Breakable commitments: present-bias, client protection and bank ownership forms

Karna Basu & Jonathan Conning*

September 6, 2016

Abstract

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 do they instead opportunistically pander to consumer biases. By how much might financial trade reduced by fear of such opportunism? We study a consumer protection problem that survives even when consumers are sophisticated and fully informed. The consumer would like to commit her future selves to a balanced path of saving accumulation *cum* debt repayment via a commitment contract, but her future selves can be tempted to raid savings or take on new debt. Investor-owned banks find it costly to credibly commit to not engage in such opportunistic 'exploitation,' leading to lower ex-ante consumer welfare and/or bank profits and lost trade. In such contexts strategic adjustments to bank ownership/governance forms may allow more credible (or renegotiation-proof) commitment contracts at lower cost, because limitations on the ability to distribute profits can signal weaker contract renegotiation incentives. This forms the basis for a theory of commercial non-profits and endogenous client protection similar to Hansmann (1996) but on new behavioral micro-foundations. Depending on market concentration and the legal environment, non-profit or 'hybrid' ownership forms (e.g. social investors tempering purely commercial objectives) may be able to offer commitment services at lower cost. The framework helps understand the evolution of ownership and contracts in consumer banking, microfinance, and mortgage and payday

*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 comments and suggestions, we thank Abhijit Banerjee, Partha Deb, Eric Van Tassel, and seminar and conference participants at the 2016 NBER Summer Institute Development Workshop, Michigan State University, IIES Stockholm, Delhi School of Economics, Queens College, Indian School of Business, Hunter College, Florida Atlantic University, CIDE-ThReD Conference on Development Theory, and NEUDC. Kwabena Donkor provided excellent early research assistance.

lending and for evaluating proposed consumer protection measures against excessive refinancing and 'overindebtedness' in these sectors. JEL Codes: O16, D03, D18

1 Introduction

Hyperbolic discounters – consumers with present-biased preferences – are those who struggle to stick to their long-term asset accumulation or debt management plans and for this reason may benefit from the 'commitment services' or designed restrictions that a financial intermediary may build into multi-period financial contracts. Long-term consumption smoothing goals are more likely to be achieved where customers can enlist financial intermediaries to act as partners in resisting temptations to raid savings and/or increase debt in ways that could undermine responsible long-term plans. The properties of *costlessly*-enforced optimal commitment banking contracts have been examined in several theoretical papers (Laibson (1997)) and a number of randomized controlled trials and other empirical studies have found quite large impacts from introducing new commitment service savings products (see for example Thaler & Benartzi (2004), Ashraf *et al.* (2006) and the studies mentioned in the survey by Bryan *et al.* (2010)).

The assumption of costless commitment services is hard to sustain. Commitments can be broken and the same person that at first demanded a commitment contract from a bank may later want to renegotiate its terms. A bank's ex-post profits may well be increased by pandering to such demands. For example, a present-biased consumer who takes out a loan that promises to balance repayments and expected consumption across future periods may at a later period find themselves unable to resist the temptation to refinance to take out an expensive (but at this later date attractive) additional consumption loan that drives up present consumption but wrecks earlier laid consumption smoothing plans.

This type of problem raises obvious consumer protection concerns in the case of naive or partially naive present-biased consumers as these are consumers who, by definition, do not understand how their own future preferences will evolve to make them vulnerable to such opportunistic exploitation by a bank. In the absence of consumer protection a bank could attempt to lure customers into contracts that it knows can be later renegotiated for additional profit.

In the case of sophisticated present-biased consumers consumer protection issues also arise but take the less obvious and harder to measure form of lost or reduced financial trade. Sophisticated consumers anticipate the possibility of opportunistic contract renegotiation and therefore will be cautious to minimize their vulnerability by insisting on renegotiation-proof contracts that must be endogenously enforced. But imposing any such additional constraints can only reduce the feasible financial contract space, reducing gains of trade and therefore lowering bank profits and/or consumer welfare or lost trade. The practical result in either case will be equilibria with less than optimal consumption smoothing with consumes saving 'too little' and/or borrowing 'too much' (in the estimation of their earlier period selves) in some periods compared to situations where full-commitment contracts could be costlessly offered.

Concerns over excessive refinancing and 'over-indebtedness' have been raised repeatedly in recent years in both the developed and the developing world. On the eve of the mortgage banking crisis in 2007, over 70 percent of new subprime mortgage loans were refinances of existing mortgages and approximately 84 percent of these were 'cash out' refinances (Demyanyk & Van Hemert (2011)). In the market for payday loans the concern of many economists and regulatory observers is not so much that fees are high for one-time short-term loans (typical cost is 15% of the amount borrowed on a 2 week loan) but rather that over 80 percent of payday loans are in fact 'rolled over' or renewed rather than paid off, incurring new fees each time resulting in very high total loan costs and placing many people into very difficult debt management situations (DeYoung (2015).[1]

Similar concerns have been raised in microfinance markets in developing countries but in such markets where government consumer protection laws are weak to non-existent, the issues take on an interesting extra dimension. To state our hypothesis briefly and bluntly: microfinance markets around the world have been dominated by a mix of mission-driven and commercial non-profits and 'hybrid' firms – for example for-profit firms partly or entirely owned and controlled by not-for-profit foundations or governments (Conning & Morduch (2011); Cull *et al.* (2009)). We argue that this is because, compared to investor-led for-profit firms hybrid and non-profit firms can make more credible commitments to not engage in the type of opportunism described above, because their incentive to do so is muted by the fact that they cannot easily distribute any profits. This is a variation on Henry Hansmann's theory of commercial non-profits but set on new behavioral micro-foundations that we will model explicitly below. In the early stages of new microfinance markets one or more non-profit firms will dominate and successfully provide the commitment services in savings and credit contracts. However as market competition intensifies the capturable advantages of strategic non-profit status are diminished and this can lead to a 'commercialization' stage and the conversion of many commercial from non-profit to for-profit form. This will be associated with a rise of of the type of consumer protection issues identified above and concomitant collapse of borrower and saver discipline. This helps make sense of the string of repayment crises in places as far flung as Morocco, Bosnia, Nicaragua, and Andra Pradesh in India where previously very successful microfinance markets followed similar trajectories

In this paper we view the provision of commitment as an important element of consumer

[1] Approximately half payday loans are rolled over in sequences lasting 7 or more turns

protection in banking. Our goal is to provide a simple dynamic framework for analyzing real-world settings where commitment is demanded but cannot be credibly provided at zero cost. We work with a quite general three-period consumption smoothing model for a sophisticated 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. They correctly anticipate that their later period-one 'One self' will have a change of preferences that will lead them 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 One 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 the first part of the paper we characterize the contracts associated with three broad scenarios: (1) **'Own smoothing' contracts** where the consumer saves on their own without a bank that might offer commitment services; (2) **Full-commitment contracts** where we assume the Zero self consumer can enter into multi-period contract that binds their One self to not renegotiate the contract terms under the assumption that this bank commitment can be costlessly enforced; and (3) **renegotiation-proof contracts** where the zero self consumer can enter into a multi-period contract to bind their One self to not renegotiate the contract terms but where the bank's commitment must be endogenously enforced by making sure that bank's ex-post profits from breaking their commitment fall short of any direct renegotiation costs to the bank summarized by a parameter κ. We describe the terms of renegotiation-proof contracts and the distribution of gains to trade between bank and consumer and how these are parameterized by κ. At high enough values the bank can credibly commit to not profit from renegotiation and the first-best costless full-commitment contract can be achieved. In the second part of the paper, described below, we explore how a bank might make costly strategic changes to its bank ownership and governance structures as another way to credibly commit to not renegotiate.

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. Separately we also characterize the terms of the contract that would be offered to a naive present-biased consumer by a bank that would strategically renegotiate the contract.

With sophisticated hyperbolic discounters, the bank wishes it could credibly commit to not renegotiate in the future. Its choice of governance and ownership structures may help it do so. A nonprofit firm, for example, faces legal restrictions on its enjoyment of earnings. So, if the bank operates as a nonprofit, it can demonstrate to the consumer that it has less to gain from future renegotiation. This improves its ability to commit to a commitment contract, which allows it to extract more surplus from the consumer. Whether it chooses to adopt nonprofit status depends on the extent to which legal restrictions improve its ability to commit while protecting its ability to enjoy the additional surplus extracted through commitment. 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.

Our model considers nonprofits more broadly than in the above example, and analyzes governance choices across consumer types and market settings. A monopolist bank prefers to operate as a nonprofit if consumers need relatively small adjustments to their consumption. In competitive markets, exclusivity plays an important role–if contracts are exclusive, firms operate as nonprofits; otherwise, for-profits dominate.

Banks, under both monopoly and competition, 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 more broadly 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. We show how the details of the trade-off depend on the governance choices available to the bank and the consumer's reservation outcomes.

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 trade-off 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 non-profit microfinance, which is rigid, and say commercial credit card lending which offers refinancing flexibility (credit card punishments gain salience because they are *less* strict, not more).

The model is indeed stylized and limited to one of many mechanisms, but we argue that it makes a number of compelling points relevant to ongoing policy debates and is able to explain some stylized facts.

1.1 Context and Related Literature

1.1.1 Commitment as Consumer Protection

Problems of consumer protection are typically analyzed through two channels: naive or uneducated consumers and their failure to correctly anticipate fees and punishments (see Gabaix & Laibson (2006), Armstrong & Vickers (2012), and Akerlof & Shiller (2015) for related arguments), and bank's moral hazard (see Dewatripont & Tirole (1999) and Oak & Swamy (2010)). We argue that, given the growing evidence of time-inconsistent preferences,[2] 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.[3] 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

[2]See, for example, Laibson *et al.* (2003), Ashraf *et al.* (2006), Gugerty (2007), and Tanaka *et al.* (2010).

[3]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.

the United States. In each of these cases the issue of refinancing or the taking of loans from multiple lenders emerges. For example, by value, more than two-thirds of the subprime mortgage loans outstanding in 2008 were made up of refinanced loans (Demyanyk & Van Hemert (2011)).

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 the use of coercive loan recovery methods.

We place consumers' struggles with intertemporal self-control issues at the center of the 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.[4]

A sophisticated hyperbolic discounter understands that his 'future selves' will attempt to take out new loans on top of old ones, or renegotiate the terms of existing loans to further defer debt repayment. If she is a saver, she will try to withdraw more rapidly in the future than she would have liked, or will deposit less than ideal. That such consumers should be and are willing to pay for commitment has been demonstrated in several theoretical and empirical papers (for examples of commitment, see Ashraf *et al.* (2006), Carrillo & Dewatripont (2008), Bryan *et al.* (2010), Fischer & Ghatak (2010), Banerjee & Duflo (2011), and Basu (2011)). By entering into a contract that commits them to a specific time-path of repayments or deposits, the consumer attempts to ensure that future selves will not skew consumption patterns to privilege instant gratification in the following periods. Viewed this light, fees and punishments for failing to adhere to a schedule are not inherently undesirable to the consumer–for sophisticated hyperbolic discounters, threats of punishment can indeed serve as useful commitment devices.

Nevertheless, the fact that consumers value commitment does not automatically imply that firms will provide it in equilibrium. In fact, there remains an open question about whether markets can be relied upon to supply commitment. The key consideration in this paper is the following: if a hyperbolic discounter is willing to pay to commit her future

[4] Bond *et al.* (2009) discuss evidence of predatory lending in the context of mortgages.

selves, her future selves are willing to pay to undo this commitment. 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. Gottlieb (2008) shows how competition leads to inefficient outcomes in immediate rewards goods. Heidhues & Koszegi (2010) study the mistakes of partially naive borrowers in competitive credit markets. Mendez (2012) analyzes predatory lending with naive consumers. Our emphasis is on the terms of any generic banking contract, analyzed broadly: under both competition and monopoly (the latter being particularly relevant to informal banking in developing economies), for both sophisticated an naive consumers, and for multiple governance and ownership structures.

1.1.2 Commercial Nonprofits

The idea that firm ownership might be strategically chosen to solve or ameliorate 'contract failure' problems dates back at least to Arrow (1963) and is one that has been articulated most clearly in the work of Henry Hansmann (1996). 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 be different from for-profit firms.[5] In this view "[a] nonprofit organization is, in

[5]Hence we abstract away from other considerations for nonprofits, as in Besley & Ghatak (2005), McIntosh & Wydick (2005), and Guha & Roy Chowdhury (2013). Nonetheless our modeling framework can be adapted to include these considerations and is the focus of related work.

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."[6] Glaeser & Shleifer (2001) 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 non-profit 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. Bubb & Kaufman (2013) 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 (Cull *et al.* (2009); Conning & Morduch (2011)). 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).

1.1.3 Market Structure and Governance Choice

Commenting upon a major microfinance crisis in the state of Andhra Pradesh in India, veteran microfinance analyst Elizabeth Rhyne (Rhyne (2011)) 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."

Her analysis, like that of many others, ultimately lays the blame on "poor governance structures." In her view, Indian MFIs might have avoided their problems and followed the

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

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. However, as our model demonstrates, these governance choices are highly dependent on market structure, and nonprofits may survive better under monopoly than under competition.

2 The Model: Setup

2.1 Consumers and financial intermediaries

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

$$u(c_t) = \begin{cases} \frac{c^{1-\rho}}{1-\rho} & \text{if } \rho > 0 \text{ and } \rho \neq 1 \\ ln(c) & \text{if } \rho = 1 \end{cases} \tag{1}$$

Given a consumption stream $C_t = (c_t, c_{t+1}, ..., c_2)$, 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 (Laibson (1997)) with a hyperbolic discount factor $\beta \in (0, 1]$. In any period t, the individual, or more accurately their period-t self, places greater relative weight on present-period gratification than her earlier selves would have done. The consumer could be sophisticated or naive about understanding the time-inconsistency of her preferences (O'Donoghue & Rabin (2001)).

It will be useful to establish simple labels to distinguish the consumer's different selves. We refer to the period-zero consumer as the 'Zero-self' or simply 'Zero' and her period-one incarnation as her 'One-self' or 'One'. Zero-self starts with claims to an arbitrary income stream over the three periods, $Y^0 = (y_0, y_1, y_2)$. Her objective will be to rearrange this into a utility maximizing consumption stream using what financial contracting savings and/or borrowing strategies are available.

The consumer has the option of contracting with one or many risk-neutral banks, depending on the market structure which could be monopolized or competitive. Each bank has access to funds that can be withdrawn from other investments at an interest rate r, which for most of the analysis we normalize to $r = 0$. A bank will participate and offer a contract if and only if it can earn non-negative profits. Because the remainder of the analysis will be focused on contracts that attempt to establish self-control, or that limit the consumer's later period selves' ability to recontract with the same or new financial intermediary, it will be important to distinguish legal or institutional environments that allow banks to establish exclusive contracts with customers and those that do not. We will analyze these cases in turn.

2.2 Full-commitment contracts under Competition

A consumer with time-inconsistent preferences cannot trust her later selves to stick to her preferred consumption plans. In this simple three-period setting Zero's concern is that her later One self will try to divert resources earmarked for period 2 consumption to boost period 1 consumption instead. Like a Stackelberg-Cournot game, Zero's stategic contracting choices are affected by her anticipation of One's best response. A bank may be able to act as a strategic partner to Zero by offering contracts with commitment services to help restrict or otherwise control the customer's later self(ves) recontracting possibilities.

In this section we study the benchmark competitive full-commitment contract where we simply assume a bank can offer the Zero consumer a multi-period contract that *costlessly* binds the consumer's latter self(ves) to not renegotiate its terms with the same bank or other banks. Although we are not making the exact mechanisms explicit yet the bank's ability to credibly commit to not renegotiate such a contract – even though *ex-post* the consumer is willing to pay the bank to do so – must rest on the assumption that the bank can credibly bond itself to incurring a high renegotiation penalty (call it κ) in the event of renegotiation that is sufficiently high to deter the bank.[7] In the next and later sections of the paper we will examine what happens when κ falls to zero and exogenously sustained commitment contracts must be replaced by self-enforcing renegotiation-proof contracts. We then study intermediate scenarios where κ may larger than zero but less than what is required to sustain full commitment. In later sections we delve deeper into what determines κ and how it might be endogenously determined via choices of bank ownership and governance forms, and shaped by the nature of the market structure.

With an exclusive full-commitment contract the consumer faces no self control problem. Zero chooses a contract in period 0 that costlessly commits her period-1 and period-2 selves to follow the chosen consumption plan. This is contract design problem is solved as a

[7]We label this the costless full-commitment contract but as this discussion suggests it rests upon the deterrence achieved by threatening very costly punishments to the bank. When the severity of the punishments that can be credibly threatened are less severe the costs of contracting will rise as the bank and consumer have to distort the consumption smoothing contract to make no-renegotiation commitments self-enforcing.

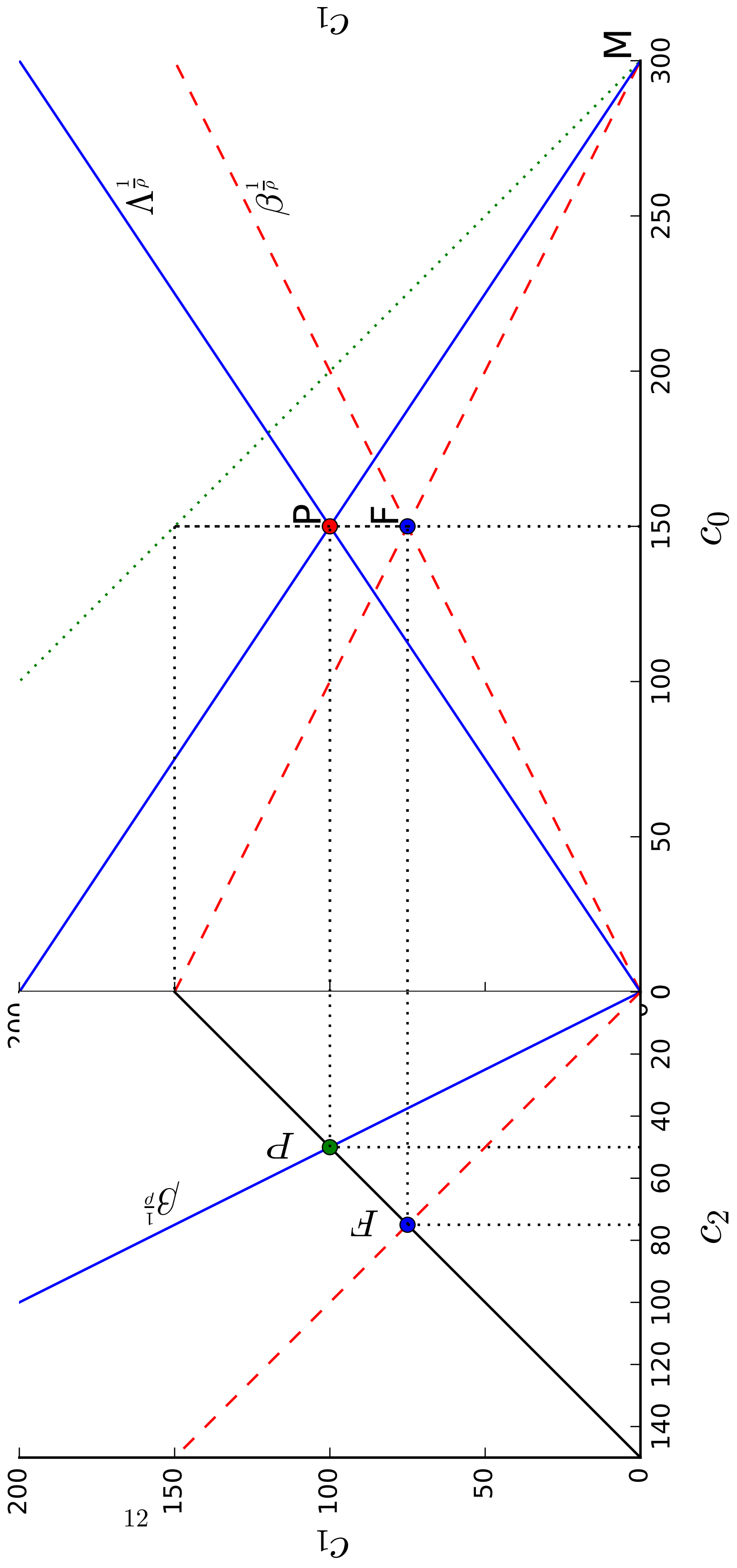

Figure 1: Self-control and the demand for commitment

standard utility maximization problem subject to an inter-temporal budget constraint (or, same thing, subject to a financial intermediary's zero-profit condition). Zero self chooses contract C^0 to solve:

$$\max_{C^0} U_0(C^0)$$

$$\Pi^0(C^0, Y^0) = \sum_{i=t}^{2} \frac{(y_t - c_t)}{(1+r)^i} \geq 0$$

The familiar first-order necessary conditions are:

$$u'(c_0) = \beta\delta(1+r)u'(c_1) = \beta\delta^2(1+r)^2 u'(c_2) \tag{3}$$

Along with a binding budget constraint these equations allow us to solve for the optimal competitive 'full-commitment' contract C^F so called because latter period selves are committed to not change it. Conceptually the equilibrium contract will be found at the tangency between the highest iso-utility surface just touching the budget hyper-plane. A 3D graphical analysis would be possible but is cumbersome so we draw insights from 2D contour projections instead.

With CRRA utility these three equations can be rewritten:

$$c_1 = [\beta\delta(1+r)]^{\frac{1}{\rho}} c_0 \tag{4}$$

$$c_2 = [\delta^2(1+r)^2]^{\frac{1}{\rho}} c_1 \tag{5}$$

This last equation and the budget constraint can be combined to give:

$$c_1 = \frac{\sum y - c_0}{1 + [\delta(1+r)]^{\frac{2}{\rho}}} \tag{6}$$

Equations 4 and 6 allow us to solve for c_0 and then the remaining periods. An increase or decrease to the term $\delta(1+r)$, which enters each expression above, essentially 'tilts' consumption to be more generally rising or falling over time as $\delta \gtreqless 1/(1+r)$. As this across-the-board level of tilt will not alter key tradeoffs of interest (unlike the degree of present-bias β parameter which will) we shall, without loss of generality, impose the assumption that $\delta = \frac{1}{1+r}$ for the remainder of the analysis. This will unclutter the math. Equations 4 to 6 now become simply

$$c_1 = \beta^{\frac{1}{\rho}} c_0,\ c_2 = c_1 \text{ and } c_1 = \frac{\sum y - c_0}{2} \tag{7}$$

In words: Zero self wants to indulge her present bias to tilt consumption toward period zero consumption but to then allocate to assure even consumption across the remaining two periods. A closed form solution for C_0 is easily found:

$$c_0^{0F} = \frac{\sum y}{1 + 2\beta^{\frac{1}{\rho}}} \tag{8}$$

$$c_1^{0F} = c_2^{0F} = \beta^{\frac{1}{\rho}} c_0^{0F}$$

These relationships and Zero's preferred contract solution can be seen in the two-panel diagram of Figure 1, drawn for a consumer with $\beta = 0.5$ and $\rho = 1$ and a present value of income $y = 300$ spread over the three periods (we are assuming $r = 0$).[8] The right hand is drawn in $c_0 - c_1$ contract space. Points along the ray OF satisfy first-order equation 4 while points along the ray YF satisfy the second tangency equation 5 and the budget line (i.e. they satisfy equation 6).[9] The intersection at F identifies the coordinates (c_0^F, c_1^F) of Zero's preferred contract, also given by equation 8. We've labeled this with subscript F to indicate it's essential similarity to the full-commitment contract that might be offered by a competitive bank.[10] The left hand panel depicts the situation in $c_1 - c_2$ space of the sliced plane at $c_0 = c_0^F$). This last diagram can be seen, now rotated 90 degrees to place c_1 on the horizontal axis, in Figure 2. Points along ray OF satisfy the Euler equation 5, while points along the line passing through DF satisfy the period 1 budget $c_1 + c_2 = \sum y - c_0$ where as drawn $c_0 = c_0^F$. Point F at the intersection of these lines identifies the coordinates (c_1^F, c_2^F) of the optimal contract.

[8]These parameter values are chosen for expositional purposes. In particular $\rho = 1$ implies that period zero consumption will remain the same with or without self-control which slightly simplifies the graphical exposition (by placing points F and P along the same first period budget line) but the analysis is easily adapted to other cases.

[9]A note on diagram interpretations may be helpful. Points that satisfy the first-order conditions lie along a ray from the origin that extends in the $(1, \beta^{\frac{1}{\rho}}, \beta^{\frac{1}{\rho}})$ direction in 3D space. The ray through OF in the right panel of Figure 1 is the projection of this ray in $c_1 - c_2$ space. As contracts are consumption bundles over three periods an iso-utility surface associated with a particular level of inter-temporal utility level given by $u(c_0) + \beta[u(c_1) + u(c_2)] = \hat{U}$. Projected onto $c_0 - c_1$ space this would be described by an entire family of contour lines, each contour line associated with a slice of the surface at a particular value of c_2. A contour line further from the origin in this $c_0 - c_1$ projection does no correspond to higher utility since every contour line in this projection is associated with the same level of utility. The left hand panel of the figure – shown in greater detail in Figure 2 – on the other hand depicts the particular subgame reached when period 0 consumption was already set at $c_0 = c_0^F$. The contour lines drawn in this figure are interpreted as conventional indifference curves with higher levels of utility as we move further from the origin. For example Zero's indifference curve $u(c_0^F) + \beta[u(c_1) + u(c_2)] = U(C^F)$ is the locus of all points (c_1, c_2) that yield the same level of utility to Zero as the full-commitment contract when $c_0 = c_0^F$. Zero's indifference curve passing through point P on the other hand also keeps $c_0 = c_0^F$ but offers an everywhere lower level of utility to Zero. The steeper indifference curves on the same diagram correspond to One's more present-biased preferences $u(c_0^F) + u(c_1) + \beta u(c_2)$

[10]This will be exactly the same full-commitment contract that a competitive bank intermediary would be led to offer if they faced an opportunity cost of funds r identical to what we have assumed here for the informal economy (normalized to $r = 0$ for convenience only): However since it is realistic to assume these rates will not be the same and that in practice consumers in the informal economy may face lower, possibly negative rates of return on informal savings and higher rates (and/or constrained access to loans.

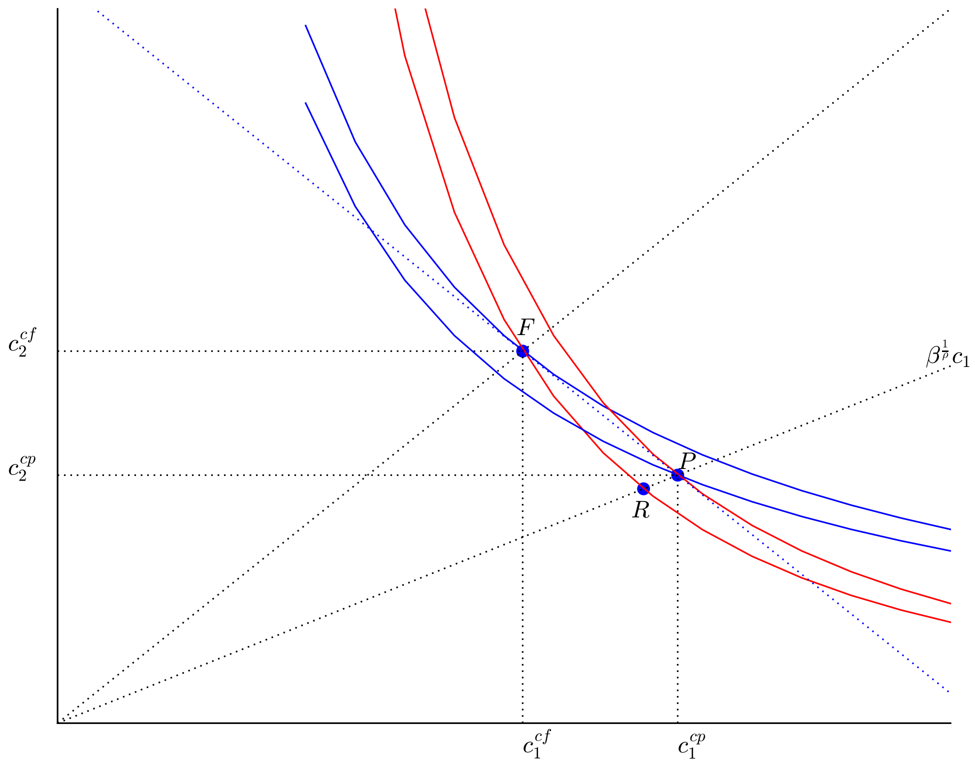

Figure 2: Full-commitment and renegotiation-proof contracts under competition (drawn at $c_0 = c_0^F, \beta = \frac{1}{2}, \rho = 1, \kappa = 0$)

Zero's preferred contract at these parameters is $C_0^F = (150, 75, 75)$. How much the consumer borrows or saves (or pays down debts) in any given period depends on the value of her income in that period relative the optimal consumption level in that period. If total income of 300 were evenly divided across periods as $Y_0 = (100, 100, 100)$ then this consumption plan would be achieved by borrowing $c_0 - y_0 = 50$ in period 0 and repaying 25 in each of periods 1 and 2 (recall $r = 0$). If the stream were instead $Y_0 = (200, 50, 50)$ the same consumption plan would be constructed by saving 50 in period 0 to be evenly allocated to raise consumption in periods 1 and 2.

2.3 The Renegotiation Problem

One would like higher period 1 consumption than Zero's preferences built into the contract. This will mean that are potential gains to trade to be split between the bank the One self by breaking earlier commitments. In the last section we simply assumed there were sufficiently high costlessly enforced penalties that deterred he bank from breaking such commitments.

Consider now then a competitive situation where there no penalties to the bank from breaking commitments and/or a situation where non-exclusivity clauses cannot be enforced so nothing stops the consumer from refinancing with another bank, for example taking out a second loan.

When period 1 arrives One-self will want to indulge her present-bias and increase period 1 consumption relative to Zero's preferences. This can be seen in Figure 2. Assume for the sake of argument that the consumer had (naively as it will turn out) accepted the contract C_0^F in period zero. At the start of period 1 this contract satisfies Zero self's optimality condition $u'(c_1^F) = u'(c_2^F)$ but from the standpoint of One's preferences it involves not enough period 1 consumption as $u'(c_1^{0F}) \geq \beta u'(c_2^{0F})$. This can be seen at point F, where Zero self's indifference curve is tangent to the budget line but One self's indifference curve is steeper than the budget line. The One self can gain by recontracting to any new tangency point above where their indifference curve through F cuts the $c_2 = \beta^{\frac{1}{\rho}} c_1$ ray (at point R – the point they'd be driven to if in period 1 they could only refinance with a monopoly bank) and below the budget line (at point P – the point they'd recontract to with competiive choices).[11] As the gains to trade are positive One can entice the bank to refinance by offering to compensate the bank for any renegotiation costs κ (at the moment assumed to be zero) less than that amount. From Zero's standpoint the bank would be opportunistically profiting by pandering to One self's impulse to 'raid savings' or 'roll over or pile on new debt.'

[11]When $\rho = 1$ (as is drawn in Figure 2) Zero will want to choose the same period 0 consumption level as the full-commitment contract. When $\rho \neq 1$ this will not be the case and the renegotiation-proof contract P will no longer lie exactly on the budget line defined by $c_0 + c_1 = \sum y - c_0^F$. The analysis is readily adapted.

As a sophisticate Zero anticipates how any such recontracting could lower their intertemporal welfare[12]. From the diagram it should be clear that Zero can do no better to limit the damage than to accept the contract at P (as any other preferred contract is not subgame-perfect and would be recontracted to P anyways). The subgame-perfect 'renegotiation-proof' contract will involve less net saving (or same thing, more net debt) in period 1 than Zero self would like (as can be seen by $u'(c_1^P) < u'(c_2^P)$) and lower Zero self welfare compared to the full-commitment contract.

The problem is like a Stackelberg game with Zero self moving first to choose a contract in anticipation of One self's best response recontracting. Consider One self's best response in a subgame defined by Zero self's choice of contract $C^0 = (c_0^0, c_1^0, c_2^0)$. One self can either accept the remaining consumption plan (c_1^0, c_2^0) or exchange it for a new contract to solve

$$\max_{c_1, c_2} u(c_1) + \beta u(c_2)$$

$$s.t.\ c_1 + c_2 \leq c_1^0 + c_2^0$$

As Zero self's budget constraint must bind at an optimum (otherwise consumption and utility could be increased) the period 1 budget can be rewritten

$$c_1 + c_2 \leq \sum y - c_0^0 \tag{9}$$

From which it is clear that all that One self needs to react to is Zero's period zero consumption c_0^0 which establishes the remaining resources to be divided between period 1 and 2. From equation 6 we know that Zero wants each extra unit that it passes forward into period 1 to be, on the margin, shared equally between period 1 and period 2 consumption (i.e. the slope of the line through YF is $\frac{dc_1}{dc_0} = -\frac{1}{2}$).

One wants to instead set $c_2 = \beta^{\frac{1}{\rho}} c_1$. Substituted this into the binding period one budget we can solve for One's best response or reaction function to Zero's period 0 consumption choice:

$$c_1^1(c_0^0) = \frac{\sum y - c_0^0}{1 + \beta^{\frac{1}{\rho}}} \tag{10}$$

$$c_2^1(c_0^0) = \beta^{\frac{1}{\rho}} \frac{\sum y - c_0^0}{1 + \beta^{\frac{1}{\rho}}} \tag{11}$$

One's reaction function is depicted by the line passing through YP in the left quadrant of Figure **??**fig:twoquad. As long as $\beta < 1$, it is easy to establish that line YP is everywhere steeper than line MF. For any given period 0 consumption level, One wants to consume

[12] A naive hyperbolic discounter, on the other hand, may believe she can achieve optimal utility though this will not be realized. See below.

more in period 1 and less in period 2 compared to Zero's preferences. Being a sophisticate, Zero anticipates her future self's reaction and chooses c_0 strategically to solve:

$$\max_{c_0} u(c_0) + \beta[u(c_1^1(c_0) + u(c_2^1(c_0))] \tag{12}$$

subject to the bank's period 0 participation or competitive zero-profit constraint.

$$c_0 + c_1^1(c_0) + c_2^1(c_0) = \sum y \tag{13}$$

From the fact that One will set $c_2^1 = \beta^{\frac{1}{\rho}} c_1^1$ and other substitutions the first-order condition for this problem can be written:

$$\begin{aligned} u'(c_0) &= \Lambda u'(c_1) \\ c_1 &= \Lambda^{\frac{1}{\rho}} c_0 \end{aligned} \tag{14}$$

where

$$\Lambda = \frac{(\beta + \beta^{\frac{1}{\rho}})}{1 + \beta^{\frac{1}{\rho}}} \tag{15}$$

Similar to what happens in a classic Stackelberg Cournot duopoly game our Zero self chooses the consumption contract that puts her on the highest iso-utility surface along One's reaction function. Equilibrium contract P in Figure satisfies both equation 10 (the line through MP) and equation 14 (the line through OP). These two equations can be solved to give a closed-form solution to the optimal contract in the absence of self-control

$$c_0^P = \frac{\sum y}{1 + \Lambda^{\frac{1}{\rho}}(1 + \beta^{\frac{1}{\rho}})} \tag{16}$$

with $c_1^P = \Lambda^{\frac{1}{\rho}} c_0^P$ and $c_2^P = \beta^{\frac{1}{\rho}} \Lambda^{\frac{1}{\rho}} c_0^P$.

As Figure 2 makes clear, Zero self has little choice but to accommodate to the fact One will grab consumption resources that Zero self would have preferred to earmarked for period 2 consumption. It is easy to verify that $\Lambda^{\frac{1}{\rho}} > \beta^{\frac{1}{\rho}}$ for all $0 < \beta < 1$ and hence that line OP is everywhere steeper than line OF. We can combine this with our earlier observation that line YP is everywhere steeper than line YF to conclude that $c_1^P > c_1^F$. In words, the absence of self control always leads to higher period one consumption compared to a situation where self-control can be exercised.

With a little extra algebra we can also show that $c_0^P \gtreqless c_0^F$ as $\rho \lesseqgtr 1$ for $0 < \beta < 1$ but also that $c_2^P < c_2^F$. That is, when $\rho < 1$ Zero responds to the self-control problem by reducing savings in period 0. When $\rho > 1$ she responds by increasing savings. However any increase in period 0 savings is never enough to lead to a net increase in period 2 consumption.

Indeed period 0 consumption adjustments are small relative to the increase in period 1 consumption. In other words even though they have a first-mover advantage, Zero can do little other than to largely accommodate to the consumption pattern that One self wants to impose. There is little Zero can do to ameliorate the self-control problem on their own – hence her demand for commitment services from the bank.

To illustrate, at our earlier parameterization $\beta = 0.5$ and $\rho = 1$ the best contract in the absence of self-control will be $C_0^P = (150, 100, 50)$ which offers considerably less consumption smoothing in later periods compared to the contract with self-control $C_0^F = (150, 75, 75)$.[13] If the consumer has income stream $Y_0 = (100, 100, 100)$ then we can think of the consumer with self-control as sticking to a balanced repayment program to keep consumption steady in the last two periods. Compared to this the consumer without self-control rolls over rather than repays the portion of the debt that they would have paid off in period one. The entire burden of repayment of the debt that Zero took out in period 0 now falls in period 2, whereas Zero would have preferred the burden to be shared equally between periods 1 and 2. If the income stream were instead $Y_0 = (200, 50, 50)$ then a consumer without self-control would be viewed as raiding savings in period 1 that an otherwise identical consumer with self-control would have earmarked for period 2 consumption. In summary, consumers that can solve the self control problem through full-commitment contracts will save more/borrow less in period 1 and consume more in period 2. The inability to solve this self-control problem will lead to lower welfare for Zero under competition.

2.3.1 The Bank's no-renegotiation constraint

There are potential gains to trade to be shared between the bank and the period 1 self consumer from renegotiating the contract. The bank's promise in period zero to not renegotiate the contract in period 1 will be credible only so long as any the gains from replacing the remaining terms (c_1, c_2) of the contract C_0 by a lower-cost renegotiated contract (c_1^1, c_2^1) do not exceed the costs of renegotiating κ:

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

Or more simply:

$$c_1 + c_2 - (c_1^1 + c_2^1) \leq \kappa$$

In this competitive setting the renegotiated terms (c_1^1, c_2^1) are given by the consumer's best reaction to (c_1, c_2) described in equations 10 and 11. If in period 1 the consumer can shop around before renegotiating the contract (i.e. if contracts are non-exclusive and competitive)

[13] For these parameters expression 16 becomes $c_0^P = \frac{\sum y}{1+2\beta}, c_1^P = \beta \cdot c_0^P$ and $c_1^P = \beta^2 \cdot c_0^P$

then the consumer will capture the gains to trade. If instead the consumer has no choice in period 1 but to renegotiate with the same bank under conditions of monopoly (which might be the case even there had been competition in period 0 for enforceable exclusive contracts) then these gains to trade will be captured by the monopolist and (c_1^1, c_2^1) will be replaced by best response functions (c_1^{m1}, c_2^{m2}) described below.

We did not make this no-renegotiation constraint explicit before only because under the full-commitment scenario we implicitly assume that κ is sufficiently large as to make te constraint be satisfied at the optimum. Under the polar opposite competitive "no self-control" scenario without exclusivity banks renegotiate contracts at zero profit so with $\kappa = 0$ the constraint is just satisfied.

The more interesting case to analyze is when $\kappa > 0$ but is not so large as to allow the contract to be satisfied at the full-commitment optimum. The core of the remainder of the paper concentrates on this more interesting case. Before we analyze this in more detail, we must first expand the framework to monopoly contracts.

2.4 Full-commitment contracts under Monopoly

This is similar to the earlier competitive full-commitment contact but now the optimand and the participation constraints are reversed. A monopolist bank aims to rearrange consumption in a way that maximizes profits. This will require pressing the consumer up against her participation constraint (PC), leaving her no worse off than under autarky. The bank solves the following problem:

$$\begin{aligned} &\max_{C_0} \Pi_0 (C_0; Y_0) \\ &s.t.\ U_0 (C_0) \geq U_0^A \end{aligned} \tag{18}$$

where U_0^A is the reservation utility the consumer would enjoy if they did not accept the contract. For the moment we will adopt the simple assumption that without a bank contract the consumer has access to no good saving or borrowing opportunities and therefore has no choice but to consume their autarky income stream so $U_0^A = u(y_0) + \beta[u(y_1) + u(y_2)]$. We will relax this assumption to explore how equilibrium contracts may change when the consumer has access to some consumption smoothing opportunities under autarky. The no-renegotiation constraint is assumed to be automatically satisfied because κ is large enough.

The first-order tangency conditions are just as expressions 3 from the competitive case. With $\delta = 1/(1+r)$ and $r = 0$ these are just

$$u'(c_0) = \beta u'(c_1) = \beta u'(c_2) \tag{19}$$

and with CRRA utility simply $c_1 = c_2 = \beta^{\frac{1}{\rho}} c_0$. With the Zero's participation constraint – which must bind at a monopoly optimum – these equations can be solved for the optimum monopoly full-commitment contract C_0^{mF} and corresponding bank profits $\Pi_0\left(C_0^{mF}; Y_0\right)$. Closed form solutions for the CRRA utility case appear as appendix equations 37 and 38, respectively. Conceptually the optimum contract is found at the tangency point where the highest iso-profit plane still touches the iso-utility surface associated with Zero's reservation utility. The Figure 3 depicts the equilibrium contract at point mF in the $c_1 - c_2$ sliced plane given by $c_0 = c_0^{mF}$.

The terms of the optimal monopoly contract, and hence also the level of bank profits achieved will be dependent on the consumer's autarky utility. C_0^{mF} rises and profits fall with U_0^A. The bank makes non-negative profits.

2.4.1 Renegotiation-proof monopoly contracts

The renegotiation-problem in the monopoly bank case is very similar but now the inability to solve the consumer's self-control problem – the inability of the bank to credibly commit to not opportunistically renegotiate the contract – is goint to cost the bank. To see this, assume again for the sake of argument that Zero had agreed to the full-commitment contract indicated by point mF in the figure. At the start of period 1 this contract satisfies Zero self's optimality condition $u'(c_1^{mF}) = u'(c_2^{mF})$ but from the standpoint of One's preferences involves not enough period 1 consumption as $u'(c_1^{mF}) \geq \beta u'(c_2^{mF})$. The monopolist Bank can gain by pandering to One's present bias recontracting to point R along the $c_2 = \beta^{\frac{1}{\rho}} c_1$ ray. If such a recontracting took place the bank would have succeeded in actually pushing Zero's intertemporal utility *below* her reservation utility (which is just met at F). A sophisticated customer of course anticipates this sort of opportunistic renegotiation and will only agree to 'renegotiation-proof' contracts in period 0 that protect her against this type of opportunistic recontracting. The renegotiation-proof contract will lie along the $c_2 = \beta^{\frac{1}{\rho}} c_1$ ray (as Zero understands that One would otherwise try to renegotiate to a point on this ray anyway) at a point where the ray hits Zero's autarky iso-utility surface. The monopolist pays the price for his inability to commit to not renegotiating the contract as profits can only have been lowered by displacing the equilibrium contract from the optimum at mF.

Consider the renegotiation problem at the start of a subgame in period 1. A customer who has agreed to contract C_0 in period 0 enters period 1 with claims to the remaining consumption stream C_1. This determines the new reservation utility. If profitable to the

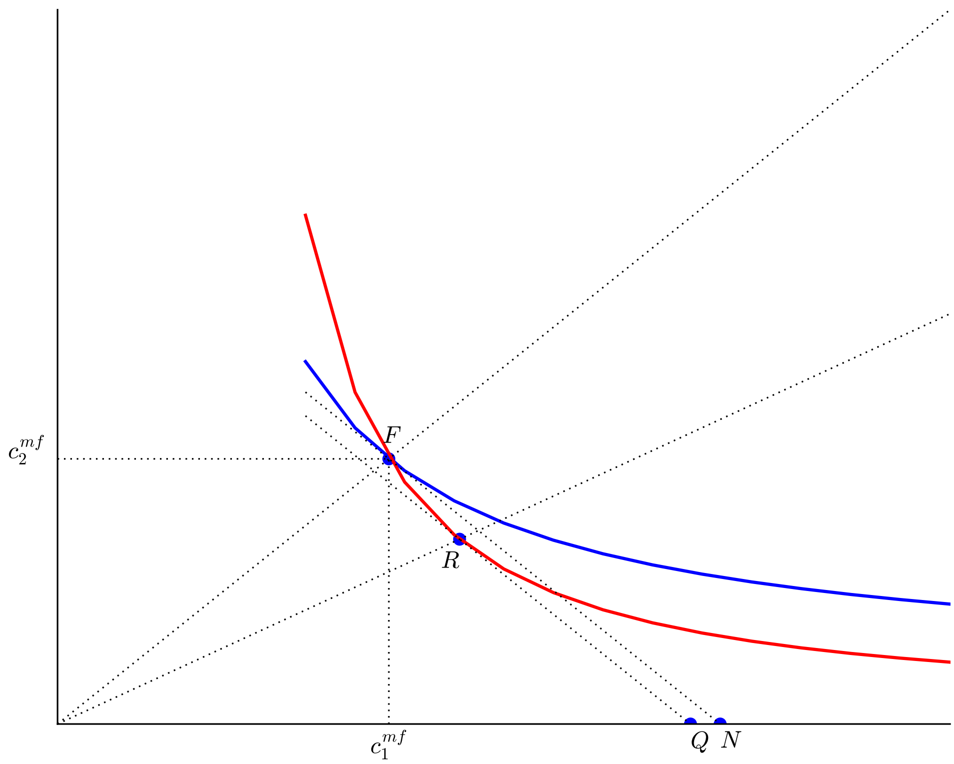

Figure 3: The monopoly renegotiation problem

bank, the optimal renegotiated contract will be chosen in period 1 to solve:

$$\max_{C_1^{m,r}} \Pi_1\left(C_1^{m,r}; C_1\right)$$
$$s.t. U_1\left(C_1^r\right) \geq U_1\left(C_1\right) \tag{20}$$

The bank offers $C_1^{m,r}(C_1) \equiv \left(c_1^{m,r}(C_1), c_2^{m,r}(C_1)\right)$ (equation 40) to replace[14] the contractually agreed C_1. Since the first-order conditions imply $u'\left(c_1^{m,r}\right) = \beta u'\left(c_2^{m,r}\right)$, the contract will set $c_1^{m,r} > c_2^{m,r}$. This also determines the maximum possible profits from renegotiation, $\Pi_1\left(C_1^{m,r}(C_1); C_1\right)$ (equation 41).

Point R in Figure 3 is the renegotiated contract that would be offered in the very special case of a customer who had agreed to a full-commitment contract in period 0 only to find herself genuinely and unexpectedly surprised to find the bank reneging on the promise by offering to renegotiate in period 1. Such commitment-breaking will be weakly profitable unless the contract satisfies the renegotiation-proof constraint

$$\Pi_1\left(C_1^{m,r}(C_1); C_1\right) \leq \kappa \tag{21}$$

where κ are non-monetary renegotiation costs. In Figure 1, the bank will find it profitable to offer to renegotiate from the continuation of the full commitment contract depicted by point F to the new contract depicted by R so long as profits gains NQ on the diagram are greater than κ.

We show in the appendix that, if the original contract offers full commitment ($c_1 = c_2$), profits from renegotiation rise strictly in C_1. The larger the consumption in periods 1 and 2, the greater the scope for rearranging consumption profitably. Also, since C_1^{mF} rises in U_0^A, full commitment becomes less sustainable if U_0^A is sufficiently large.

2.4.2 Competitive renegotiation proof contracts with and without exclusivity

We've analyzed competitive contracts but now make renegotiation costs explicit and allow for the possibility of non-exclusive contracts. The consequences of renegotiation in period 1 depend on the exclusivity of contracts. If contracts are exclusive consumers can be kept from contracting with other financial intermediaries in period 1. The analysis is then identical to the monopoly case for any given C_1. If contracts are non-exclusive, banks can again compete in period 1, so the renegotiated contract satisfies a fresh zero-profit constraint. If

[14] The bank in effect swaps net income stream $(y_1 - c_1, y_2 - c_2)$ for $(y_1 - c_1^r, y_2 - c_2^r)$.

renegotiation is feasible, it satisfies:

$$\underset{C_1^{c,r}}{max} U_1\left(C_1^{c,r}\right)$$
$$s.t. \Pi_1\left(C_1^{c,r}; C_1\right) \geq \kappa \tag{22}$$

This yields a renegotiated contract that can be denoted $C_1^{c,r}\left(C_1\right)$, which returns all surplus to the period 1 agent, resulting in higher consumption than if contracts were exclusive.

A full-commitment contract will survive if and only if there is no way to offer period 1 a new contract that (a) leaves her with at least as much discounted utility as in the original contract, and (b) generates additional profits of at least κ to the bank offering the new contract. We can make two observations about these conditions. First, they apply identically whether contracts are exclusive or not. Second, since full-commitment contracts under competition do not vary with autarky utility, the survival of full-commitment is also independent of autarky utility (unlike under monopoly). So there will be some $\bar{\kappa}$ such that full-commitment contracts will not be renegotiated if and only if $\kappa \geq \bar{\kappa}$ (equation 45).[15]

Proposition 1. *If $\kappa \geq \bar{\kappa}$, then full-commitment contracts survive under both monopoly and competition. If $\kappa < \bar{\kappa}$, then:*

(a) There is some $U^m < U_0^*$ such that the monopolist full commitment contract $C_0^{m,fc}$ will be renegotiated if and only if $U_0^A \epsilon (U^m, U_0^*)$.

(b) The competitive full commitment contract $C_0^{c,fc}$ will be renegotiated at any U_0^A.[16]

An implication is that, under monopoly, consumers with already relatively smooth consumption are more likely to get full-commitment contracts that can be sustained. The comparison between monopoly and competition is perhaps more subtle than it appears. Intuitively, one might expect monopoly to sustain commitment better than competition because of differential abilities to commit–the monopolist can promise not to renegotiate while no such promise is possible under competition. In the language of our model, this would be equivalent to competitive firms facing a lower κ than the monopolist. That may indeed be the case, but our result points to another mechanism that survives even when the costs of renegotiation are identical across market structures. Here, the monopolist's superior ability to commit comes from the fact that its commitment contract is itself less susceptible to renegotiation. Having, at the outset, offered the consumer a contract with the lowest possible consumption, there is relatively less for the firm to gain by renegotiating in period 1.

[15] We assume contracts are renegotiated only if strictly preferred by the bank (monopoly) or consumer (competition).

[16] All proofs are in the appendix.

Note that in the special case of $\kappa = 0$, *any* full-commitment contract would be subject to renegotiation.

To make the continuing analysis interesting, we assume that $\kappa < \bar{\kappa}$.

2.5 Naivete vs Sophistication

The discussion above makes no distinction between sophistication and naivete since we have so far been concerned only with the *possibility* of full-commitment, conditional on a consumer's outside option. The only relevance of naivete, as in Section 3.3, is in the possibly higher perceived autarky utility. If the naif is indeed unrealistically optimistic about her autarky utility, a monopolist would be forced to give her a contract with higher consumption, which in turn would be more likely to be renegotiated. So, a naif is more likely to see a full-commitment contract being renegotiated than is the equivalent sophisticated consumer, not just because of a failure to be strategic (as we will see below) but also because her full-commitment contract makes renegotiation relatively more valuable to the bank.

Under competition, such considerations do not exist since the full-commitment contract is not constrained by the consumer's autarky utility.

3 Renegotiation-Proof Contracts

Having characterized the renegotiation problem, we now examine contracts that take this into account. For sophisticated hyperbolic discounters, contracts will be informed by their ability to anticipate possible renegotiation. For naifs, banks will capitalize on the consumer's failure to do the same.

3.1 Sophisticated Hyperbolic Discounters

3.1.1 Monopoly

A sophisticated consumer will rationally anticipate the bank and her own later self's temptation to break promises and therefore would reject the contract C_0^{mf} if it fails the renegotiation-proofness conditions laid out in Section 4. Since C_0^{mf} was designed to leave the period 0 consumer with autarky utility, any renegotiation would make her strictly worse off than in autarky. The bank will want to avoid renegotiation costs κ since it can always do better by simply offering the renegotiated contract as the original contract. So the only contracts that will be offered must be renegotiation-proof: contracts that the bank will not find profitable to renegotiate.

The bank must therefore offer a contract that satisfies period 0's participation constraint and an additional renegotiation-proofness constraint. The resulting contract must limit period 1 self's (and hence also the bank's) potential gains from renegotiation. There are two ways the bank could achieve this. It could maintain $c_1 = c_2$, but lower the values enough that renegotiation is just barely unprofitable. Or, it could increase c_1 relative to c_2 so as to lower what the period 1 self is willing to pay for renegotiation to just below the bank's cost, κ , as to keep renegotiation credibly unprofitable to the bank. As long as κ is small enough for the constraint to bind, this must lower bank profits relative to the costless full-commitment case because the monopolist must now compensate the consumer for a contract that delivers less consumption smoothing across periods 1 and 2.

The monopolist solves the following problem:

$$\begin{aligned} &\max_{C_0} \Pi_0 \left(C_0; Y_0\right) \\ s.t. &U_0 \left(C_0\right) \geq U_0^A && (23) \\ &\Pi_1 \left(C_1^{m,r}\left(C_1\right); C_1\right) \leq \kappa && (24) \end{aligned}$$

Let us denote the solution $C_0^{m,rp}$ (the superscript is for 'monopolist, renegotiation-proof'). The first restriction is the same participation constraint as before. The second (24) is a no-renegotiation constraint that states that to be credibly committed any increase in profits that the bank can gain from renegotiating the continuation contract from C_1 to $C_1^{m,r}(C_1)$ must fall short of the cost of renegotiation cost κ. The renegotiation-proof contract is explicitly derived in the special case of $\kappa = 0$ (Equation 60).

For, $\kappa > 0$, the contract cannot be explicitly derived, but its key properties can be established. A convenient way to interpret the renegotiation-proofness constraint is this: given any amount s to be split across periods 1 and 2, the constraint specifies, at a minimum, how much must be allocated to period 1. When binding, the renegotiation-proofness constraint can be thought of as yielding functions $c_1(s)$ and $c_2(s) = s - c_1(s)$. If c_1 is sufficiently large relative to c_2 (i.e. the allocation is sufficiently close to period 1's optimum), the allocation will not be renegotiated by the bank. We show in the appendix that $c_1(s) > \frac{1}{2}$ and $\frac{dc_1(s)}{ds} > \frac{c_1}{s}$. The first is obvious–when the constraint binds, the bank can no longer offer full-commitment ($c_1 = \frac{s}{2}$) and to prevent renegotiation it must allow period 1 to consume more than period 2 does. The second result states that, as more is consumed in periods 1 and 2 combined, the ratio of consumption is forced to skew further in period 1's favor. This follows from the fact that, if c_1 and c_2 were to rise in fixed proportions, the bank's profits from renegotiation would rise. So to prevent renegotiation as s grows, c_1 would need to grow disproportionately.

Proposition 2. *If $U_0^A \leq U^m$, the bank offers the full-commitment contract. If $U_0^A > U^m$:*

(a) $\Pi_0\left(C_0^{m,rp};Y_0\right) < \Pi_0\left(C_0^{m,fc};Y_0\right)$

(b) There is some $\bar{U}\epsilon\left(U^m,U_0^*\right)$ *such that, if* $U_0^A > \bar{U}$*, the bank would earn negative profits and therefore will not offer a contract.*

(c) $c_0^{m,rp} > c_0^{m,fc}$.

Proposition 2 compares the renegotiation-proof contract to the full-commitment contract when the renegotiation-proofness constraint binds. Part (a) states that bank profits are lower than under full-commitment. The bank wishes it could promise to not renegotiate but it cannot make such a promise credible without giving up some profits. The problem here is not one of cheating or contract failure (as examined for example by Hansmann (1980) and Glaeser-Shleifer (2001)), it is the possibility of a legitimate renegotiation (a voluntary agreement to tear up the old contract) between the consumer and the firm. The monopolist would gain from having higher renegotiation costs since in equilibrium renegotiation does not take place.

Part (b) follows from part (a). If the bank were able to provide full commitment, it could offer a profit-making contract to any individual with even minimal smoothing needs. Now however, for individuals whose autarky utility is close enough to U_0^*, the bank would make negative profits and therefore no contract is offered. This is because the renegotiation-proofness constraint may require even greater imbalance in consumption across periods 1 and 2 than under autarky.

Part (c) is about the terms of the contract itself–when full-commitment is not feasible, the renegotiation-proof contract involves higher consumption in period 0 (i.e. either a smaller loan or less savings than under full commitment). The following is a sketch of the argument. Let the full-commitment contract be described by some $c_0 = c_0^{m,fc}$ (period 0 consumption) and $s = s^{m,fc}$ (sum of c_1 and c_2, which are equal in size). Since the full-commitment contract lies at an optimum, it must be true that at the levels of consumption specified by the contract, the marginal utility of present consumption is equal to the discounted marginal utility of future consumption:

$$\frac{du\left(c_0^{m,fc}\right)}{dc_0} = \frac{d\left(\beta u\left(\frac{s^{m,fc}}{2}\right) + \beta u\left(\frac{s^{m,fc}}{2}\right)\right)}{ds} \tag{25}$$

Now suppose $c_0^{m,rp} = c_0^{m,fc}$, so that period 0 consumption in the renegotiation-proof contract is held the same. Since any future consumption will be split unevenly, in order to continue to satisfy the consumer's period 0 participation constraint, it must be true that $s^{m,rp} > s^{m,fc}$. We show in the appendix that $s^{m,rp}$ will be large enough that, regardless of the parameters of the consumer's utility function,

$$\frac{du\left(c_0\right)}{dc_0} > \frac{d\left(\beta u\left(c_1\left(s\right)\right) + \beta u\left(c_2\left(s\right)\right)\right)}{ds} \tag{26}$$

So the bank can 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.

To summarize: the requirement that contracts be renegotiation-proof results in higher period 0 consumption and lower bank profits, and the denial of service to consumers whose smoothing needs are relatively small.

3.1.2 Competition

As mentioned earlier, under competition, the terms of a renegotiated contract depend on exclusivity. It follows that a renegotiation-proof contract under period 1 exclusivity must satisfy $\Pi_1\left(C_1^{c,r}\left(C_1\right);C_1\right) \leq \kappa$ and a renegotiation-proof contract under non-exclusivity must satisfy $\Pi_1\left(C_1^{c,m}\left(C_1\right);C_1\right) \leq \kappa$. These conditions are equivalent: each states that there is no way to simultaneously satisfy the consumer's participation constraint and the bank's 'κ-profit constraint'. So, regardless of exclusivity, the equilibrium contract under competition is given by:[17]

$$\begin{aligned} &\max_{C_0} U_0\left(C_0\right) \\ &s.t.\Pi_0\left(C_0;Y_0\right) \geq 0 &(27)\\ &\Pi_1\left(C_1^{c,r}\left(C_1\right);C_1\right) \leq \kappa &(28)\end{aligned}$$

This yields a contract $C_0^{c,rp}$. Given our assumption that $\kappa < \bar{\kappa}$, the renegotiation-proofness constraint (28) binds at any autarky utility. The properties of the equilibrium contract are summarized in the next proposition, which is structured like Proposition 2.

Proposition 3. *(a) At any autarky utility, full commitment is infeasible, so* $U_0\left(C_0^{c,rp}\right) < U_0\left(C_0^{c,fc}\right)$.

(b) If $U_0^A > \bar{U}$ *(as defined in Proposition 2), the period 0 consumer would do worse than in autarky and therefore will not accept the contract.*

(c) The relationship between $c_0^{c,rp}$ *and* $c_0^{c,fc}$ *is ambiguous. If* $\rho \leq \hat{\rho}$*, then* $c_0^{c,rp} > c_0^{c,fc}$*. If* $\rho > \hat{\rho}$*, then there are parameter values under which* $c_0^{c,rp} < c_0^{c,fc}$.

[17] A similar analysis could be conducted even if the costs of renegotiating another bank's contract are different from κ, the costs of renegotiating one's own contract.

The first two parts of the proposition are intuitively similar to the case of monopoly. First, the additional constraint results in a contract that cannot deliver the optimal utility to the period 0 consumer. Second, if the consumer's autarky utility is high, there cannot be renegotiation-proof contracts in equilibrium since the consumer could do better on her own. In fact, the cutoff autarky utility above which contracts are not offered is the same as under monopoly–this is the parameter region where it is impossible to simultaneously satisfy the consumer's participation constraint and the banks' zero-profit constraint.

Part (c) is a deviation from the results under monopoly. Under competition, we find that period 0 consumption under renegotiation-proof contracts may be bigger or smaller than under full-commitment. Again, we provide the intuition here (the actual proof involves a few additional steps). The competitive full-commitment contract must satisfy: $\frac{du\left(c_0^{c,fc}\right)}{dc_0} = \frac{d\left(\beta u\left(\frac{s^{c,fc}}{2}\right)+\beta u\left(\frac{s^{c,fc}}{2}\right)\right)}{ds}$. If the renegotiation-proof contract were to have the same c_0, it must also have the same s (to continue satisfying the zero-profit constraint), but consumption will be split in period 1's favor. So, the marginal utility of future consumption becomes: $\frac{d\left(\beta u\left(c_1\left(s^{c,fc}\right)\right)+\beta u\left(c_2\left(s^{c,fc}\right)\right)\right)}{ds}$. 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:

$$\frac{d\left(\beta u\left(c_1\left(s^{c,fc}\right)\right)+\beta u\left(c_2\left(s^{c,fc}\right)\right)\right)}{ds} < \frac{d\left(\beta u\left(\frac{s^{c,fc}}{2}\right)+\beta u\left(\frac{s^{c,fc}}{2}\right)\right)}{ds} = \frac{du\left(c_0^{c,fc}\right)}{dc_0} \tag{29}$$

In such a case, the renegotiation-proof contract must involve higher period 0 consumption than the full-commitment contract. If, on the other hand, the utility function is highly convex (high ρ), then an imbalanced split has a higher marginal utility, so the renegotiation-proof contract will have lower period 0 consumption than under full-commitment.[18] This can be seen more explicitly in the clase of $\kappa = 0$ (Equation 61).

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). The key reason that the latter possibility does not exist under monopoly is the following: under monopoly, a switch from full-commitment to renegotiation-proofness while maintaining the same c_0 would require such a large jump in future total consumption (to maintain the same discounted utility under imbalanced consumption) that the marginal utility would necessarily fall.

[18] The actual argument, in the appendix, is a little more complicated since we must consider the effect of a change in s not just on utilities, but also on the relative ratios of c_1 and c_2.

3.2 Naive Hyperbolic Discounters

For naive agents, the problem of renegotiation does not 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 period 0 consumer.[19]

3.2.1 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, renegotiation generates additional profits for the bank. Notice that, unlike with sophisticates, service is not denied to any naif since the consumer would, at the very least, be willing to accept the full-commitment contract (since she would not anticipate renegotiation).

With a naive hyperbolic discounter, the bank must choose between a renegotiation-proof contract and one that will be renegotiated upon. If the consumer's autarky utility is very low, then the initial contract can extract so much surplus that there is little to gain from renegotiation. But when autarky utility is high, the consumer must be offered a contract with high consumption in each period. It is such consumers, the ones who have relatively less need for banking, who will find their contracts renegotiation. In such cases, the bank solves the following problem:[20]

$$\begin{aligned} \underset{C_0}{max} \quad & \Pi_0\left(C_0;Y_0\right)+\Pi_1\left(C_1^r\left(C_1\right);C_1\right)-\kappa \\ s.t. \quad & U_0\left(C_0\right)\geq U_0^A \end{aligned} \tag{30}$$

Let the solution be denoted $C_0^{m,n}$ (the superscript is for 'monopoly, naive'). This is explicitly derived in the appendix (67, 68). The bank maximizes profits by offering a contract that divides future consumption as much in favor of period 2 as possible. 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

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

[20] 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.

infinity.[21] In each case, the greater the imbalance between the contracted c_1 and c_2, the greater the bank's profits from renegotiation.

This contract can be compared to the full-commitment contract and to the renegotiation-proof contract for sophisticates. In particular, it will involve lower period 0 consumption than under full-commitment or 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. *(a) A monopoly contract will be accepted at any autarky utility.*

(b) There is some $U^n < U^m$ (U^m as defined in Proposition 1) such that, if $U_0^A \leq U^n$, the naive agent will receive the monopoly full commitment contract and it will not be renegotiated.

(c) If $U_0^A > U^n$, the monopoly contract will satisfy $c_0^{m,n} < c_0^{m,fc} < c_0^{m,rp}$ (either explicitly or in the limit), and will be renegotiated in period 1.

3.2.2 Competition

Under competition too, contracts will be renegotiated and firms must account for renegotiation. First, note that if contracts are not exclusive, the equilibrium contract must be identical to the full-commitment contract. This is because the firm offering the contract in period 0 does not expect to benefit from renegotiation.

Under exclusive contracts, anticipated profits from future renegotiation will be returned to the consumer through more favorable initial contracts. The equilibrium contract satisfies:

$$\begin{aligned} &\underset{C_0}{max} U_0\left(C_0\right) \\ &s.t. \Pi_0\left(C_0; Y_0\right) + \Pi_1\left(C_1^r\left(C_1\right); C_1\right) \geq \kappa \end{aligned} \tag{31}$$

Let the solution be denoted $C^{c,n}$ ('competition, naive'). As under monopoly, first-order conditions lead to a corner solution where contracts favor period 2 relative to period 1. This

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

maximizes the potential gains from renegotiation.

Unlike under monopoly, these anticipated gains must be returned to the consumer. Some of these gains are returned to period 0, so there is no clear prediction about whether period 0 consumption will be lower or higher than under full commitment. This is formalized in Proposition 5.

Proposition 5. *(a) The competitive contract will be accepted at any autarky utility and will be renegotiated in period 1.*

(b) The non-exclusive competitive contract will be identical to the full-commitment contract, $c_0^{c,fc}$.

(c) Under exclusive contracts, the relationship between $c_0^{c,n}$ *and* $c_0^{c,fc}$ *is ambiguous. If* $\rho < 1$, $c_0^{c,n} < c_0^{c,fc}$. *If* $\rho > 1$, *then there are parameter values under which* $c_0^{c,n} > c_0^{c,fc}$.

4 Nonprofits

Suppose a firm has the possibility of operating as a non-profit. Narrowly stated, non-profit status means that the firm cannot tie managers' compensation or outside shareholders dividends to firm profits because, technically speaking, a non-profit firm has no shareholders. As described in the introduction we prefer a more elastic and encompassing definition of the non-profit term to include firms that may have shareholders but who have adopted ownership or governance structures that place credible and visible constraints on the distribution of profits that can be made to managers or investor shareholders. The principals of such firms may also care about social objectives such as the welfare of their customers directly, and not just about profits. This broader interpretation allows the non-profit firm category to include cooperatives as well as social enterprises and 'hybrid' firms which might be incorporated as for-profit firms but are owned and controlled by social investors.[22]

To model these ideas in a simple yet still rich manner, assume that we can classify a firms' nonprofit orientation by a simple parameter $\alpha \epsilon (0, 1]$, which can be viewed as the degree of 'non-profitness'. A lower α indicates a firm that because of its ownership and governance structure places the welfare of its clients ahead of that of its 'owners' and makes the capture of profits by principals (managers, outside investors) more difficult.

α affects the firm's maximization problem in two ways. First, it reduces the firm's ability to capture its raw profits, so that if a firm earns profits Π, its principals get to enjoy only a fraction $\alpha\Pi$. Glaeser & Shleifer (2001) adopt this approach suggesting that it captures the idea of how the principals of a nonprofit, though legally barred from paying themselves cash

[22]For example, most of the commercial firms that one finds in modern microfinance (including poster-child firms of he 'commercialization' revolution in microfinance including Bancosol of Bolivia or Compartamos of Mexico) are incorporated as for-profit corporations but on close inspection turn out to be majority-owned and controlled by 'social investors' that are themselves non-profit foundations.

profits, might capture profits imperfectly via the consumption of perquisites or 'dividends in kind.' Second, α possibly alters the renegotiation costs that the firm will incur when it breaks its promises to customers (e.g. shame, regret, loss of social reputation). The cost of renegotiation is now $\eta(\alpha)$, with η rising weakly in α and $\eta(1) = \kappa$.

Our analysis below focuses on firms that are purely self-interested. We derive conditions under which the pursuit of profits leads a firm to voluntarily switch to a form of governance with $\alpha < 1$.

Before proceeding, it should be noted that firms never have an incentive to switch to nonprofit status when consumers are naive. Since the consumer does not perceive a need for commitment, the nonprofit's promise of superior commitment is of no value to her.

4.1 Monopoly

The firm establishes its type α via the adoption of legal non-profit status and/or by choosing credible and stable ownership and governance structures that commit it to those limitations. When facing a sophisticated hyperbolic discounter, a monopoly firm of type α designs a renegotiation-proof contract to solve

$$\underset{C_0}{max} \alpha \Pi_0 \left(C_0; Y_0 \right)$$

$$s.t. U_0 \left(C_0 \right) \geq U_0^A \tag{32}$$

$$\alpha \Pi_1 \left(C_1^r \left(C_1 \right); C_1 \right) \leq \eta \left(\alpha \right) \tag{33}$$

The first constraint is the customer participation constraint as before while the second constraint is a modified no-renegotiation constraint. This states that the value of captured profits from not renegotiating the contract should exceed captured profits from renegotiation net of renegotiation costs. Notice that if we define $\kappa(\alpha) \equiv \frac{\eta(\alpha)}{\alpha}$, we can rewrite the no-renegotiation constraint as

$$\Pi_1(C_1^r(C_1); C_1) \leq \kappa(\alpha) \tag{34}$$

This makes the no-renegotiation constraint look just like the earlier constraint (24) except that κ is now a function of α. Indeed the earlier pure for-profit monopoly renegotiation problem is just a special case of the above program with $\alpha = 1$.

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. Because the non-profit can more credibly commit to not renegotiate contracts that offer greater consumption smoothing in periods 1 and 2, period 0 becomes more willing to pay for this smoothing service.

The captured-profits maximizing solution gives a contract $C_0^{m,np}$ ('monopoly, non-profit'). The no-renegotiation constraint is now relaxed compared to the earlier pure for-profit case. With a relaxed renegotiation-proof constraint $\Pi_0(C_0^{m,np};Y) > \Pi_0(C_0^{m,rp};Y)$ but whether or not it will be in the bank principals' best interest to strategically convert to non-profit status depends on whether the profits they can capture under non-profit status exceed the profits they could earn as a pure for-profit, in other words on whether $\alpha\Pi_0(C_0^{m,np};Y) \geq \Pi_0(C_0^{m,rp};Y)$. 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.

Proposition 6 describes conditions under which a firm will operate as a non-profit. Since the for-profit firm's no-renegotiation constraint binds, the nonprofit can offer greater commitment and thereby extract greater surplus from the consumer through the contract signed in period 0. The question is: does the rise in extracted surplus outweigh the fact that all profits are now discounted? 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.

If the right-hand side of the no-renegotiation constraint remains unchanged (as in the proposition), the firm is forced to address the tradeoff–a lower α means both better commitment and reduced ability to capture profits.

If the consumer's autarky consumption bundle is sufficiently close to optimal to start with then even a nonprofit may be unable to offer sufficiently smoother consumption that allows it to cover costs. For intermediate levels of autarky utility, the nonprofit is able to more credibly promise it will not renegotiate, so it earns positive profits where the pure for-profit would have earned small or negative profits. Here, 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 α nonprofit status dominates.

Finally, for autarky bundles far from the optimal, the for-profit firm would anyway be making substantial profits. In this case, the nonprofit'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.

Proposition 6. *Consider any $\bar{\alpha} < 1$ and a corresponding $\eta(\bar{\alpha}) = \kappa$. There is some $\underline{U}^{np} \epsilon \left(U^m, \bar{U}\right)$ and $\bar{U}^{np} \epsilon (\bar{U}, U_0^*]$ such that, if $U_0^A \epsilon \left(\underline{U}^{np}, \bar{U}^{np}\right)$, the bank strictly prefers $\alpha = \bar{\alpha}$ over $\alpha = 1$.*

In Figure 4 we illustrate the case where non-pecuniary costs to breaking a promise not to

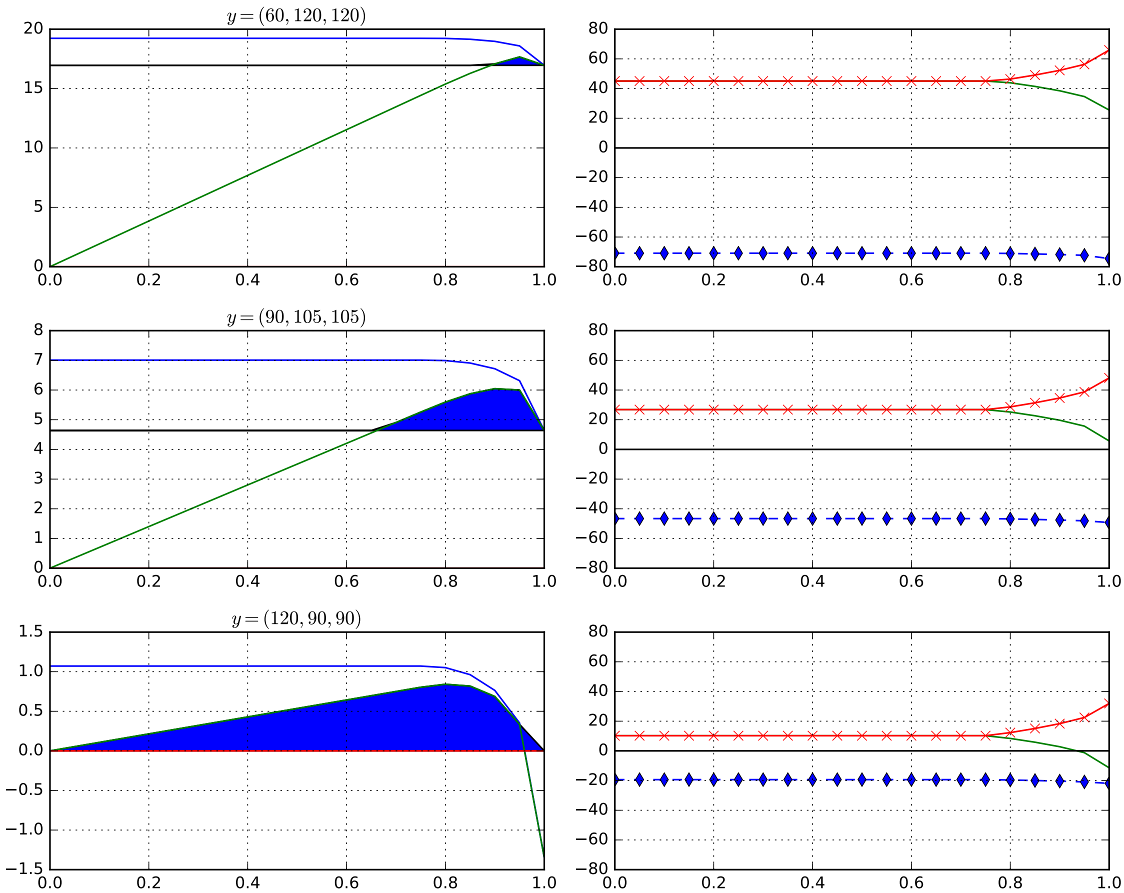

Figure 4: Choice of non-profit status and initial endowment income

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 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 $\Pi_0(C_0^{m,\alpha}; Y)$ and the lower curve captured profits $\alpha\Pi_0(C_0^{m,\alpha}; Y)$. A horizontal line has been drawn in to indicate the level of profits $\Pi_0(C_0^{m,1}; Y)$ 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 no 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 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 $\alpha = 0.7$.

We conclude this section with a note on the nature of profit-capture restrictions faced by a nonprofit. Our assumption that the firm can capture a fixed fraction of raw profits was useful for exposition but is perhaps not realistic, and is not necessary for our results. We might imagine that, at low levels of profits, the nonprofit can capture most of the profits, and that as profits rise so do the restrictions on the firm's ability to capture them. In other words, a nonprofit captures $f(\Pi)$ of profits, where $f(0) = 0$, $0 < f'(\Pi) < 1$, and $f''(\Pi) < 0$. In such cases, we can again clearly see how non-profit status could be attractive to the firm: the concavity of f can 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).

4.2 Competition

4.2.1 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 period 1 self goes to the bank (this grants the bank monopoly power in period 1). A firm of type α will be led to offer contract terms to solve

$$\max_{C_0} U_0\left(C_0\right)$$

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

$$\Pi_0(C_1^r(C_1); C_1) \leq \kappa(\alpha) \tag{36}$$

Here, as before, $\kappa(\alpha) = \frac{\eta(\alpha)}{\alpha}$, captures the idea that the principals of a firm that adopts non-profit status commit themselves to capturing a smaller share of raw profits and may also suffer greater direct disutility from breaking promises to customers. Let the contract that solves this program be denoted $C_0^{c,e,np}$ ('competition, exclusive, nonprofit').

Consider first a field where all firms start as pure for-profits ($\alpha = 1$). If the no-renegotiation constraint binds, consumer welfare must be lower than that when the firms can commit to not renegotiate since an additional constraint is imposed. Starting from this situation consider now one firm's strategic choice of whether to adopt non-profit status (i.e. to change its ownership and governance structures to an $\alpha = \bar{\alpha} < 1$ relaxing the no-renegotiation constraint. One firm deviating into nonprofit status in this way can make positive profits. So, if the borrowers are sophisticated hyperbolics, in equilibrium all firms become nonprofit. As proven in the appendix, competition will ensure however that in equilibrium the principals of all firms are capturing zero profits.

4.2.2 Non-Exclusive Contracts

In the previous section, we had a setting with competition in period 0 but exclusive contracting and monopoly power in period 1. 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

Figure 5: Summary of results

firms alone.[23]

Proposition 7. *Consider any $\bar{\alpha} < 1$ and a corresponding $\eta(\bar{\alpha}) = \kappa$.*

(a) In a competitive banking market with exclusive contracts, all active firms will be nonprofits (contracts will be offered for $U_0^A \leq \bar{U}^{np}$, with $\bar{U}^{np} > \bar{U}$ as defined in Proposition 6).

(b) In a competitive banking market with non-exclusive contracts, for-profits must exist in equilibrium (contracts will be offered for $U_0^A \leq \bar{U}$, with $\bar{U}$ as defined in Proposition 2).

5 Discussion

5.1 Summary

The model above formalizes the renegotiation problem faced by banks that contract with hyperbolic discounters, and shows how the problem is addressed in equilibrium contracts. Figure 5 summarizes the key results of Sections 4-6. We show how contracts depend on relative distances from the optimal autarky utility, U_0^*. The results, taken together, generate some natural yet novel empirical predictions that are in principle testable.

We first discuss naive hyperbolic discounters. Under monopoly, contracts will be subject to renegotiation if the consumer is close enough to optimal autarky; if not, the full-commitment contract leaves the consumer with so little consumption that there is no point renegotiating. The parameter region in which contracts are renegotiated is relatively large, and even includes cases where the sophisticate would be offered a full-commitment contract. The renegotiable contract involves less period-0 consumption that the full-commitment contract as this allows the bank to exploit renegotiation possibilities most comprehensively.

Under competition, any contract will be renegotiated, even for consumers whose autarky outcomes are very poor. This is because competitive full-commitment contracts always leave consumers with relatively high levels of future consumption. The implications for contract terms are ambiguous.

Next, we turn to sophisticated hyperbolic discounters under monopoly. If the consumer's smoothing needs are large (i.e. autarky utility is low), full-commitment is feasible since the contract terms leave little that is susceptible to renegotiation. If smoothing needs are moderate, the consumer is offered a renegotiation-proof contract which has a larger period-0 consumption than the full-commitment contract (this serves to reduce the contract's

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

susceptibility to renegotiation). If smoothing needs are small, the consumer is better off in autarky than in any contract that satisfies the renegotiation-proof constraint. As a result of the problem of renegotiation, such consumers will not be offered contracts.

Under competition, sophisticates will not be offered full-commitment contracts. Since full-commitment would entail high consumption levels in periods 1 and 2 regardless of autarky utility, the renegotiation-proofness constraint must always bind. As under naivete, the implications for contract terms are ambiguous.

Finally, we derive conditions under which banks will operate as nonprofits. A monopoly will switch to a nonprofit if, as a for-profit its profits were close to zero (above or below). In these cases, the bank is willing to forgo some enjoyment of its profits in exchange for a loosened renegotiation-proofness constraint. So, nonprofits will be able to serve some consumers relatively closer to optimal autarky utility, ones who would be unbanked under for-profit monopoly.

Under competition, banks should operate as nonprofits if contracts are exclusive, as they can capture the surplus from improved contracts possible under nonprofit status. If contracts are non-exclusive, no bank can benefit from operating as a nonprofit, and for-profit banks prevail.

5.2 Additional Considerations

Our model delivers predictions about how renegotiation concerns affect commitment contracts, and about parameter regions in which these concerns actually matter. In particular, we generate comparative statics over autarky utilities. Autarky utility is not informative in isolation, but in conjunction with total income serves as an indicator of the extent of smoothing that remains to be provided by a bank. As we show, contract terms depend in particular ways at different degrees of smoothing needs.

The sizes of relevant parameter regions discussed above will vary according to other parameter values such as the cost of renegotiation, κ, and total income, y. For example, as κ rises there will be an expansion of the parameter region in which full commitment survives. On the other hand, as y rises, commitment in general will be harder to sustain since contracts terms must allow for higher consumption in periods 1 and 2.

While the focus of our paper is on contracts, a few observations on welfare can be made. Under hyperbolic discounting, there is no obvious notion of welfare, and for our purpose we take it to the discounted utility of the period 0 self. Clearly, for sophisticated hyperbolic discounters under monopoly, welfare remains constant regardless of renegotiation concerns and bank governance–the consumer is always left with autarky utility. Under competition, welfare is lower under renegotiation-proof contracts relative to full-commitment (when the constraint binds), and nonprofits serve to raise welfare.

Finally, recall that our model assumes contracts can only be initiated in period 1. As a result, the consumer knows that the alternative to a period-0 contract is autarky, and the bank knows that if a contract isn't signed in period 0 there are no remaining opportunities to contract with the consumer. This determines participation constraints for both the consumer and the bank. The model could easily be extended to allow for contracts that are initiated in period 1. While this would alter participation constraints, the qualitative results of our model would continue to hold.

Consider how participation constraints change if contracts could also be initiated in period 1. For the consumer, the possibility of a period 1 contract could either tighten or loosen the participation constraint in period 0 (if the period 1 contract further skews consumption away from period 2, the constraint loosens; otherwise it tightens). For the bank, the possibility of a period 1 contract must weakly tighten its participation constraint. The problem can be analyzed in the same way as in our original model, but the parameter values under which a bank will switch from for-profit to nonprofit status will change. For example, consider the monopolist bank. If, as a for-profit, period 0 profits are nearly 0, its decision about whether to switch to nonprofit status becomes more complicated–it could turn into a nonprofit and thereby earn higher profits in period 0, but it must also consider the possibility of remaining a for-profit and simply offering a contract in period 1 rather than in period 0.

6 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 affects different types of consumers 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.

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?

A Appendix: CRRA Derivations and Proofs

A.1 Full-Commitment

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

$$C_0^{mF} = \left(1, \beta^{\frac{1}{\rho}}, \beta^{\frac{1}{\rho}}\right) \cdot \left(\frac{U_0^A (1-\rho)}{1+2\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}} \tag{37}$$

$$\Pi_0\left(C_0^{mF}; Y_0\right) = y - \left(U_0^A (1-\rho)\right)^{\frac{1}{1-\rho}} \left(1+2\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}} \tag{38}$$

For the competitive banks that offer full-commitment, the solutions is determined by the first-order condition and the bank's participation constraint:

$$C_0^{F} = \left(1, \beta^{\frac{1}{\rho}}, \beta^{\frac{1}{\rho}}\right) \cdot \left(\frac{y}{1+2\beta^{\frac{1}{\rho}}}\right) \tag{39}$$

A.2 Renegotiation

Given an existing contract C_1, a monopolist bank that renegotiates in period 1 will offer the following new contract:

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

The corresponding profit gains from renegotiation are:

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

An alternate way to restate the above is the following: Let s and α be defined such that $c_1 = \alpha s$ and $c_2 = (1-\alpha)s$. Then:

$$C_1^r(C_1) = \left(s, s\beta^{\frac{1}{\rho}}\right) \cdot \left(\frac{\alpha^{1-\rho} + \beta(1-\alpha)^{1-\rho}}{1+\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}} \tag{42}$$

Profit gains from renegotiation become:

$$\Pi_1\left(C_1^r\left(C_1\right);C_1\right)=(s)\left(1-\left(\alpha^{1-\rho}+\beta\left(1-\alpha\right)^{1-\rho}\right)^{\frac{1}{1-\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right) \tag{43}$$

By construction, profits from renegotiation are strictly positive (and increasing in s), except in the special case where C_1 is optimal from period 1's perspective ($(1-\alpha)=\beta^{\frac{1}{\rho}}\alpha$), in which case they are 0. It can also easily be confirmed that profits from renegotiation fall in α as long as the allocation is such that period 1 would like a larger α than the current contract offers.

Proof of Proposition 1: (a) In any full-commitment contract, $\alpha=\frac{1}{2}$. Inserting this into (43), the following must be satisfied for a full-commitment contract to survive:

$$k\geq s\left(1-\frac{1}{2}\left(1-\beta\right)^{\frac{1}{1-\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right) \tag{44}$$

Substituting for s from the competitive full-commitment contract, we can rewrite the above condition as:

$$k\geq\frac{2\beta^{\frac{1}{\rho}}y}{1+2\beta^{\frac{1}{\rho}}}\left(1-\frac{1}{2}\left(1-\beta\right)^{\frac{1}{1-\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right)\equiv\bar{\kappa} \tag{45}$$

Since in any full-commitment contract (monopoly and competition), s will be no larger than $\frac{2\beta^{\frac{1}{\rho}}y}{1+2\beta^{\frac{1}{\rho}}}$, no full-commitment contract will be renegotiated if $\kappa\geq\bar{\kappa}$.

(c) If $\kappa<\bar{\kappa}$, condition 45 fails, so the competitive full-commitment contract cannot survive.

(b) The monopolist full-commitment contract will not survive if s is sufficiently large. Since $s^{m,fc}$ is exponentially increasing in U_0^A, there must be some U^m such that the contract will not survive if and only if $U_0^A>U^m$. Since the contract cannot survive at U_0^* (here, the contract is identical to the competitive contract) , $U^m<U_0^*$. $\square$

A.3 Renegotiation-Proof Contracts

A.3.1 Sophisticated Hyperbolic Discounters

When the renegotiation-proofness constraint binds, consumption in periods 1 and 2 must satisfy:

$$(s)\left(1-\left(\alpha^{1-\rho}-\beta\left(1-\alpha\right)^{1-\rho}\right)^{\frac{1}{1-\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right)=\kappa \tag{46}$$

For any s, there may be two values of α that satisfy the constraint with equality–one with too little consumption relative to period 1's optimal, one with too much consumption relative to period 1's optimal. The relevant value for us is the first. This defines a continuous function $\alpha(s)$.

$$\alpha(s) = min\left\{\alpha : (s)\left(1-\left(\alpha^{1-\rho}+\beta(1-\alpha)^{1-\rho}\right)^{\frac{1}{1-\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{-\rho}{1-\rho}}\right)=\kappa\right\} \tag{47}$$

As s rises, to continue satisfying the constraint we must have $\alpha(s)$ rising too (if fractions stayed constant, profits from renegotiation would rise).

We can also rewrite the first-order condition of the bank's maximization problem using the new notation. For any s and α, let $V(s,\alpha) = u(\alpha s) + u((1-\alpha)s)$. This is the discounted utility over periods 1 and 2, from period 0's perspective. The solution, $C_0 = (c_0, \alpha s, (1-\alpha)s)$, must satisfy:

$$\frac{du(c_0)}{dc_0} = \beta\frac{dV(s,\alpha)}{ds} \tag{48}$$

In other words, at the profit-maximizing contract the marginal dollar should be equally valuable whether consumed immediately or distributed across future periods.

Proof of Proposition 2: (a) Since the 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.

(b) Clearly, $\Pi_0(C_0^{m,rp};Y_0)$ falls strictly in U_0^A (if autarky utility falls, the bank can always do better, at least by simply lowering c_0). Since at $U_0^A = U^m$, $\Pi_0(C_0^{m,rp};Y_0) = \Pi_0\left(C_0^{m,fc};Y_0\right) > 0$ and at $U_0^A = U^*$, $\Pi_0(C_0^{m,rp};Y_0) < \Pi_0\left(C_0^{m,fc};Y_0\right) = 0$, there must be some intermediate autarky utility above which the bank's maximized profits will be negative.

(c) consider any $c_0 \le c_0^{m,fc}$ and s such that $U_0\left(c_0, \frac{s}{2}, \frac{s}{2}\right) = U_0^A$. We can find the corresponding $\bar{s}$ that, while satisfying the participation constraint, gives the same utility from period 0's perspective:

$$V\left(s,\frac{1}{2}\right) = \bar{V}(\bar{s},\alpha(\bar{s})) \tag{49}$$

$$\Rightarrow 2\frac{\left(\frac{1}{2}s\right)^{1-\rho}}{1-\rho} = \frac{(\alpha(\bar{s})\bar{s})^{1-\rho}}{1-\rho}+\frac{((1-\alpha(\bar{s}))\bar{s})^{1-\rho}}{1-\rho} \tag{50}$$

$$\Rightarrow \bar{s} = s\left(\frac{2\left(\frac{1}{2}\right)^{1-\rho}}{\alpha(\bar{s})^{1-\rho}+(1-\alpha(\bar{s}))^{1-\rho}}\right)^{\frac{1}{1-\rho}} \tag{51}$$

From this, we get the following inequality:

$$\frac{dV\left(\bar{s},\alpha\left(\bar{s}\right)\right)}{ds} = \bar{s}^{-\rho}\left(\alpha\left(\bar{s}\right)^{1-\rho}+\left(1-\alpha\left(s\right)\right)^{1-\rho}\right)+\frac{d\alpha\left(\bar{s}\right)}{ds}\bar{s}^{1-\rho}\left(\alpha\left(\bar{s}\right)^{-\rho}-\left(1-\alpha\right)^{-\rho}\right) \tag{52}$$

$$< \bar{s}^{-\rho}\left(\alpha\left(\bar{s}\right)^{1-\rho}+\left(1-\alpha\left(s\right)\right)^{1-\rho}\right) \tag{53}$$

$$= s^{-\rho}\left(2\left(\frac{1}{2}\right)^{1-\rho}\right)\left(\frac{2\left(\frac{1}{2}\right)^{1-\rho}}{\alpha\left(\bar{s}\right)^{1-\rho}+\left(1-\alpha\left(\bar{s}\right)\right)^{1-\rho}}\right)^{\frac{-1}{1-\rho}} \tag{54}$$

$$< s^{-\rho}\left(2\left(\frac{1}{2}\right)^{1-\rho}\right)=\frac{dV\left(s,\frac{1}{2}\right)}{ds} \tag{55}$$

The first line above splits the effect of s on V into two–the first term represents the change in utility holding α constant, and the second term represents the (negative) effect of the further skewing of consumption that results from a rise in s. The final inequality follows from the fact that, since $\bar{s} > s^{m,fc}$, the renegotiation constraint must bind so that $\alpha\left(\bar{s}\right) > \frac{1}{2}$. Finally, the following inequality holds:

$$\frac{du\left(c_0\right)}{dc_0} \geq \frac{du\left(c_0^{m,fc}\right)}{dc_0} = \beta\frac{dV\left(s^{m,fc},\frac{1}{2}\right)}{ds} \tag{56}$$

$$\geq \beta\frac{dV\left(s,\frac{1}{2}\right)}{ds} > \beta\frac{dV\left(\bar{s},\alpha\left(\bar{s}\right)\right)}{ds} \tag{57}$$

We have shown that at any $c_0 \leq c_0^{m,fc}$, for a contract that satisfies the renegotiation-proofness constraint, the marginal utility of period 0 consumption will be higher than the discounted marginal utility of future consumption, so the bank could earn strictly higher profits by raising c_0 and lowering s further. Therefore, in the renegotiation-proof contract, $c_0^{m,rp} > c_0^{m,fc}$. □

Proof of Proposition 3: (a) We know that $U_0\left(C_0^{c,fc}\right) = U_0^*$. By assumption, since the renegotiation-proofness constraint is binding, the renegotiation-proof contract cannot offer the optimal consumption path. Therefore $U_0\left(C_0^{c,rp}\right) < U_0\left(C_0^{c,fc}\right)$.

(b) Consider $\bar{U}$, as constructed in Proposition 2. If $U_0^A > \bar{U}$, it is impossible to construct a contract that earns nonnegative profits and gives the consumer at least autarky utility. Therefore, any contract that earns zero profits would give the period 0 consumer less than autarky utility. (As an aside, observe that $\bar{U} = U_0\left(C_0^{c,rp}\right)$.)

(c) At the full-commitment contract:

$$\frac{du\left(c_0^{c,fc}\right)}{dc} = \beta\frac{dV\left(s^{c,fc},\frac{1}{2}\right)}{ds} = \left(s^{c,fc}\right)^{-\rho}\left(2\left(\frac{1}{2}\right)^{1-\rho}\right) \tag{58}$$

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

$$\begin{aligned}\frac{dV\left(s^{c,fc},\alpha\left(s^{c,fc}\right)\right)}{ds} &= \left(s^{c,fc}\right)^{-\rho}\left(\alpha\left(s^{c,fc}\right)^{1-\rho}+\left(1-\alpha\left(s^{c,fc}\right)\right)^{1-\rho}\right)\\ &+\frac{d\alpha\left(s^{c,fc}\right)}{ds}\left(s^{c,fc}\right)^{1-\rho}\left(\alpha\left(s^{c,fc}\right)^{-\rho}-\left(1-\alpha\left(s^{c,fc}\right)\right)^{-\rho}\right)\end{aligned}\tag{59}$$

The first term–the direct effect of a change in s–is weakly less than $\frac{dV\left(s^{c,fc},\frac{1}{2}\right)}{ds}$ 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 $\rho < 1$, $\frac{dV\left(s^{c,fc},\alpha\left(s^{c,fc}\right)\right)}{ds} < \frac{dV\left(s^{c,fc},\frac{1}{2}\right)}{ds} = \frac{du\left(c_0^{c,fc}\right)}{dc}$, so the renegotiation-proof contract must satisfy $c_0^{c,rp} > c_0^{c,fc}$.

Next, we consider the case when $\rho > 1$. We can make the following observations about $\alpha(s)$. First, $\lim_{\kappa\to 0}\alpha(s) = \frac{\beta^{\frac{-1}{\rho}}}{1+\beta^{\frac{-1}{\rho}}}$ (this follows from the fact that at $\kappa = 0$, the contract must satisfy $u'(c_1) = \beta u'(c_2)$). Second, implicitly differentiating equation 47 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 59 will be sufficiently small in magnitude that $\frac{dV\left(s^{c,fc},\alpha\left(s^{c,fc}\right)\right)}{ds} > \frac{dV\left(s^{c,fc},\frac{1}{2}\right)}{ds} = \frac{du\left(c_0^{c,fc}\right)}{dc}$. In this case, the renegotiation-proof contract must satisfy $c_0^{c,rp} < c_0^{c,fc}$. $\square$

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:

$$C_0^{m,rp} = \left(\left(\frac{U_0^A(1-\rho)}{1+\beta^{\frac{1}{\rho}}\left(\frac{\left(1+\beta^{\frac{1-\rho}{\rho}}\right)^{\frac{1}{\rho}}}{\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{1-\rho}{\rho}}}\right)}\right)^{\frac{1}{1-\rho}}, \left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{\rho}}c_0^{m,rp}, \beta^{\frac{1}{\rho}}\left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{\rho}}c_0^{m,rp}\right)\tag{60}$$

$$C_0^{c,rp} = \left(\frac{y}{1+\beta+\beta^{\frac{1}{\rho}}}, \left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}}\right)c_0^{c,rp}, \beta^{\frac{1}{\rho}}\left(\frac{\beta+\beta^{\frac{1}{\rho}}}{1+\beta^{\frac{1}{\rho}}}\right)c_0^{c,rp}\right)\tag{61}$$

It can easily be established that $c_0^{m,rp} > c_0^{m,fc}$, $c_0^{c,rp} > c_0^{m,fc}$ if $\rho > 1$, and $c_0^{c,rp} < c_0^{m,fc}$ if $\rho < 1$.

A.3.2 Naive Hyperbolic Discounters

Suppose the monopolist intends to renegotiate the contract. The maximization problem, combined with the expression for $C_1^r(C_1)$ (42), simplifies to:

$$\max_{c_0,c_1,c_2} y - c_0 - \frac{\left(c_1^{1-\rho} + \beta c_2^{1-\rho}\right)^{\frac{1}{1-\rho}}}{\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{\rho}{1-\rho}}} - \kappa \tag{62}$$

$$s.t. \frac{c_0^{1-\rho}}{1-\rho} + \beta\frac{c_1^{1-\rho}}{1-\rho} + \beta\frac{c_2^{1-\rho}}{1-\rho} \geq U_0^A \tag{63}$$

The partial derivatives of the resulting Lagrangian are:

$$\frac{\partial \mathcal{L}}{\partial c_0} = -1 - \lambda c_0^{-\rho} \tag{64}$$

$$\frac{\partial \mathcal{L}}{\partial c_1} = c_1^{-\rho}\left[-\left(\frac{c_1^{1-\rho} + \beta c_2^{1-\rho}}{1+\beta^{\frac{1}{\rho}}}\right)^{\frac{\rho}{1-\rho}} - \lambda\beta\right] \tag{65}$$

$$\frac{\partial \mathcal{L}}{\partial c_2} = c_2^{-\rho}\left[-\beta\left(\frac{c_1^{1-\rho} + \beta c_2^{1-\rho}}{1+\beta^{\frac{1}{\rho}}}\right)^{\frac{\rho}{1-\rho}} - \lambda\beta\right] \tag{66}$$

An interior solution, with $\frac{\partial \mathcal{L}}{\partial c_1} = 0$ and $\frac{\partial \mathcal{L}}{\partial c_2} = 0$ does not exist since (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 $\rho < 1$:

$$C_0^{m,n} = \left(\left(\frac{U_0^A(1-\rho)}{2+\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}}, 0, \left(\frac{1+\beta^{\frac{1}{\rho}}}{\beta}\right)^{\frac{1}{1-\rho}}\left(\frac{U_0^A(1-\rho)}{2+\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}}\right) \tag{67}$$

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

$$C_0^{m,n} = \left(\left(\frac{U_0^A(1-\rho)}{1+\left(1+\beta^{\frac{1}{\rho}}\right)\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}}, \beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{1}{1-\rho}}\left(\frac{U_0^A(1-\rho)}{1+\left(1+\beta^{\frac{1}{\rho}}\right)\beta^{\frac{1}{\rho}}}\right)^{\frac{1}{1-\rho}}, \infty\right) \tag{68}$$

Proof of Proposition 4: (a) At any autarky utility, the monopolist could at least offer the full-commitment contract.

(b) The bank must choose between a renegotiation-proof contract and a renegotiable contract (67, 68). By construction of U^m, the following must be true at any $U_0^A \geq U^m$:

$$\Pi_0\left(C_0^{m,rp};Y_0\right) \leq \Pi_0\left(C_0^{m,fc};Y_0\right) \leq \Pi_0\left(C_0^{m,fc};Y_0\right) + \Pi_1\left(C_1^r\left(C_1^{m,fc}\right);C_1^{m,fc}\right) - \kappa \quad (69)$$

Since $C_0^{m,n}$ is uniquely determined and $C_0^{m,n} \neq C_0^{m,fc}$, profits from the best renegotiable contract must be strictly higher than profits from the renegotiation-proof contract at any $U_0^A \geq U^m$.

The following can be verified from the explicit derivations of $C_0^{m,fc}$ and $C_0^{m,n}$. First, if U_0^A is sufficiently small, $\Pi_0\left(C_0^{m,fc};Y_0\right) > \Pi\left(C_0^{m,n};Y_0\right) + \Pi_1\left(C_1^r\left(C_1^{m,n}\right);C_1^{m,n}\right) - \kappa$. Second,

$$\frac{d}{dU_0^A}\Pi_0\left(C_0^{m,fc};Y_0\right) > \frac{d}{dU_0^A}\left[\Pi\left(C_0^{m,n};Y_0\right) + \Pi_1\left(C_1^r\left(C_1^{m,n}\right);C_1^{m,n}\right) - \kappa\right] \quad (70)$$

It follows that there is some $U^n < U^m$ such that if $U_0^A \leq U^n$, the naive agent will receive the monopoly full commitment contract, which will be renegotiation-proof.

(c) This can be confirmed from the explicit formulations of $C_0^{m,fc}$ (37) and $C_0^{m,n}$ (67, 68). $\square$

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

$$\max_{c_0,c_1,c_2} \frac{c_0^{1-\rho}}{1-\rho} + \beta\frac{c_1^{1-\rho}}{1-\rho} + \beta\frac{c_2^{1-\rho}}{1-\rho} \quad (71)$$

$$s.t. y - c_0 - \frac{\left(c_1^{1-\rho} + \beta c_2^{1-\rho}\right)^{\frac{1}{1-\rho}}}{\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{\rho}{1-\rho}}} - \kappa \geq 0 \quad (72)$$

The first-order conditions are the same as under monopoly (64, 65, 66). Combining these with the zero-profit constraint, we get the following solution. If $\rho < 1$:

$$C_0^{c,n} = \left(\frac{y-\kappa}{2+\beta^{\frac{1}{\rho}}}, 0, \left(\frac{1+\beta^{\frac{1}{\rho}}}{\beta}\right)^{\frac{1}{1-\rho}}\left(\frac{y-\kappa}{2+\beta^{\frac{1}{\rho}}}\right)\right) \quad (73)$$

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

$$C_0^{c,n} = \left(\frac{y-\kappa}{1+\beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)}, \beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)^{\frac{1}{1-\rho}}\left(\frac{y-\kappa}{1+\beta^{\frac{1}{\rho}}\left(1+\beta^{\frac{1}{\rho}}\right)}\right), \infty\right) \quad (74)$$

Proof of Proposition 5: (a) Banks can at least offer the consumer the full-commitment contract, so a contract is feasible at any autarky utility. Since, from the consumer's perspective, any renegotiation-proof contract is strictly dominated by the full-commitment contract (which will be renegotiated), in equilibrium she will be offered a contract that will be renegotiated.

(b) 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 identical to the full-commitment contract.

(c) Suppose $\rho < 1$. Comparing $C_0^{c,fc}$ (39) to $C_0^{c,n}$ (73), it is clear that $c_0^{c,n} < c_0^{c,fc}$. Suppose $\rho > 1$. If κ is small enough, $c_0^{c,n} > c_0^{c,fc}$. $\square$

A.4 Nonprofits

Lemma 1. $\Pi_0 \left(C_0^{m,rp}; Y_0\right)$ *and* $\bar{\alpha}\Pi_0 \left(C_0^{m,np}; Y_0\right)$ *are continuously decreasing in* U_0^A.

Proof of Lemma 1: We prove the above for renegotiation-proof contracts of for-profit banks. The same argument applies to nonprofit banks. First, it is clear that profits are strictly decreasing in U_0^A: If autarky utility drops from $U_0^A = U$ to $U_0^A \bar{=} U$, at $\bar{U}$ the bank can always do better than offering the contract it offered at U.

Next, we prove right-continuity at any $U_0^A = U$. Let the maximized profits at U be $\Pi_0 \left(C_0; Y_0\right)$, where $C_0 = (c_0, c_1, c_2)$. This contract must satisfy $U_0 \left(C_0\right) = U$. For any $\bar{U} > U$, profits must be lower, and bounded below by $\Pi_0 \left(\bar{C}_0; Y_0\right)$, with the contract defined as $\bar{C}_0 = (c_0 + x, c_1, c_2)$ where x satisfies $U_0 \left(\bar{C}_0\right) = \bar{U}$. Since $\underset{\bar{U} \to U^+}{lim} \Pi_0 \left(\bar{C}_0; Y_0\right) = \Pi_0 \left(C_0; Y_0\right)$, the profit function is right-continuous.

Finally, we prove left-continuity at at any $U_0^A = U$. For any $\bar{U} < U$, denote maximized profits $\Pi_0 \left(\bar{C}_0; Y_0\right)$, where $\bar{C}_0 = (\bar{c}_0, \bar{c}_1, \bar{c}_2)$. These contracts must satisfy $U_0 \left(\bar{C}_0\right) = \bar{U}$. At U, profits must be lower, and bounded below by $\Pi_0 \left(C_0; Y_0\right)$, with the contract defined as $C_0 = (\bar{c}_0 + x, \bar{c}_1, \bar{c}_2)$ where x satisfies $U_0 \left(C_0\right) = U$. Since $\underset{\bar{U} \to U^-}{lim} \Pi_0 \left(\bar{C}_0; Y_0\right) = \Pi_0 \left(C_0; Y_0\right)$, the profit function is left-continuous. $\square$

Proof of Proposition 6: Let $\eta\left(\bar{\alpha}\right) = 1$, to minimize the attractiveness of the nonprofit. Consider $U_0^A = \bar{U}$. Since the for-profit's renegotiation-proofness binds and leaves the firm with zero profits, and since the non-profit's renegotiation-proofness constraint is looser than the for-profit's, we know that $\bar{\alpha}\Pi \left(C_0^{m,np}; Y_0\right) > 0 = \Pi \left(C_0^{m,rp}; Y_0\right)$. Since profits must be continuously decreasing in U_0^A, and since $\bar{\alpha}\Pi \left(C_0^{m,np}; Y_0\right) < \Pi \left(C_0^{m,rp}, ; Y\right)$ at $U_0^A = U^m$ (where the for-profit's renegotiation-proofness constraint no longer binds) and $\bar{\alpha}\Pi \left(C_0^{m,np}; Y_0\right) \leq 0$ at $U_0^A = U_0^*$, there must exist autarky utility values as described in the proposition statement such that, if $U_0^A > \underline{U}^{np}$ the bank strictly prefers to operate as a nonprofit relative to a for-profit, and if $U_0^A \geq \bar{U}^{np}$ it weakly prefers to not offer a contract. $\square$

Proof of Proposition 7: (a) Suppose all firms are for-profit. There is some ε_1 and ε_2

satisfying $0 < \varepsilon_2 < \varepsilon_1$ and a corresponding $\hat{C}_0 = (c_0^{c,e,np}, c_1^{c,e.np} - \varepsilon_1, c_2^{c,e,np} + \varepsilon_2)$ such that $U_0\left(C_0^{c,e,np}\right) = U_0\left(\hat{C}_0\right)$ and $\Pi_0(C_1^r\left(\hat{C}_1\right); \hat{C}_1) \leq \kappa(\bar{\alpha}) < \kappa(1)$. So, any firm can make positive profits by operating as a non-profit. Therefore, in equilibrium, consumers will borrow only from non-profit firms. Given the construction of $\bar{U}^{np}$, firms can make nonnegative profits while satisfying the participation constraint only if $U_0^A \leq \bar{U}^{np}$.

(b) 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

Akerlof, George A., & Shiller, Robert J. 2015. *Phishing for phools: the economics of manipulation and deception*. Princeton: Princeton University Press.

Armstrong, Mark, & Vickers, John. 2012. Consumer Protection and Contingent Charges. *Journal of Economic Literature*, **50**(2), 477–493.

Arrow, Kenneth J. 1963. Uncertainty and the Welfare Economics of Medical Care. *American Economic Review*, **53**(5), 941–973.

Ashraf, N., Karlan, D., & Yin, W. 2006. Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines. *The Quarterly Journal of Economics*, **121**(2), 635–672.

Banerjee, Abhijit V., & Duflo, Esther. 2011. *Poor economics: a radical rethinking of the way to fight global poverty*. 1st ed edn. New York: PublicAffairs.

Basu, Karna. 2011. Hyperbolic Discounting and the Sustainability of Rotational Savings Arrangements. *American Economic Journal: Microeconomics*, **3**(4), 143–171.

Besley, Timothy, & Ghatak, Maitreesh. 2005. Competition and Incentives with Motivated Agents. *American Economic Review*, **95**(3), 616–636.

Bond, Philip, Musto, David K., & Yilmaz, Bilge. 2009. Predatory mortgage lending. *Journal of Financial Economics*, **94**(3), 412–427.

Bryan, Gharad, Karlan, Dean, & Nelson, Scott. 2010. Commitment Devices. *Annual Review of Economics*, **2**(1), 671–698.

Bubb, Ryan, & Kaufman, Alex. 2013. Consumer biases and mutual ownership. *Journal of Public Economics*, **105**(Sept.), 39–57.

Carrillo, Juan D., & Dewatripont, Mathias. 2008. Promises, Promises, . . . *The Economic Journal*, **118**(531), 1453–1473.

Conning, Jonathan, & Morduch, Jonathan. 2011. Microfinance and Social Investment. *Annual Review of Financial Economics*, **3**(1), 407–434.

Cull, Robert, Demirguc-Kunt, Asli, & Morduch, Jonathan. 2009. Microfinance Meets the Market. *Journal of Economic Perspectives*, **23**(1), 167–192.

Demyanyk, Yuliya, & Van Hemert, Otto. 2011. Understanding the Subprime Mortgage Crisis. *Review of Financial Studies*, **24**(6), 1848–1880.

Dewatripont, Mathias, & Tirole, Jean. 1999. *The Prudential Regulation of Banks*. 2. print edn. The Walras-Pareto lectures, no. 1. Cambridge, Mass.: MIT Press.

DeYoung, et al. 2015 (Oct). *Reframing the Debate about Payday Lending*.

Fischer, Greg, & Ghatak, Maitreesh. 2010. Repayment Frequency in Microfinance Contracts with Present-Biased Borrowers. *Economic Organisation and Public Policy Discussion Papers Series*, **21**.

Gabaix, X., & Laibson, D. 2006. Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *The Quarterly Journal of Economics*, **121**(2), 505–540.

Glaeser, Edward L., & Shleifer, Andrei. 2001. Not-for-profit entrepreneurs. *Journal of Public Economics*, **81**(1), 99–115.

Gottlieb, Daniel. 2008. Competition over Time-Inconsistent Consumers. *Journal of Public Economic Theory*, **10**(4), 673–684.

Gugerty, Mary Kay. 2007. You Can't Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya. *Economic Development and Cultural Change*, **55**(2), 251–282.

Guha, Brishti, & Roy Chowdhury, Prabal. 2013. Micro-finance competition: Motivated micro-lenders, double-dipping and default. *Journal of Development Economics*, **105**(Nov.), 86–102.

Hansmann, Henry. 1996. *The ownership of enterprise*. 1. harvard univ. press paperback ed edn. Cambridge, MA: Harvard University Press.

Heidhues, Paul, & Koszegi, Botond. 2010. Exploiting Naivete about Self-Control in the Credit Market. *American Economic Review*, **100**(5), 2279–2303.

Laibson, D. 1997. Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics*, **112**(2), 443–478.

Laibson, David, Repetto, Andrea, & Tobacman, Jeremy. 2003. A Debt Puzzle. *Pages 228–266 of:* Aghion, Philippe, Frydman, Roman, Stiglitz, Joseph, & Woodford, Michael (eds), *Knowledge, Information, and Expectations in Modern Economics: In Honor of Edmund S. Phelps*. Princeton, NJ: Princeton University Press.

McIntosh, Craig, & Wydick, Bruce. 2005. Competition and microfinance. *Journal of Development Economics*, **78**(2), 271–298.

Mendez, Rodrigue. 2012. Predatory Lending: A Model of Behavioral Pricing. *Paris School of Economics working paper*.

Oak, Mandar, & Swamy, Anand. 2010. Only Twice as Much: A Rule for Regulating Lenders. *Economic Development and Cultural Change*, **58**(4), 775–803.

O'Donoghue, T., & Rabin, M. 2001. Choice and Procrastination. *The Quarterly Journal of Economics*, **116**(1), 121–160.

Rhyne, Elisabeth. 2011. On Microfinance: Who's to Blame for the Crisis in Andhra Pradesh? *The Huffington Post*, May.

Tanaka, Tomomi, Camerer, Colin F, & Nguyen, Quang. 2010. Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American Economic Review*, **100**(1), 557–571.

Thaler, Richard H., & Benartzi, Shlomo. 2004. Save more tomorrowâ: Using behavioral economics to increase employee saving. *Journal of political Economy*, **112**(S1), S164âS187.