It is assumed that $\alpha = 0.9$.

Figure 4.3: For smaller α, the JL contract offers larger loans than the IL contract, while for larger α, the IL contract offers larger loans than the JL contract.

4.5 Flexible Joint Liability (FJL) Contracts

The grim-trigger strategy played in JL lending may be too harsh. Consider a situation in which some borrowers default strategically, but it is still beneficial for the other group members to take care of the entire repayment and qualify for a new loan next round. Moreover, the strategically defaulting members may have good explanations and the other group members may want to punish them with less than grim-trigger. In this section, we examine joint liability under a strategy that is less severe than grim-trigger. In this type of lending, players start in the lending phase and they cooperate until someone defaults strategically. The players then proceed to a punishment phase and exclude the defaulter for T periods of receiving loans. Later, when punishment is served, they let allow him to reenter the game. The timing is the same as before. The lender's problem in this case is similar to JL lending case with a change in the incentive constraint, (4.3). Each successful borrower must have a correct incentive not to default strategically, and to repay not only for himself, but also for all the defaulting peers (even for the entire group, in the worst case). Thus we must have

$$Y^H + \delta^{T+1} V^R_{FJL} < Y^H - nR + \delta V^R_{FJL},$$

that can be simplified to

$$nR < \left(1 - \delta^T\right) \delta V^R_{FJL}. \tag{4.14}$$

Clearly, if T is a large constant, then $1 - \delta^T \to 1$, and (4.14) will be equal to the incentive constraint of the JL contract—i.e., (4.3). Note that (4.14) is more difficult to satisfy compared to (4.3), as it is less costly to default strategically under the FJL contract.

Proposition 4.4. *There exist* $\tilde{\delta}_{FJL}\left(n,\alpha,\delta,T,Y^H,Y^L\right)$ *and* $\tilde{L}_{FJL}\left(n,\alpha,Y^H\right)$ *and* $\hat{L}_{FJL}\left(n,\alpha,\delta,T,Y^H,Y^L\right)$ *such that:*

a) For any $\delta \geq \tilde{\delta}_{FJL}\left(n,\alpha,\delta,T,Y^H,Y^L\right)$, the FJL contract is feasible if and only if $L \leq \tilde{L}_{FJL}\left(n,\alpha,Y^H\right)$.

b) For any $\delta \leq \tilde{\delta}_{FJL}\left(n,\alpha,\delta,T,Y^H,Y^L\right)$, the FJL contract is feasible if and only if $L \leq \hat{L}_{FJL}\left(n,\alpha,\delta,T,Y^H,Y^L\right)$.

Moreover, whenever the FJL contract is feasible, for any $\alpha \neq 0$, the lender demands the optimal repayment $R_{FJL} = \frac{L+\varepsilon}{[1-(1-\alpha)^n]}$ from each borrower, and the expected lifetime utility for each borrower will amount to $V^R_{FJL} = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\delta[1-(1-\alpha)^n]}$.

Proof. In the case of FJL lending, the optimal contract (L,R) is a solution to the problem below. Adopting the same method used in the proof of Proposition 4.1, we solve the following problem:

$$\begin{aligned}
\underset{L,R}{maximize} \quad V^R_{FJL} &= \frac{\mathbb{E}(Y)-R\left[1-(1-\alpha)^n\right]}{1-\delta\left[1-(1-\alpha)^n\right]} && (4.15)\\
s.t. \quad nR &\leq \left(1-\delta^T\right)\delta V^R_{FJL} && (4.16)\\
nR &\leq Y^H && (4.17)\\
R &\geq \frac{L+\varepsilon}{1-(1-\alpha)^n}. && (4.18)
\end{aligned}$$

It is clear that V^R_{FJL} is decreasing in R, so the bank would like to set R as low as possible. Constraint (4.18) gives the minimum R required for breaking even, which is $R_{FJL} = \frac{L+\varepsilon}{[1-(1-\alpha)^n]}$. As long as this R is below the upper limit expressed in (4.16) and (4.17), lending is feasible and otherwise, it is not. Replacing R_{FJL} in the above problem, we will have:

$$\begin{aligned}
\underset{L,R}{maximize} \quad V^R_{FJL} &= \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\delta\left[1-(1-\alpha)^n\right]}\\
L &\leq \frac{\delta\left(1-\delta^T\right)\mathbb{E}(Y)\left[1-(1-\alpha)^n\right]}{n-\left[n\delta-\delta\left(1-\delta^T\right)\right]\left[1-(1-\alpha)^n\right]} - \varepsilon \equiv \hat{L}_{FJL}\left(n,\alpha,\delta,T,Y^H,Y^L\right)\\
L &\leq \frac{Y^H\left[1-(1-\alpha)^n\right]}{n} - \varepsilon \equiv \tilde{L}_{FJL}\left(n,\alpha,Y^H\right).
\end{aligned}$$

A feasible L must satisfy $L \leq Min\left\{\hat{L}_{FJL},\tilde{L}_{FJL}\right\}$ in order to satisfy both constraints. There are two cases: either the case that $\hat{L}_{FJL} \leq \tilde{L}_{FJL}$ or the case that $\tilde{L}_{FJL} > \hat{L}_{FJL}$.

i) $\hat{L}_{FJL} \leq \tilde{L}_{FJL}$ if and only if

$$\delta \leq \frac{Y^H}{\left(1-\delta^T\right)\mathbb{E}(Y) + \frac{n-\left(1-\delta^T\right)}{n}\left[1-(1-\alpha)^n\right]Y^H} \equiv \tilde{\delta}_{FJL}\left(n,\alpha,\mu,Y^H,Y^L\right).$$

ii) $\tilde{L}_{FJL} \leq \hat{L}_{FJL}$ if and only if $\delta \geq \tilde{\delta}$.

Therefore, if $0 \leq \delta \leq \tilde{\delta}_{FJL}$, the FJL contract is feasible for any $L \leq \hat{L}_{FJL}$, and if $\tilde{\delta}_{FJL} \leq \delta < 1$, the FJL contract is feasible for any $L \leq \tilde{L}_{FJL}$. □

It is suggested by the proof of Proposition 4.4 that $\tilde{L}_{FJL} = \tilde{L}_{JL}$, and for large T, also $\tilde{\delta}_{FJL} = \tilde{\delta}_{JL}$ and $\tilde{L}_{FJL} = \tilde{L}_{JL}$. That is, when the punishment phase is very long, we are again in the same situation as the JL contract in which borrowers played a grim-trigger strategy against their peers. Moreover, a comparison between Proposition 4.4 and Proposition 4.1 shows that the optimal repayment amount and borrowers' welfare are the same with the JL contract and the FJL contract, as $R_{FJL} = R_{JL}$ and $V^R_{FJL} = V^R_{JL}$, regardless of the magnitude of T.

We continue this section by investigating if and how the length of the punishment phase can affect the maximum feasible loan granted to each borrower and whether these changes affect the optimal group size. To this end, first in Lemma 4.3, we look at the changes of $\hat{L}_{FJL}$ and $\tilde{\delta}_{FJL}$ with respect to changes of n and α when other parameters $\left(\alpha, \delta, Y^H, Y^L\right)$ are given. Next, in Lemma 4.4, we compare $\tilde{L}_{FJL}$ and $\tilde{\delta}_{FJL}$ with their counterparts, $\tilde{L}_{JL}$ and $\tilde{\delta}_{JL}$.

Lemma 4.3. *Assume $\hat{L}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$, $\tilde{\delta}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ are functions defined in Proposition 4.4.*

1) There is a $\hat{\delta}_{FJL}(n, \alpha)$ such that:

i. For any $0 < \delta < \hat{\delta}_{FJL}(n, \alpha)$, $\hat{L}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ is strictly decreasing in n.

ii. For any $\hat{\delta}_{FJL}(n, \alpha) < \delta < \tilde{\delta}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$, $\hat{L}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ is strictly increasing in n.

2) $\tilde{\delta}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ is strictly decreasing in both n and α.

3) For any given n and α, $\dfrac{n}{(2n-1)-(n-1)\delta^T} < \tilde{\delta}_{FJL} < \dfrac{Y^H}{(1-\delta^T)Y^L}$.

Lemma 4.3 shows that the behavior of $\hat{L}_{FJL}$ and $\tilde{\delta}_{FJL}$ in response to changes of n and α are similar to what we saw in Lemmas 4.1 and 4.2 for $\hat{L}_{JL}$ and $\tilde{\delta}_{JL}$. Note that neither of $\hat{\delta}$ or $\tilde{L}$ depend on T, and they are the same for the JL and the FJL contracts. We recall from Lemmas 4.1 and 4.2 that $\tilde{L}$ is strictly decreasing in n, $\hat{\delta}$ is strictly increasing in both n and α, and for any given n and α, $\hat{\delta} < 1$.

Lemma 4.4. *Assume $\hat{L}_{JL}\left(n, \alpha, \delta, Y^H, Y^L\right)$ and $\tilde{\delta}_{JL}\left(n, \alpha, Y^H, Y^L\right)$ are functions defined in Proposition 4.1, and $\hat{L}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ and $\tilde{\delta}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ are functions defined in Proposition 4.4.*

1) $\hat{L}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right) \leq \hat{L}_{JL}\left(n, \alpha, \delta, Y^H, Y^L\right)$.

2) $\hat{L}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ is strictly decreasing in T.

3) $\tilde{\delta}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right) \geq \tilde{\delta}_{JL}\left(n, \alpha, Y^H, Y^L\right)$.

4) $\tilde{\delta}_{FJL}\left(n, \alpha, \delta, T, Y^H, Y^L\right)$ is strictly decreasing in T.

Lemma 4.4 shows that the interval $\left(\hat{\delta}, \tilde{\delta}_{FJL}\right)$ is wider than $\left(\hat{\delta}, \tilde{\delta}_{JL}\right)$; that is, $\left(\hat{\delta}, \tilde{\delta}_{FJL}\right)$ can allow for larger group sizes than was possible under the JL contract. However, this lemma argues that $\tilde{\delta}_{FJL}$ is strictly decreasing in the length of the punishment period, T. Intuitively, when the punishment phase is not that long and it is less costly for members to default strategically, we need a larger group to assure repayment. Proposition 4.5 discusses that the FJL contract have a disadvantage compared to the JL contract in terms of the maximum loan that can be offered to borrowers.

Proposition 4.5. *When both the JL and the FJL contracts are feasible, the FJL contract is feasible for, at most, the same range of loans as the JL contract.*

Proof. First, we recall from Propositions 4.1 and 4.4, the maximum loan that can be offered to each borrower each time is the minimum of $\hat{L}$ and $\tilde{L}$, and it is dependent on the magnitude of $\tilde{\delta}$. We also recall from Lemma 4.4 that $\tilde{\delta}_{JL} \leq \tilde{\delta}_{FJL}$. In order to compare the feasible loans that can be offered under the FJL and the JL contracts, we need to look at three different cases:

i. For any $\delta < \tilde{\delta}_{JL}$, the JL contract is feasible for all $L \leq \hat{L}_{JL}$, and the FJL contract is feasible for all $L \leq \hat{L}_{FJL}$. According to Lemma 4.4, $\hat{L}_{FJL} \leq \hat{L}_{JL}$. That is, for any $\delta < \tilde{\delta}_{JL}$, the JL contract can offer larger range of loans to each borrower than the FJL contract.

ii. For any $\tilde{\delta}_{JL} \leq \delta \leq \tilde{\delta}_{FJL}$, the FJL contract is feasible for all $L \leq \hat{L}_{FJL}$ and the JL contract is feasible for all $L \leq \tilde{L}$. From the proof of Proposition 4.4, we know that for any $\delta \leq \tilde{\delta}_{FJL}$, $\hat{L}_{FJL} \leq \tilde{L}$. Therefore, for any $\tilde{\delta}_{JL} \leq \delta \leq \tilde{\delta}_{FJL}$, the JL contract can offer a larger range of loans to each member than the FJL contract.

iii. For any $\delta > \tilde{\delta}_{FJL}$, both the JL and the FJL contracts are feasible for any $L \leq \tilde{L}$. That is, for any $\delta > \tilde{\delta}_{FJL}$, both the JL and the FJL contracts can offer a similar range of loans to each borrower. □

Thus, clearly reducing the length of the punishment phase creates a disadvantage for joint liability contracts in terms of the maximum feasible loan, how much is the disadvantage exactly? By reducing the punishment, are we reducing the advantage of joint liability compared to individual liability to the extent that the IL contract becomes better than the FJL contract? In Proposition 4.6, we try to answer this question by reestablishing results similar to Proposition 4.3.

Proposition 4.6. *The following statements hold when both the IL and the FJL contracts are feasible.*

1) Each borrowers' repayment amount is lower and his welfare is higher under the FJL contract than the IL contract.

2) There exists $\underline{\alpha}'$, such that for $\alpha < \underline{\alpha}'$, the FJL contract is feasible for a larger loan than the IL contract only if at least for some n,

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha (n-1)(1-\delta^T)[1-(1-\alpha)^n]} < \delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha \mathbb{E}(Y)};$$

for any $\alpha > \underline{\alpha}'$, the IL contract is feasible for a larger loan than the FJL contract only if $Y^L < 2\left(1-\delta^T\right)Y^H$.

3) For very large n, the IL contract can always offer a larger maximum loan than the FJL contract.

Proof. 1) When both the IL and the FJL contracts are feasible, each borrower is then supposed to repay $R_{FJL} = \frac{L+\varepsilon}{[1-(1-\alpha)^n]}$ under the FJL contract and $R_{IL} = \frac{L+\varepsilon}{\alpha}$ under the IL contract. Clearly, the borrower pays less under the FJL contract. Each borrower's expected lifetime utility under the FJL contract is $V_{JL}^R = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\delta[1-(1-\alpha)^n]}$ and under the IL contract is $V_{IL}^R = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\alpha\delta}$. Obviously, the borrower's expected lifetime utility of is higher under the FJL contract.

2) Is feasibility easier to achieve through the FJL contract or the IL contract? There are two cases:

i. If $\delta < \tilde{\delta}_{FJL}$, then the FJL contract is feasible for all $L \leq \hat{L}_{FJL} = \alpha\delta\mathbb{E}(Y)\times\frac{(1-\delta^T)[1-(1-\alpha)^n]}{\alpha[n-\delta(1-\delta^T)(n-1)[1-(1-\alpha)^n]]} - \varepsilon$, and the IL contract is feasible for all $L \leq \alpha\delta\mathbb{E}(Y) - \varepsilon$. Thus, the FJL contract can offer a larger range of loans to borrowers than the IL contract only if

$$\frac{[1-(1-\alpha)^n]}{\alpha[n-\delta(1-\delta^T)(n-1)[1-(1-\alpha)^n]]} > 1,$$

which can be rewritten as

$$\delta\left(1-\delta^T\right) > \frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)[1-(1-\alpha)^n]}.$$

Otherwise, the IL contract can offer a larger range of loans to borrowers than the FJL contract. Since we already have an upper bound for δ, it must be true that

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)(1-\delta^T)[1-(1-\alpha)^n]} < \delta < \tilde{\delta}_{FJL}.$$

ii. If $\delta > \tilde{\delta}_{FJL}$, the FJL contract is then feasible for all $L \leq \tilde{L}_{FJL} = \tilde{L}_{JL}$, and according to the proof of Proposition 4.3, the FJL contract can offer a larger range of loan to each member than the IL contract if and only if

$$\tilde{\delta}_{FJL} < \delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}.$$

Comparing the results of parts (a) and (b), the FJL contract is feasible for a larger amount of loans than the IL contract only if

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)(1-\delta^T)[1-(1-\alpha)^n]} < \delta < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}. \tag{4.19}$$

What are the circumstances for which (4.19) holds? A necessary condition to satisfy (4.19) is

$$\frac{\alpha n - [1-(1-\alpha)^n]}{\alpha(n-1)(1-\delta^T)[1-(1-\alpha)^n]} < \frac{Y^H[1-(1-\alpha)^n]}{n\alpha\mathbb{E}(Y)}.$$

Both the RHS and the LHS of (4.19) are strictly monotonic in $\alpha \in (0,1)$. If α is very large, it is easy to see that (4.19) never holds because

$$\lim_{\alpha \to 1} \frac{\alpha n - [1-(1-\alpha)^n]}{\alpha (n-1)(1-\delta^T)[1-(1-\alpha)^n]} = \frac{1}{(1-\delta^T)}$$

$$\lim_{\alpha \to 1} \frac{Y^H [1-(1-\alpha)^n]}{n\alpha \mathbb{E}(Y)} = \frac{1}{n}.$$

If α is very small, (4.19) holds if $Y^L < 2\left(1-\delta^T\right) Y^H$ because

$$\lim_{\alpha \to 0} \frac{\alpha n - [1-(1-\alpha)^n]}{\alpha (n-1)(1-\delta^T)[1-(1-\alpha)^n]} = \frac{1}{2(1-\delta^T)}$$

$$\lim_{\alpha \to 0} \frac{Y^H [1-(1-\alpha)^n]}{n\alpha \mathbb{E}(Y)} = \frac{Y^H}{Y^L}.$$

Therefore, we should conclude that there exists a critical $\underline{\alpha}'$ such that for any $\alpha < \underline{\alpha}'$, the necessary condition holds only if $Y^L < 2\left(1-\delta^T\right) Y^H$. For any $\alpha > \underline{\alpha}'$, (4.19) never holds and so the IL contract performs better than the FJL contract in terms of larger feasible maximum loans.

4) For very large n, (4.19) never holds:

$$\lim_{n \to \infty} \frac{\alpha n - [1-(1-\alpha)^n]}{\alpha (n-1)(1-\delta^T)[1-(1-\alpha)^n]} = \frac{1}{(1-\delta^T)}$$

$$\lim_{n \to \infty} \frac{Y^H [1-(1-\alpha)^n]}{n\alpha \mathbb{E}(Y)} = 0.$$

□

Proposition 4.6 shows that although the FJL contract can handle smaller maximum feasible loans than the JL contract, it can still be better than the IL contract. More specifically, the FJL contract, similar to the JL contract, can have a positive effect on borrowers' welfare and the repayment rate compared to the IL contract and support at least the same maximum loan as the IL contract.

Proposition 4.6 suggests a necessary condition—i.e., (4.19)—under which the FJL contract can offer a larger maximum loan than the IL contract. The necessary condition suggested by Proposition 4.6 is another version of the relative feasibility condition suggested by Proposition 4.3, except that this new version is tighter and more difficult to satisfy than the original version. Proposition 4.6 also verifies that (4.19) can hold for small enough α, that is, the maximum loan that can be offered to borrowers under the FJL contract can be higher than under the IL contract, when the chance of project success is small.

From Proposition 4.5 and Proposition 4.6, we should conclude that from the borrower's point of view the FJL and the JL contracts are almost similar, as they both improve borrower's welfare and can offer a larger maximum loan to the borrower compared to the IL contract. This gives the borrowers the opportunity to decide how much flexibility they want to show regarding their group member's strategic default.

4.6 Project Correlation

Although most of the literature on microfinance group lending assumes independence between projects,[3] it may be more realistic to assume that projects are correlated. Correlation can come from the macro environment, such as whether condition, the externalities formed through joint liability, or because the businesses are linked to each other. For example, when a harvest is good in a region, all agriculture-related businesses are affected positively. In village economies, there is an additional reason to assume that projects are correlated. In these economies each small business becomes more attractive as money increases in the society. Therefore, any successful project can boost other businesses in the neighborhood.

In this section, we examine whether larger groups are always better and have higher outcomes when projects are correlated. We include correlation of projects in our model by assuming that α is not a constant but an increasing function of n. This is a realistic assumption, because being in charge of each other's repayments would increase cooperation among group members, and thus group members may actually contribute to each other's chance of project success. We only consider the case of the FJL contract, as the JL contract can be seen as special cases of the FJL contract, in which the punishment phase is very long—i.e., $T \to \infty$. We assume $\alpha(n)$ is a concave function. We consider this to be a realistic assumption because it might be the case that a very large group deals with higher tensions, which could have adverse effects on a project's success. We also assume that $Y^H = 1$ and $Y^L = 0$ for simplicity of calculations. Note that we do not lose generality with this latter assumption, as we have already assumed that an outcome could only be high or low and nothing in between.

Since assuming project correlation does not affect our model, except for the fact that α depends on n, the results of Proposition 4.4 stay valid. That is, when project correlation exists, for the same $\tilde{\delta}_{FJL}\left(n,\alpha(n),\delta,T,Y^H,Y^L\right)$, $\tilde{L}_{FJL}\left(n,\alpha(n),Y^H\right)$ and $\hat{L}_{FJL}\left(n,\alpha(n),\delta,T,Y^H,Y^L\right)$, which were assumed in Proposition 4.4, the following statements are correct. For any $\delta \geq \tilde{\delta}_{FJL}$, the FJL contract is feasible if and only if $L \leq \tilde{L}_{FJL}$, and for any $\delta \leq \tilde{\delta}_{FJL}$, the FJL contract is feasible if and only if $L \leq \hat{L}_{FJL}$. Moreover, whenever the FJL contract is feasible, for any $\alpha \neq 0$, the lender demands the optimal repayment $R_{FJL} = \frac{L+\varepsilon}{[1-(1-\alpha(n))^n]}$ from each borrower, and the expected lifetime utility for each borrower will amount to $V^R_{FJL} = \frac{\mathbb{E}(Y)-(L+\varepsilon)}{1-\delta[1-(1-\alpha(n))^n]}$.

[3]There are some exceptions that consider project correlations (see e.g., Laffont (2003) and Ahlin and Townsend (2007)).

In what follows, we examine the effect of project correlation on the maximum feasible loan under the FJL contract. The following lemma provides us with some insight on where the maximum feasible loan can be increasing in the group size.

Lemma 4.5. *Assume $\tilde{\delta}'_{FJL}(n,\alpha(n),\delta,T)$, $\hat{L}'_{FJL}(n,\alpha(n),\delta,T)$ and $\tilde{L}'_{FJL}(n,\alpha(n))$ are functions defined in Proposition 4.4 with the only difference that α is a function of n, and it is assumed that $Y^H=1$ and $Y^L=0$.*

1) If $\alpha'(n) < \frac{\alpha(n)}{n-(1-\delta^T)}$, then there exists $\hat{\delta}'_{FJL}(n,\alpha(n),\alpha'(n),\delta,T)$ such that

i. For any $0<\delta<\hat{\delta}'_{FJL}(n,\alpha(n),\alpha'(n),\delta,T)$, $\hat{L}'_{FJL}(n,\alpha(n),\delta,T)$ is strictly decreasing in n.

ii. For any $\hat{\delta}'_{FJL}(n,\alpha(n),\alpha'(n),\delta,T)<\delta<\tilde{\delta}'_{FJL}(n,\alpha,\delta,T)$, $\hat{L}'_{FJL}(n,\alpha(n),\delta,T)$ is strictly increasing in n.

2) If $\alpha'(n) < \frac{1-(1-\alpha(n))^n+(1-\alpha(n))^n\ln(1-\alpha(n))^n}{n(1-\alpha(n))^{n-1}}$, then $\tilde{L}'_{FJL}(n,\alpha(n))$ is strictly decreasing in n.

Lemma 4.5 shows that when projects are weakly correlated (i.e., correlation degree is lower than boundaries suggested in Lemma 4.5), the determinants of the maximum feasible loan under the FJL contract—i.e., $\hat{L}'_{FJL}(n,\alpha(n),\delta)$ and $\tilde{L}'_{FJL}(n,\alpha(n))$—still show behaviors similar to the case in which projects are independent. Therefore, similar to the case of independent projects, the maximum feasible loan could be increasing in group size for some mid-range discount factor, more specifically for any $\delta\in\left(\hat{\delta}'_{FJL},\tilde{\delta}'_{FJL}\right)$. Therefore, in case projects are weakly correlated, whether the maximum loan is increasing in n depends on $\left(\hat{\delta}'_{FJL},\tilde{\delta}'_{FJL}\right)$ being non-empty. However, contrary to the case in which projects are independent, this is not the only time that the maximum loan could increase in n.

From Lemma 4.5, it could also be inferred that if the projects are correlated (i.e., the correlation degree is higher than boundaries suggested in Lemma 4.5), the behavior of $\hat{L}'_{FJL}$ and $\tilde{L}'_{FJL}$ will then be reversed. In this case, for any $0<\delta<\hat{\delta}'_{FJL}$, $\hat{L}'_{FJL}$ will be strictly increasing in n, for any $\hat{\delta}'_{FJL}<\delta<\tilde{\delta}'_{FJL}$, $\hat{L}'_{FJL}$ will be strictly decreasing in n, and moreover, $\tilde{L}'_{FJL}$ will be always strictly increasing in n. To put it concisely, when projects are strongly correlated, the determinants of the maximum feasible loan under the FJL contract, and consequently the maximum feasible loan itself, are always increasing in group size except for some mid-range discount factor, $\delta\in\left(\hat{\delta}'_{FJL},\tilde{\delta}'_{FJL}\right)$. Proposition 4.7 proves that $\left(\hat{\delta}'_{FJL},\tilde{\delta}'_{FJL}\right)$ is non-empty when projects are weakly correlated and it is empty when they are strongly correlated.

Proposition 4.7. *When projects are correlated, feasibility of the FJL contract is always increasing in n.*

Proof. 1) When projects are weakly correlated (i.e., correlation degree is lower than boundaries suggested in Lemma 4.5), feasibility of the FJL contract is increasing in n as far as

$$\hat{\delta}'_{FJL}\left(n,\alpha(n),\alpha'(n),\delta,T\right)<\tilde{\delta}'_{FJL}(n,\alpha(n),\delta,T).$$

We know from Lemma 4.5 that $\alpha' < \frac{\alpha}{n-(1-\delta^T)}$, thus the above inequality can be written as

$$\alpha' > \frac{\alpha\left[1-(1-\alpha)^n\right]+\alpha(1-\alpha)^n\ln(1-\alpha)^n-\dfrac{\alpha\left[1-(1-\alpha)^n\right]^2}{\alpha\left(1-\delta^T\right)+\frac{n-1+\delta^T}{n}\left[1-(1-\alpha)^n\right]}}{n\left[1-(1-\alpha)^n\right]+n^2\alpha(1-\alpha)^{n-1}-\dfrac{\left(n-1+\delta^T\right)\left[1-(1-\alpha)^n\right]^2}{\alpha\left(1-\delta^T\right)+\frac{n-1+\delta^T}{n}\left[1-(1-\alpha)^n\right]}}. \tag{4.20}$$

For large n, the RHS of (4.20) is zero:

$$\lim_{n\to\infty}\frac{\alpha\left[1-(1-\alpha)^n\right]+\alpha(1-\alpha)^n\ln(1-\alpha)^n-\dfrac{\alpha\left[1-(1-\alpha)^n\right]^2}{\alpha\left(1-\delta^T\right)+\frac{n-1+\delta^T}{n}\left[1-(1-\alpha)^n\right]}}{n\left[1-(1-\alpha)^n\right]+n^2\alpha(1-\alpha)^{n-1}-\dfrac{\left(n-1+\delta^T\right)\left[1-(1-\alpha)^n\right]^2}{\alpha\left(1-\delta^T\right)+\frac{n-1+\delta^T}{n}\left[1-(1-\alpha)^n\right]}}$$

$$=\lim_{n\to\infty}\frac{1+0-\dfrac{1}{2-\delta^T}}{n+0-\dfrac{n-1+\delta^T}{2-\delta^T}}=0.$$

For small n, the RHS of (4.20) is smaller than zero:

$$\lim_{n\to 1}\frac{\alpha\left[1-(1-\alpha)^n\right]+\alpha(1-\alpha)^n\ln(1-\alpha)^n-\dfrac{\alpha\left[1-(1-\alpha)^n\right]^2}{\alpha\left(1-\delta^T\right)+\frac{n-1+\delta^T}{n}\left[1-(1-\alpha)^n\right]}}{n\left[1-(1-\alpha)^n\right]+n^2\alpha(1-\alpha)^{n-1}-\dfrac{\left(n-1+\delta^T\right)\left[1-(1-\alpha)^n\right]^2}{\alpha\left(1-\delta^T\right)+\frac{n-1+\delta^T}{n}\left[1-(1-\alpha)^n\right]}}$$

$$=\lim_{n\to 1}\frac{(1-\alpha)\ln(1-\alpha)}{2-\delta^T}<0.$$

Thus, the RHS of (4.20) is always negative, while the LHS is always positive (we assume that α is increasing in n). Therefore, (4.20) is always valid.

2) From our previous discussion in part 1 of this proposition, it can be understood that when projects are strongly correlated (i.e., correlation degree is higher than boundaries suggested in Lemma 4.5), feasibility of the FJL contract is increasing in $n$if and only if

$$\hat{\delta}'_{FJL}\left(n,\alpha(n),\alpha'(n),\delta,T\right) > \tilde{\delta}'_{FJL}\left(n,\alpha(n),\delta,T\right).$$

As projects are strongly correlated, $\alpha' > \frac{\alpha}{n-(1-\delta^T)}$. Thus, the above inequality can be also rewritten as (4.20) that we have already proved to hold always. □

Proposition 4.7 suggests that when projects are correlated, the maximum loan that can be offered under the FJL contract and consequently under the JL contract is higher in larger groups.

4.7 Conclusion

The original question that motivates our study is how group size affects the efficiency of microcredit lending under joint liability contracts and determining the group optimal size. We analyze a model of repeated microcredit lending in which a microcredit is given to a group of n members by a benevolent lender. Borrowers invest the loan on their separate projects and the lender does not observe the outcome of the borrowers' risky projects. Therefore, if there is some default on borrowers' loan repayments, the lender is not able to identify whether the default is strategic or not.

Most of the existing theoretical papers in microfinance discuss joint liability lending in groups of only two members, although experimental and empirical studies emphasize the importance of group size (Abbink et al., 2006 and Galak et al., 2011). Our results show that group size can be an influential factor in improving lending efficiency and the assumption $n = 2$, which is largely used in the literature, may result in neglecting the effect of group size. In particular, we calculate the optimal group size endogenously while deriving the optimal contract that maximizes the borrowers' welfare subject to the following conditions: repayment being affordable for each borrower, repayment being better than default for each borrower, and the lender breaking even.

Our results show that although joint liability contracts are feasible under a smaller set of parameter values than individual lending contracts, the feasibility of joint liability contracts can increase in the group size for riskier projects, meaning that larger groups are more reliable in repaying the loan when projects are risky, but the group size cannot become too large. Our findings confirm the results of Conning (2004) and Ahlin (2015). These authors, although focused on different biases (free-riding, and local borrower information respectively), also conclude that borrowers' groups cannot become too large. Intuitively, there are costs and benefits in being a member of a large group of borrowers. On the one hand, it enhances the chance of assured repayment for a defaulting member. On the other hand, there is also a higher threat of repayment for other defaulting group members. While for risky projects the insurance provided by larger groups becomes increasingly more attractive, it is less necessary that low-risk projects are insured by multiple group members. In this paper, a characterization of the optimal group size is suggested given the maximum and minimum returns of projects, the discount factor of borrowers for future loans, and the chance of project success.

Furthermore, assuming that group members play a grim-trigger strategy against any strategically defaulting member, we discuss that joint liability, when feasible, has a positive effect on borrowers' welfare and the repayment rate compared to individual liability. We demonstrate that joint liability can also outperform individual liability in terms of the maximum loan that can be offered to borrowers. This is contrary to the existing literature, which concludes that a lender may be better off by choosing individual liability over joint liability (Besley and Coate, 1995, Armendáriz de Aghion, 1999), unless borrowers are able to impose strong social sanctions on each other. The reason that we obtain a different result is because we assume that borrowers play grim-trigger against any strategically defaulting member of the group, that is, they exclude the strategically defaulting member from the lending game at least for a T period of time and maybe even forever.

Besides comparing joint liability and individual liability contracts, our study engages in comparing flexible joint liability with joint liability contracts and seeks to learn how much remission is possible for the defaulting borrowers. We discuss that from the borrowers' point of view, flexible joint liability and joint liability contracts are the same, as their lifetime welfare and repayment amount remains the same under both contracts. Therefore, they can decide upon their level of flexibility against strategically defaulting peers. Moreover, in order to enhance the likelihood of repayment, larger groups should be formed when the punishment phase is not overly long. Finally, we investigate the effect of project correlation on the feasibility of joint liability contracts, and we argue that the maximum loan that can be offered to joint liability groups could increase in the group size, and this result is robust to changes in the degree of correlation and borrowers' discount factor.

Larger groups would be more attractive when borrowers are loss averse, since they can assure a borrower's payoff through insurance provided by the group members. Hence, a large group can also be seen as a way of reducing the risk linked with group lending (Stiglitz, 1990). We leave this issue for further research.

4.8 Appendix

Proof of Lemma 4.1. 1) The derivative of $\hat{L}_{JL}$ with respect to n (or $\frac{\partial \hat{L}_{JL}}{\partial n}$) after simplification is

$$\frac{\delta \mathbb{E}(Y)\left[-(1-\alpha)^n \ln(1-\alpha)^n + \delta\left[1-(1-\alpha)^n\right]^2 - \left[1-(1-\alpha)^n\right]\right]}{\left[n-\delta(n-1)\left[1-(1-\alpha)^n\right]\right]^2}.$$

$\frac{\partial \hat{L}_{JL}}{\partial n} < 0$ if and only if

$$-(1-\alpha)^n \ln(1-\alpha)^n + \delta\left[1-(1-\alpha)^n\right]^2 - \left[1-(1-\alpha)^n\right] < 0,$$

that is, $\frac{\partial \hat{L}_{JL}}{\partial n} < 0$ if and only if

$$\delta < \frac{\left[1-(1-\alpha)^n\right] + (1-\alpha)^n \ln(1-\alpha)^n}{\left[1-(1-\alpha)^n\right]^2} \equiv \hat{\delta}_{JL}(n,\alpha).$$

Thus, for any $0 < \delta < \hat{\delta}_{JL}$, $\hat{L}_{JL}$ is strictly decreasing in n, and if for some δ, $\hat{\delta}_{JL} < \delta < \tilde{\delta}_{JL}$, then $\hat{L}_{JL}$ will be strictly increasing in n.

2) The derivative of $\tilde{L}_{JL}$ with respect to n (or $\frac{\partial \tilde{L}_{JL}}{\partial n}$) after simplification is

$$\frac{(1-\alpha)^n\left[1-\ln(1-\alpha)^n\right] Y^H - Y^H}{n^2}.$$

$\frac{\partial \tilde{L}_{JL}}{\partial n} < 0$ because for any $\alpha \neq 0$, $(1-\alpha)^n [1-\ln(1-\alpha)^n] < 1$. The simple proof of this claim is as follows. Assume, $\rho = (1-\alpha)^n$, then $H(\rho) = \rho[1-\ln\rho]$. Since $\frac{\partial H}{\partial \rho} = -\ln\rho$ is positive for any $0 < \rho < 1$, zero for $\rho = 1$, and negative for any $\rho > 1$, then $H(\rho)$ has a maximum at $\rho = 1$. Therefore, $H(\rho) < H(\rho = 1)$ or $H(\rho) < 1$.

Thus, $\tilde{L}_{JL}$ is strictly decreasing in $n\square$

Proof of Lemma 4.2. 1) Derivatives of $\hat{\delta}_{JL}$ with respect to n and α, $\frac{\partial \hat{\delta}_{JL}}{\partial n}$ and $\frac{\partial \hat{\delta}_{JL}}{\partial n}$, after simplification are

$$\frac{\partial \hat{\delta}_{JL}}{\partial n} = (1-\alpha)^n \ln(1-\alpha)^n \times \frac{\ln(1-\alpha)^n [1+(1-\alpha)^n] + 2[1-(1-\alpha)^n]}{[1-(1-\alpha)^n]^3}$$

$$\frac{\partial \hat{\delta}_{JL}}{\partial \alpha} = -n(1-\alpha)^{n-1} \times \frac{\ln(1-\alpha)^n [1+(1-\alpha)^n] + 2[1-(1-\alpha)^n]}{[1-(1-\alpha)^n]^3}.$$

$\hat{\delta}_{JL}$ is strictly increasing in both n and α if and only if

$$\ln(1-\alpha)^n [1+(1-\alpha)^n] + 2[1-(1-\alpha)^n] < 0,$$

which can be rewritten as

$$K(n,\alpha) \equiv \ln(1-\alpha)^n + \frac{2[1-(1-\alpha)^n]}{[1+(1-\alpha)^n]} < 0.$$

It can be proved that for any $0 < \alpha < 1$, $K(n,\alpha) < 0$ as follows. Assume $\rho = (1-\alpha)^n$, then $K(\rho) = \ln\rho + \frac{2[1-\rho]}{[1+\rho]}$. Since for any $0 < \rho < 1$, $\frac{dK(\rho)}{d\rho} = \frac{[1-\rho]^2}{\rho[1+\rho]^2} > 0$, $K(\rho)$ is increasing in ρ. On the other hand, $\lim_{\rho\to 1} K(\rho) = 0$. Therefore, for any $0 < \rho < 1$, $K(\rho) < 0$, and so for any $0 < \alpha < 1$, $K(n,\alpha) < 0$.

Thus, $\hat{\delta}_{JL}$ is strictly increasing in both n and α.

2) $\hat{\delta}_{JL}$ is strictly increasing in α, thus $\lim_{\alpha\to 0} \hat{\delta}_{JL} < \hat{\delta}_{JL} < \lim_{\alpha\to 1} \hat{\delta}_{JL}$, that is, $-\infty < \hat{\delta}_{JL} < 1$.

3) $\tilde{\delta}_{JL}$ is strictly decreasing in both n and α, simply because $[1-(1-\alpha)^n]$ is strictly increasing in both n and α.

4) $\tilde{\delta}_{JL}$ is strictly decreasing in α, thus $\lim_{\alpha\to 1} \tilde{\delta}_{JL} < \tilde{\delta}_{JL} < \lim_{\alpha\to 0} \tilde{\delta}_{JL}$, that is $\frac{n}{2n-1} < \tilde{\delta}_{JL} < \frac{Y^H}{Y^L}$.
$\square$

Proof of Lemma 4.3. 1)

$$\frac{\partial \hat{L}_{FJL}}{\partial n} = \frac{\delta\left(1-\delta^T\right)\mathbb{E}(Y)\left[-(1-\alpha)^n \ln(1-\alpha)^n + \delta[1-(1-\alpha)^n]^2 - [1-(1-\alpha)^n]\right]}{[n-\delta[n-(1-\delta^T)][1-(1-\alpha)^n]]^2}.$$

$\frac{\partial \hat{L}_{FJL}}{\partial n} < 0$ if and only if

$$-(1-\alpha)^n \ln(1-\alpha)^n + \delta\left[1-(1-\alpha)^n\right]^2 - \left[1-(1-\alpha)^n\right] < 0.$$

Thus, $\frac{\partial \hat{L}_{FJL}}{\partial n} < 0$ if and only if

$$\delta < \frac{(1-\alpha)^n \ln(1-\alpha)^n + \left[1-(1-\alpha)^n\right]}{\left[1-(1-\alpha)^n\right]^2} \equiv \hat{\delta}(n,\alpha).$$

Therefore, $\hat{L}_{FJL}$ is strictly decreasing in n for any $0 < \delta < \hat{\delta}(n,\alpha)$, and $\hat{L}_{FJL}$ is strictly increasing in n for any $\hat{\delta}(n,\alpha) < \delta < \tilde{\delta}_{FJL}$.

2) $\tilde{\delta}_{FJL}$ is strictly decreasing in both n and α, simply because its denominator is strictly increasing in both n and α.

3) Since $\tilde{\delta}_{FJL}$ is strictly decreasing in α, we have $\lim_{\alpha\to 1} \tilde{\delta}_{FJL} < \tilde{\delta}_{FJL} < \lim_{\alpha\to 0} \tilde{\delta}_{FJL}$, that is

$$\frac{n}{(2n-1)-(n-1)\delta^T} < \tilde{\delta}_{FJL} < \frac{Y^H}{(1-\delta^T)Y^L}. \quad \square$$

Proof of Lemma 4.4. 1) $\hat{L}_{FJL} \leq \hat{L}_{JL}$ if and only if

$$\frac{\delta\left(1-\delta^T\right)\mathbb{E}(Y)\left[1-(1-\alpha)^n\right]}{n-\delta\left[n-(1-\delta^T)\right]\left[1-(1-\alpha)^n\right]} \leq \frac{\delta\mathbb{E}(Y)\left[1-(1-\alpha)^n\right]}{n-\delta(n-1)\left[1-(1-\alpha)^n\right]},$$

which can be simplified to $\left(1-\delta^T\right) \leq 1$ that is always true.

2)
$$\frac{\partial \hat{L}_{FJL}}{\partial T} = \frac{n\delta^{T+1}(\ln\delta)\mathbb{E}(Y)\left[1-(1-\alpha)^n\right]\left(\delta\left[1-(1-\alpha)^n\right]-1\right)}{\left(n-\delta\left[n-(1-\delta^T)\right]\left[1-(1-\alpha)^n\right]\right)^2}.$$

$\frac{\partial \hat{L}_{FJL}}{\partial T}$ is always positive.

3) $\tilde{\delta}_{FJL} \geq \tilde{\delta}_{JL}$ if and only if

$$\left(1-\delta^T\right)\mathbb{E}(Y) + \frac{n-\left(1-\delta^T\right)}{n}\left[1-(1-\alpha)^n\right]Y^H \leq \mathbb{E}(Y) + \frac{n-1}{n}\left[1-(1-\alpha)^n\right]Y^H,$$

which can be simplified to
$$\frac{\left[1-(1-\alpha)^n\right]Y^H}{n} \leq \mathbb{E}(Y),$$

which is always true.

4)
$$\frac{\partial \tilde{\delta}_{FJL}}{\partial T} = \frac{Y^H\delta^T \ln\delta\left[\frac{1}{n}\left[1-(1-\alpha)^n\right]Y^H - \mathbb{E}(Y)\right]}{\left[(1-\delta^T)\mathbb{E}(Y) + \frac{n-\left(1-\delta^T\right)}{n}\left[1-(1-\alpha)^n\right]Y^H\right]^2}.$$

$\frac{\partial \hat{L}_{FJL}}{\partial T} < 0$ if and only if $\frac{1}{n}\left[1-(1-\alpha)^n\right]Y^H - \mathbb{E}(Y) < 0$ which is always true. □

Proof of Lemma 4.5.

1)

$$\begin{aligned}\frac{\partial \hat{L}'_{FJL}}{\partial n} =& \frac{\delta\alpha'\left(1-\delta^T\right)\left[1-(1-\alpha)^n\right]\left(n-\delta\left[n-\left(1-\delta^T\right)\right]\left[1-(1-\alpha)^n\right]\right)}{\left(n-\delta\left[n-(1-\delta^T)\right]\left[1-(1-\alpha)^n\right]\right)^2} \\ &+ \frac{n\delta\alpha\left(1-\delta^T\right)\left[n\alpha'(1-\alpha)^{n-1}-(1-\alpha)^n\ln(1-\alpha)\right]}{\left(n-\delta\left[n-(1-\delta^T)\right]\left[1-(1-\alpha)^n\right]\right)^2} \\ &- \frac{\delta\alpha\left(1-\delta^T\right)\left[1-(1-\alpha)^n\right]\left(1-\delta\left[1-(1-\alpha)^n\right]\right)}{\left(n-\delta\left[n-(1-\delta^T)\right]\left[1-(1-\alpha)^n\right]\right)^2}.\end{aligned}$$

$\frac{\partial \hat{L}'_{FJL}}{\partial n} < 0$ if and only if

$$\begin{aligned}&\alpha'\left[1-(1-\alpha)^n\right]\left(n-\delta\left[n-\left(1-\delta^T\right)\right]\left[1-(1-\alpha)^n\right]\right) \\ &+n\alpha\left[n\alpha'(1-\alpha)^{n-1}-(1-\alpha)^n\ln(1-\alpha)\right] \\ &-\alpha\left[1-(1-\alpha)^n\right]\left(1-\delta\left[1-(1-\alpha)^n\right]\right) \\ &< 0.\end{aligned}$$

Assuming that $n\alpha' - \alpha'\left(1-\delta^T\right) - \alpha < 0$ or $\alpha' < \frac{\alpha}{n-(1-\delta^T)}$, the above inequality can be simplified to

$$\delta < \frac{(n\alpha'-\alpha)\left[1-(1-\alpha)^n\right] + n\alpha\left[n\alpha'(1-\alpha)^{n-1}-(1-\alpha)^n\ln(1-\alpha)\right]}{\left(n\alpha'-\alpha'(1-\delta^T)-\alpha\right)\left[1-(1-\alpha)^n\right]^2} \equiv \hat{\delta}'_{FJL}\left(n,\alpha,\alpha',\delta,T\right).$$

Thus, for any $0 < \delta < \hat{\delta}'_{FJL}$, $\hat{L}'_{FJL}$ is strictly decreasing in n, and for any $\hat{\delta}'_{FJL} < \delta < \tilde{\delta}'_{FJL}$, $\hat{L}'_{FJL}$ is strictly increasing in n.

2)

$$\frac{\partial \tilde{L}'_{FJL}}{\partial n} = \frac{n\alpha'(1-\alpha)^{n-1}-(1-\alpha)^n\ln(1-\alpha)^n-\left[1-(1-\alpha)^n\right]}{n^2}.$$

$\frac{\partial \tilde{L}'_{FJL}}{\partial n} < 0$ if and only if

$$\alpha' < \frac{1-(1-\alpha)^n+(1-\alpha)^n\ln(1-\alpha)^n}{n(1-\alpha)^{n-1}}.$$

Thus, if α is such that the above inequality is satisfied, $\tilde{L}'_{FJL}$ is strictly decreasing in n. □

Chapter 5

Conclusion

This dissertation is written as a collection of papers and each chapter provides its own conclusion. In this chapter, we present a summary of the findings discussed in the previous chapters and some possible directions for further research. The first two studies of this dissertation are concentrated on the impact of reference-dependent preferences and loss aversion on market outcome. In particular, we present two theoretical models that introduce reference-dependent preferences to classical models of monopoly pricing and labor market signaling, and examine how reference-dependent preferences affect the existence and selection of equilibria. The third study of this dissertation analyzes a theoretical model of a repeated lending situation and discusses the effect of group size on the lending outcome when groups of borrowers are held jointly liable for the total loan repayment of the group.

5.1 Summary of Findings

In recent years, the prospect theory framework has increasingly replaced the conventional expected utility theory framework, as the former has proved successful in explaining different phenomena that existing theories could not. Theoretical behavioral economics literature on adverse selection that adopts assumptions of prospect theory focuses on one central question: How do reference-dependent preferences and loss aversion affect the existence of separating equilibria? In Chapters 2 and 3 of this dissertation, we try to shed more light on how reference-dependent preferences and loss aversion influence the market equilibria and their selection in hidden information situations. We deviate from the existing literature by arguing that reference-dependent preferences may give rise to separating equilibria, whereas most of the literature shows how reference-dependent preferences lead to pooling equilibria or bunching equilibria. In Chapter 4 of the dissertation, we use the conventional expected utility theory as our modeling framework and concentrate on a different problem. In this chapter, we investigate the role of group size in microcredit contracts with joint liability when the threat of ex-post moral hazard exists.

In Chapters 2 Chapter 3, we incorporate loss-averse agents with reference-dependent preferences into standard monopoly pricing and signaling models. In Chapter 2, we compare the case in which consumers have reference-dependent preferences with the case in which they have standard preferences. We show that when consumers have reference-dependent preferences, menu pricing becomes easier in two ways. First, with reference-dependent preferences, even when the consumers' intrinsic utility does not satisfy the single-crossing property, menu pricing is possible. That is, we can have menu pricing even when richer consumers do not have a higher marginal utility of quality. If rich consumers expect to buy high quality at a high price, and poor consumers expect to buy low quality at a low price, loss-averse consumers with reference-dependent preferences self-select their types naturally. This is because buying the low-quality option gives a feeling of loss in quality to the high-valuation consumer, who expected to gain higher quality; and similarly, buying the high-quality option gives a feeling of loss in price to the low-valuation consumer, who expected to pay less. Second, with standard preferences, it may be the case that the monopolist prefers offering a single high price and thus selling only to the rich consumers rather than menu pricing, whereas with reference-dependent preferences, the monopolist prefers menu pricing when the portion of high-valuation consumers is not very high among consumers. However, this result is conditioned on both types having the same marginal utility of quality. In this chapter, we also prove that self-selection happens only if the monopolist offers high quality at a discount. However, the monopolist does not need to offer a very high discount when consumers have reference-dependent preferences because a loss-averse rich consumer is less inclined to choose the low-quality option in any case.

In Chapter 3, we show that in a signaling game, loss-averse workers, who have heterogeneous expectations, are naturally inclined to signal credibly. The reason that a loss-averse, high-productivity worker acquires a high education is that he strongly dislikes the low wages that are associated with a low education. Although acquiring a low education leads to some gains in the cost of education, this gain does not compensate for the loss he experiences in wages. Analogously, the reason that a loss-averse, low-productivity worker acquires a low education is that he strongly dislikes the cost of acquiring a high education. For this type of worker, the higher wage that is associated with a higher education does not compensate for the loss that he experiences paying for a high education. In this chaper, we formally prove that when senders have reference-dependent preferences, and their utility function contains a gain-loss element, separating equilibria exist for a larger variety of productivity differences compared to the case in which senders have standard preferences. We also prove that when the degree of loss aversion is sufficiently high, separating equilibrium may exist even when the single-crossing property is violated.

In both models discussed in Chapter 2 and Chapter 3, multiple personal equilibria may arise. This happens because if agents have reference-dependent preferences, the same contract can support several expectation sets. In a monopoly pricing model, for the same price and quality offered, it may be a personal equilibrium that the consumer buys the good because he expects to buy it, and it may also be a personal equilibrium that the consumer does not buy the good because he does not expect to buy it. Similarly, in the signaling model, for the same wage scheme, it may be a personal equilibrium that the sender acquires a high education because he expect to acquire a high education, and at the same time, it may be a personal equilibrium that he does not acquire a high education because he does not expect to.

We employ different methods of equilibrium selection in Chapter 2 and Chapter 3, in the case of multiple personal equilibria. In the monopoly pricing model of Chapter 2, we use *strategic framing* as a modeling assumption. We assume that the monopolist can ensure that the consumer forms expectations such that he expects to buy a particular variety of the good. This may take place if the monopolist is able to influence the consumer expectations and thereby his reference point. In the signaling model of Chapter 3, we utilize the Kőszegi and Rabin's (2006; 2007) notion of *preferred personal equilibrium*, which is the personal equilibrium most preferred by senders. The idea of the preferred personal equilibrium is based on the assumption that senders are able to anticipate the outcomes of their expectations, and they always pick the expectation that creates a higher level of outcome for them. In the case of multiple personal equilibria, we prove that the separating personal equilibrium can be the preferred personal equilibrium under some parameter settings even when the single-crossing property is violated. There may be situations where the separating personal equilibrium exists, but it is not preferred to the pooling personal equilibrium, as long as loss aversion is sufficiently high, that is, the case in which "the pooling personal equilibrium is always the preferred personal equilibrium" will not occur for a sufficiently high degree of loss aversion. The reason that we do not use strategic framing in the signaling model is that in this game, the informed player with expectation-based loss aversion moves first, and the other player has no chance to manipulate his expectations before signing a contract.

Chapter 4 deals with the importance of group size in group lending contracts under joint liability. Agents of our model are subject to ex-post moral hazard (i.e., they may default strategically on their loan repayments) and unlike the previous two chapters, are assumed to have standard preferences. This study calculates the optimal group size endogenously while deriving the optimal contract that maximizes the borrower's welfare and lets the lender break even. The results show that although joint liability contracts are feasible under a smaller set of parameter values than individual liability contracts, the feasibility of joint liability contracts can increase in the group size for riskier projects, meaning that larger groups are more reliable in repaying the loan when projects are risky. However, it is important that group size cannot grow too large. Intuitively, there are costs and benefits in being a member of a large group of borrowers. On the one hand, it increases the repayment insurance for a defaulting member. On the other hand, it also increases the risk of having ti repay for other defaulting members of the group. Although larger groups have a higher performance in high-risk projects, in low-risk projects, lending to smaller groups and sometimes even to individuals, can be optimal. In this chapter, we calculate the optimal group size given the minimum and maximum return of projects, the discount factor of borrowers for future loans, and the chance of project success. Furthermore, we prove that joint liability lending has a positive effect on borrowers' welfare and the repayment rate compared to individual liability lending. If the group size is not too large, joint liability lending can also outperform individual liability lending in terms of the maximum loan that can be offered to the borrower.

Later in this chapter, we take our model one step further by looking at two cases. First, we relax the grim-trigger strategy played by the group borrowers against each other, and assume a less severe punishment against strategically defaulting borrowers. More specifically, we assume that the punishment phase lasts only for T periods of the lending game instead of infinite periods. We prove that the shorter the length of the punishment phase, the larger the group should be in order to ensure the total repayment. Second, we assume that projects are correlated, which is a more realistic assumption than considering disjointed projects, as microcredit borrowers often live close to each other and are exposed to the same bad luck or negative shocks. We prove that if projects are correlated, larger groups should be formed, as group members have higher incentives to help each other towards success.

5.2 Directions for Further Research

The studies in this dissertation have their limitations and could be extended in various ways. The literature on the importance of loss aversion and reference-dependent preferences is still in its early stage, and many relevant questions remain.

In our stylized models in Chapters 2 and Chapter 3, we make some assumptions that are not very realistic, but make the models tractable and the analysis straightforward. This, as first step, helped us to focus on the effect of loss aversion and reference-dependent preferences on the existence of equilibria, and in particular, separating equilibria. For example, we assume in Chapter 2 that all consumers are identical in terms of the weight they put on a gain-loss function or in their degree of loss aversion. One possible extension of our monopoly pricing model is to look at the case in which consumers are not identical. Also, in Chapter 3, we assume senders are of binary types while in reality, there are many types of senders in a job market. An immediate possible extension of our study in Chapter 3 is to consider senders of multiple types or even a continuum of types. We expect to see similar results after extending this model to multiple types or a continuum of types.

A large literature emphasizes that, “how a reference point is formed” greatly affects equilibria (Tversky and Kahneman, 1981; Pope and Schweitzer, 2011). As we show in these two chapters, and as also proved by Carbajal and Ely (2016), “when a reference point is formed”, is also an influential factor in determining equilibria. This subject is almost neglected by the existing literature and deserves more attention from researchers. Testing the validity of the results experimentally is another direction for further research. Ultimately, we are seeking theories that explain the complex reality of life, and experimental economics has proved its ability to put a higher standard on our theories by questioning their actualities. Another possible extension of our studies in Chapter 2 and Chapter 3 is to examine each of them experimentally.

An interesting extension of our lending model in Chapter 4 would be to take behavioral biases into account. For example, larger groups may have an additional attraction when borrowers are loss averse. Intuitively, the insurance provided by group members on each member's payoff would then be evaluated additionally positive by loss-averse borrowers. A group of borrowers from the same village usually are not complete strangers and have ties to each other. The question of whether and how family ties and friendships of group members affect the outcome of microlending is another interesting direction for further research.

Bibliography

Abbink, K., B. Irlenbusch, and E. Renner (2006), Group size and social ties in microfinance institutions, *Economic Inquiry*, *44*(4), 614–628.

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011), Reference points and effort provision, *The American Economic Review*, *101*(2), 470–492.

Ahlin, C. (2015), The role of group size in group lending, *Journal of Development Economics*, *115*, 140–155.

Ahlin, C., and R. M. Townsend (2007), Selection into and across credit contracts: theory and field research, *Journal of Econometrics*, *136*(2), 665–698.

Akerlof, G. A. (1970), The market for "lemons": Quality uncertainty and the market mechanism, *The Quarterly Journal of Economics*, *84*(3), 488–500.

Armantier, O., and A. Boly (2012), Framing of incentives and effort provision, *Working Paper*.

Armendáriz de Aghion, B. (1999), On the design of a credit agreement with peer monitoring, *Journal of Development Economics*, *60*(1), 79–104.

Armendáriz de Aghion, B., and C. Gollier (2000), Peer group formation in an adverse selection model, *The Economic Journal*, *110*(465), 632–643.

Armstrong, M., and S. Huck (2010), Behavioral economics as applied to firms: a primer, *Competition Policy International*, *6*(1), 2–45.

Armstrong, M., and J. Vickers (2001), Competitive price discrimination, *RAND Journal of Economics*, *32*(4), 579–605.

Ausubel, L. M., P. Cramton, and R. J. Deneckere (2002), Bargaining with incomplete information, *Handbook of Game Theory with Economic Applications*, *3*, 1897–1945.

Bagwell, L. S., and B. D. Bernheim (1996), Veblen effects in a theory of conspicuous consumption, *The American Economic Review*, *86*(3), 349–373.

Baker, M., B. Mendel, and J. Wurgler (2016), Dividends as reference points: A behavioral signaling approach, *Review of Financial Studies*, *29*(3), 697–738.

Baland, J.-M., R. Somanathan, and Z. Wahhaj (2013), Repayment incentives and the distribution of gains from group lending, *Journal of Development Economics*, *105*, 131–139.

Banerjee, A. V. (2013), Microcredit under the microscope: what have we learned in the past two decades, and what do we need to know?, *Annual Review of Economics*, *5*(1), 487–519.

Bell, D. E. (1985), Disappointment in decision making under uncertainty, *Operations Research*, *33*(1), 1–27.

Berg, I. (1970), Education for jobs; the great training robbery., *ERIC*.

Berge, L. I. O., K. Bjorvatn, and B. Tungodden (2014), Human and financial capital for microenterprise development: Evidence from a field and lab experiment, *Management Science*, *61*(4), 707–722.

Berle Jr, A. A., and G. C. Means (1930), Corporations and the public investor, *The American Economic Review*, *20*(1), 54–71.

Bernheim, B. D. (1994), A theory of conformity, *Journal of Political Economy*, *102*(5), 841–877.

Besley, T., and S. Coate (1995), Group lending, repayment incentives and social collateral, *Journal of Development Economics*, *46*(1), 1–18.

Bewley, T. F. (1998), Why not cut pay?, *European Economic Review*, *42*(3), 459–490.

Bhole, B., and S. Ogden (2010), Group lending and individual lending with strategic default, *Journal of Development Economics*, *91*(2), 348–363.

Blinder, A., E. R. Canetti, D. E. Lebow, and J. B. Rudd (1998), *Asking about prices: a new approach to understanding price stickiness*, Russell Sage Foundation.

Bolton, P., and M. Dewatripont (2005), *Contract theory*, MIT press.

Bourjade, S., and I. Schindele (2012), Group lending with endogenous group size, *Economics Letters*, *117*(3), 556–560.

Bradbury, J. W., and L. Vehrencamp (1998), *Principles of animal communication*, Sinauer Associates Sunderland.

Cai, H., J. Riley, and L. Ye (2007), Reserve price signaling, *Journal of Economic Theory*, *135*(1), 253–268.

Carbajal, J. C., and J. C. Ely (2016), A model of price discrimination under loss aversion and state-contingent reference points, *Theoretical Economics*, *11*(1555-7561), 455–485.

Chu, M. (2015), Edyficar and mibanco: The emergence of m&a in microfinance, *Harvard Business School Case 315-030*.

Chu, M., and G. A. Herrero (2009), Mibanco: Meeting the mainstreaming of microfinance, *Harvard Business School Case 309-095*.

Cohen, A. (2005), Asymmetric information and learning: Evidence from the automobile insurance market, *Review of Economics and Statistics*, *87*(2), 197–207.

Conning, J. (2004), Monitoring by delegates or by peers? joint liability loans under moral hazard, *Hunter College, Working Paper*.

Crawford, V. P., and J. Meng (2011), New york city cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income, *The American Economic Review*, *101*(5), 1912–1932.

De Mel, S., D. McKenzie, and C. Woodruff (2008), Returns to capital in microenterprises: evidence from a field experiment, *The Quarterly Journal of Economics*, *123*(4), 1329–1372.

Dixit, A. K., and J. E. Stiglitz (1977), Monopolistic competition and optimum product diversity, *The American Economic Review*, *67*(3), 297–308.

Ellison, G. (2006), Bounded rationality in industrial organization, *Econometric Society Monographs*, *42*, 142.

Galak, J., D. Small, and A. T. Stephen (2011), Microfinance decision making: A field study of prosocial lending, *Journal of Marketing Research*, *48*(Special Issue), 130–137.

Genesove, D., and C. Mayer (2001), Loss aversion and seller behavior: Evidence from the housing market.

Ghatak, M. (2000), Screening by the company you keep: Joint liability lending and the peer selection effect, *The Economic Journal*, *110*(465), 601–631.

Ghatak, M., and T. W. Guinnane (1999), The economics of lending with joint liability: theory and practice, *Journal of Development Economics*, *60*(1), 195–228.

Gill, D., and V. Prowse (2012), A structural analysis of disappointment aversion in a real effort competition, *The American Economic Review*, *102*(1), 469–503.

Glazer, A., and K. A. Konrad (1996), A signaling explanation for charity, *The American Economic Review*, *86*(4), 1019–1028.

Gneezy, A., U. Gneezy, and D. O. Lauga (2014), A reference-dependent model of the price-quality heuristic, *Journal of Marketing Research*, *51*(2), 153–164.

Grafen, A. (1990), Biological signals as handicaps, *Journal of Theoretical Biology*, *144*(4), 517–546.

Guinnane, T. W. (2001), Cooperatives as information machines: German rural credit cooperatives, 1883–1914, *The Journal of Economic History*, *61*(2), 366–389.

Gul, F. (1991), A theory of disappointment aversion, *Econometrica*, *59*(3), 667–686.

Hahn, J.-H., J. Kim, S.-H. Kim, and J. Lee (2014), Price discrimination with loss averse consumers, *Seoul National University, Working Paper*.

Heidhues, P., and B. Kőszegi (2004), The impact of consumer loss aversion on pricing, *Wissenschaftszentrum Berlin für Sozialforschung (WZB), Working Paper No. SP II, 17*.

Heidhues, P., and B. Kőszegi (2008), Competition and price variation when consumers are loss averse, *The American Economic Review*, *98*(4), 1245–1268.

Herweg, F., and K. Mierendorff (2013), Uncertain demand, consumer loss aversion, and flat-rate tariffs, *Journal of the European Economic Association*, *11*(2), 399–432.

Herweg, F., D. Müller, and P. Weinschenk (2010), Binary payment schemes: Moral hazard and loss aversion, *The American Economic Review*, *100*(5), 2451–2477.

Kahneman, D., and A. Tversky (1979), Prospect theory: An analysis of decision under risk, *Econometrica*, *47*(2), 263–291.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1991), Anomalies: The endowment effect, loss aversion, and status quo bias, *The Journal of Economic Perspectives*, *5*(1), 193–206.

Karle, H. (2013), Creating attachment through advertising: Loss aversion and pre-purchase information, *CER-ETH–Center of Economic Research at ETH Zurich, Working Paper No. 13/177*.

Katzur, T., and R. Lensink (2012), Group lending with correlated project outcomes, *Economics Letters*, *117*(2), 445–447.

Kim, S.-H., and J. Lee (2014), Firm behavior under consumer loss aversion, *Seoul Journal of Economics*, *27*, 171–186.

Kőszegi, B. (2010), Utility from anticipation and personal equilibrium, *Economic Theory*, *44*(3), 415–444.

Kőszegi, B. (2014), Behavioral contract theory, *Journal of Economic Literature*, *52*(4), 1075–1118.

Kőszegi, B., and M. Rabin (2006), A model of reference-dependent preferences, *The Quarterly Journal of Economics*, *121*(4), 1133–1165.

Kőszegi, B., and M. Rabin (2007), Reference-dependent risk attitudes, *The American Economic Review*, *97*(4), 1047–1073.

Kőszegi, B., and M. Rabin (2009), Reference-dependent consumption plans, *The American Economic Review*, *99*(3), 909–936.

Kreps, D. M., and R. Wilson (1982), Reputation and imperfect information, *Journal of Economic Theory*, *27*(2), 253–279.

Lacalle-Calderón, M., and S. Rico-Garrido (2006), Los microcréditos: una oportunidad de inclusión laboral, *Finance and the Common Good*, *25*, 94–100.

Laffont, J.-J. (2003), Collusion and group lending with adverse selection, *Journal of Development Economics*, *70*(2), 329–348.

Leland, H. E., and D. H. Pyle (1977), Informational asymmetries, financial structure, and financial intermediation, *The journal of Finance*, *32*(2), 371–387.

Littlefield, E., J. Morduch, and S. Hashemi (2003), Is microfinance an effective strategy to reach the millennium development goals?, *Focus Note*, *24*(2003), 1–11.

Loomes, G., and R. Sugden (1986), Disappointment and dynamic consistency in choice under uncertainty, *The Review of Economic Studies*, *53*(2), 271–282.

Lopes, L. L., and G. C. Oden (1999), The role of aspiration level in risky choice: A comparison of cumulative prospect theory and sp/a theory, *Journal of Mathematical Psychology*, *43*(2), 286–313.

Mahmood, M., and M. Hui (2007), Kashf foundation: A pakistani microfinance organization gears up for dramatic growth, *INSEAD Case 807-048-9*.

Marketing-News (1985), Panelists offer pricing strategy advice for consumer and industrial products, *Marketing News*, *19*(3).

Marr, A. (2003), A challenge to the orthodoxy concerning microfinance and poverty reduction, *Journal of Microfinance/ESR Review*, *5*(2), 7–42.

Mas, A. (2006), Pay, reference points, and police performance, *Quarterly Journal of Economics*, *121*(3), 783–821.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995), *Microeconomic Theory*, vol. 1, Oxford University Press.

Milana, C., and A. Ashta (2012), Developing microfinance: A survey of the literature, *Strategic Change*, *21*(7-8), 299–330.

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

Narayanan, V., and V. Kasturi Rangan (2010), Equitas microfinance: The fastest-growing mfi on the planet, *Harvard Business School Case 510-104*.

Odean, T. (1998), Are investors reluctant to realize their losses?, *The Journal of finance*, *53*(5), 1775–1798.

Piccione, M., and R. Spiegler (2012), Price competition under limited comparability, *The Quarterly Journal of Economics*, *127*(1), 97–135.

Pope, D. G., and M. E. Schweitzer (2011), Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes, *The American Economic Review*, *101*(1), 129–157.

Prior, F., and A. Argandoña (2009), Credit accessibility and corporate social responsibility in financial institutions: the case of microfinance, *Business Ethics: A European Review*, *18*(4), 349–363.

Puto, C. P. (1987), The framing of buying decisions, *Journal of Consumer Research*, *14*(3), 301–315.

Quelch, J. A., and N. Laidler (2003), Accion international, *Harvard Business School Case 304095*.

Riley, J. G. (1979), Informational equilibrium, *Econometrica*, *47*(2), 331–359.

Riley, J. G. (2001), Silver signals: Twenty-five years of screening and signaling, *Journal of Economic literature*, *39*(2), 432–478.

Rothschild, M., and J. Stiglitz (1976), Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, *The Quarterly Journal of Economics*, *90*(4), 629–649.

Salant, Y., and R. Siegel (2013), Contracts with framing, *Northwestern University, Working Paper*.

Samuelson, W. (1984), Bargaining under asymmetric information, *Econometrica*, *52*(4), 995–1005.

Samuelson, W., and R. Zeckhauser (1988), Status quo bias in decision making, *Journal of Risk and Uncertainty*, *1*(1), 7–59.

Scully, N. D. (2004), Microcredit no panacea for poor women, *Global Development Research Centre, Washington, DC*.

Shapiro, D., and A. Zhuang (2015), Dividends as a signaling device and the disappearing dividend puzzle, *Journal of Economics and Business*, *79*, 62–81.

Spence, M. (1973), Job market signaling, *The Quarterly Journal of Economics*, *87*(3), 355–374.

Spence, M. (1977), Nonlinear prices and welfare, *Journal of Public Economics*, *8*(1), 1–18.

Spence, M. (1980), Multi-product quantity-dependent prices and profitability constraints, *The Review of Economic Studies*, *47*(5), 821–841.

Spiegler, R. (2011), *Bounded rationality and industrial organization*, Oxford University Press.

Spiegler, R. (2014), Competitive framing, *American Economic Journal: Microeconomics*, *6*(3), 35–58.

Stiglitz, J. E. (1975), The theory of "screening" education, and the distribution of income, *The American Economic Review*, *65*(3), 283–300.

Stiglitz, J. E. (1990), Peer monitoring and credit markets, *The World Bank Economic Review*, *4*(3), 351–366.

Stiglitz, J. E., and A. Weiss (1981), Credit rationing in markets with imperfect information, *The American Economic Review*, *71*(3), 393–410.

Stuart, G., R. Campbell, and A. Saich (2008), Ngo microfinance in the tibet autonomous region, *Kennedy School of Government Case Program CR16-08-1913.0, Harvard University*.

Tedeschi, G. A. (2006), Here today, gone tomorrow: Can dynamic incentives make microfinance more flexible?, *Journal of Development Economics*, *80*(1), 84–105.

Tversky, A., and D. Kahneman (1981), The framing of decisions and the psychology of choice, *Science*, *211*(4481), 453–458.

Tversky, A., and D. Kahneman (1992), Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and uncertainty*, *5*(4), 297–323.

Von Neumann, J., and O. Morgenstern (1944), *Theory of games and economic behavior*, Princeton university press.

Zahavi, A. (1975), Mate selection - a selection for a handicap, *Journal of Theoretical Biology*, *53*(1), 205–214.

Nederlandse Samenvatting

Drie hoofdstukken van dit proefschrift, nl. hoofdstukken 2, 3 en 4, zijn geschreven als afzonderlijke papers. Hieronder volgt een korte samenvatting van elk hoofdstuk.

Hoofdstuk 2 gaat in op de optimale prijsstrategie van een monopolist die te maken heeft met consumenten die heterogene waarderingen hebben, die refentie-afhankelijke voorkeuren hebben, en die onderhevig zijn aan verliesaversie. Er is informatie-asymmetrie, doordat de monopolist de waarderingen van de consument niet observeert. Met de aannames dat de monopolist bij de consument verwachtingen kan induceren om de gewenste variatie van het product te kopen, en dat deze verwachtingen bepalend zijn voor het referentiepunt van de klant, komen we tot de volgende twee resultaten voor een model met verlies-aversie gebaseerd op verwachtingen. Ten eerste is wordt menu prijsstelling mogelijk zelfs wanneer er sprake is van een overtreding van de single-crossing voorwaarde (consumenten met een hoge waardering ervaren niet persé een hoger marginaal nut van kwaliteit dan consumenten met een lagere waardering). Ten tweede kan menu prijsstelling wenselijker worden, in vergelijking met het enkel verkopen aan consumenten met een hoge waardering.

In hoofdstuk 3 wordt een signaalmodel ontwikkeld met afzenders die referentie-afhankelijke voorkeuren hebben en onderhevig zijn aan verliesaversie. Zoals algemeen bekend is de single-crossing voorwaarde, die zegt dat hogere types lagere marginale signaalkosten hebben, een noodzakelijke voorwaarde voor het bestaan van scheidende evenwichten in signaalspellen. Deze voorwaarde beperkt echter het toepassingsgebied van het signaalmodel. In dit paper laten we zien dat, met de aanname dat afzenders onderhevig zijn aan verlies-aversie gebaseerd op verwachtingen, de single-crossing voorwaarde niet langer noodzakelijk is voor het bestaan van scheidende evenwichten.

In hoofdstuk 4 ontwikkelen we een model van herhaalde microfinanciering om te onderzoeken hoe groepsgrootte van invloed is op optimale groepsleencontracten met gemeenschappelijke aansprakelijkheid. In de bestudeerde situatie verzorgt een welwillende uitlener microkrediet aan een groep leners om in projecten te investeren. De uitkomst van elk risicovol project is niet zichtbaar voor de uitlener. Daardoor kan de uitlener niet achterhalen of er een strategische reden is wanneer een groep leners verzuimt te betalen. De groep leners heeft recht op een nieuwe lening wanneer de vorige lening is terugbetaald. We karakteriseren het optimale contract en bepalen de optimale groepsgrootte van de leners endogeen. Uit dit model blijkt dat, ondanks het feit dat gemeenschappelijke aansprakelijkheidscontracten alleen uitvoerbaar zijn voor een kleinere set parameters dan individuele contracten, gemeenschappelijke aansprakelijkheidscontracten een positief effect hebben op het welzijn van en terugbetaling door de leners. Daarnaast tenderen onze resultaten naar een toename van de groepsgrootte wanneer het projectrisico toeneemt. Tot slot analyseren we in dit hoofdstuk het effect van minder strenge sancties en projectcorrelatie op de uitvoerbaarheid en karakteristieken van gemeenschappelijke verantwoordelijkheidscontracten. Uit onze resultaten blijkt dat minder strenge sancties geen effect hebben op de mate van terugbetaling of het welzijn, maar het leenplafond van gemeenschappelijke aansprakelijkheid omlaag brengt. Echter, dit negatieve effect kan teniet gedaan worden door grotere groepen te vormen. Uit onze resultaten blijkt ook dat projectcorrelatie een hoger leenplafond toestaat in grotere groepen.

Curriculum Vitae

Bahar (Najmeh) Rezaei was born in Tehran, where she completed her B.Sc. and M.Sc. degrees in Applied Mathematics. After graduation, she worked as junior university lecturer of mathematics and statistics in Tehran and also as researcher in computer science related projects in Germany. In January 2011, Bahar undertook a PhD study at the chair of Theoretical Microeconomics at Utrecht University School of Economics (USE), where she completed this dissertation. During her PhD, she spent two years at University of California in Los Angeles (UCLA) as a visiting graduate researcher.

TKI Dissertation Serries

USE 001 **Bastian Westbrock** (2010): Inter-firm Networks: Economic and Sociological Perspectives.
USE 002 **Yi Zhang** (2011): Institutions and International Investments: Evidence from China and Other Emerging Markets.
USE 003 **Ryan van Lamoen** (2011): The Relationship between Competition and Innovation: Measuring Innovation and Causality.
USE 004 **Martijn Dröes** (2011): House Price Uncertainty in the Dutch Owner-Occupied Housing Market.
USE 005 **Thomas van Huizen** (2012): Behavioural Assumptions in Labour Economics: Analysing Social Security Reforms and Labour Market Transitions.
USE 006 **Martijn Boermans** (2012): International Entrepreneurship and Enterprise Development.
USE 007 **Joras Ferwerda** (2012): The Multidisciplinary Economics of Money Laundering.
USE 008 **Federico D'Onofrio** (2013): Knowing to Transform: Three ways for agricultural economists to observe Italy 1900-1940.
USE 009 **Saraï Sapulete** (2013): Works Council Effectiveness: Determinants and Outcomes.
USE 010 **Britta Hoyer** (2013): Network Formation under the Threat of Disruption.
USE 011 **Coen Rigtering** (2013): Entrepreneurial Orientation: Multilevel Analysis and Consequences.
USE 012 **Beate Cesinger** (2013): Context and Complexity of International Entrepreneurship as a Field of Research.
USE 013 **Jan de Dreu** (2013): Empirical Essays on the Governance of Financial Institutions.
USE 014 **Lu Zhang** (2013): Industrial Specialization: Determinants, Processes and Consequences.
USE 015 **Matthias Filser** (2013): Strategic Issues in Entrepreneurship and Family Business Research.
USE 016 **Mikko Pohjola** (2013): A Compilation of Studies on Innovation in Firms: Capabilities, Strategies, and Performance.
USE 017 **Han-Hsin Chang** (2013): Heterogeneity in Development.
USE 018 **Suzanne Heijnen** (2014): Analyses of Sickness Absence.

USE 019 **Mark Kattenberg** (2014): The Economics of Social Housing: Implications for Welfare, Consumption, and Labor Market Composition.
USE 020 **Daniel Possenriede** (2014): The Economics of Temporal and Locational Flexibility of Work.
USE 021 **Dirk Gerritsen** (2014): The Relevance of Security Analyst Opinions for Investment Decisions.
USE 022 **Shiwei Hu** (2014): Development in China and Africa.
USE 023 **Saara Tamminen** (2014): Heterogeneous Firms, Mark-Ups, and Income Inequality.
USE 024 **Marcel van den Berg** (2014): Does Internationalization Foster Firm Performance?
USE 025 **Emre Akgündüz** (2014): Analyzing Maternal Employment and Child Care Quality.
USE 026 **Jasper Lukkezen** (2014): From Debt Crisis to Sovereign Risk.
USE 027 **Vesile Kutlu** (2015): Essays on Subjective Survival Probabilities, Consumption, and Retirement Decisions.
USE 028 **Brigitte Crooijmans** (2015): Leiden Fusies tot Efficiëntere Woningcorporaties? Een Exploratieve Studie naar Schaalvoordelen in de Sociale Huisvesting.
USE 029 **Andrej Svorenčík** (2015): The Experimental Turn in Economics: A History of Experimental Economics.
USE 030 **Seçil Hülya Danakol** (2015): Foreign Direct Investment, Foreign Aid and Domestic Entrepreneurship.
USE 031 **Ioana Deleanu** (2015): Anti-Money Laundering Efforts: Failures, Fixes and the Future.
USE 032 **Jaap Oude Mulders** (2016): Organizations, managers, and the employment of older workers after retirement.
USE 033 **Malka de Castro Campos** (2016): Private Consumption-Savings Behavior and Macroeconomic Imbalances.
USE 034 **Tahereh Rezai Khavas** (2016): Fairness concerns and cooperation in context.
USE 035 **Joyce Delnoij** (2016): Auctions with competing sellers and behavioral bidders.
USE 036 **Krista Bruns** (2017): Emergence and Diffusion of Institutions and their Effect on Economic Growth.
USE 037 **Daan van der Linde** (2017): Democracies under Rising Inequality: New Tests of the Redistributive Thesis.
USE 038 **Swantje Falcke** (2017): On the move: Analyzing immigration determinants and immigrant outcomes.
USE 039 **Joep Steegmans** (2017): House Prices and Household Mobility in The Netherlands: Empirical Analyses of Financial Characteristics of the Household.
USE 040 **Najmeh Rezaei Khavas** (2017): Essays in Information Economics.