The costs of commitment: present-bias, consumer protection, and firm ownership forms

Karna Basu & Jonathan Conning*

December 22, 2019

Abstract

Hyperbolic discounters value consumption-smoothing commitment contracts for saving or borrowing, but these may be renegotiated by future selves and banks. This generates a consumer protection problem even for informed and sophisticated consumers. We study how such renegotiation threats shrink enforceable contracts and distorts equilibrium contract terms (monopolists, for example, must lend more). Costly strategic ownership/governance choices (e.g. mutual forms) can expand contracting by constraining banks ability to profit from renegotiation. The model explains contracts, ownership forms and market structure patterns in consumer banking and microfinance, and helps frame policy debates over "excessive refinancing" and "overindebtedness" in these sectors. JEL Codes: O16, D03, D18

*Department of Economics, Hunter College & The Graduate Center, City University of New York. Email: kbasu@hunter.cuny.edu, jconning@hunter.cuny.edu. We are grateful to the Roosevelt House Public Policy Institute and a Small Business Administration research grant for support. For detailed comments and suggestions, we thank Temisan Agbeyegbe, Abhijit Banerjee, Partha Deb, Maitreesh Ghatak, Alexander Karaivanov, Igor Livshits, and Eric Van Tassel; conference participants at the NBER Development Economics Summer Institute, 'Economics of Social Sector Organizations' (Chicago Booth), CIDE-ThReD Conference on Development Theory, ACEGD (ISI Delhi) and NEUDC (Harvard Kennedy School); and seminar participants at Michigan State University, IIES Stockholm, Delhi School of Economics, Queens College, Indian School of Business, Hunter College, Florida Atlantic University, Pontificia Universidad Javeriana, and The Graduate Center CUNY. Kwabena Donkor provided excellent early research assistance. Replication files for figures and other supplemental materials at: `https://github.com/jhconning/commitment`.

1 Introduction

When do financial intermediaries provide the commitment services to help present-biased consumers stick to long-term savings accumulation and/or debt management plans? When instead might they try to opportunistically profit by pandering to those same consumer biases? Hyperbolic discounters – consumers with present-biased and dynamically inconsistent preferences – struggle to stick to long-term plans. **?** was first to formalize the idea that sophisticated hyperbolic consumers – those who correctly understand how their own changing preferences would lead future selves to try to undo earlier laid consumption plans – might demand and benefit from contracts and other commitment devices to constrain their future choices.[1] Features of financial arrangements from automatic payroll deduction savings plans to fixed payment amortization mortgages as well as the high-frequency repayment and joint-liability provisions of microcredit loans have all been interpreted as commitment mechanisms that purposefully make making changes to long-term contract planning costly. Randomized controlled trial field experiments in both developed and developing countries have demonstrated possible positive demand take-up and asset accumulation impacts to the introduction of new financial commitment products in several different contexts[2] although this same evidence also highlights the apparent puzzle of why, if such interventions offer such apparent benefits, they weren't offered more widely by the market in the first place.

The *demand* for commitment services is one thing, but contracting for its *supply* can be difficult and costly. In particular, why should a financial intermediary's own promise to help the consumer remain committed to the terms of a contract be believed and not itself be renegotiated? The bank understands that the consumer who demands commitment contracts in one period may, in later periods and with new preferences, be willingly to pay to refinance or renegotiate its terms – with the same bank or a new one. Pandering to such changing preferences may increase bank profits and most courts would judge such renegotiations as voluntary and legal.[3] Credible commitment contracts must therefore "tie-the-hands" of both the intermediary as well as the consumer they serve.

The widespread perception that failures of commitment (or opportunistic contract renegotiation) can distort or spoil the operation of markets is implied by the loaded language often used to describe many financial behaviors in both academic and popular press writings. Commentators for example often speak of 'overindebtedness' in the market for microcredit, of 'debt traps' and 'excessive' debt rollovers in payday lending, and of behaviors such as

[1]For related reasons, naive consumers who underestimate how their future preferences will change, may also be advantaged by certain public regulations or certain forms of private paternalism of organizations that constrain individuals actions (**?**).

[2]See for example **?**, **?**, **?**, **?**. **?** survey this literature.

[3]In most countries including the United States courts will not penalize voluntary renegotiation, on the principle that there is no injured promisee (see discussion in **?**, p448)

the 'raiding of equity' and 'excessive refinance' in hommortgage markets. Blame for such perceived problems is variously placed on the consumer and/or the financial intermediary: on consumers for supposedly being present biased and exhibiting weak self-control and on financial intermediaries for failing to restrain their impulse to profit from pandering to those same biases in possibly deceptive and socially destructive ways. Such mis-behavior by financial intermediaries is in turn often attributed to 'failures of governance' and/or the failure of regulation to offer sufficient consumer protections. Although many such behavioral analyses are framed in terms of failure to protect 'naive' hyperbolic discounters who would fail to understand how their own changing future preferences leave them vulnerable to exploitation, similar issues may also affect sophisticates, as we shall see.

In this paper, we take seriously the threat of opportunistic contract renegotiations that cannot be completely deterred via exogenous mechanisms. This leads us to study how endogenous commitment mechanisms, implemented via costly distortions to contract terms and/or the firm's own costly surrender of rights, may adapt to and complement available imperfect exogenous commitment mechanisms. This offers a lens for understanding when firms do or do not have the incentive to provide 'market consumer protection' to substitute for or replace weak or missing government consumer protection. We show how firms may offer protections against opportunistic pandering to present-biased consumers via costly strategic ownership and capital structure decisions in ways that may, and be shaped by, the character and structure of the consumer finance sector.

These contract renegotiation considerations are analyzed within a simple canonical consumption smoothing contract design problem to derive a parameterized spectrum of endogenous commitment contract forms and firm ownership structures for both naive and sophisticated consumers and under conditions of competition or monopoly. The relative parsimony of this framework allows us to provide analytical clarification of some general mechanisms and contract design features that may have been missed or partly obscured in prior work, as well as derive some known and new results.

To see what is gained, consider the classic **?** study of how savings placed into illiquid lower return investments might serve as a costly strategy for sophisticated hyperbolic discounters to constrain their future selves' ability to raid saved resources. The credibility of such commitment depends on two interrelated factors. First, it should be costly to undo the plan (in Laibson's model liquidation is costly to the consumer because illiquid assets can only be accessed with a delay) and second and more subtly, the consumption-savings path must also be sufficiently adjusted to accommodate the consumer's future selves' preferences that undoing the plan does not provide them such gains as to make them want to incur those liquidation costs. This second endogenous enforcement element is required because the exogenous enforcement mechanism (the liquidation penalty that the consumer and/or the bank would incur withdrawing illiquid savings periods) is not sufficient (or too costly) by itself to deter deviations from the consumer's earlier period self's desired plans. Achieving 'self-control' is hence costly to the consumer and/or the bank as it requires distortion to the first-best smoothing contract to remain self-enforcing. If a perfect unbreakable commitment technology were available then the first-best smoothing commitment contract could be achieved at no cost to the parties, but our focus is on understanding the more interesting and general case where such external commitment technologies are limited and therefore commitments are potentially breakable.

Implicit in the setup is a tradeoff between the cost of breaking commitment and the need to accommodate future selves' preferences. Unfortunately, the relatively elaborate way the illiquid savings commitment technology is modeled in Laibson's 1997 paper make these tradeoffs make the equilibrium contracts difficult to characterize without resorting to numerical simulation or imposing restrictive parameter assumptions. By working with a canonical consumption smoothing model with a simply parameterized 'renegotiation cost' technology we can study savings as well as borrowing behaviors and more clearly pinpoint these tradeoffs.

In contrast to much of the literature on commitment, we take seriously the fact that after a contract is signed, future action sets include the rewriting of contracts. Any credible contract must satisfy 'no-renegotiation' constraint(s) that bind *all* parties (the consumer and a bank). We derive and compare equilibrium contract terms (and later in the paper, firm endogenous ownership/governance structures) under the assumption of competition, monopoly or bargaining settings using analytical and graphical methods.[4]

We start by exploring the feasibility of first-best 'full-smoothing commitment contract' and how this depends on market structure and a sufficiently high external 'cost of renegotiation'.[5] Next, we establish a number of non-obvious properties of the second-best 'imperfect-smoothing commitment contract' that will be implemented when the first-best cannot be credibly sustained. This in turn leads us to show how firms may choose to make changes to firm ownership and capital structure as a costly strategy to provide endogenous consumer protection to expand captured profits and trade with sophisticated present-biased consumers. We show how these choices depend crucially on market structure. This argument is similar to the theory of commercial non-profits based on asymmetric information due to **?** and formalized by **?** and others, but set on new behavioral micro-foundations and with no need for asymmetric information. Our paper is therefore a complement to the paper by

[4]In **?** the market for later-period contracts is competitive (new or existing banks can offer new contracts) so only the consumer's own renegotiation costs matter. We extend the analysis to also study situations where later-period contracts remain exclusive or monopolized.

[5]Under hyperbolic discounting, there is no obvious measure of welfare. We define 'full smooting' or 'first-best' as the contract that maximizes the discounted utility of the initial signatory.

? which also builds a model of endogenous firm ownership structure as a form of consumer protection, but the underlying behavioral stories are quite different. In their analysis firms hide non-contractible penalties in loan contracts and opportunistically charge such fees on a mix of suspecting (sophisticates) and unsuspecting (naive) risk-neutral customers. Their model is not so much a model of conflicting selves as a model of hidden penalties and it requires a population of exploitable naive borrowers in order to produce an inefficiency. Our model by contrast is built upon a contract-renegotiation problem due to time-inconsistent preferences for saving and/or borrowing contracts by risk-averse customers that survives even if there are only sophisticated customers.

Our framework moves beyond the simple 'for-profit/non-profit' dichotomy of some of the earlier commercial non-profit literature to explore a whole spectrum of 'hybrid' ownership firms (e.g. for-profit financial intermediaries variously owned and controlled by 'social' as well as private investors). This helps make sense of some of the ownership and capital structures observed historically in consumer banking intermediaries in the United States and other now developed countries as well as in microfinance in developing countries to this day. The sensitivity of equilibrium contracts to changes in market structure can also help make sense of recent episodes where rising 'commercialization' and intense competition appear to have been associated with periods of rising refinancing, multiple borrowing, indebtedness, and complaints of insufficient consumer protection. In some instances such episodes led to financial crash and political backlash, as in the case of the 2010 microfinance crisis of Andra Pradesh or the 2008 sub-prime loan financial crisis.

Analyses have typically divvied up blame between consumers for their supposed present bias and lack of self-control and at financial intermediaries for failing to protect their consumers from their own loan officers who aggressively pandered to those same consumer biases. Microfinance industry funded campaigns such as the 'Smart Campaign'[6] have now launched with slogans such as "[p]rotecting clients is not only the right thing to do; it's the smart thing to do." These campaigns seek to get financial intermediaries' to publicly adhere to consumer protection principles to prevent aggressive loan sales and protect against client 'overindebtedness.'

The paper makes clear how the above results depend crucially on consumer type (sophisticated or naive), market structure (monopoly or competition), and costs of renegotiation. While stylized and limited to one of many mechanisms the model makes a number of compelling points relevant to ongoing policy debates and is able to explain some stylized facts while generating sometimes counterintuitive empirical predictions. Our goal is to provide a novel streamlined framework that is useful for and compatible with extensions that introduce additional real-world factors.

[6] See http://smartcampaign.org/

1.1 Outline of arguments

We work with a quite general three-period consumption smoothing model for a present-biased consumer with quasi-hyperbolic preferences that allows for saving (repayment) or borrowing (dissaving) in each period. In each contracting scenario the consumer's period zero-self (henceforth 'Zero-self') has a bias for present consumption but wants to smooth future consumption across periods one and two. She correctly anticipates that her later period-one 'One-self' will have a change of preferences that will lead her to want to 'raid savings' and/or take on new debt to drive up period one consumption at the expense of period two consumption, thereby undoing Zero-self's early intent to balance consumption across the two periods. In every case the equilibrium contract will be the subgame perfect Nash equilibrium of a game where Zero-self chooses a contract first anticipating One-self's reactions, possibly limited by the Bank's exogenously or endogenously enforced commitment to agree to not renegotiate with One-self.

In Section 2 we describe a consumer who faces an income stream that, in the absence of a bank ('autarky'), can be rearranged to provide imperfect consumption smoothing at best. We describe banks that have access to funds at a competitive interest rate, and can offer the Zero-self consumer a 3-period contract. The extent to which contract terms can be enforced in future periods depends on some non-pecuniary renegotiation cost κ, which is borne by the bank and can be interpreted as a concern for the consumer's well-being or own reputation.

We then build a framework for analyzing equilibrium contracts as the outcome of a Stackelberg-type game where Zero-self moves first while anticipating One-self's best response. As an example, we derive the equilibrium contract when the banks have high renegotiation costs (so that contract terms are always respected). This yields the first-best contract from Zero-self's perspective–she is able to allocate consumption across periods 1 and 2 in accordance with her own preferences without conceding to One-self's present-bias. We label this contract the 'full-smoothing commitment contract,' though it should be clear that the 'full-smoothing' part is judged against Zero's preferences. We compare full-smoothing commitment contracts under competition and monopoly.

Section 3 formalizes the renegotiation problem. If κ is small, neither a monopolist nor competitive banks can credibly offer full-smoothing commitment contracts. This is because any sophisticated Zero-self consumer will understand that their future One-self and the bank stand to share gains from breaking any would-be commitment full-consumption smoothing contract. The only commitment contracts that will be considered credible and therefore capable of expanding the contract space are those that can satisfy a 'no-renegotiation' constraint that makes them self-enforcing.

We show that if renegotiation costs κ are below particular cutoff levels $\bar{\kappa}$, full-smoothing commitment contracts will not be offered in equilibrium. The cutoffs are more stringent under competition than under monopoly. This difference–the relatively greater feasibility of full-commitment under monopoly–is not due to any superior ability to commit on the part of a monopolist. It is due rather to the fact that monopolies offer less consumption than do competitive contracts to begin with. At lower levels of consumption, the potential period 1 gains from renegotiation are also lower, thus making renegotiation less profitable and hence commitment more feasible.

In Section 4, we derive contracts when the 'no-renegotiation' constraint binds. We first focus on sophisticated hyperbolic discounters. The Zero-self can enter into a multi-period contract that helps bind her One-self to contract terms only to the extent that the bank's commitment can be endogenously enforced (i.e. the bank's ex-post gain in profits from breaking their commitment must fall short of any direct renegotiation costs κ). The 'imperfect-smoothing commitment contract' represents a compromise between Zero-self's and One-self's preferences–consumption allocations between periods 1 and 2 must be tilted sufficiently in favor One self as to make renegotiation unprofitable but the extent of this tilt will be governed by the renegotiation cost. This reduces the potential gains to trade between consumers and banks, and contracts will result in lower bank profits (monopoly) or lower consumer discounted utility (competition).

We characterize the shapes of contracts.[7] All else equal, under monopoly imperfect-smoothing commitment contracts involve larger loans (or reduced savings) compared to full-smoothing. Under competition, the comparison is ambiguous. We explain this contrast between monopoly and competition using the intuition of income and substitution effects (from the consumer's perspective, a weakening of commitment has only substitution effects under monopoly while it has both substitution and income effects under competition).

Section 4 finally turns to naive hyperbolic discounters who fail to anticipate the extent to which contracts may be renegotiated. Now, Zero-self is offered a contract in which consumption in periods 1 and 2 strongly tilted towards period 2. This maximizes the potential gains from renegotiation. Under monopoly, this is achieved through a small loan (or high savings) since the consumer believes her future to be better than it will turn out to be. So, the naive consumer is not targeted with large loans; instead, she is offered a small teaser loan that will subsequently be rolled over in a manner that resembles some aspects of payday lending. Under competition, again initial loan/savings sizes are ambiguous since anticipated gains from renegotiation must be distributed back to the consumer.

[7] We are able to provide quite complete characterizations of optimal contracting scenarios under the assumption of monopoly or competition in the market for period-zero banking contracts with or without the assumption of enforceable exclusive contracts in later periods. We can provide exact closed-form solutions for contract terms for CRRA utility functions for most of these cases including renegotiation-proof contracts when $\kappa = 0$. For the $\kappa > 0$ cases where closed form solutions cannot be directly obtained we can nonetheless characterize some important contract properties and solve for contracts numerically.

In Section 5, banks may explore commercial nonprofit status as a mechanism to more credibly commit to not opportunistically exploiting the weaknesses of its sophisticated time-inconsistent clients. By operating as a nonprofit (or as a 'hybrid' bank), the bank agrees to face legal or governance restrictions on how any profits generated from any such opportunistic renegotiation can be distributed and enjoyed. The bank can now credibly convince the sophisticated consumer that it will be less likely to renegotiate the contract in the future. This allows the bank to offer the consumer an initial contract that maintains the restrictions on future consumption patterns that the consumer demands, raising the contracting surplus and therefore how much can be ultimately extracted by the bank's stakeholders.

A firm's decision about whether to adopt nonprofit status rests on a trade-off. As a non-profit, the firm has an opportunity to extract greater surplus from the consumer (by providing commitment), but now faces restrictions on the ability of managers and shareholders to enjoy this surplus. In the case of monopoly, the bank will adopt nonprofit status if the following is true: non-profit restrictions should be sufficiently severe that the bank is able to extract more surplus from the consumer, but should not be so severe that it is unable to enjoy the surplus. That nonprofit firms may survive even in the absence of motivated agents or asymmetric information is, to the best of our knowledge, a novel result.

This trade-off is also sensitive to market structure. Under competition, a lender's ability to provide effective commitment through non-profit status depends on the exclusivity of contracts. When long-term contracts can be made exclusive, the tradeoff disappears and all active firms function as non-profits. This is because of the zero-profit condition–since firms do not make profits anyway, there is nothing to lose from switching to non-profit status. On the other hand, there are profits to be gained–if all other firms are for-profit, a firm could make positive profits by offering superior commitment as a non-profit (this is valuable even if its enjoyment of these profits is limited).

When contracts are not exclusive, commitment generated through non-profit status becomes impossible to achieve. Since non-profit firms would make zero profits anyway, each firm has an incentive to switch to for-profit status so it can take advantage of the opportunity to re-finance *other* banks' loans. As a result, for-profit firms must be active in equilibrium, and their presence will eliminate the possibility of non-profit commitment.

This can partly explain a key difference between traditional monopolistic non-profit microfinance, which is rigid, and say competitive commercial credit card lending which offers refinancing flexibility (credit card punishments gain salience because they are *less* strict, not more).

1.2 Context and Related Literature

1.2.1 Commitment as a form of Consumer Protection

Concerns about excessive refinancing and 'over-indebtedness' have been raised especially in the lead up and wake of financial crises. On the eve of the mortgage banking crisis in 2007, over 70 percent of all new subprime mortgage loans were refinances of existing mortgages and approximately 84 percent of these were 'cash out' refinances (**?**). In the market for payday loans in the United States economists and regulatory observers express concern not so much that fees are high (the typical cost is 15% of the amount borrowed on a 2 week loan) but rather that 4 out of 5 payday loans are 'rolled over' or renewed rather than paid off resulting in very high total loan costs and placing many people into very difficult debt management situations (**?**).

Problems of consumer protection are typically analyzed through two channels: naive or uneducated consumers and their failure to correctly anticipate fees and punishments (see **?**, **?**, and **?** for related arguments), and bank's moral hazard (see **?** and **?**). We argue that, given the growing evidence of time-inconsistent preferences,[8] a bank's ability to provide credible commitment should also fall under this umbrella–sometimes consumers *want* punishments or fees to limit renegotiation.

In recent years, especially in light of crises in consumer credit markets, there has been renewed emphasis on consumer protection and better governance and regulation in banking.[9] One particular outcome of concern has been borrower over-indebtedness, an issue that has been at the center of recent microfinance repayment crises in places as far-flung as Morocco, Bosnia, Nicaragua and India, as well as the 2008 mortgage lending crisis in the United States. In each of these cases the issue of refinancing or the taking of loans from multiple lenders emerges.

Journalistic and scholarly analyses of such situations, including the recent mortgage crisis in the United States, have often framed the issues as problems of consumer protection, suggesting that many lenders designed products to purposefully take advantage of borrowers who have limited financial literacy skills and are naive about their self-control problems. Informed by such interpretations, new regulations introduced in the wake of these crises have swung toward restricting the terms of allowable contracts, for example by setting maximum interest rates and limiting the use of coercive loan recovery methods.

We place consumers' struggles with intertemporal self-control issues at the center of the

[8] See, for example, **?**, **?**, **?**, and **?**.

[9] In the US, the Consumer Financial Protection Bureau was set up in 2011 under the Dodd–Frank Wall Street Reform and Consumer Protection Act. In India, the far-reaching Micro Finance Institutions Development and Regulation Bill of 2012 was designed to increase government oversight of MFIs in response to the credit crisis in the state of Andhra Pradesh, and the perception that lax consumer protection and aggressive lending practices had led to rising over-indebtedness and stress.

analysis, but argue that borrowers may be more sophisticated in their understanding of their own time-inconsistency than is often assumed. From this perspective, 'predatory lending' is not primarily about tricking naive borrowers into paying more than they signed up for with hidden penalties or misleading interest rates quotes, but about offering excessive flexibility and refinancing of financial contracts in ways that limit or undermine the commitments to long term consumption and debt management paths that borrowers themselves may be attempting to put in place.[10]

Here, a bank that promises to be rigid and is then flexible could be seen as hurting, rather than helping, the consumer. We take seriously the bank's ex-post considerations and derive conditions under which it would renegotiate.

In this sense, our paper complements some others that demonstrate how commitment can be undone in related settings. **?** shows how competition leads to inefficient outcomes in immediate rewards goods. **?** study the mistakes of partially naive borrowers in competitive credit markets. **?** analyzes predatory lending with naive consumers.

1.2.2 Commercial Non-profits in finance

The idea that firm ownership might be strategically chosen to solve or ameliorate 'contract failure' problems dates back at least to **?** and is one that has been articulated most clearly in the work of Henry **?**. Hansmann argued that in markets where the quality of a product or service might be difficult to verify, clients may rationally fear that investor-led firms will be tempted to opportunistically skimp on the quality of a promised product or service, or reveal a hidden fee, and this can greatly reduce or even eliminate contracting. In such circumstances becoming a 'commercial non-profit' may be a costly but necessary way to commit the firm to not act opportunistically, hence enabling trade.

Hansmann gives as a primary historical example the development of consumer saving, lending and insurance products in the United States and Europe. Life insurance in the United States for example has until quite recently always been dominated by mutuals. Rate payers could not trust investor-led firms to not act opportunistically by, for example, increasing premiums or by skimping or reneging on death benefit payouts. Mutuals on the other hand had little incentive to cheat clients to increase shareholder dividends as the clients themselves are the only shareholders. Mutuals therefore enjoyed a distinct competitive advantage until sufficient state regulatory capacity developed.

In the present analysis we begin by following Hansmann in defining nonprofits by the legal restrictions faced by them, setting aside other ways (such as motivation) in which they might

[10]**?** discuss evidence of predatory lending in the context of mortgages. In 2016 the Consumer Financial Protection Bureau put forth a proposal to protect payday loan consumers including limits on the number and frequency of re-borrowings, available at `http://files.consumerfinance.gov/f/documents/CFPB_Proposes_Rule_End_Payday_Debt_Traps.pdf`.

be different from for-profit firms.[11] In this view "[a] nonprofit organization is, in essence, an organization that is barred from distributing its net earnings, if any, to individuals who exercise control over it, such as members, officers, directors, or trustees."[12] **?** have formalized Hansmann's central argument to show that when a firm cannot commit to maintaining high quality, it might choose to operate as a commercial nonprofit rather than as an investor-led for-profit in order to credibly signal that it has weaker incentives to cheat the consumer on aspects of unobserved product quality. As Hansmann describes it, firm ownership form adapts endogenously as a "crude form of consumer protection" in unregulated emerging markets where asymmetric information problems are rife. **?** modify this model so that the non-contractible quality issue is on hidden penalties, which are incurred with certainty by some borrowers. All of these models are built rely on some form of asymmetric information or contract verification problem.

A contribution of our paper is to argue that a theory of ownership form can be built on behavioral micro-foundations even in environments with no asymmetric information and with sophisticated forward-looking agents. We believe this is an important element for understanding the development of consumer finance in developed countries historically as well as the current shape of microfinance today where non-profit and 'hybrid' forms still dominate the sector in most developing countries (**??**). Hybrid ownership forms include the many microfinance firms that, though technically incorporated as for-profit financial service providers, are in fact dominated by boards where, by design, social investors or client representatives exert substantial governance control. Hybrid forms such as these would appear to confer many of the benefits of non-profit status (specifically, credible commitment to consumer protection) with fewer of the costs (in particular, unlike a pure non-profit they can and do issue stock to outside investors although usually in a manner that does not lead to challenge control).

1.2.3 Market Structure and Governance Choice

Commenting upon a major microfinance crisis in the state of Andhra Pradesh in India, veteran microfinance investor and market analyst Elizabeth **?** describes the build up of "rising debt stress among possibly tens of thousands of clients, brought on by explosive growth of microfinance organizations . . ." fueled by the rapid inflow of directed private lending and new equity investors who, because they "paid dearly for shares in [newly privatized] MFIs . . . needed fast growth to make their investments pay off ."

[11] Hence we abstract away from other considerations for nonprofits, as in **?**, **?**, and **?**. Nonetheless our modeling framework can be adapted to include these considerations.

[12] In practice, nonprofit firms also enjoy certain benefits that are denied to for-profit firms (see, for example, Cohen, 2015). But for the purposes of Hansmann's (and our) argument, it is the *restrictions*, not benefits, that generate improved outcomes.

She goes on to lay the blame on "poor governance frameworks" for behaviors that included "loan officers [that] often sell loans to clients already indebted to other organizations." In her view, Indian MFIs might have avoided their problems and followed the model of leading microfinance organizations in other countries like Mibanco (Peru) and Bancosol (Bolivia) which "were commercialized with a mix of owners including the original non-governmental organization (NGO), international social investors (including development banks), and some local shareholders. The NGOs kept the focus on the mission, while the international social investors contributed a commercial orientation, also tempered by social mission." These are the types of hybrid ownership forms, along with nonprofit firms, that we argue can provide surplus building consumer protection through a reduced incentive to renegotiate. Rhyne's argument is that a number of Indian state regulations made it difficult for such hybrid ownership forms to rise organically in India. As our model makes clear, these governance choices are highly dependent on market structure, and nonprofits may survive better under monopoly than under competition.

2 The model: setup

There are three periods, $t \in \{0, 1, 2\}$. In any period t, the consumer's instantaneous utility from consumption level c_t is given by a CRRA function defined over all non-negative consumption:

$$u(c_t) \equiv \frac{c_t^{1-\rho}}{1-\rho} \tag{1}$$

with some $\rho > 0$ as the coefficient of relative risk aversion.[13]

We model the consumer 'as a sequence of temporal selves ... indexed by their respective periods of control over the consumption decision **?**, p.451'. Given a consumption stream $C_t \equiv (c_t, .., c_2)$, the period-$t$ self's discounted utility is:

$$U_t(C_t) \equiv u(c_t) + \beta \sum_{i=t+1}^{2} \delta^{i-t} u(c_i) \tag{2}$$

This describes quasi-hyperbolic preferences, with a standard exponential discount factor $\delta \in (0, 1]$ and a hyperbolic discount factor $\beta \in (0, 1)$. In any period t, the individual (henceforth referred to as the "t-self") discounts the entire future stream of utilities by β. As a result, when faced with any tradeoff between consumption in periods t and $t+x$, the t-self places greater relative weight on period-t consumption than her earlier selves would have done. The consumer could be sophisticated (her time-inconsistency is common knowledge

[13] When $\rho = 1$ the function becomes $u(c_t) = ln(c_t)$.

across all t-selves) or naive (she believes her future selves to be exponential discounters with a discount factor of δ). (**?**).

The Zero-self begins with an endowment of claims to an arbitrary positive income stream over the three periods, $Y_0 \equiv (y_0, y_1, y_2)$. Her objective is to or rearrange this into a preferable consumption stream C_0 to maximize $U_0(C_0)$ in (2) using what financial contracting or other saving/borrowing strategies as may be available.

In the absence of access to the financing and commitment-services offered by a bank the consumer can only achieve an autarky consumption stream which delivers a corresponding autarky utility denoted U_0^A. The simplest assumption is that this autarky consumption stream corresponds to the endowment income stream. More realistically, the autarky consumption stream is what might be achieved via the more limited financing and commitment services available via informal banking or self-commitment strategies.

Section 2.1.1 describes the benchmark optimal consumption smoothing stream C_0^F and associated utility level U_0^F that Zero-self could achieve if she had perfect access to borrowing and/or saving at competitive interest rates with all the commitment required to make sure the contract is not renegotiated. There are many reasons why in practice autarky consumption plans might fall short of this optimum. For example, if the consumer's income is back-heavy, borrowing constraints might mean she must consume income as it arrives. If her income is front-heavy she may be able to construct a somewhat smoothed consumption stream but there may be technological restrictions to saving that place the return to savings well below the market rate – the insecurity of storing cash at home being one obvious explanation. More pivotal to our analysis, however, is that even with access to perfectly secure savings or borrowing, a consumer with time-inconsistent preferences cannot trust her later selves to follow her optimal consumption path. While remaining deliberately agnostic about autarky technologies, the rest of the paper focuses on the reasonable and interesting case $U_0^A < U_0^F$ where there are potential gains to financial contracting with a new intermediary.

The consumer will have the option to contract with one or many risk-neutral banks, depending on whether the period 0 market structure is monopolized or competitive. Each bank can access funds at interest rate opportunity cost r. At this market interest rate, the present value of the consumer's income stream is:

$$y \equiv \sum_{i=0}^{2} \frac{y_i}{(1+r)^{i-t}} \tag{3}$$

A period 0 financial contract allows the consumer to exchange income stream Y_0 for a new smoother consumption path C_0. A bank will participate if and only if it can expect to earn non-negative profits $\Pi_0(C_0; Y_0)$, where profits are defined as:

$$\Pi_t(C_t; Y_t) \equiv \sum_{i=t}^{2} \frac{(y_i - c_i)}{(1+r)^{i-t}} \tag{4}$$

The optimal contract may involve borrowing (dissaving) or savings (repaying debt) in period t depending on whether $(c_t > y_t)$ or $(c_t > y_t)$, respectively. We begin by assuming contracts can only be initiated in period 0.[14] Contracts may however be renegotiated by the consumer and the original bank or possibly a new one in period 1. If this happens, we assume the bank incurs a non-monetary cost, $\kappa \geq 0$. We could interpret this to include a concern for reputation or some other impact on the social preferences of its owners.[15]

2.1 Optimal commitment contracts

We first characterize optimal consumption-smoothing contracts when the consumer can perfectly and costlessly bind their latter selves to not renegotiate the term of the contracts with the same bank or other banks. We do this for the case of competition and monopoly. The credibility of a bank's own commitment to not renegotiate the contract with the consumer's future selves ultimately must rest on the assumption that the bank would face a credible deterrent penalty in the event of renegotiation. It will be informative to derive expressions for the minimum deterrent penalties required to sustain optimal consumption smoothing in different settings.

2.1.1 Optimal commitment contracts under competition

A consumer with time-inconsistent preferences cannot trust her latter selves to stick to her preferred consumption plans. In this three-period setting Zero-self's concern is that her later One-self will attempt to divert resources earmarked for period 2 consumption to boost period 1 consumption. Similar to a Stackelberg-leader in a Cournot game, Zero-self's strategic saving/borrowing choices are affected by her anticipation of One-self's and the banks' best response. The challenge will be to choose contract terms that keep the bank engaged as a credible strategic partner to Zero-self, delivering the (possibly imperfect) commitment services that help restrict or otherwise control the consumer's later selves' best responses.

If banks can be assumed to credibly and costlessly commit to never renegotiate[16] then

[14] This assumption is explored and lifted in Section 6.

[15] We discuss the source and nature of such costs in depth in section 5. The bank could incur monetary costs in addition to the non-pecuniary ones but we assume these to be 0 as they can be netted out and do not affect the analysis in any important way.

[16] For the moment this also means that contracts are 'exclusive' in the sense that no new bank can in later periods enter to 'buy out' or renegotiate a contract or, equivalently, that they too are dissuaded from it by

the sophisticated consumer's self-control problem is removed. Competition for period 0 contracts ensures that Zero-self will, in effect, choose a preferred contract that commits her One- and Two-selves to follow the chosen consumption plan. This is a standard utility maximization problem subject to an inter-temporal budget constraint (i.e. to the financial intermediary's zero-profit condition). Zero-self chooses C_0 to solve:

$$\max_{C_0} U_0(C_0) \tag{5}$$

$$\text{s.t.} \quad \Pi_0(C_0; Y_0) \geq 0 \tag{6}$$

The familiar first-order necessary conditions are:

$$u'(c_0) = \beta\delta(1+r)u'(c_1) = \beta\delta^2(1+r)^2 u'(c_2) \tag{7}$$

An increase or decrease to the term $\delta(1+r)$ essentially 'tilts' the consumption profile to generally rising or falling over time as $\delta \gtrless \frac{1}{1+r}$. As this across-the-board tilt will not alter key tradeoffs of interest (unlike the degree of present-bias β parameter which does) we shall impose the assumption that $\delta = \frac{1}{1+r} = 1$ for the remainder of the analysis. This is without loss of generality and greatly unclutters the math. The simplified first-order conditions are:

$$u'(c_0) = \beta u'(c_1) = \beta u'(c_2) \tag{8}$$

The binding bank zero profit constraint and first-order conditions allow us to solve for the competitive efficient smoothing full commitment contract C_0^F. A closed form solution for C_0^F is easily found in the CRRA case (37).[17] In this case the FOCs can be rewritten:

$$c_1^F = c_2^F = \beta^{\frac{1}{\rho}} c_0^F \tag{9}$$

Zero-self indulges her present bias (by tilting consumption toward herself) and then allocates remaining resources to be consumed evenly across the remaining two periods.

Consider a simple example where $\beta = 0.5$, $\rho = 1$ and endowment income has present value $\sum y_t = 300$. Zero-self's preferred commitment contract will be $C_0^F = (150, 75, 75)$. If the total income arrives evenly across periods as $Y_0 = (100, 100, 100)$ then this consumption plan would imply borrowing of $c_0^F - y_0 = 50$ in period 0 to be repaid as equal installments of 25 in periods 1 and 2. Had the stream instead been $Y_0 = (200, 50, 50)$ the consumer would save 50 in period 0 to raise consumption by 25 in each of periods 1 and 2. We return to

a credible threat of prohibitive penalties.

[17] All CRRA derivations and closed-form solutions are in the appendix.

these simple numerical examples below to illustrate why when commitment becomes costly and imperfect One-self may carry 'too much debt' or 'not save enough' relative to Zero's preferred choices.[18]

2.1.2 Optimal commitment contracts under Monopoly

When the bank has monopoly power in period 0 the optimal contract will maximize bank profits subject to a consumer participation constraint:

$$\max_{C_0} \Pi_0 \left(C_0; Y_0\right) \tag{10}$$

$$s.t.\ U_0 \left(C_0\right) \geq U_0^A \tag{11}$$

The first-order tangency conditions are again given by expressions 8. Substituting these into Zero-self's binding participation constraint which must bind at a monopoly optimum, allows us to solve for the optimal contract C_0^{mF} and corresponding bank profits $\Pi_0 \left(C_0^{mF}; Y_0\right)$. Closed form solutions for the CRRA case appear as appendix equations 38 and 39. Consumption C_0^{mF} rises and profits fall with the consumer's autarky utility U_0^A.

Conceptually, the equilibrium contract under competition will be found at the tangency between the highest iso-utility surface just touching the budget hyper-plane. Under monopoly, the optimum contract will be at the tangency point where the highest iso-profit plane just touching the iso-utility surface associated with Zero-self's reservation utility. Since the bank fully captures the gains to trade under monopoly, consumption in each period will be lower than under competition.

3 The Renegotiation Problem

Now to questions at the heart of the paper: when is commitment credible, how is it sustained, and at what cost? At issue is the fact that One-self always prefers higher period 1 consumption than what Zero-self wants to build into a contract, so there may be tempting gains to trade from breaking earlier contract commitments. The credibility of the bank's commitment must in turn rest on the threat of a sufficiently high cost to renegotiation costly punishment κ to deter the bank from engaging in such renegotiation. Below we will derive the minimum deterrent punishment required to sustain efficient contracting and the contract adaptations required when the available deterrent falls short of this threshold.

The fraught nature of this potential renegotiation problem is depicted in Figure 1 for the

[18] These parameter values are chosen for expositional purposes. In particular $\rho = 1$ implies that period zero consumption will be the same with or without commitment but the analysis is easily adapted to other values.

case where renegotiation costs are set to $\kappa = 0$ which is to say where One-self and a bank can rewrite any contract at zero penalty. Assume – just for the sake of argument now – that the consumer had (naively as it will turn out) accepted a full-smoothing commitment contract C_0^* in period zero (or C_0^{m*} in the monopoly case), associated with the continuation contract at point F in the $c_1 - c_2$ plane. This contract satisfies Zero-self's optimality condition $u'(c_1) = u'(c_2)$ as indicated by the fact that Zero-self's indifference curve is tangent to the bank's iso-profit line. Since One-self discounts period 2 utility more heavily, this bundle however provides 'too much' to period 2 consumption as $u'(c_1) > \beta u'(c_2)$. This is reflected in the fact that at F One-self's indifference curve is steeper than the bank's iso-profit line. There are mutual gains-to-trade that can be shared by recontracting from F to any new tangency point (along the $c_2 = \beta^{\frac{1}{\rho}} c_1$ ray where One-self's first-order conditions are met) between point $M_{\bar{\kappa}}$ which is the renegotiated contract least favorable to One-self (chosen if the bank could act as monopolist in period 1) and point C_0 which is the renegotiated contract most favorable to One-self (chosen under competitive renegotiation).

Being a sophisticate, Zero-self of course anticipates the problem and will only agree to contracts that satisfy a no-renegotiation constraint to deter the bank(s) and her future self from such harmful renegotiations, as described below. The addition of a new binding no-renegotiation constraint however can only reduce the feasible contract space, reducing consumer welfare and/or bank profits and trade. When the market for period 0 contracts is competitive, consumer welfare will be reduced. If instead the market is monopolized bank profits will suffer as the bank can no longer offer the least cost (most profitable) smoothing contract.

3.1 Renegotiated contracts

To derive a no-renegotiation constraint we must first understand the terms of renegotiated contracts even if, in equilibrium, no such renegotiations will take place. Consider a contract $C_0^0 = (c_0^0, c_1^0, c_2^0)$ and the period 1 subgame determined by its associated continuation contract $C_1^0 = (c_1^0, c_2^0)$. A renegotiation takes place when One-self and a bank agree to replace continuation contract (c_1^0, c_2^0) by a new contract (c_1, c_2).

First consider the case when period 1 banks compete to replace contract C_1^0 with renegotiated contract $C_1^1(C_1^0)$. The contract renegotiation problem becomes:

$$C_1^1(C_1^0) = \underset{C_1}{argmax}\ U_1(C_1) \tag{12}$$

$$\text{s.t. } \Pi_1(C_1; C_1^0) \geq \kappa \tag{13}$$

Figure 1: Optimal competitive contract and renegotiation threat

where the bank participation constraint can be stated as $(c_1^0 + c_2^0) - (c_1 + c_2) \geq \kappa$. To entice a bank to participate the renegotiated contract must increase bank profits (reduce contract expenses) by an amount that equals or exceeds the bank renegotiation cost. Competition insures this constraint exactly binds. As long as U_1 is well behaved and κ is not so high as to make renegotiation infeasible, this can be solved for an interior $C_1^1(C_1^0)$ using the first-order condition $u'(c_1^1) = \beta u'(c_2^1)$ and binding condition 13.[19] For example with zero renegotiation costs ($\kappa = 0$) contract F in Figure 1 would be renegotiated to $C_F(0)$. For positive κ (but less than $\bar{\kappa}$ in the figure) the consumer will surrender just enough surplus to the bank as to get them to participate such as at contract $R_F(\kappa)$ in the figure.

If a bank is a monopolist in period 1, and κ is not so high as to make renegotiation infeasible, the renegotiated contract would solve:

$$C_1^{1m}(C_1^0) = \arg\max_{C_1} \ \Pi_1(C_1; C_1^0) - \kappa \tag{14}$$

$$U(C_1) \geq U(C_1^0) \tag{15}$$

In Figure 1 the monopolist would renegotiate contract F to a point just *above* $R_F(\bar{\kappa})$ which just entices One-self to participate and captures all the gains to renegotiation for the monopolist. Appendix equation (41) shows the monopolist's preferred renegotiated contract from any given continuation contract $C_1^0$and associated profit gains reduced by κ.

These renegotiations will not happen in equilibrium but are important for determining the path of equilibrium play.

3.2 The 'no-renegotiation' condition

When will a contract *not* be renegotiated in period 1? Assuming a tie-breaking rule in favor of Zero-self's preferences, depending on whether the market structure in period 1 is monopolized or competitive, the conditions for no renegotiation in period 1 can be described by:

$$U_1\left(C_1^1(C_1^0)\right) \leq U_1(C_1^0) \tag{16}$$

$$\Pi_1(C_1^{1m}\left(C_1^0\right); C_1^0) \leq \kappa \tag{17}$$

These are in fact two ways of expressing the same thing: a period 0 contract will be renegotiation-proof if and only if it is *not* possible in period 1 for a bank and One-self to agree to a new contract that simultaneously (a) leaves One-self with at least as much

[19] Given CRRA utility the contract is renegotiated to $c_1 = \frac{c_1^0 + c_2^0 - \kappa}{1+\beta^{\frac{1}{\rho}}}$ and $c_2 = \beta^{\frac{1}{\rho}} c_1$

Figure 2: Monopoly commitment contract with $\kappa < \bar{\kappa}$

discounted utility as the original contract, and (b) generates additional profits of at least κ to the bank. In short, the contract is renegotiation-proof so long as renegotiation costs are large enough to exhaust any potential gains to trade between the two parties. These requirements can be expressed as a single no-renegotiation condition. For the CRRA case:

$$u(c_1^0) + \beta u(c_2^0) \geq u\left(\frac{c_1^0 + c_2^0 - \kappa}{1+\beta^{\frac{1}{\rho}}}\right)(1+\beta^{\frac{1}{\rho}}) \tag{18}$$

The determination of the boundary between renegotiation-proof and non-renegotiation proof contracts in $c_1 - c_2$ space is illustrated in Figure 2. Consider a candidate continuation contract (c_1^0, c_2^0) such as at point C. One-self would renegotiate this to any contract lying above their indifference curve running through C (not drawn). The bank however will only agree to a proposed renegotiation if it can lower contract costs to below existing costs net of the renegotiation cost κ. In diagram terms, the bank will agree only if the renegotiated contract falls *below* the lower of the drawn isocost lines with period 1 cost $c_1^0+c_2^0-\kappa$. Contract C is renegotiation-proof because even the most generous renegotiation that One-self can offer the bank fails to meet this condition. Contract F, which smooths consumption efficiently between period 1 and period 2 under Zero-self's preferences, was not renegotiation proof for the given renegotiation cost κ illustrated in the earlier Figure 1. Contract F does become renegotiation-proof however if external renegotiation costs rise to $\bar{\kappa}$ or larger, a threshold we determine mathematically in the next section. The no-renegotiation constraint line (for a given κ) is drawn in Figure 1 as the boundary line that partitions contracts between those that are and are not renegotiation-proof. This boundary is described by a binding equation (18). Point P_κ is a contract along this boundary line[20]: the contract is not renegotiated because One-self's discounted utility at this contract (the left hand side of (18)) lies just above the discounted utility even at the most efficient renegotiated contract (given by the right hand side of (18)).[21]

With this understanding we can now turn back to the period 0 optimization problem. When that market is competitive the optimal contract (given any κ) will be determined by the maximization problem described by Zero self's objective (5) subject to both the bank participation constraint (6) and no-renegotiation constraint (18). When the market for period 0 contracts is monopolized the bank maximizes (10) subject to Zero-self's participation constraint (6) and the same no-renegotiation constraint (18). We study these problems in

[20]Although the constraint seems approximately linear and parallel to the $c_2 = \beta^{\frac{1}{\rho}}$ line it is in fact non-linear (with slope $\frac{dc_2}{dc_1}$ determined by implicitly differentiation of the no-renegotiation constraint). When $\kappa = 0$ it coincides with the $c_2 = \beta^{\frac{1}{\rho}}$ line. See online appendix.

[21]Recall that by assumption ties are broken in favor of Zero-self's preferences so no-renegotiation is weakly preferred here.

section 4 below.

3.3 When are efficient smoothing commitment contracts credible?

What is the minimum renegotiation cost sufficient to deter the renegotiation of full-smoothing commitment contracts? This is found by setting $c_1^0 = c_1^F = c_2^0$ in the no-renegotiation condition (18) and solving for κ. A competitive full-smoothing commitment contract will survive if and only if

$$\kappa \geq \bar{\kappa} \equiv c_1^F \cdot \Upsilon \tag{19}$$

while a monopolistic full-smoothing commitment contract will survive if and only if:

$$\kappa \geq \bar{\kappa}^m \equiv c_1^{mF} \cdot \Upsilon \tag{20}$$

where Υ is the constant in 8.2.2.

Here $\bar{\kappa}$ and $\bar{\kappa}^m$ are the threshold minimum renegotiation costs required to deter the renegotiation of the first-best efficient smoothing commitment contract. The greater the consumption levels in the efficient contract (the greater the scope for profitable contract rearrangements in period 1), the more costly it becomes to deter renegotiation.

Under competition c_1^F is independent of autarky utility (given a fixed value of y) so $\bar{\kappa}$ does not depend on how close or far from optimal consumption smoothing the consumer is in autarky. With monopoly in period 0 the threshold $\bar{\kappa}^m$ which rises linearly with c_1^{mF} will be non-decreasing in autarky utility U_0^A (see 39). Since $c_1^F > c_1^{mF}$ for any initial Y_0 we must also always have $\bar{\kappa}^m < \bar{\kappa}$. Proposition 1 summarizes:

Proposition 1. *Given threshold renegotiation costs $\bar{\kappa}$ and $\bar{\kappa}^m$ as defined in Conditions 19 and 20.*

(a) *The competitive full-smoothing commitment contract survives if and only if $\kappa \geq \bar{\kappa}$.*

(b) *The monopolistic full-smoothing commitment contract survives if and only if $\kappa \geq \bar{\kappa}^m$ with $\bar{k}^m$ strictly rising in the consumer's autarky utility.*

(c) $\bar{\kappa}^m < \bar{\kappa}$.

An implication of statement (b) in the proposition, is that under monopoly, consumers with better autarky options are less likely to get full-smoothing commitment contracts that could be sustained. A consumer with higher autarky utility must be offered higher consumption by the monopolist, and the no-renegotiation condition is harder to satisfy at higher levels of consumption. This serves to dampen the advantages of improved outside options for sophisticated hyperbolic discounters contracting with monopoly banks.

A monopolist is relatively better at delivering efficient-smoothing commitment contracts than under competition ($\bar{\kappa}^m < \bar{\kappa}$), but this is not because monopolists are inherently better at committing; rather, this follows from the fact that having at the outset extracted surplus by offering the consumer a contract with the lowest possible consumption, there is relatively less surplus left to be captured via renegotiation in period 1.[22]

4 Imperfect-Smoothing Commitment Contracts

When bank renegotiation costs are not high enough to sustain efficient-smoothing commitment contracts, that is where $\kappa < \bar{\kappa}$ under competition or $\kappa < \bar{\kappa}^m$ under monopoly, commitment or renegotiation-proofness must be partly self-enforcing but this generally will require contract distortions which we now characterize.

A bank that contracts with naifs will capitalize on the consumer's failure to anticipate harmful future renegotiations (Section 4.3). A sophisticated consumer is wise to the problem and will only agree to renegotiation-proof contracts (Sections 4.2 and 4.3). In the absence of sufficiently high external renegotiation penalties however the parties will resort to additional endogenous enforcement mechanisms by shifting the terms of continuation contracts closer to One-self's preferred choices as a costly strategy to reduce the gains to renegotiation. We label these 'imperfect-smoothing commitment' contracts. They are still technically 'full commitment' contracts in the sense that renegotiation is avoided in equilibrium but they generally provide less than perfect or efficient consumption smoothing from Zero-self's perspective compared to contracts with stronger external enforcement penalties.

For expositional convenience, we first discuss the monopoly case.

4.1 Monopoly

When the market for period 0 contracts is monopolized the bank will want to maximize multi-period profits subject to Zero-self's participation and the no-renegotiation constraint

$$\max_{C_0} \Pi_0 \left(C_0; Y_0\right) \tag{21}$$

$$s.t. U_0 \left(C_0\right) \geq U_0^A \tag{22}$$

$$\Pi_1 \left(C_1^1 \left(C_1\right); C_1\right) \leq \kappa \tag{23}$$

A sophisticate consumer rationally anticipates how her later self may be tempted to

[22]There may be other reasons outside of this model that make monopolists better at committing (i.e. having a higher κ). Our point is that this is not necessary for monopolists to offer better smoothing in more circumstances than competitive firms.

renegotiate and will insist on renegotiation-proofness. When $\kappa < \bar{\kappa}^M$ however this will imply distortion away from efficient smoothing which can only harm bank profits since this raises the contract cost of keeping the consumer at their reservation utility. Hence the monopoly bank itself will insist on renegotiation-proof contracts and, as we shall see, may be willing to spend to improve externally imposed renegotiation penalties.

The bank wants to search for the most profitable renegotiation-proof contract that lies on Zero's participation constraint (11). Consider a candidate level of period 0 consumption c_0^0. The associated continuation contract C_1^0 must lie along Zero-self's autarky utility surface which can be projected as indifference curve $\beta\left[u(c_1^0) + u(c_2^0)\right] = U_0^A - u(c_0^0)$ in $c_1 - c_2$ space. Note this indifference curve shifts down or up as we increase or decrease c_0^0, which for the moment we take as given. Many continuation contracts are both renegotiation-proof and satisfy Zero-self's participation (all in the area above the indifference curve and below the no-renegotiation boundary) but the most profitable amongst these will be at point P_κ in Figure 2 at the intersection of the two constraints. This gives us the optimal renegotiation-proof continuation contract $C_1^m(c_0^0)$ from any c_0^0. The monopolist's optimal contract is then determined by choosing over c_0^0.

4.1.1 Properties of the contract

The renegotiation-proof contract can be explicitly derived for the CRRA case of $\kappa = 0$ (Equation 8.3.1). For, $0 < \kappa < \bar{\kappa}^m$, the contract cannot be derived in closed form, but its key properties can be established and it can be easily solved for numerically.[23]

Proposition 2. *Suppose $\kappa < \bar{\kappa}^m$ and the consumer is sophisticated. Under monopoly, the profit-maximizing renegotiation-proof contract (C_0^{mP}) has the following properties:*

(i) $\Pi_0\left(C_0^{mP}; Y_0\right) < \Pi_0\left(C_0^{mF}; Y_0\right)$

(ii) $c_0^{mP} > c_0^{mF}$

Proposition 2 compares the renegotiation-proof contract to the full-smoothing commitment contract when the renegotiation-proofness constraint binds. First, bank profits will be lower than under full-smoothing commitment. The bank wishes it could promise to not renegotiate but it cannot make such a promise credible without giving up some profits. The monopolist would be better off with higher external renegotiation penalties since in equilibrium renegotiation does not take place.

A related observation is that the bank will prefer not to contract with individuals who have minimal smoothing needs; for individuals whose autarky utility is close enough to U_0^F,

[23]The contract can however be easily found numerically and all relevant figures in this paper were so derived. See the online appendix for details and python code.

the bank would make negative profits under the best renegotiation-proof contract, so there will be lost trade.

The second statement of the proposition is about the terms of the contract – when full-smoothing commitment is not feasible, the renegotiation-proof contract will involve higher consumption in period 0 (i.e. either a smaller loan or less savings) compared to full-smoothing commitment. The following is a sketch of the argument (the proof in the appendix uses some additional notation for logical clarity which we describe intuitively here).[24]

Any contract C_0 can be fully described in terms of three variables, c_0, s, and α. Here, c_0 is period 0 consumption and s is the total consumption allocated to periods 1 and 2. α determines the share of s that is consumed in period 1. So, $C_0 = (c_0, \alpha s, (1-\alpha)s)$. This notation serves two purposes. First, α captures renegotiation concerns. Under full-smoothing, $\alpha = \frac{1}{2}$, as this is optimal from Zero-self's perspective. When the no-renegotiation constraint binds, we get a function $\alpha(s)$ which tells us how much larger c_1 must be relative to c_2 so that further renegotiation would be unprofitable. Intuitively, as total consumption in periods 1 and 2 s rises the parties must rely more on distorting self-enforcement mechanisms to supplement the fixed no-renegotiation penalty κ as a sufficient deterrent to renegotiation. Zero-self must in effect become more accommodating of One-self's preferences so period 1 must get a bigger share.

Define the continuation utility from Zero-self's perspective as $V(s,\alpha) \equiv \beta[u(\alpha s) + u((1-\alpha)s)]$. At any contract that constitutes an optimum, the following must be true:

$$\frac{du(c_0)}{dc_0} = \frac{dV}{ds} \tag{24}$$

Otherwise, the bank could raise profits by reallocating consumption from period 0 to the future or vice versa. This is just a restatement of the first-order condition.

Now suppose the optimal renegotiation-proof contract specified the same level of period 0 consumption as the full-smoothing contract, so $c_0^{mP} = c_0^{mF}$. Since any future consumption must be split unevenly, in order to continue to satisfy the consumer's period 0 participation constraint, it must be true that $s^{mP} > s^{mF}$. We show in the appendix that s^{mP} would have to be large enough that, at these values,

$$\frac{du\left(c_0^{mP}\right)}{dc_0} > \frac{dV}{ds} \tag{25}$$

so this contract could not be profit-maximizing. In other words, a switch from full-smoothing commitment to renegotiation-proofness while maintaining the same c_0 would require such

[24]If the reader prefers to skip this, we present a purely intuitive explanation of the results near the end of Section 4.2.2.

a large jump in future total consumption (to continue satisfying Zero-self's participation constraint) that the marginal utility of future consumption would be low. So the bank could do better by raising period 0 consumption at the expense of future consumption.

The bank limits renegotiation possibilities by transferring consumption away from the future (when renegotiation is a temptation) to the present. Relative to full-smoothing commitment, consumers get contracts with larger loans or less savings.

4.2 Competition

When the market for period 0 contracts is competitive the optimal contract solves:

$$\max_{C_0} \; U_0\left(C_0\right) \tag{5}$$

$$\Pi_0\left(C_0; Y_0\right) \geq 0 \tag{6}$$

$$\Pi_1\left(C_1^1\left(C_1\right); C_1\right) \leq \kappa \tag{17}$$

As noted in section 3.2, the no-renegotiation constraint 17 assures that gains-to-trade from renegotiation fall short of bank renegotiation costs. Even if new banks could enter in period 1 to offer part or or all of the surplus from renegotiation to One-self in period 1 the constraint deters renegotiation as long as those banks also face renegotiation cost κ.[25]

We can reuse Figure 2 to interpret the contract design. Zero-self wants to search for the most profitable renegotiation-proof contract that lies on the bank's participation constraint. Suppose Zero-self has chosen a candidate period 0 level of consumption c_0^0. To be part of an optimum renegotiation-proof contract Zero must ensure that the continuation contract is renegotiation proof and the bank participation constraint $c_1^0+c_2^0 = y-c_0^0$. Many continuation contracts are both renegotiation-proof and satisfy Zero-self's participation (all below both the zero-profit line and the no-renegotiation boundary) but the most-preferred by Zero will be at point P_κ in Figure 2 at the intersection of the two constraints.[26] This gives us the optimal renegotiation-proof continuation contract $C_1(c_0^0)$ from any c_0^0. Zero-self's optimal contract is then determined by backward induction, choosing over c_0^0.

Except for the special case where $\kappa = 0$, there is no simple closed form solution for the optimal contract even in the CRRA case. This can be seen from the fact that there are two points where the no-renegotiation condition indifference curve crosses the period 1

[25] Later we discuss the empirically relevant cases where competing banks might enter in period 1 and offer to renegotiate at lower or zero renegotiation cost).

[26] Note that while we are reusing Figure 2 to describe both the monopoly and the competitive contract design problem because, conceptually, they are very similar, optimal consumption levels will be generally higher under competition so the point P_κ is not the same in both cases.

Figure 3: Imperfect smoothing commitment contract with $\kappa < \bar{\kappa}$

zero-profit line. As described in the appendix the c_1^0 coordinates of these two roots are given by a non-linear equation (**??**).

In the special case of perfect competition with costless renegotiation ($\kappa = 0$) there will be a unique solution and a closed form. In figure 3 think of how $P(\kappa)$ slides down the budget line as κ shrinks until we get to a point where One's indifference curve is tangent to the period 1 budget line. This continuation contract is 'renegotiation-proof' only in the very narrow sense that it won't be renegotiated because it already delivers One-self's preferred consumption choice. This contract is explicitly solved in expression (8.3.1).

Consider the practical interpretation of a simple example: with $\beta = 0.5$ and $\rho = 1$ and $\kappa = 0$ the best available competitive contract $C_0^P = (150, 100, 50)$ offers considerably less consumption smoothing in later periods compared to the benchmark full-smoothing $C_0^F = (150, 75, 75)$. If the consumer's initial income stream were arranged as $Y_0 = (100, 100, 100)$ we would interpret the absence of commitment case ($\kappa = 0$) as rolling over period one debt that Zero would have preferred to have seen repaid. The entire burden of the debt that Zero took out in period 0 but would prefer to have been shared equally between periods 1 and 2 is 'rolled over' and placing the entire burden now on period 2. Had the income stream instead been $Y_0 = (200, 50, 50)$ then we might interpret the consumer in period one as 'raiding savings' that, with commitment, One self would have protected for period 2 consumption.

In this example, consumers that can obtain full-smoothing commitment contracts will save/repay more or borrow less in period 1 and consume more in period 2. The inability to commit leads to lower welfare for Zero in this competitive setting. However, this is not generally true in the monopoly case, as we show below. Credible division rules between periods 1 and 2 depend on the no-renegotiation constraint in some subtle ways.

4.2.1 Properties of the contract

Suppose contract C_0^P is the solution to the maximization problem described by 5, 6 and 17.

Proposition 3. *Suppose $\kappa < \bar{\kappa}$ and the consumer is sophisticated. Under competition, the competitive renegotiation-proof contract that maximizes Zero-self's discounted utility (C_0^P) has the following properties:*

(i) $U_0\left(C_0^P\right) < U_0\left(C_0^F\right)$

(ii) The relationship between c_0^P and c_0^F is ambiguous. There is some $\hat{\rho}$ such that: if $\rho \leq \hat{\rho}$, then $c_0^P > c_0^F$; if $\rho > \hat{\rho}$, then there are parameter values under which $c_0^P < c_0^F$.

The first statement is straightforward: Since $\kappa < \bar{\kappa}$ means the new renegotiation-proofness constraint (17) binds full-smoothing smoothing cannot be achieved and the consumer's welfare must be lower than under the first-best contract.

Now, will period 0 consumption be higher or lower than under full-smoothing commitment? The proposition is that this depends on parameter values, in particular the intertemporal elasticity of substitution $\frac{1}{\rho}$. Consider the competitive full-smoothing commitment contract C_0^F. Following the modified notation of section 4.1.1 and first-order condition (8), it must be true that:

$$\frac{du\left(c_0^F\right)}{dc_0} = \frac{\partial V\left(s^F, \frac{1}{2}\right)}{\partial s}$$

Now suppose the competitive renegotiation-proof contract C_0^P involves the same period 0 consumption as under full-smoothing commitment, so that $c_0^P = c_0^F$. By the bank's zero-profit constraint, the contract will also have $s^P = s^F$, but consumption will be split in period 1's favor. If the utility function is relatively linear (low ρ), then an imbalanced split of s results in a lower marginal utility than from a balanced split. So:

$$\frac{du\left(c_0^F\right)}{dc_0} > \frac{dV\left(s^F, \alpha\left(s^F\right)\right)}{ds}$$

In such a case, the renegotiation-proof contract must involve higher period 0 consumption than the full-smoothing commitment contract.

If, on the other hand, the utility function is highly convex (high ρ), then an imbalanced split results in higher marginal utility relative to full-smoothing. In such cases, the renegotiation-proof contract will have lower period 0 consumption than under full-smoothing commitment under certain parameter values.[27] This can be seen more explicitly in the case of $\kappa = 0$ (Equation 8.3.1).

So, under competition, the renegotiation-proofness constraint could change the contract in either direction: a larger loan (less saved) or a smaller loan (more saved). Period 2 consumption however always falls relative to the full-smoothing commitment case, even in the cases when Zero saves more/borrows less. In fact for CRRA utility the adjustment of period 0 consumption (in the absence of commitment compared to with commitment) is always relatively small while the adjustment to period 1 and period 2 consumption is relatively much larger.[28] In other words despite having a first-mover advantage, Zero can do little other than to partially accommodate to the consumption pattern that One-self wants to impose.

The contrast between monopoly and competition can be explained using the intuition of

[27] The precise construction of the cutoff value $\hat{\rho}$ is somewhat complicated, as $\frac{dV}{ds}$ depends not just on $u'(c_1)$ and $u'(c_2)$, but also on how the sharing rule, $\alpha(s)$, changes with s.

[28] To illustrate, with $\kappa = 0$ at no point does period 0 consumption rise or fall by more than six percent for any value $\rho \in (0, \infty)$ and $\beta \in (0, 1)$ but at reasonable parameter values such as $\rho = 0.5$ and $\beta = 0.5$ in the absence of commitment period 1 consumption rises to 149 percent of the level it would be with commitment, and period 2 consumption falls to just 37 percent of what it would be.

income and substitution effects. In either case, a move from full-smoothing commitment can be viewed as a rise in the "price" of future utility from Zero-self's perspective. As a result, substitution effects will lead to an increase in period 0 consumption and a drop in future consumption. Under monopoly, since the consumer is always left at her autarky utility, there are no income effects. When the renegotiation-proofness constraint binds, the price of future utility effectively rises, as a result of which substitution effects lead to greater period 0 consumption. Under competition, income and substitution effects work against each other; the net result depends on the shape of the consumer's utility function.

4.3 Contracting with Naive Hyperbolic Discounters

For naive agents, the problem of renegotiation does not generally lead to a renegotiation-proof contract. The naif believes she will not be tempted to renegotiate. Banks therefore offer contracts that take into account the potential renegotiation. Under monopoly, the bank adds to its profits by engaging in renegotiation that was not anticipated by the consumer in period 0. Under competition, banks return the potential surplus from renegotiation to the Zero-self.[29]

4.3.1 Monopoly

Relative to a sophisticated consumer, with a naive consumer the monopolist bank can make additional profits on two margins. First, since there is no perceived renegotiation problem, the consumer is willing to accept a contract that is more profitable for the bank up-front; subsequently, possible renegotiation generates additional profits for the bank.[30]

The bank must choose between a renegotiation-proof contract and one that will be renegotiated upon. If κ is sufficiently large there is little to gain from renegotiation and the consumer will be offered the full-smoothing commitment contract. But when κ is relatively small, the bank might prefer to offer a contract that will subsequently be renegotiated. In such cases, the bank solves the following problem:[31]

$$\begin{aligned} \underset{C_0}{max} \quad & \Pi_0\left(C_0; Y_0\right) + \Pi_1\left(C_1^{m1}\left(C_1\right); C_1\right) - \kappa \\ s.t. \quad & U_0\left(C_0\right) \geq U_0^A \end{aligned} \tag{26}$$

[29] A similar analysis could be carried out if consumers were misinformed not about their own preferences but about κ.

[30] There is an additional consideration – that naive hyperbolic discounters might be inaccurately optimistic about autarky outcomes because of a failure to anticipate commitment problems. This would have the interesting effect of tightening the participation constraint and reducing surplus available to the monopolist

[31] We do not need to worry about a renegotiation-proofness constraint here. Since period 0 believes her period 1 preferences are consistent with her own, she expects any renegotiation of the period 0 contract to yield the same discounted utility as the contract itself.

Let the solution be denoted C_0^{mN}. This is explicitly derived in the appendix (8.3.2, 8.3.2). As under competition, the bank maximizes profits by offering a contract that divides future consumption as much in favor of period 2 as possible. The greater the imbalance between the contracted c_1 and c_2, the greater the bank's profits from renegotiation. We show that if $\rho < 1$, the contract is at a corner solution where $c_1 = 0$. If $\rho > 1$, an explicit solution does not exist, but maximization pushes the contract to a point where c_2 approaches infinity.[32] This contract can be compared to the full-smoothing commitment contract and to the renegotiation-proof contract for sophisticates. In particular, it will involve lower period 0 consumption than under both full-smoothing commitment and renegotiation-proofness. This result appears counter-intuitive. In the case of lending, it does not reinforce the narrative of banks preying on naive consumers by offering them relatively large loans with steep repayments. Indeed, there are other considerations beyond the scope of this model, such as the possibility of collateral seizure, that could generate large loans. But our limited model helps to highlight a particular aspect of contracting with naive hyperbolic discounters: here, the bank offers them relatively *small* loans because its gains from renegotiation depend on the surplus that the initial contract delivers to periods 1 and 2. In order to fully take advantage of the consumer's naivete, the consumer must start out with sufficiently small repayments that the bank could profit from rearranging them.

The next proposition summarizes the above discussion.

Proposition 4. *Suppose the consumer is naive. Under monopoly:*

(i) If κ is sufficiently higher than $\bar{\kappa}^m$, the firm will offer the agent the full-smoothing commitment contract (C_0^{mF}) and it will not be renegotiated.

(ii) Otherwise, the contract C_0^{mN} will satisfy $c_0^{mN} < c_0^{mF} < c_0^{mP}$ (either explicitly or in the limit), and it will be renegotiated in period 1.

4.3.2 Competition

Under competition, with naive consumers contracts must account for renegotiation to have firms continue earning zero profits. First, note that if contracts are not exclusive, the equilibrium contract must be identical to the full-smoothing commitment contract. This is because the firm offering the contract in period 0 does not expect to benefit from renegotiation, so the contract gets competed down to the one that maximizes the naive Zero-self's perceived utility while delivering zero profits to the bank.

Under exclusive contracts, anticipated profits from future renegotiation will be returned to the consumer through more favorable initial contracts. If κ is sufficiently small, the

[32]This can be dealt with by a reasonable assumption of an upper bound on contract terms.

equilibrium contract involves renegotiation and satisfies:

$$\begin{aligned} &\underset{C_0}{max} U_0 \left(C_0\right) \\ &\quad s.t. \Pi_0 \left(C_0; Y_0\right) + \Pi_1 \left(C_1^{m1}\left(C_1\right); C_1\right) \geq \kappa \end{aligned} \tag{27}$$

Let the solution be denoted C_0^N. This is explicitly derived in the appendix (8.3.2 and 8.3.2). Contracts divide future consumption as much in favor of period 2 as possible. This maximizes the potential gains from renegotiation. In the context of loans, this suggests contracts where the debt burden is heaviest in the intermediate stages, resulting in renegotiation to postpone payments.

Unlike under monopoly, competition returns anticipated renegotiation gains to the consumer. Some of these gains are returned to the Zero-self, so there is no clear prediction about whether period 0 consumption will be lower or higher than under full-smoothing commitment.

Proposition 5. *Suppose the consumer is naive. Under competition:*

(a) If contracts are not exclusive: The consumer will accept the full-smoothing commitment contract, C_0^F. The contract will be renegotiated in period 1 if and only if $\kappa < \bar{\kappa}$.

(b) If contracts are exclusive:

(i) If κ is sufficiently higher than $\bar{\kappa}$, the consumer will accept the full-smoothing commitment contract (C_0^F) and it will not be renegotiated.

(ii) Otherwise, the consumer will accept a contract C_0^N with the following properties: if $\rho < 1$, $c_0^N < c_0^F$. If $\rho > 1$, then there are parameter values under which $c_0^N > c_0^F$.

5 Not for profit and hybrid ownership forms

Consider next the case of a firm that, in a pre-contract stage, has the possibility of choosing its ownership structure say by incorporating as a legal non-profit or, more broadly, by choosing a degree of 'hybrid' ownership, for example by retaining for-profit status but allowing social investors to establish considerable ownership stakes and managerial control. In the spirit of Hansmann () and the discussion in the introduction, we model this as a restriction on the firm's ability to distribute raw profits to managers and shareholders:

Definition. Given 'raw profits' Π_0, a 'nonprofit' firm retains 'captured profits' $f\left(\Pi_0\right)$, where $f(0) = 0$, $f'\left(\Pi_0\right) \in (0, 1)$, and $f''\left(\Pi_0\right) \leq 0$.

This formulation follows **?** who argued that though the principals of a non-profit may be technically legally barred from tying compensation to cash profits, they can in practice capture a fraction of those profits in costly and imperfect ways via the consumption of perquisites or 'dividends in kind' (e.g. the lavish expense account). The ability of perquisites to substitute for unrestricted consumption falls as profits get larger.

Setting aside welfare concerns that might drive firms to adopt nonprofit or hybrid status, we examine when purely profit-minded firms might make such governance choices; i.e. when can a voluntary restriction on the ability to enjoy profits make a self-interested firm better off?[33] This has parallels to the explanation for commercial nonprofits due to **?** and modeled by **?** but established on quite different behavioral grounds.[34]

At the outset, it should be noted that profit-oriented principals have no incentive to switch to hybrid/nonprofit status when consumers are naive. Since the consumer does not perceive a need for commitment, any promise of superior commitment is of no value to her. The analysis with sophisticated consumers follows.

5.1 Monopoly

In a pre-contract phase the firm now first establishes its type via the adoption of legal nonprofit status and/or by choosing credible and stable ownership and governance structures that commit it to those limitations. If the monopoly firm were to operate as a nonprofit or a hybrid, when facing a sophisticated hyperbolic discounter it would design a renegotiation-proof contract to solve:

$$\underset{C_0}{max}\ f\left(\Pi_0\left(C_0; Y_0\right)\right) \tag{28}$$

$$U_0\left(C_0\right) \geq U_0^A \tag{29}$$

$$f\left(\Pi\left(C_0; Y_0\right) + \Pi_1\left(C_1^{m1}\left(C_1\right); C_1\right)\right) - f\left(\Pi\left(C_0; Y_0\right)\right) \leq \kappa \tag{30}$$

Why might a profit-maximizing firm choose to operate as a nonprofit when that reduces its ability to capture profits? The answer lies in the loosening of the no-renegotiation constraint (30). Any gains from renegotiation are worth less than they would be to the for-profit firm. Clearly, the for-profit monopolist's contract $\left(C_0^{mP}\right)$ would now leave the no-renegotiation constraint slack. Because the non-profit can more credibly commit to not renegotiate contracts that offer greater consumption smoothing across periods 1 and 2, Zero-self becomes more willing to pay for this consumption stream.

[33] Indeed, welfare concerns could directly improve consumer outcomes by either allowing Zero-self's participation constraint to be slack or by raising the costs of renegotiation κ. We consider the latter point in an example below.

[34] In those accounts a firm delivers less than a promised quantity or quality of a good or service, unambiguously harming the time-consistent client. The client discovers this after the fact but cannot challenge the contract breach only because it is too difficult or costly. In contrast in our model the firm and the One-self customer both gain from voluntarily breaking existing contract commitments and Zero-self is no longer around to mount a challenge.

The captured-profits maximizing solution gives a contract that we denote C_0^{mNP}. If $\kappa < \bar{\kappa}^m$, with a relaxed renegotiation-proof constraint $\Pi_0(C_0^{mNP}; Y_0) > \Pi_0(C_0^{mP}; Y_0)$ but whether or not it will be in the bank principals' best interest to strategically convert to non-profit status depends on whether the captured profits under non-profit status exceed the profits they could earn as a pure for-profit, in other words on whether $f\left(\Pi_0(C_0^{mNP}; Y_0)\right) > f\left(\Pi_0(C_0^{mP}; Y_0)\right)$. The monopolist faces a tradeoff in considering non-profit status: higher raw profits (as the commitment problem is partly solved) but a diminished capture of those raw profits.

Does the rise in extracted surplus outweigh the fact that all profits are now discounted? Proposition 6 (in Section 5.2) establishes the existence of captured profit functions that would be strictly preferred to for-profit status for monopoly firms. This is easy to see: the possible concavity of f could leave the enjoyment of profits relatively unaffected while significantly loosening the no-renegotiation constraint (since renegotiation would raise profits further, and since f is concave, these additional profits would count for little).

Given particular captured profit functions, we can also ask which consumers are more likely to be served by nonprofit monopolists. If consumers are far from optimal in autarky, then the for-profit firm would anyway be making substantial profits. In this case, the non-profit's credibility advantages are not enough to outweigh the fact that it loses a significant amount of enjoyment of its profits due to legal restrictions.

However, for consumers with higher autarky utility, the gains that can be captured from nonprofit status are large relative to the profits that a for-profit would have made, so the firm prefers to operate as a nonprofit. As an example, consider an autarky consumption bundle at which the for-profit firm would earn zero profits. Now, the nonprofit firm can earn positive profits, so regardless of f nonprofit status dominates.

5.1.1 An example

Let us consider a situation where a firm may choose it's degree of hybrid-ness or for-profit orientation, indexed by a parameter $\alpha \in [0, 1]$. For a chosen α, let the captured profits function be linear:

$$f\left(\Pi_0\right) = \alpha \Pi_0 \tag{31}$$

We can interpret α as the maximum fraction of raw profits that can be distributed to managers and shareholders. An $\alpha = 1$ would represent a pure for-profit investor-led firm, $\alpha = 0$ a strictly regulated non-profit.

We can also allow α to directly affect the non-pecuniary renegotiation cost the firm's principals incur when they opportunistically break contractual promises to customers. A more hybrid or non-profit firm dominated by social investors is more likely to hire staff and managers that internalize client welfare and social investor motivations and therefore are more likely to feel non-pecuniary costs associated with guilt, shame or loss of reputation from breaking promises. If we now label the cost of renegotiation $\eta(\alpha)$ – replacing our earlier κ – this idea is captured by assuming that function η falls weakly in α. Putting both mechanisms together gives us modified no-renegotiation constraint (32) which states that the fraction of raw profits Π_1 that can be captured from renegotiating a contract must not exceed renegotiation costs:

$$\alpha\Pi_1(C_1^{m1}(C_1); C_1) \leq \eta(\alpha) \tag{32}$$

If we define $\kappa(\alpha) \equiv \frac{\eta(\alpha)}{\alpha}$, this no-renegotiation constraint can be written as

$$\Pi_1(C_1^{m1}(C_1); C_1) \leq \kappa(\alpha) \tag{33}$$

which looks like the earlier constraint (23) except κ is now a function of α. The earlier renegotiation problems were for the special case of a pure for-profit firm with $\alpha = 1$ but we can now analyze contracting, captured profits, and client welfare at any level of α and ownership choices in a strategic equilibrium.

To the extent that the loosening of the no-renegotiation constraint happens through the right-hand side (i.e. via term $\eta(\alpha)$, which represents the firm's motivation to honor the initial agreement), the firm benefits unambiguously–it is able to offer better commitment *and* fully retain the added profits.

In Figure 4 we illustrate the case where non-pecuniary costs to breaking a promise not to renegotiate fall with α according to $\eta(\alpha) = 10(1-\alpha)$ and hence that the overall cost to renegotiation varies with α according to $\kappa(\alpha) = 10(1-\alpha)/\alpha$. The plots depict captured profits that would be achieved at different levels of α starting from three different initial endowment streams. These three streams - (60, 120, 120), (90, 105, 105) and (120, 90, 90) – are equal in their present value of 300 but differ in terms of period 0 income (with remaining income allocated equally across period 1 and 2). The higher of the two curved lines represents 'raw' profits $\Pi_0(C_0^{m\alpha}; Y_0)$ and the lower curve captured profits $\alpha\Pi_0(C_0^{m\alpha}; Y_0)$. A horizontal line has been drawn in to indicate the level of profits $\Pi_0(C_0^{mP}; Y_0)$ captured by a pure for-profit ($\alpha = 1$). Consider the top panel where the customer has initial income $(60, 120, 120)$. As this type of customer wants to borrow heavily in period 0, profits to the bank are large, even in the case of renegotiation-proof contracts. Adopting non-profit status by lowering α confers limited profit gain however: the cost of lowering alpha (giving up a share of already high profits) is not compensated for by the gains from being able to credibly commit to a smoother contract. However at $(90, 105, 105)$ the tradeoff is different and profits can be increased. In the picture any non-profit with an α between approximately 0.7 and less than one captures more profits than a pure for-profit. Finally for customers with an

Figure 4: Captured rent by ownership status and endowment income

endowment $(120, 90, 90)$ are already fairly close to their preferred consumption stream so the profits to be captured even under full commitment are not that large. Indeed in this case a pure for-profit cannot earn positive profits. Here the cost of adopting non-profit status is low compared to the gains, and we the simulation reveal that any non-profit status firm captures more profits than a pure for-profit, and maximum captured profits are achieved at around $\alpha = 0.7$.

5.2 Competition

5.2.1 Exclusive contracts

Consider what would happen in the competitive market situation now if contracts can be assumed to remain exclusive, so that any new surplus in the event of a renegotiation between the bank and the One-self goes to the bank (this grants the bank monopoly power in period 1). In this setting, a nonprofit/hybrid firm will be led to offer contract terms to solve:

$$\underset{C_0}{max} U_0 \left(C_0 \right) \tag{34}$$

$$s.t. f \left(\Pi_0 (C_0; Y_0) \right) \geq 0 \tag{35}$$

$$f \left(\Pi \left(C_0; Y_0 \right) + \Pi_1 \left(C_1^{m1} \left(C_1 \right); C_1 \right) \right) - f \left(\Pi \left(C_0; Y_0 \right) \right) \leq \kappa \tag{36}$$

Let the contract that solves this program be denoted C_0^{eNP} . Consider first a field where all firms start as pure for-profits and earn zero profits. If the no-renegotiation constraint binds, Zero-self's utility must be lower than optimal. Starting from this situation consider now one firm's strategic choice of whether to adopt non-profit status. One firm deviating into nonprofit status in this way can make positive profits while offering Zero-self a contract with a higher discounted utility because of the loosened no-renegotiation constraint (36). So, if the borrowers are sophisticated hyperbolics, in equilibrium all firms become nonprofit and earn zero profits.

5.2.2 Non-Exclusive Contracts

Now, assume that exclusivity and period 1 monopoly power disappears. Firms can compete to renegotiate each other's contracts in period 1.

If there were only nonprofits in equilibrium, any one firm could make positive profits by switching to for-profit status and undoing a rival bank's contract in period 1. The advantages of undercutting other firms' contracts outweigh the benefits of promising one's own clients it will not renegotiate. As a result, equilibrium contracts will be determined by for-profit firms, and consumers will be offered lower commitment than from non-profit firms alone.[35]

[35]The same argument applies if banks can costlessly renegotiate other bank's contracts.

The above discussion is summarized in the following proposition.

Proposition 6. *(a) Suppose $\kappa < \bar{\kappa}^m$. Under monopoly, there exist captured profit functions such that the firm will operate as a nonprofit.*

(b) Suppose $\kappa < \bar{\kappa}$. Under competition: (i) If contracts are exclusive, firms will operate as nonprofits for any captured profit discount function. (ii) If contracts are not exclusive, there is no captured profit discount function under which firms will operate as nonprofits.

6 Additional considerations

Our preceding analysis made the simplifying assumption that contracts between consumers and banks could only be initiated in period 0. The alternative to a period 0 contract was autarky for the consumer and zero profits for the bank. This served to streamline the analysis. We now discuss the interesting problem of how the contract space is enriched by allowing unbanked consumers to sign two-period contracts in period 1, possibly without contracting in period 0.[36] The main change will relate to the formulation of reservation values and their implications for the shape and feasibility of the period 0 contracts. This does not change most qualitative results of the paper but the discussion raises interesting questions about circumstances where consumers could be better off under an autarky economy compared to one with a financial intermediary.

Consider the monopolist bank facing a sophisticated hyperbolic discounter. If a contract were not signed in period 0, they would meet again in period 1. In period 1, the contract must satisfy One-self's participation constraint, which would be determined by some unbanked consumption path $C_1^{A'}$. This can be stated formally. Given some consumption path $C_1^{A'}$, the bank solves:

$$\max_{C_1} \Pi_1 \left(C_1; C_1^{A'} \right)$$
$$\text{s.t. } U_1 \left(C_1 \right) \geq U_1 \left(C_1^{A'} \right)$$

Let the solution be denoted $C_1^{m'}$. Two observations can be made. First, the bank can always offer a period 1 contract that delivers nonnegative profits. This is because any contract will satisfy One-self's optimality condition, $u' \left(c_1^{m'} \right) = \beta u' \left(c_2^{m'} \right)$. So, except in the special case where the autarky consumption path satisfies this condition, the bank can make positive profits in period 1. Second, the autarky consumption path $C_1^{A'}$ might differ from C_1^A, the consumer's autarky utility in the absence of banking. In other words, C_1^A maximizes $U_0 \left(C_0^A \right)$ while $C_1^{A'}$ maximizes $U_0 \left(c_0^{A'}, c_1^{m'}, c_2^{m'} \right)$. In the latter case, period 0 anticipates that

[36] We thank Abhijit Banerjee for helpful discussions on this point.

consumption across periods 1 and 2 is guaranteed to satisfy period 1's optimality condition. We can denote $C_0^B \equiv \left(c_0^{A'}, c_1^{m'}, c_2^{m'}\right)$, which corresponds to a Zero-self utility of U_0^B.

In period 0, any contract must meet the Zero-self's reservation utility, U_0^B:

$$\underset{C_0}{max}\Pi_0\left(C_0;Y_0\right)$$
$$\text{s.t.}U_0\left(C_0\right) \geq U_0^B$$

The maximization problem looks familiar, apart from the modified reservation utility. The Zero-self's discounted utility from such a contract is no longer monotonic in her Zero-self's full-autarky utility, U_0^A. For example, consider two hypothetical consumers who in autarky must consume their income streams, which deliver the same autarky utility but through different consumption paths: consumer X has $c_1^A = c_2^A$ while consumer Y has $c_1^A > c_2^A$ in a way that satisfies period 1's optimality condition. Then, for consumer X, $U_0^B < U_0^A$ while for consumer Y, $U_0^B = U_0^A$. It follows that, since period 0 contracts depend on the distribution of future consumption, a consumer who fares relatively better in the absence of a bank may fare relatively worse under a banking contract.

Given this benchmark full-commitment contract, the renegotiation-proof contract can be solved for by adding a no-renegotiation constraint to the above maximization problem. The constraint is the same as used previously, and again narrows the set of contracts that can be offered in period 0. As in Proposition 2, the renegotiation-proof constraint results in lower profits and greater period 0 consumption relative to full-commitment. These results are independent of the period 0 reservation utility and therefore remain unchanged.

A key difference here, however, is that a renegotiation-proof contract will be offered to all consumers (unlike before, where the bank was better off not contracting with consumers whose autarky utility left them close enough to the first-best). Intuitively, this is because the alternative to a period 0 contract is not autarky; rather, it is a period 1 contract that tilts consumption in period 1's favor. Since, in period 0, the bank can at least offer the consumer a consumption path of C_0^B, it ensures that a contract will be accepted.

By opening up the possibility of period 1 contracts, we introduce an additional consideration–the same bank that offers commitment itself creates a need for commitment. By threatening to fully indulge the One-self's preferences, the bank is always able to induce the Zero-self to accept an offer of partial commitment, no matter how weak.

Finally, observe that the bank's decision about whether to operate as a nonprofit is subject to the same tradeoff between improved commitment and reduced enjoyment of profits. However, under monopoly the attractiveness of nonprofit status drops (relative to the case where period 1 contracts are disallowed) due to the fact that even the for-profit bank finds it profitable to offer contracts to consumers at all levels of autarky utility.

7 Conclusion

The starting point for this paper is the observation that the solution to any commitment problem must also address a renegotiation problem. We show how the renegotiation problem depends on costs of renegotiation and how it changes contract terms in sometimes unexpected ways. In this context, we also provide a rationalization of commercial nonprofits in the absence of asymmetric information.

We argue that the model sheds some light on trends in microfinance, payday lending, and mortgage lending. We hope this paper also offers a framework that can be built upon. The incorporation of additional 'real-world' factors could improve our understanding of particular institutions and generate empirically relevant comparative statics. Examples of these include nondeterministic incomes, private and heterogenous types, collateral and strategic default, and longer time horizons.

Furthermore, the analysis could be expanded to heterogeneous populations. For instance, how might a monopoly's governance choices be affected when it serves a market that comprises both naive and sophisticated hyperbolic discounters? With sophisticates, the firm would prefer high renegotiation costs while with sophisticates, it would prefer that the same costs be low.

Finally, the differences between monopoly and competition open up some new, potentially interesting questions. How does market structure evolve and what are the implications for commitment? And through this evolution might there emerge third parties to contracts between consumers and banks that can more effectively enforce the commitment that is sought after on both sides of the market?

8 Appendix: CRRA Derivations and Proofs

8.1 Full-commitment

8.1.1 Competition

Combining the first-order conditions (8) and the budget constraint (6) of the utility maximization problem, the competitive full-smoothing commitment contract C_0^F is:

$$C_0^F = \left(\frac{y}{1+2\beta^{\frac{1}{\rho}}}\right) \cdot \left(1, \beta^{\frac{1}{\rho}}, \beta^{\frac{1}{\rho}}\right) \tag{37}$$

8.1.2 Monopoly

For the monopolist bank that offers full-commitment, the solution is determined by the first-order condition and the consumer's participation constraint:

$$C_0^{mF} = \left(\frac{U_0^A\left(1-\rho\right)}{1+2\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}} \cdot \left(1, \beta^{\frac{1}{\rho}}, \beta^{\frac{1}{\rho}}\right) \tag{38}$$

$$\Pi_0\left(C_0^{mF}; Y_0\right) = y - \left(U_0^A\left(1-\rho\right)\right)^{\frac{1}{1-\rho}}\left(1+2\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}} \tag{39}$$

It can easily be verified that $C_0^F > C_0^{mF}$.

8.2 The no-renegotiation constraint

Consider any existing continuation contract C_1^0. The competitively renegotiated contract (most beneficial to the consumer) will be:

$$C_1^1\left(C_1^0\right) = \left(\frac{c_1^0 + c_2^0 - \kappa}{1+\beta^{\frac{1}{\rho}}}\right) \cdot \left(1, \beta^{\frac{1}{\rho}}\right) \tag{40}$$

The condition to make sure the consumer will neither propose nor accept this most favorable renegotiated is

$$U(C_1^1\left(C_1^0\right)) \leq U(C_1^0)$$

Substituting 40 into this and re-arranging allows us to write the no-renegotiation constraint as the condition:

$$u(c_1^0) + \beta u(c_2^0) \geq (1+\beta^{\frac{1}{\rho}})u\left(\frac{c_1^0 + c_2^0 - \kappa}{1+\beta^{\frac{1}{\rho}}}\right) \tag{41}$$

The same no-renegotiation constraint can be derived starting from the assumption of period 1 monopoly. The most favorable renegotiation for the monopolist is:

$$C_1^{m1}\left(C_1^0\right) = \left(\frac{(c_1^0)^{1-\rho} + \beta(c_2^0)^{1-\rho}}{1+\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}} \cdot \left(1, \beta^{\frac{1}{\rho}}\right) \tag{41}$$

The contract will not be renegotiated so long as the profits gains to this most favorable renegotiation fall short of renegotiation costs:

$$\Pi_1\left(C_1^{m1}\left(C_1^0\right); C_1^0\right) = \left(c_1^0 + c_2^0 - \kappa\right) - \left((c_1^0)^{1-\rho} + \beta(c_2^0)^{1-\rho}\right)^{\frac{1}{1-\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}} \leq \kappa \tag{41}$$

This can be rearranged to yield the same condition as (41).

8.2.1 No-renegotiation condition

Substituting from 41 in the no-renegotiation condition (17), we get the following explicit no-renegotiation condition:

$$u(c_1^0) + \beta u(c_2^0) \leq u\left(\frac{c_1^0 + c_2^0 - \kappa}{1 + \beta^{\frac{1}{\rho}}}\right)(1 + \beta^{\frac{1}{\rho}}) \tag{41}$$

This condition applies identically whether contract renegotiation happens under competition or monopoly.

8.2.2 No-renegotiation condition for full-smoothing contracts

Setting $c_1^0 = c_2^0$ in the no-renegotiation constraint (**??**) above we can re-arrange to the constraint as:

$$\kappa \geq c_1^0 \cdot \Upsilon \tag{41}$$

where

$$\Upsilon = \left[2 - \left[\frac{(1+\beta)}{\left(1 + \beta^{\frac{1}{\rho}}\right)^{\rho}}\right]^{\frac{1}{1-\rho}}\right] \tag{41}$$

8.3 Imperfect-Smoothing Commitment Contracts

Redefine any consumption stream in the following manner:

$$C_0 = (c_0, c_1, c_2) \equiv (c_0, \alpha s, (1 - \alpha)\, s) \tag{41}$$

so that c_1 and c_2 are expressed as shares of total future consumption s. Since the no-renegotiation constraint places restrictions on the relative values of c_1 and c_2, we can rewrite the constraint (**??** using the new notation to get a continuous function $\alpha\,(s)$, which determines the mimimum fraction of any s that must be offered to One-self to prevent renegotiation:

$$(s)\left(1 - \left(\alpha^{1-\rho} + \beta\,(1-\alpha)^{1-\rho}\right)^{\frac{1}{1-\rho}}\left(1 + \beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right) \leq \kappa \tag{41}$$

Observe that at One-self's optimal division of s, $\left(c_2 = \beta^{\frac{1}{\rho}} c_1 \Longleftrightarrow \alpha = \frac{1}{1+\beta^{\frac{1}{\rho}}}\right)$, there cannot be profit gains from renegotiation so the constraint will be slack. For any s, there may be two values of α that satisfy the constraint with equality–one with α smaller than One-self would like (lower boundary), and another with α larger than One-self would like (upper boundary). Assuming the full-smoothing contract does not satisfy the constraint, the second-best contract must lie on the lower boundary. This defines a continuous function $\alpha(s)$, which determines the mimimum fraction of any s that must be offered to One-self to prevent renegotiation.

$$\alpha(s) = min\left\{\alpha : (s)\left(1 - \left(\alpha^{1-\rho} + \beta(1-\alpha)^{1-\rho}\right)^{\frac{1}{1-\rho}}\left(1 + \beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right) = \kappa\right\} \tag{41}$$

It can easily be verified that $\alpha'(s) > 0$ (profits from renegotiation rise in s, so if s rises there must be an increase in the share allocated to 1-self to compensate). Implicitly differentiating the binding no-renegotiation constraint by s, we have:

$$\frac{d\alpha}{ds} = \left(\frac{k}{s^2}\right)\left(\frac{1+\beta^{\frac{1}{\rho}}}{\alpha^{1-\rho} + \beta(1-\alpha)^{1-\rho}}\right)^{\frac{\rho}{1-\rho}}\left(\frac{1}{\alpha^{-\rho} - \beta(1-\alpha)^{-\rho}}\right) \tag{41}$$

The terms in the first two sets of parentheses are always positive. The last term is positive when the no-renegotiation constraint is binding (One-self would ideally like $\alpha^{-\rho} = \beta(1-\alpha)^{-\rho}$ but if $\kappa > 0$ she has to settle for $\alpha^{-\rho} > \beta(1-\alpha)^{-\rho}$.

Finally, for any s and α, let

$$V(s,\alpha) \equiv \beta\left[u(\alpha s) + u((1-\alpha)s)\right] \tag{41}$$

This is the discounted utility over periods 1 and 2, from period 0's perspective. It will be useful to note that the first-order conditions of the full-smoothing contract problems (competition and monopoly) can be written as:

$$\frac{du(c_0)}{dc_0} = \frac{dV\left(s,\frac{1}{2}\right)}{ds} \tag{41}$$

8.3.1 Sophisticated Hyperbolic Discounters

Proof of Proposition 2: (i) Since the full-commitment profit-maximizing contract was uniquely determined, and since it does not satisfy the renegotiation-proofness constraint, the renegotiation-proof contract must yield lower profits than the full-commitment contract does.

(ii) Using the modified notation, the full-smoothing contract terms are c_0^{mF} and s^{mF}, with $\alpha^{mF} = \frac{1}{2}$. The imperfect-smoothing contract terms are c_0^{mP} and s^{mP}, with $\alpha^{mP} =$

$\alpha\left(s^{mP}\right)$. Suppose $c_0^{mP} \leq c_0^{mF}$. Then, to satisfy Zero-self's participation constraint,

$$V\left(s^{mP}, \alpha\left(s^{mP}\right)\right) \geq V\left(s^{mF}, \frac{1}{2}\right) \tag{41}$$

$$\Rightarrow s^{mP} \geq s^{mF}\left[\frac{\left(\frac{1}{2}\right)^{1-\rho}+\left(\frac{1}{2}\right)^{1-\rho}}{\left(\alpha^{mP}\right)^{1-\rho}+\left(1-\alpha^{mP}\right)^{1-\rho}}\right]^{\frac{1}{1-\rho}} \tag{41}$$

Differentiating $V\left(s^{mP}, \alpha^{mP}\right)$, we get the following inequalities:[37]

$$\begin{aligned}
\frac{dV\left(s^{mP}, \alpha^{mP}\right)}{ds} &= \frac{\partial V\left(s^{mP}, \alpha^{mP}\right)}{\partial s} + \frac{\partial V\left(s^{mP}, \alpha^{mP}\right)}{\partial \alpha}\frac{d\alpha^{mP}}{ds} && (41)\\
&< \frac{\partial V\left(s^{mP}, \alpha^{mP}\right)}{\partial s} && (41)\\
&= \beta\left(s^{mP}\right)^{-\rho}\left[\left(\alpha^{mP}\right)^{1-\rho}+\left(1-\alpha^{mP}\right)^{1-\rho}\right] && (41)\\
&\leq \beta\left(s^{mF}\right)^{-\rho}\left[\frac{\left(\frac{1}{2}\right)^{1-\rho}+\left(\frac{1}{2}\right)^{1-\rho}}{\left(\alpha^{mP}\right)^{1-\rho}+\left(1-\alpha^{mP}\right)^{1-\rho}}\right]^{\frac{-\rho}{1-\rho}}\left[\left(\alpha^{mP}\right)^{1-\rho}+\left(1-\alpha^{mP}\right)^{1-\rho}\right] && (41)\\
&= \beta\left(s^{mF}\right)^{-\rho}\left[\left(\frac{1}{2}\right)^{1-\rho}+\left(\frac{1}{2}\right)^{1-\rho}\right]\left[\frac{\left(\alpha^{mP}\right)^{1-\rho}+\left(1-\alpha^{mP}\right)^{1-\rho}}{\left(\frac{1}{2}\right)^{1-\rho}+\left(\frac{1}{2}\right)^{1-\rho}}\right]^{\frac{1}{1-\rho}} && (41)\\
&< \beta\left(s^{mF}\right)^{-\rho}\left[\left(\frac{1}{2}\right)^{1-\rho}+\left(\frac{1}{2}\right)^{1-\rho}\right] && (41)\\
&= \frac{dV\left(s^{mF}, \alpha^{mF}\right)}{ds} && (41)\\
&= \frac{du\left(c_0^{mF}\right)}{dc_0^{mF}} && (41)\\
&\leq \frac{du\left(c_0^{mP}\right)}{dc_0^{mP}} && (41)
\end{aligned}$$

Since $\frac{dV\left(s^{mP},\alpha^{mP}\right)}{ds} < \frac{du\left(c_0^{mP}\right)}{dc_0^{mP}}$, this contract cannot be profit maximizing for the monopolist (it could do better by reallocating consumption away towards Zero-self). This contradiction implies that our assumption is incorrect. It must be true that at the profit-maximizing imperfect-smoothing contract, $c_0^{mP} > c_0^{mF}$. $\square$

Proof of Proposition 3: (i) We know that $U_0\left(C_0^F\right) = U_0^F$. By assumption, since the

[37] An explanation of the steps: Line 8.3.1 follows from the fact that $\alpha(s)$ rises in s (derived from Equation 8.3) and V falls as α rises, making the allocation worse from Zero-self's perspective. Line 8.3.1 follows from Inequality 8.3.1. Line 8.3.1 follows from the FOC of the monopolist's profit-maximization problem with full-smoothing contracts.

renegotiation-proofness constraint is binding, the renegotiation-proof contract cannot offer the optimal consumption path. Therefore $U_0\left(C_0^P\right) < U_0\left(C_0^F\right)$.

(ii) At the full-commitment contract:

$$\frac{du\left(c_0^F\right)}{dc_0} = \frac{dV\left(s^F, \frac{1}{2}\right)}{ds} = \left(s^F\right)^{-\rho}\left(2\left(\frac{1}{2}\right)^{1-\rho}\right) \tag{41}$$

Consider a renegotiation-proof contract with $c_0 = c_0^F$. To keep bank profits zero, this contract would also have $s = s^F$. But in the renegotiation-proof contract, s must be divided according to $\alpha\left(s^F\right)$. So:

$$\begin{aligned}\frac{dV\left(s^F, \alpha\left(s^F\right)\right)}{ds} &= \left(s^F\right)^{-\rho}\left(\alpha\left(s^F\right)^{1-\rho} + \left(1-\alpha\left(s^F\right)\right)^{1-\rho}\right) \\ &+ \frac{d\alpha\left(s^F\right)}{ds}\left(s^F\right)^{1-\rho}\left(\alpha\left(s^F\right)^{-\rho} - \left(1-\alpha\left(s^F\right)\right)^{-\rho}\right)\end{aligned} \tag{40}$$

The first term–the direct effect of a change in s–is weakly less than $\frac{dV\left(s^F,\frac{1}{2}\right)}{ds}$ if $\rho \leq 1$ and strictly greater if $\rho > 1$. The second term–the component of $\frac{dV}{ds}$ that is driven by the change in α–is strictly negative. Therefore, if $\rho < 1$, $\frac{dV\left(s^F,\alpha\left(s^F\right)\right)}{ds} < \frac{dV\left(s^F,\frac{1}{2}\right)}{ds} = \frac{du\left(c_0^F\right)}{dc}$, so the renegotiation-proof contract must satisfy $c_0^P > c_0^F$.

Next, we consider the case when $\rho > 1$. We can make the following observations about $\alpha(s)$. First, $\lim_{\kappa\to 0}\alpha(s) = \frac{\beta^{\frac{-1}{\rho}}}{1+\beta^{\frac{-1}{\rho}}}$. Second, implicitly differentiating equation 8.3 with respect to s, and combining it with the previous limit result, we get $\lim_{\kappa\to 0}\frac{d\alpha(s)}{ds} = 0$. Therefore, if $\rho > 1$ and κ is small enough, the second term in Equation 8.3.1 will be sufficiently small in magnitude that $\frac{dV\left(s^F,\alpha\left(s^F\right)\right)}{ds} > \frac{dV\left(s^F,\frac{1}{2}\right)}{ds} = \frac{du\left(c_0^F\right)}{dc}$. In this case, the renegotiation-proof contract must satisfy $c_0^P < c_0^F$. $\square$

If $\kappa = 0$, the renegotiation-proof contracts can be explicitly derived since in any contract it must be true that $c_2 = \beta^{\frac{1}{\rho}}c_1$. Solving the respective maximization problems, we get the following equilibrium contracts for monopoly and competition, respectively:

$$C_0^{mP} = \left(\left(\frac{U_0^A (1-\rho)}{1+\beta^{\frac{1}{\rho}} \left(\frac{\left(1+\beta^{\frac{1-\rho}{\rho}}\right)^{\frac{1}{\rho}}}{\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{1-\rho}{\rho}}} \right)} \right)^{\frac{1}{1-\rho}}, \left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}} \right)^{\frac{1}{\rho}} c_0^{mP}, \beta^{\frac{1}{\rho}} \left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}} \right)^{\frac{1}{\rho}} c_0^{mP} \right) \tag{40}$$

$$C_0^P = \left(\frac{y}{1+\beta+\beta^{\frac{1}{\rho}}}, \left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}} \right) c_0^P, \beta^{\frac{1}{\rho}} \left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}} \right) c_0^P \right) \tag{40}$$

It can easily be established that $c_0^{mP} > c_0^{mF}$, $c_0^P > c_0^{mF}$ if $\rho > 1$, and $c_0^P < c_0^{mF}$ if $\rho < 1$.

8.3.2 Naive Hyperbolic Discounters

Suppose the monopolist intends to renegotiate the contract. The maximization problem, combined with the expression for $C_1^{m1}(C_1)$ (41), simplifies to:

$$\underset{c_0,c_1,c_2}{max}\ y - c_0 - \frac{\left(c_1^{1-\rho} + \beta c_2^{1-\rho}\right)^{\frac{1}{1-\rho}}}{\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{\rho}{1-\rho}}} - \kappa \tag{40}$$

$$s.t. \frac{c_0^{1-\rho}}{1-\rho} + \beta \frac{c_1^{1-\rho}}{1-\rho} + \beta \frac{c_2^{1-\rho}}{1-\rho} \geq U_0^A \tag{40}$$

The partial derivatives of the resulting Lagrangian are:

$$\frac{\partial \mathcal{L}}{\partial c_0} = -1 - \lambda c_0^{-\rho} \tag{40}$$

$$\frac{\partial \mathcal{L}}{\partial c_1} = c_1^{-\rho} \left[-\left(\frac{c_1^{1-\rho} + \beta c_2^{1-\rho}}{1+\beta^{\frac{1}{\rho}}} \right)^{\frac{\rho}{1-\rho}} - \lambda \beta \right] \tag{40}$$

$$\frac{\partial \mathcal{L}}{\partial c_2} = c_2^{-\rho} \left[-\beta \left(\frac{c_1^{1-\rho} + \beta c_2^{1-\rho}}{1+\beta^{\frac{1}{\rho}}} \right)^{\frac{\rho}{1-\rho}} - \lambda \beta \right] \tag{40}$$

An interior solution, with $\frac{\partial \mathcal{L}}{\partial c_1} = 0$ and $\frac{\partial \mathcal{L}}{\partial c_2} = 0$ does not exist (on a $c_1 - c_2$ plot, the two first-order conditions do not intersect). If $\rho < 1$, the Lagrangian is maximized at a corner solution with $c_1 = 0$. If $\rho > 1$, the Lagrangian is maximized at the limit as c_2 approaches infinity. Using this, the maximization problem can be re-solved. If $\rho < 1$:

$$C_0^{mN} = \left(\left(\frac{U_0^A (1-\rho)}{2+\beta^{\frac{1}{\rho}}} \right)^{\frac{1}{1-\rho}}, 0, \left(\frac{1+\beta^{\frac{1}{\rho}}}{\beta} \right)^{\frac{1}{1-\rho}} \left(\frac{U_0^A (1-\rho)}{2+\beta^{\frac{1}{\rho}}} \right)^{\frac{1}{1-\rho}} \right) \tag{40}$$

If $\rho > 1$, the solution is undefined, but in the limit is given by:

$$C_0^{mN} = \left(\left(\frac{U_0^A (1-\rho)}{1+\left(1+\beta^{\frac{1}{\rho}}\right)\beta^{\frac{1}{\rho}}} \right)^{\frac{1}{1-\rho}}, \beta^{\frac{1}{\rho}} \left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{1}{1-\rho}} \left(\frac{U_0^A (1-\rho)}{1+\left(1+\beta^{\frac{1}{\rho}}\right)\beta^{\frac{1}{\rho}}} \right)^{\frac{1}{1-\rho}}, \infty \right) \tag{40}$$

Let us define profits from such a contract as:

$$\Pi_0^{mN} \equiv \Pi\left(C_0^{mN}; Y_0\right) + \Pi_1\left(C_1^{m1}\left(C_1^{mN}\right); C_1^{mN}\right) - \kappa$$

Proof of Proposition 4: (i) and (ii) are simultaneously established through the following observations. First, Π_0^{mN} is strictly falling in κ while $\Pi_0\left(C_0^{mF}; Y_0\right)$ is invariant in κ. Second, at $\kappa = \bar{\kappa}^m$,

$$\Pi_0\left(C_0^{mF}; Y_0\right) = \Pi_0\left(C_0^{mF}; Y_0\right) + \Pi_1\left(C_1^{m1}\left(C_1^{mF}\right); C_1^{mF}\right) - \kappa < \Pi_0^{mN} \tag{40}$$

Third, if κ gets indefinitely large, $\Pi_0\left(C_0^{mF}; Y_0\right) > \Pi_0^{mN}$. Finally, it can be verified from the explicit derivations that $c_0^{mN} < c_0^{mF}$. □

We now derive equilibrium contracts for naive consumers under perfect competition. Suppose contracts are exclusive. Then, a contract that is renegotiated satisfies:

$$\max_{c_0,c_1,c_2} \frac{c_0^{1-\rho}}{1-\rho} + \beta\frac{c_1^{1-\rho}}{1-\rho} + \beta\frac{c_2^{1-\rho}}{1-\rho} \tag{40}$$

$$s.t. y - c_0 - \frac{\left(c_1^{1-\rho} + \beta c_2^{1-\rho}\right)^{\frac{1}{1-\rho}}}{\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{\rho}{1-\rho}}} - \kappa \geq 0 \tag{40}$$

The first-order conditions are the same as under monopoly (8.3.2, 8.3.2, 8.3.2). Combining these with the zero-profit constraint, we get the following solution. If $\rho < 1$:

$$C_0^N = \left(\frac{y-\kappa}{2+\beta^{\frac{1}{\rho}}}, 0, \left(\frac{1+\beta^{\frac{1}{\rho}}}{\beta} \right)^{\frac{1}{1-\rho}} \left(\frac{y-\kappa}{2+\beta^{\frac{1}{\rho}}} \right) \right) \tag{40}$$

If $\rho > 1$, the solution is undefined, but in the limit is given by:

$$C_0^N = \left(\frac{y-\kappa}{1+\beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)}, \beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{1}{1-\rho}}\left(\frac{y-\kappa}{1+\beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)}\right), \infty \right) \tag{40}$$

Proof of Proposition 5:

(a) Under non-exclusive contracts, firms offering period 0 contracts do not benefit from renegotiation (profits from renegotiation will equal κ). So the equilibrium contract is the one that is arrived at without taking renegotiation into account–i.e. the full-commitment contract. If $\kappa < \bar{\kappa}$, the gains from renegotiation exceed the transaction costs, so the contract will be renegotiated.

(b) The following observations establish part (b). First, $U_0\left(C_0^N\right)$ is strictly falling in κ while $U_0\left(C_0^F\right)$ is invariant in κ. Second, at $\kappa = \bar{\kappa}$, $U_0\left(C_0^N\right) > U_0\left(C_0^F\right)$ (this must be true by construction of C_0^N). Third, if κ gets indefinitely large, $U_0\left(C_0^N\right) < U_0\left(C_0^F\right)$, so Zero-self will prefer the full-smoothing commitment contract over the renegotiable contract.

Suppose $\rho < 1$. Comparing C_0^F (37) to C_0^N (8.3.2), it is clear that $c_0^N < c_0^F$. Suppose $\rho > 1$. If κ is small enough, $c_0^N > c_0^F$. □

8.4 Nonprofits

Proof of Proposition 6: (a) A non-profit will earn higher raw profits Π than a for-profit. If $f\left(\Pi\left(C_0^{mNP}; Y_0\right)\right) \geq \Pi\left(C_0^{mP}; Y_0\right)$ (i.e. if the captured profit function has a slope sufficiently close to 1 up to $\Pi\left(C_0^{mNP}; Y_0\right)$), the firm will choose to operate as a nonprofit.

(b) (i) Suppose all firms are for-profit and offer the renegotiation-proof contract C_0^P. There is some ε_1 and ε_2 satisfying $0 < \varepsilon_2 < \varepsilon_1$ and a corresponding $\hat{C}_0 = \left(c_0^P, c_1^P - \varepsilon_1, c_2^P + \varepsilon_2\right)$ such that $U_0\left(\hat{C}_0\right) = U_0\left(C_0^P\right)$ and

$$f\left(\Pi_0\left(\left(\hat{C}_0\right); Y_0\right) + \Pi_1\left(C_1^{m1}\left(\hat{C}_1\right); \hat{C}_1\right)\right) < \kappa$$

So, any firm can make positive profits by operating as a non-profit. Therefore, in equilibrium, consumers will borrow only from non-profit firms.

(ii) If all firms are nonprofit, an individual firm has a strict incentive to switch to for-profit status, and make profits in period 1. Therefore, there must be for-profits in equilibrium, and equilibrium contracts will be constrained by their presence. □

References