Long-Term Contracts, Commitment, and Optimal Information Disclosure*

Alessandro Dovis
University of Pennsylvania
and NBER
adovis@upenn.edu

Paolo Martellini
University of Wisconsin-Madison
martellini@wisc.edu

March 2024

Abstract

This paper studies optimal information disclosure in dynamic insurance economies with income risk in which an incumbent firm acquires more information about an agent's persistent type than the rest of the market. If the incumbent can commit to long-term contracts but the agent can walk away, the optimal disclosure prescribes no information revelation to minimize the high-income type's outside option and maximize cross-subsidization. If the incumbent lacks commitment, low-income consumers receive their full information allocation and no cross-subsidization is feasible for any public information disclosure because of adverse selection. Information design aims at implementing intertemporal consumption smoothing between the first period and the second period for high type consumers. The optimal information disclosure entails a bad signal and a partially informative good signal. Next, we allow for idiosyncratic motives to switch firms. Idiosyncratic motives reduce adverse selection and allow for the possibility of cross-subsidization for low types. Even in this case, disclosing some information may be optimal, but full information disclosure never is. Lastly, we show that when the gains from cross-subsidization are larger than the losses from adverse selection distortions, long-term relations are harmful to consumers. Our results can be used to analyze the consequences of policy proposals like open banking.

First version: March 2024. We thank Natalia Kovrijnykh, Venky Venkateswaran, and Alexander Von Hafften for useful comments.

1 Introduction

In many markets, firms engage in long-term relationships with consumers. Observing the history of transactions, incumbent firms acquire information about consumers that competitors do not have access to. For example, health and car insurance companies gradually learn about consumers' health or driving record; credit card companies infer their customers' repayment probability by observing their spending pattern; in the labor market, employers possess information about their employees that is hidden from the public record.[1] This ex-post informational monopoly provides incumbent firms with an advantage relative to the competition.

This paper asks to what extent incumbent firms should be forced to share information with competitors. This question is important for the pervasiveness of this type of asymmetric information and because of the emergence of an increasing number of policies aimed at regulating information disclosure. One prominent example is the recent trend in many countries toward the adoption of so-called *open banking*, a set of regulations that compel banks to make data on their customers' history available to competitors. Another prominent example, in the opposite direction, is represented by laws that forbid employers from asking workers about their past wages. Our main objective is to characterize the design of optimal information disclosure, taking into account the equilibrium response of incumbents and outsiders to the resulting amount of information available to the public.

We answer these questions in a simple two-period insurance economy where firms compete to attract consumers and the incumbent firm learns the consumer's type over time.[2] Consumers lack commitment and always have the option to switch to a new firm in the second period. We have three main results. First, if firms can commit to the terms of the contract, the optimal information disclosure provides no information to minimize the value of the outside option for high-type consumers and maximize ex-post cross-subsidization of low-types. Second, if firms cannot commit, then disclosing some information may improve welfare. In this case, no cross-subsidization is possible in the second period. Information design aims at implementing intertemporal consumption smoothing between the first period and the second period for high type consumers. Finally, we consider a version of the model where agents have idiosyncratic motives to switch firms in the second period. Idiosyncratic motives reduce adverse selection among switchers and allow for the possibility of cross-subsidization for low types. Even in this case, disclosing some information may be optimal, but full information disclosure never is. We find that the optimal amount of information is not monotone in the fraction of switchers in the

[1] For example, Kahn (2013) find supporting evidence for the labor market, Cohen (2012) for the auto insurance market, and Ioannidou and Ongena (2010) for the corporate loan market.

[2] One can reinterpret our model as one where lenders learn about the default probability of a borrower.

economy.

We study a simple two period economy where risk-averse consumers seek insurance against income fluctuations. For simplicity, we assume that income can take two values. In the first period, consumers and firms have the same information. Firms compete to attract consumers by offering insurance contracts. At the end of the first period, the consumer and the incumbent firm learn the income realization. Competing firms (outsiders) do not observe the realization of income but only observe a public signal from the disclosure policy. Outsiders also observe the set of contracts offered by the incumbent firm that acts as a Stackelberg leader, but not the contract offered to each individual consumer. Thus, in the second period the incumbent has an information advantage relative to its competitors as it can condition the contract to the previous income realization while outsiders' offers must satisfy an incentive constraint. In the second period, outsiders offer contracts to consumers that decide whether to move or not. In order to ensure equilibrium existence (in pure strategies) and the possibility of cross-subsidization -even in the absence of the incumbent- we assume that after all contracts are posted, firms can pay an arbitrarily small cost to withdraw their contract, as in the game described by Netzer and Scheuer (2014) for a Rothschild and Stiglitz (1976) economy.

Our goal is to characterize the public disclosure policy that maximizes ex-ante welfare. A public disclosure policy is a map from the individual income realization in the first period–our proxy for a consumer's type–to a signal that is observed by everyone in the economy. The public disclosure policy effectively controls the composition of the pool of consumers with a particular signal, hence determining the maximal amount that outsiders are willing to offer to such consumers in the second period. Consider for example the high-income consumers. If the disclosure policy fully reveals information, then the outsiders can offer these consumers a constant consumption profile at their expected income level. If instead high-income consumers receive the same signal as some low-income ones, outsider firms must either distort the consumption profile of the high-income types not to attract low-types or transfer resources to the low-income types. Thus, the maximal value that outsiders can offer is increasing in the share of high-income with a given signal. The value of the outside option for high-type affects the long-term contract that the incumbent firm can offer in the first period because consumers always have the option to leave the incumbent.

We consider two forms of firms' commitment power. When firms can commit to long-term contracts, we find no information disclosure to be optimal. Commitment allows incumbent firms to deliver as much intertemporal consumption smoothing and cross-subsidization of low-type consumers as permitted by the ex-post participation constraint of high-type consumers. Information disclosure tightens such constraint hence reducing the amount of insurance that can be sustained. Notwithstanding the presence of private information, the equilibrium consumption profile features the same insurance pattern of the canonical models by Harris and Holmstrom (1982) and Thomas and Worrall (1988): consumption is smoothed over time unless the ex-post participation constraint is binding and the consumption profile is *back-loaded*.

We next study the case in which also the firms cannot commit to the terms of the contract beyond the current period. Consider the continuation equilibrium outcome in the second period for an arbitrary disclosure policy. High-income consumers are offered full insurance at a value that matches the best offer that could be made by the outsiders. Because outsiders lack information about the consumer's type, they must offer a value lower than the one under full information because of the need to separate high-types from low-types (conditional on a given public signal). Thus, incumbents make profits on the high-type consumers in the second period because they can deliver the same value without distorting the allocation to make it incentive compatible–as they can exclude low-type consumers. The amount of such profits is decreasing in the amount of information provided by the signal.

Regardless of the public disclosure policy, in the second period low-income consumers never receive transfers in equilibrium–they consume their expected income in the second period. Incumbents know consumers' history hence they never make negative profits on any given type. The lack of cross-subsidization for low-type consumers is in contrast to the version of the model where all firms are equally uninformed about the consumers' type as shown in Netzer and Scheuer (2014). For that to happen, firms must also serve high types–and make profit on them. The presence of an informed incumbent prevents outsiders from attracting high-type consumers since the incumbent can offer them a higher value given that it does not need to apply an adverse selection discount. Thus, in equilibrium they know that if they attract any customers, they must be the low-type, even if they have a good signal.

Competition among firms in the first period ensures that these expected profits from the high-income consumers in the second period are rebated back to consumers in period 1. Thus, typically the optimal contract is *front-loaded*.

We next turn to the characterization of the optimal public disclosure policy. Because low-income consumers' consumption in the second period is independent of the information structure, the information design is exclusively driven by intertemporal motives between first period consumption and second period consumption for high-income consumers. Under mild conditions, we show that the optimal disclosure policy takes the form of a '*bad news*' system in which all high-income consumers receive a good signal but only some low-income ones do. The more low-income consumers are pooled with high-income ones, the lower the outside option for the latter in the second period, and the higher the ex-post profits of the firm. Firms' profits are rebated as first period consumption given ex-ante perfect competition among firms. At the optimal policy, equilibrium consumption features an inverse of the Harris and Holmstrom (1982) profile: same consumption in the first period and in the second period for high-types, no consumption insurance for low-types.

The ability of firms to commit – say because of reputation considerations – is also a critical factor for the optimal amount of information to disclose. We show that the time series of the contract's terms can help to disentangle whether a firm can commit or not. If terms are front-loaded, this is an indicator of firms' inability to commit and so more information must be provided. If instead terms are back-loaded, this is an indicator that firms can commit and so less (no) information should be disclosed to competitors.

In the last part of the paper, we introduce idiosyncratic motives to induce consumers to switch firms, hence allowing for cross-subsidization by the outsiders given the lower informativeness of the switching decision. We show that, as the share of switchers increases, the optimal disclosure policy does not always provide less information because more information might enable – instead of hindering – cross-subsidization among switchers. This is because for the outsiders to offer a transfer to the low-income consumers, the pool of high-type switchers and low-type must have a sufficiently high share of the former. Thus, the need to have enough information.[3]

Optimal information disclosure weighs intertemporal smoothing against cross-subsidization. When the share of switchers increases, the latter motive eventually dominates and information must increase in order to trigger a transfer to low-types by outsiders–who would otherwise offer the least-cost-separating contract. Once the contract offered by outsiders entails cross-subsidization, the optimal amount of information decreases in the share of switcher in order to maximize the number of low-types that receive a transfer.

An implication of our analysis is that optimal information disclosure crucially depends on the composition of the population of consumers and on the strength of idiosyncratic switching motives. When switching motives are weak, so that consumers respond primarily to the terms of the contracts, optimal information disclosure is increasing in the fraction of high-types in the population since incumbents make second period profits on them and, absent information disclosure, they excessively front load consumption. When switching motives are strong, so that the pool of switchers is less affected by adverse selection, optimal information disclosure is decreasing in the fraction of high-types in the population since less information is needed in order to trigger transfers to low-types by outsiders.

[3]Note the difference with the case in which incumbent can commit to long-term contracts. There cross-subsidization was provided by the incumbent and it was maximal at no information because it minimized the high-type's outside option. Here are the outsiders that can give a transfer to the low-type if the pool of switchers has sufficiently many high types. Thus the need to provide some information.

Finally, our analysis shed lights on the value of long-term relationships. With commitment on the firm side, having long-term relationship is always preferable to a sequence of spot contracts. When the firm cannot commit there is a trade-off: long-term relationship reduces the cost of providing incentives and allow for some intertemporal smoothing but prevent cross-subsidization. In particular, long-term relationships are preferable when the share of high-income consumer in the economy is low because the benefits of cross-subsidization with spot contracts are limited.

Related literature

The closest paper to our is de Garidel-Thoron (2005) that also studies the role of information sharing in an insurance economy with long-term contracts where incumbent gains an informational advantage relative to competition. His main analysis is under commitment on the firm side and he compares welfare under two polar opposites: full information disclosure and no information disclosure. de Garidel-Thoron (2005)'s main result is that the value under no information disclosure is higher. The key contribution of our paper is to study the optimal information disclosure showing that some information disclosure is optimal when firms lack commitment. This is true also in the case with idiosyncratic switching motives where the ex-post adverse selection problem is less severe.

At a broad level, our paper is related to the growing literature on information design: Kamenica and Gentzkow (2011), Bergemann et al. (2015), Bergemann and Morris (2019), Mathevet et al. (2020). In this literature the closest paper are Garcia and Tsur (2021) and Immorlica et al. (2022) who also study optimal information disclosure in adverse selection economies. Both papers consider static economy and only allow for pooling contract (this is without loss in Immorlica et al. (2022)). Our main innovation is the dynamic analysis that naturally leads to asymmetric information among firms and to dynamic consumption smoothing motives that are key drivers of the optimal disclosure policy, but are absent in their analysis. Calzolari and Pavan (2006) study the problem of information disclosure in a multiple principals setting. Guerrieri (2008), Guerrieri et al. (2010) and Lester et al. (2019) study the role of competition in frictional markets characterized by adverse selection. The critical difference with our work is that we study an economy where firms enters a repeated relationship with the agent and compete with other firms in the second period after gaining an informational advantage.

More broadly, this paper contributes to the literature on dynamic long-term contracts where agents cannot commit that follows Harris and Holmstrom (1982) and Thomas and Worrall (1988) seminal contributions. We contribute to this literature by studying how asymmetric information that arises between the incumbent firm and its competitors affect the terms of the contract and by studying how public information disclosures can manipulate the value of the agent's outside options and therefore the equilibrium contract.[4] Sharpe (1990) studies a model with asymmetric learning between firms and long-term lending relationships. The equilibrium dynamics–for a given disclosure policy–is similar to the one in our economy, but he does not study the optimal information disclosure policy.

This paper is related to the literature that studies the effect on lending conditions of the implementation of open banking policies, that is, a set of regulations that compel banks to make data on their customers' history available to competitors. See for example Babina et al. (2024), Di Maggio and Yao (2021), He et al. (2023). This literature considers a static environment where absent open banking the traditional banking sector has more information than the competition. One of the themes in this literature is the role of information sharing to stimulate the entry of new fintech firms. These static analyses ignore that the possibility of realizing profits in later periods because of informational monopoly incentivizes entry. This aspect is analyzed by Jin and Vasserman (2021) in the context of car insurance markets. They study how the adoption of technologies to monitor driving habits affects the dynamics of prices when such information is proprietary. They show in a counterfactual of their estimated model that a policy forcing firms to share data will reduce the incentives of firms to elicit such data and lead to lower welfare in equilibrium. This conclusion is in line with our result that full information sharing is never optimal.

2 Insurance economy

Consider a pure exchange economy that lasts for two periods, $t = 1, 2$. There are two types of agents: consumers and insurance companies. The consumers are risk averse and demand insurance against fluctuations to their income. They have common preferences over the consumption good given by $u(c)$ with u strictly concave. They discount the future with β and we assume that $\beta = q$ where q is the price of period 2 consumption in terms of period 1 consumption.

Consumers are uncertain about their income in period 1 and 2. Income can take on two values, $y_t \in \{y_L, y_H\}$ with $y_H > y_L$. We assume that $y_1 \sim \pi_1(y_1)$ and $y_2 \sim \pi_2(y_2|y_1)$. Thus, period 1 income is useful to forecast period 2 income. We define Y_{2H} and Y_{2L} to be the expected income in period 2 conditional on having a high or low income realization in period 1,

$$Y_{2H} \equiv \sum_{y_2} \pi_2(y_2|y_H)\, y_2 > Y_{2L} \equiv \sum_{y_2} \pi_2(y_2|y_L)\, y_2.$$

We assume that $Y_{2H} > Y_{2L}$ so a higher income realization in period 1 forecast a higher

[4] Another strand of the literature allows for asymmetric information to arise about the outside offer that the agent receives. See for example Hopenhayn and Werning (2008).

average income in period 2. We also defined the (unconditional) expected income in period 1, $Y_1 = \sum_s \pi_1(y_s) y_s$,and in period 2, $Y_2 = \sum_s \pi_1(y_s) Y_{2s}$.

In the appendix, we show how this restriction arises in an economy where consumers can be of one of two unknown types, $\theta \in \{\theta_L, \theta_H\}$, that affect the probability distribution of income $p(y|\theta)$, with $p(y_H|\theta_H) > p(y_H|\theta_L)$, and agents learn about the type by observing the income realizations. Such economy can be exactly mapped into our economy with $Y_1 = Y_2$, so expected income in period 1 and unconditional expected income in period 2 are the same. We choose to illustrate our results with a simple pure-exchange economy to minimize the required notation.

One can interpret this setting as a labor market application by letting income be output and c being the compensation paid to the worker. In the appendix, we show how one can reinterpret this model as one where lenders learn about the default probability of a borrower.

Information and market structure Firms compete to attract consumers by offering contracts that specify consumption levels conditional on all the available information at the time. At the beginning of period 1, all agents share the same information. Firms simultaneously offers contracts to consumers and consumers chooses which contract to sign among the offered ones. We will call the firm chosen by the consumer in period 1 the *incumbent*. At the end of period 1, income y_1 is realized with probability $\pi_1(y_1)$ and it is observed by the consumer and the incumbent. Payments $c_1(y_1)$ for period 1 consumption are made. The *outsiders*–i.e the other firms–do not observe y_1 directly but they observe a public signal m in some signal space M. Public signals are distributed according to some distribution $\mu \in \Delta(M)$ where $\mu(m|y_1)$ denotes the share of consumers with income y_1 that receive a signal m. We will refer to (M, μ) as the *public disclosure policy*.

At the beginning of period 2, the incumbent acts as a Stackelberg leader and offers a contract conditional on the consumer's history (y_1, m). Outsiders observe the menu of contracts offered by the incumbent and can poach consumers by offering a menu of insurance contracts $x^o(m)$ conditional on publicly available information only, m. Critically, outsiders do not observe the realization of y_1 and the contract offered by the incumbent to a particular consumer. Firms observe the contracts offered and decide whether to withdraw their contract with a cost $\varepsilon \geqslant 0$. Finally, consumers choose the contract that maximizes their utility among those that are left after the withdrawal stage. If there are contracts that offer the same value, we consider the following tie-breaking rule: if the incumbent is among the contract with highest value then consumer stays with the incumbent; if not they are evenly split among the outsiders with the same offer.[5]

[5]The selection rule is without loss of generality because the incumbent has an informational advantage and can always attract all consumers by offering a little more consumption and still making positive profits.

In setting up the interaction among firms in period 2, we follow Netzer and Scheuer (2014) and allow for a stage in which firms observe other firms' contract and can decide to withdraw their contracts with a small cost $\varepsilon > 0$. This assumption guarantees that there exist a unique continuation equilibrium in pure strategies in period 2 for any signal m and contract $c_2(y_1, m, y_2)$ offered by the incumbent. In particular, the ability to withdraw the offered menu of contracts x^o guarantees the existence of equilibrium outcomes where there is cross-subsidization among contracts offered in a menu by the outsiders. The possibility of withdrawing the contract after observing the set of contracts offered in equilibrium rules out deviations in which competitors offer a cream-skimming contract to the high-income consumers. These deviations are at the root of the possibility of inexistence of pure strategy equilibria in Rothschild and Stiglitz (1976). This is similar to the logic in Hellwig (1987).[6] The presence of a strictly positive cost of withdrawing guarantees uniqueness of continuation equilibrium outcomes by ruling out non-competitive behavior that can arise as explained by Netzer and Scheuer (2014). [7]

We will consider two cases. First, as a benchmark, we assume that the incumbent can commit to the terms of the contract in period 2 but the consumers can walk away from the incumbent and sign with another firm. Second, we assume that also the insurance firm cannot commit to the terms of the contract in period 2. Under both assumptions on commitment, we characterize the equilibrium outcome for a given public disclosure policy and then characterize the one that maximizes ex-ante welfare.

As a warm-up, note that if both the incumbent and the consumer can commit to the terms of the contract in period 2 and to stay with the firm respectively, then the public disclosure policy is irrelevant and the initial contract provides perfect insurance with cross-subsidization across ex-post types i.e. for all t, y^t, $c_t\left(y^t\right) = \frac{Y_1+qY_2}{1+q}$. The public disclosure policy affects equilibrium outcomes when agents are not committed to their actions in period 2 and the information available to the outsiders can affect the outside options for the consumers and the incumbents.

3 Equilibrium outcome in period 2

We start characterizing the equilibrium by studying the continuation equilibrium in period 2 given the incumbent's offered contract, $c_2(y_1, m, y_2)$ and withdrawal strategy $\delta(x^o)$ where x^o is the set of contracts offered by the outsiders in period 2. This characterization

Our selection rule effectively makes the incumbent choice set closed.

[6] Hellwig (1987) allows firms to offer only one contract, not menus.

[7] An alternative in our setup would be to select – among the set of possible equilibrium outcomes – the one that minimizes the value for the incumbent. This "robust" selection will deliver the same equilibrium outcome as in the presence of a positive withdrawal cost.

does not depend on the incumbent ability to commit and gives the outside options that equilibrium contracts must satisfy to be immune from poaching in equilibrium.

Consider agents with a signal m, and let V_L and V_H be the value offered by the incumbent contract to low and high-income consumers respectively. Denote by s the fraction of consumers that have$y_1 = y_H$:

$$s = \frac{\mu(m|y_H)\,\pi_1(y_H)}{\sum_{y_1}\mu(m|y_1)\,\pi_1(y_1)}. \tag{1}$$

To characterize the continuation equilibrium, it is useful to define the maximal value that can be offered by the outsider to a consumer with history (y_H, m):

$$V^o(s) = \max_{c_H^o(y_2), V_L^o} \sum_{y_2} \pi_2(y_2|y_H)\,u\left(c_H^o(y_2)\right) \tag{2}$$

subject to the outsider's non-negative profit condition,

$$s\sum_{y_2}\pi_2(y_2|y_H)\left(y_2 - c_H^o(y_2)\right) + (1-s)\left[\sum_{y_2}\pi_2(y_2|y_L)\,y_2 - C\left(V_L^o\right)\right] \geqslant 0,$$

where $C = u^{-1}$, the incentive compatibility constraint,

$$V_L^o \geqslant \sum_{y_2}\pi_2(y_2|y_L)\,u\left(c_H^o(y_2)\right),$$

and the participation constraint for the low-income consumers,

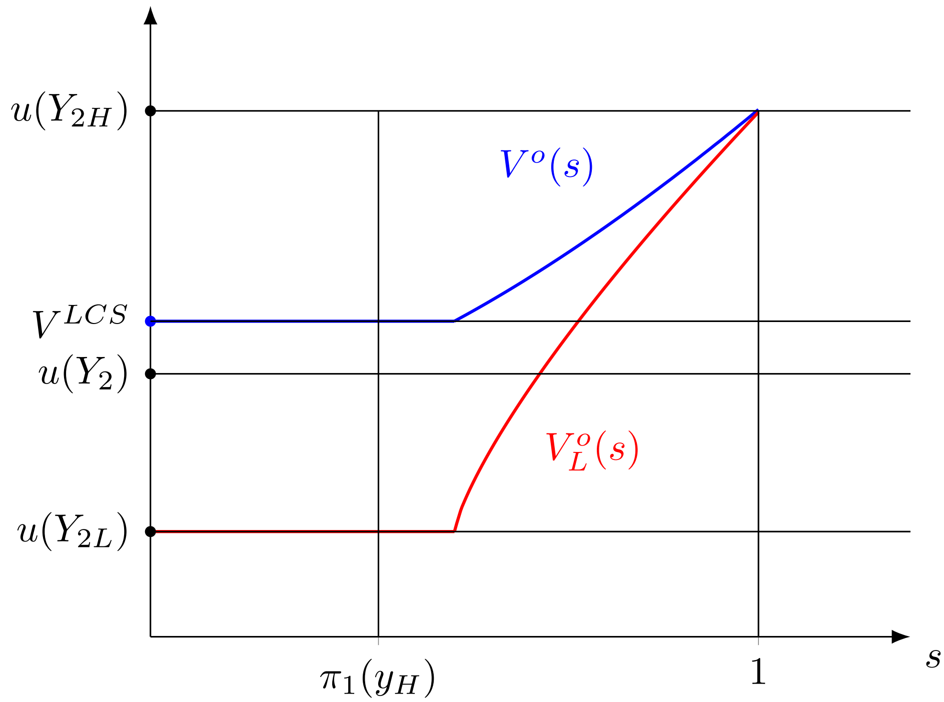
$$V_L^o \geqslant u(Y_{2L}). \tag{3}$$

Our description of (2) imposes that low-income consumers receive the same consumption in both states. This restriction is without loss of generality since it is never optimal to deliver V_L^o in a distorted manner, so we impose it purely to simplify the characterization of the solution to the problem.

The program (2) is what Netzer and Scheuer (2014) term the Miyazaki-Wilson Program after Miyazaki (1977) and Wilson (1977). Without the incumbent and its informational advantage, Netzer and Scheuer (2014) show that the solution to this program characterizes the unique equilibrium outcome of the game in period 2. If the share of high income consumers is low, the participation constraint for the low-income consumer (3) is binding and the optimal solution is the *least-cost-separating* allocation with no cross-subsidization across types and value V^{LCS} for the high-income type.[8] If instead the share

[8]Formally, the least-cost-separating allocation solves a restricted version of (2) where the participation

Figure 1: Outside option for the high-type



s is sufficiently high, the solution has cross-subsidization, $V_L > u(Y_{2L})$, as illustrated in Figure 1. This is because it is cheaper to provide a subsidy to the few low-income consumers than to distort the allocation for high-income income consumers to satisfy the incentive compatibility constraint.

The ability to withdraw contracts allows cross-subsidization to be a feature of the equilibrium outcome as mentioned above. Moreover, it also preempts the outsiders from offering a *cream-skimming* contract that only attracts the high-income agent without attracting the low-income type. This cream-skimming contract delivers a maximal value $V^{cs}(V_L)$ to the high-income type where

$$V^{cs}(V_L) = \max_{c(y_2)} \sum_{y_2} \pi_2(y_2|y_H)\, u(c(y_2)) \tag{4}$$

subject to the outsider's non-negative profit condition,

$$\sum_{y_2} \pi_2(y_2|y_H)(y_2 - c(y_2)) \geqslant 0,$$

constraint (3) must hold with equality.

and the incentive compatibility constraint for the low-income consumers,

$$V_L \geqslant \sum_{y_2} \pi_2 (y_2|y_L)\, u (c (y_2)) .$$

This contract can offer a higher value than V^o because the outsiders do not have to subsidize the low-income consumers to relax the incentive constraint but they rely on the incumbent offering them a higher value than $u (Y_{2L})$.[9] If the incumbent withdraws its offer (to the low-income consumers) the outsiders does not find it profitable to offer such contract.

The next lemma characterizes the continuation equilibrium in period 2 for consumers with signal m given a couple of values offered by the incumbent, (V_L, V_H), and the share of high-income consumers with signal m.

Lemma 1. *Given s, the incumbent's offer (V_L, V_H) and withdrawal policy:*

1. *If $V_H \geqslant V^o (s)$, $V_L \geqslant u (Y_{2L})$, and the incumbent withdraw its offers if the outsiders' offer the cream-skimming contract (4) then both low and high-income consumers stay with the incumbent and their value is (V_L, V_H) respectively;*

2. *If $V_H < V^o (s)$ and $V_L < V^o_L (s)$ then the outsiders will offer the Miyazaki-Wilson contract and attract both the low and the high-income consumers;*

3. *If $V_H < V^{cs} (V_L)$, $V_L \geqslant u (Y_{2L})$ and the incumbent does not withdraw its offers if the outsiders' offer the cream-skimming contract (4) to the high-income consumers, the low-income consumers stay with the incumbent and the high-income consumersaccept the cream-skimming contract.*

The proof is in Appendix. The main conclusion is that the incumbent firm retains high-income consumers if and only if it offers them a value above $V^o (s)$ and it withdraws its offer if $V^{cs} (V_L)$ is offered. Thus, the value $V^o (s)$ imposes a minimal continuation value for the high-income consumer. Absent the option to withdraw its contract, the incumbent firm would have to offer at least $V^{cs} (V_L)$ to high-income consumers in order to prevent them from being cream-skimmed by outsiders. As we show in the next sections, the incumbent's withdrawal option does facilitate cross-subsidization whenever the incumbent firm offers $V_L > u (Y_{2L})$ to low-income consumers, as under one-sided commitment–but not under two-sided lack of commitment.

[9] Note that $V^{LCS} = V^{cs} (u (Y_{2L}))$.

4 One-sided commitment

We now consider the long-term equilibrium contract and the optimal public disclosure policy when firms can commit to the continuation contract but the consumer can walk away from the incumbent in period 2 and accept an outsider's offer. We show that the optimal disclosure aims at minimizing the outside option for the borrower that draws y_H in period 1 to maximize the degree of cross-subsidization between types by relaxing the ex-post participation constraint for the y_H-type. To do so, the best disclosure policy reveals no information.

Optimal contract First, we characterize the equilibrium contract for a given (M, μ). Firms in period 1 offer a state-contingent long-term contract $\{c_1(y_1), c_2(y_1, m, y_2)\}$ and a withdrawal policy to attract consumers. Competition among firms implies that the equilibrium contract must maximize consumer's expected utility subject to a zero-profits condition for the firm. The continuation equilibrium described in Lemma 1 imposes further restrictions to the set of contracts offered. In particular:

Lemma 2. *The optimal contract offered in period 1 satisfies*

$$\sum_{y_2} \pi_2(y_2|y_H)\, u(c_2(y_H, m, y_2)) \geqslant V^o(s(m)), \tag{5}$$

$$\sum_{y_2} \pi_2(y_2|y_L)\, u(c_2(y_L, m, y_2)) \geqslant u(Y_{2L}). \tag{6}$$

Moreover, the incumbent commits to withdraw its offer if the outsiders offers a cream-skimming contract intended for the high-type.

The lemma states that we are always in case 1 of Lemma 1. In fact, if (5) and (6) are violated, we can find an alternative contract for the incumbent that satisfies (5) and (6), retains all consumers by delivering the same amount of utility, and makes (weakly) positive profits because the incumbent can use its informational advantage relative to the outsiders to offer the same value without having to distort allocations to satisfy the incentive compatibility constraints. These extra profits can then be used to increase consumption in period 1. Finally, committing to withdrawing the offer if the outsiders offer the cream-skimming contract is optimal because it relaxes constraint (5) as competitors cannot poach high-income types without also attracting low-types.

Given a public disclosure policy (M, μ), the optimal contract offered in period 1 solves

$$\max_{c_1, c_2} \sum_{y_1} \pi_1(y_1) \left[u(c_1(y_1)) + \sum_{m} \mu(m|y_1) \sum_{y_2} \pi_2(y_2|y_1)\, \beta u(c_2(y_1, m, y_2)) \right] \tag{7}$$

subject to the firm's non-negative profits condition,

$$\sum_{y_1} \pi_1 (y_1) \left[y_1 - c_1 (y_1) + q \sum_{m} \mu (m|y_1) \sum_{y_2} \pi_2 (y_2|y_1) (y_2 - c_2 (y_1, m, y_2)) \right] \geqslant 0,$$

and the ex-post participation constraints for high and low-income consumers (5) and (6) respectively.

It is clear that the equilibrium contract offers full consumption insurance against income fluctuations in period 1, $c_1 (y_L) = c_1 (y_H) = c_1$, and against fluctuations in income in period 2 contingent on (y_1, m) i.e. $c_2 (y_1, m, y_L) = c_2 (y_1, m, y_H) = c_2 (y_1, m)$ for all (y_1, m). The participation constraint for the low-income consumer (6) is going to be slack because the incumbent wants to subsidize consumption of the low-income type in period 2.

Optimal disclosure policy We now tun to the design of the optimal public disclosure policy. The next proposition shows that the value in (7) is maximized by a disclosure policy that provides no information under the following assumption:

Assumption 1. $K (s) \equiv C (V^o (s))$ *is convex.*

That is, the incumbent's cost of providing utility $V^o (s)$ in period 2, $K (s) = C (V^o (s))$, is convex in the share of high-income consumers. Assumption 1 guarantees that the maximal profits the incumbent firm makes on high-income consumers do not increase if those consumers are associated to more than one signal. If this assumption is violated, it may be optimal to provide some information to increase the amount of resources that can be extracted from high-income consumers and reallocated to the first period and/or to the low-income consumers in the second period, potentially increasing consumers' ex-ante utility if the ex-post participation constraint for the high-income consumers (5) is binding.

Throughout, we will assume that $K (s)$ is convex. In the Appendix, we provide sufficient conditions for this to be the case for an economy with log utility. We show that this is the case if $Y_{2H} - Y_{2L}$ and $\pi_2 (y_H|y_L)$ are sufficiently small. In all our numerical examples we find that the function $K (s)$ is convex, that is, the sufficient conditions are not necessary.

Proposition 1. *Under Assumption 1, the optimal disclosure policy reveals no information when the firm has commitment.*

Proof. Following the logic in Kamenica and Gentzkow (2011), we can equivalently write the problem of choosing the optimal disclosure policy μ as one of choosing the optimal distribution of shares s subject to the Bayesian plausibility constraint,

$$\sum_s p(s)\, s = \pi_1(s)\,. \tag{8}$$

That is:

$$\max_{p\in\Delta([0,1])} \max_{c_1, V_L(s), V_H(s)} u(c_1) + \beta \sum_s p(s)\left[sV_H(s) + (1-s)\, V_L(s)\right] \tag{9}$$

subject to

$$c_1 = Y_1 + q\left[Y - \sum_s p(s)\left[sC(V_H(s)) + (1-s)\, C(V_L(s))\right]\right] \geqslant 0$$

$$V_H(s) \geqslant V^o(s)\,, \tag{10}$$

and (8).

Suppose by way of contradiction that the optimal disclosure policy induces a distribution $p \neq p^*$ where p^* is such that $p^*(\pi_1(y_H)) = 1$ and 0 otherwise. If $\bar{V}_H = \sum_s p(s)\, sV_H(s) \geqslant V^o(\pi_1(y_H))$, then by the convexity of C it is possible to deliver $\bar{V}_H$ at a lower cost with no information. Thus, by providing no information it is possible to improve the ex-ante utility by increasing consumption in period 1 (or by increasing $V_L(s)$).

If instead $\bar{V}_H = \sum_s p(s)\, sV_H(s) < V^o(\pi_1(y_H))$ note that

$$\sum_s p(s)\, sC(V_H(s)) \geqslant \sum_s p(s)\, sC(V^o(s)) = \sum_s p(s)\, sK(s) > \pi_1(y_H)\, K(\pi_1(y_H))$$

where the first inequality follows from (10), the second from the definition of K and the last from the convexity of K. Thus, providing no-information and delivering $V_H = V^o(\pi_1(y_H))$ increases the continuation value for high-income consumers and it lowers it cost. Thus providing information cannot be optimal. Q.E.D.

The optimal public disclosure policy aims at minimizing the outside option for high-income consumers in period 2 to maximize the degree of cross-subsidization between types. Revealing no information is optimal because it maximizes the distortions that the outsiders must impose on the high-income consumers to separate them from low-income ones. These high distortions lower the value for the high-income consumers that can be offered by the outsiders and allow for greater cross-subsidization.

We can further characterize the optimal allocation under one-sided lack of commitment. The optimal allocation has

$$c_1 = c_2(y_L) \leqslant c_2(y_H) = C(V^o)\,. \tag{11}$$

where the inequality is strict if the ex-post participation constraint (5) is slack. As in Harris and Holmstrom (1982) and Thomas and Worrall (1988), the equilibrium allocation has perfect consumption smoothing between period 1 and 2 after a low income realization, but consumption must be increased after a high income realization if (5) is binding to retain the high-income consumers.

5 Two-sided lack of commitment

We next consider the case in which also the firms cannot commit to the terms of the contract beyond the current period. Here we show that in the twice repeated economy it is not possible to cross-subsidize the low type in period 2 and, no matter what the public disclosure policy is, we have that $c_2(y_L, m, y_2) = Y_{2L}$. Thus, the public information disclosure does not hinder (or enhance) the cross-subsidization that the low-income consumer can receive in the second period. However, information disclosure does determine the extent of consumption smoothing between period 1 and period 2 conditional on being a high-income consumer. The informational advantage of the incumbent affects the profits it can extracts in period 2 on high-income consumers but, because of competition in period 1, these profits are rebated to the consumer in period 1. Intuitively, the more information is disclosed, the lower the incumbent's ex-post profits, and the more consumption is tilted toward the second period.

The optimal disclosure policy can be described by a two-signal system, $m \in \{g, b\}$. All high income consumers received the good signal g. Low-income consumer receive the good signal with probability $1 - \mu$ and a bad signal with probability μ. Thus, μ here measures the degree of informativeness of the disclosure policy. If $\mu = 1$ there is perfect information while if $\mu = 0$ the outsiders have no information in addition to their prior. We show that the optimal disclosure policy spans from no information to full information according to the income distribution in the first period and the value offered by outsiders to high income consumers absent information, with more disclosure being optimal the higher the first and the lower the second.

Outcome in period 2

We characterize the outcome by backward induction starting from the terminal period 2. Agents have history y_1 and a signal $m \sim \mu(y_1)$. We model the insider as a Stackelberg leader that offers its contract to existing consumers, mimicking the timing in the case with one-sided commitment. The timing is the same as in the one-side commitment case.

The next lemma characterizes the unique continuation equilibrium outcome for a given public disclosure policy:

Lemma 3. *For any signal m, all consumers are fully insured against income fluctuations in period 2. There is no cross-subsidization from the high income consumers to the low income consumers and the latters always consume $c_2(y_L, m, y_2) = Y_{2L}$. The consumption of high income agents is*

$$c_2(y_H, m, y_2) = C(V^o(s(m))).$$

Proof. The logic of Lemma 2 implies that the incumbent will always offer contracts that satisfy (5) and (6). We are going to show that the following is the unique equilibrium outcome: the incumbent offers a contract with full insurance and value $V^o(s(m))$ to the consumers with history (y_H, m) and a contract with full insurance and consumption level Y_{2L} to the consumers with history (y_L, m).

Suppose first that in equilibrium low-income consumers receive a payoff of $u(Y_{2L})$. It is then clear that the insider will offer a contract with full insurance and value $V^o(s(m))$ to the consumers with history (y_H, m). Full insurance is optimal to minimize the cost of delivering such level of utility. Note that with full insurance the incumbent is making positive profits because $C(V^o(s(m))) \leqslant Y_{2H}$, with equality only if the signal is fully revealing. Offering a value $V^o(s(m))$ is optimal because if the incumbent offers less then the outsiders can attract all high-income consumers and erase all the incumbent's profits. Offering more is not optimal because it only reduces profits.

We are left to show that it is optimal to offer a full insurance contract with value $u(Y_{2L})$ to the low income consumers. Clearly, offering less is not feasible because any outsider can always offer a full-insurance contract with value $u(Y_{2L})$. Offering more has no advantages as the outside options for the high-income type is constant in the value offered by the incumbent to the low-income type. Q.E.D.

As described in the proof, the incumbent must offer a value of $V_H = V^o(s(m))$ to agents with history (y_H, m) to prevent competitors from poaching them away. Note that, unless the signal is perfectly informative, firms make profits on the high income consumers because

$$C(V^o(s(m))) < Y_{2H}.$$

This is because the incumbent knows the income realization in period 1 and the contract does not need to satisfy the incentive compatibility constraint in order to exclude low income consumers. Thus, the incumbent can offer the same value as the outsiders, but it does so by offering full insurance against period 2 income fluctuations, hence economizing on costs.

Consumers with low income in period 1 always consume their expected value of income Y_{2L}. Thus, firms make no profits on them. Even if consumers with $y_1 = y_L$ may receive a good signal and be pooled with the agents with $y_1 = y_H$, because of adverse selection they still get to consume Y_{2L}. Thus, their value is independent from the signal and from the disclosure policy.

Outcome in period 1

In period 1, insurance companies are competing for consumers. The equilibrium contract maximizes the consumer's expected utility subject to the dynamic zero profits condition and anticipating future contracts. Firms anticipate that they are going to make profits in period 2 on the consumers with a high realization in period 1 (unless the signals fully reveal the agent's type). Because of ex-ante competition, they are distributing such profits to consumers in period 1. The optimal level of consumption offered in period 1 solves

$$V_1 = \max_{c_1} \sum_{y_1} \pi_1(y_1) \left[u(c_1(y_1)) + \sum_m \mu(m|y_1) V^o(s(m)) \right]$$

subject to

$$\sum_{y_1} \pi_1(y_1) \left[y_1 - c_1(y_1) + q \sum_m \mu(m|y_1) \sum_{y_2} \pi_2(y_2|y_1)(y_2 - C(V^o(s(m)))) \right] \geqslant 0.$$

It is clear that the optimal contract offers perfect insurance statically, i.e.

$$c_1(y_1) = c_1 = Y_1 + q\pi_1(y_H) \left[Y_{2H} - \sum_m \mu(m|y_s) C(V^o(s(m))) \right]. \tag{12}$$

We summarize the characterization of the equilibrium outcome given a public disclosure policy (M, μ) in the next lemma:

Lemma 4. *Given a public disclosure policy* (M, μ)*, the equilibrium outcome has*

$$c_1(y_1) = Y_1 + q\pi_1(y_H) \sum_m \mu(m|y_H) \Pi(m) \tag{13}$$

$$c_2(y_L, m, y_2) = Y_{2L} \tag{14}$$

$$c_2(y_H, m, y_2) = Y_{2H} - \Pi(m) \tag{15}$$

where $\Pi(m) = Y_{2H} - C(V^o(s(m)))$.

Summing up, the equilibrium outcome does not have cross-subsidization in the second period between consumers with high and low income in the first period. No matter what the public disclosure policy is, a consumer with low income realization is offered a contract with full consumption insurance in the second period but no subsidy. There is though potentially cross subsidization across periods between period 1 and period 2 after a good realization in period 1, $y_1 = y_H$: the higher the profits that the incumbent makes on high-income consumers in period 2, the higher the consumption in period 1. Moreover, other than in the case with perfect information revelation, there is essentially no mobility in period 2: all high-income consumers stay with the incumbent because the incumbent makes strictly positive profits on them due to its informational advantage. Low-income consumers are indifferent between staying with the incumbent or moving to an outsiders.

6 Optimal public disclosure policy

We now characterize the optimal public disclosure policy. Given the characterization in Lemma 4, the choice of the disclosure policy does not affect the consumption of the low income consumer that always consumes Y_{2L} in the second period but it can affect the consumption in period 2 for high income consumers and consequently in period 1. In fact, by affecting the share of high-income consumers with a given signal, the information designer can affect the outside option for the high-income consumer in period 2 and therefore the value the incumbent offers them to retain them. Such values can range from the value of the least-cost-separating contract to $u(Y_{2H})$ because that is the range of V^o as illustrated in Figure 1.

Formally, using the characterization in Lemma 4, we can write the problem of choosing the optimal public disclosure policy as

$$\max_{c_1,(\mu,M),s(m)} u(c_1) + \beta\pi_1(y_H)\sum_{m\in M}\mu(m|y_H)\,V^o(s(m)) + \beta\pi_1(y_L)\,u(Y_{2L}) \tag{16}$$

subject to the intertemporal zero-profit condition

$$Y - c_1 + q\pi_1(y_H)\sum_{m\in M}\mu(m|y_H)\left[Y_{2H} - C(V^o(s(m)))\right] \geqslant 0$$

and the share of y_H type with signal m is

$$s(m) = \frac{\pi_1(y_H)\,\mu(m|y_H)}{\pi_1(y_H)\,\mu(m|y_H) + (1-\pi_1(y_H))\,\mu(m|y_L)}.$$

The optimal disclosure policy depends on whether the following condition is satisfied:

$$C(V^o(\pi_1(y_H))) \leqslant Y_1 + q\pi_1(y_H)\left(Y_{2H} - C(V^o(\pi_1(y_H)))\right) \tag{17}$$

That is, if under no-information disclosure, consumption in period 2, $C(V^o(\pi_1))$, is lower than consumption in period 1.

For low values π_1, condition (17) does not hold, the lower bound on high income consumers' value is high and incumbent's profits in period 2 are too low to attain consumption equalization. Thus, the best that can be done is to minimize such profits. When the incumbent's cost of providing utility in period 2, $K(s) = C(V^o(s))$, is convex in the share of high-income consumers with a given signal, then the optimal way to do so is to provide no information.

For intermediate values of $\pi_1(y_H)$, condition (17) is satisfied and the disclosure policy is designed to perfectly smooth consumption between period 1 and period 2 after a high income realization in period 1,

$$c_1 = c_2(y_H). \tag{18}$$

This can be achieved by considering two signals only, $M = \{g, b\}$ (good or bad) with a *bad-signal* structure: All high-income consumers receive a good signal together with a fraction of low-income individuals. Only low income consumers receive the bad signal. It is optimal to have some low income consumers with a good signal–even if, as we have shown, this does not affect their consumption–to manipulate the value for the high income consumers and equalize their consumption between period 1 and 2. This is attained for an intermediate value of signal informativeness between full information revelation and no-information.

For values of $\pi_1(y_H)$ sufficiently high (close to 1) condition (17) holds but it is not possible to find a disclosure policy such that consumption in period 1 and 2 are equated. In fact, for $\pi_1(y_H) = 1$, even under full information disclosure we have $c_2(y_H) = Y_{2H} < y_H = Y_1 = c_1$ so consumption is front-loaded. [10]The firm should commit to delayed consumption profile (saving on the behalf of the consumer) but it cannot commit to do so. Thus, for π_1 large enough it is optimal to offer full information and consumption is still front-loaded as $c_1 = Y_1 > Y_{2H} = c_2(y_H)$.

The following proposition characterizes the optimal information disclosure:

Proposition 2. *There exists two cutoffs π^* and π^{**} with $0 < \pi^* < \pi^{**} \leqslant 1$ such that: i) If $\pi_1(y_H) \leqslant \pi^*$ then $c_1 < c_2(y_2)$ and it is optimal to provide no information; ii) If $\pi^* \leqslant \pi_1(y_H) \leqslant \pi^{**}$ then consumption is equalized between period 1 and 2 after a high income realization and the optimal disclosure policy has a bad-signal structure i.e. $M = \{g, b\}$ (good or bad) and $\mu(g|y_H) = 1$ and $\mu(b|y_L) \in (0, 1)$ so a bad signal fully reveals the period 1 income; iii) if $\pi_1(y_H) > \pi^{**}$ then $c_1 > c_2(y_H)$ and the optimal disclosure policy provides full information.*

[10]This is true unless income is perfectly persistent and $\pi_2(y_H|y_H) = 1$. In this case $Y_{2H} = y_H$ and the consumption is constant between period 1 and period 2 after a good income realization. This knife-edge case is not interesting as it is easy for the outsiders to separate consumers in period 2 and there are no adverse selection problems.

The formal proof is in the appendix. We first show that under the convexity of $K(s) = C(V^o(s))$ it is optimal to assign the same signal to all high-income consumers. This is because introducing dispersions for high-income types simply increases the cost of providing utility to such consumers. Thus, without loss of generality, we can consider two signals $M = \{g, b\}$ and have $\mu(g|y_H) = 1$. We are only left to choose the fraction of consumers with low income in period 1 that receive the same "good" signal. This is not going to affect their consumption but it affects the composition of the pool of agents with a good signal and therefore what incumbent must offer to retain the high income consumers. The information design problem can then induce any continuation values for the high income consumer in the range $[V^o(\pi_1(y_H)), V^o(1)]$.

We can then write problem (16) as

$$\max_{V_H} u(c_1(V_H)) + \beta\pi_1(y_H) V_H + \beta\pi_1(y_L) u(Y_{2L}) \tag{19}$$

subject to

$$c_1(V_H) = Y_1 + q\pi_1(y_H)[Y_{2H} - C(V_H)]$$

$$V_H \in [V(\pi_1(y_H)), V^o(1)]$$

and then recover $s(g)$ and $\mu(b|y_L)$ from

$$V_H = V^o(s(g)) \text{ and } s(g) = \frac{\pi_1(y_H)}{\pi_1(y_H) + \pi_1(y_L)(1 - \mu(b|y_L))}. \tag{20}$$

The optimal disclosure policy can then be represented by the share of high-income types with good signal, $s(g)$. In the appendix, we show that for low levels of $\pi_1(y_H)$ condition (17) does not hold or equivalently the constraint $V_H \geqslant V^o(\pi_1(y_H))$ is binding. It is therefore optimal to provide no information so $s(g) = \pi_1(y_H)$ as illustrated in Figure 2. This is because even by maximizing period 2 profits on high-income consumers, by minimizing their values in equilibrium, we have that $c_1 < c_2(y_H)$.

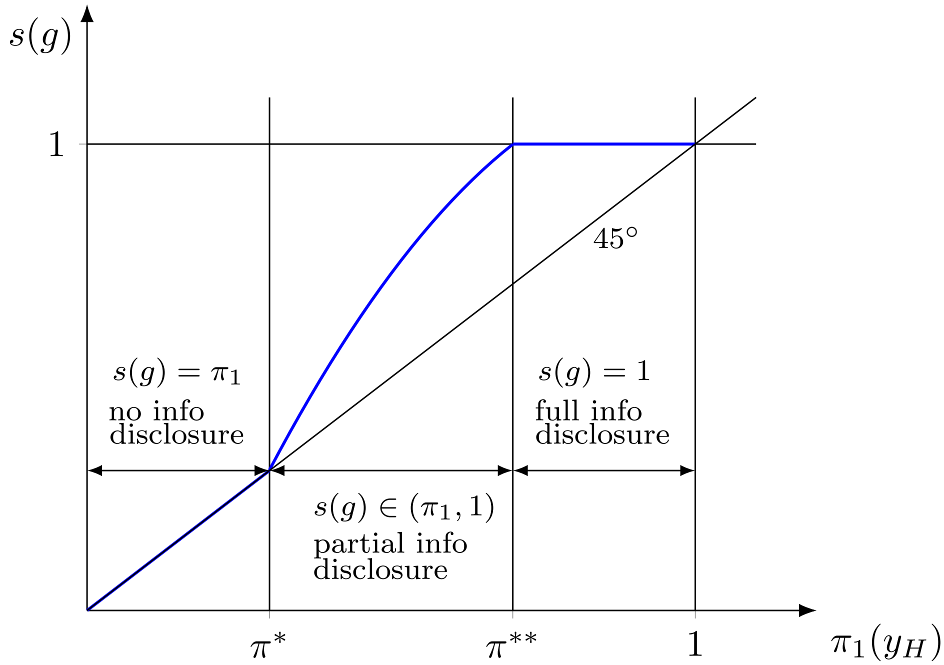
For intermediate values of $\pi_1(y_H)$ the value of expected profits under no-information are then higher (and also Y_1 is higher) so condition (17) is satisfied and V_H that solves (19) is interior. It follows that

$$c_1 = c_2(y_H) = \frac{Y_1 + q\pi_1(y_H) Y_{2H}}{1 + q\pi_1(y_H)}$$

and $s > \pi_1(y_H)$, as shown in Figure 2. Moreover, note since in this region $c_2(y_H)$ is increasing in $\pi_1(y_H)$ and $V^o(s)$ is increasing in s, then s must be increasing in $\pi_1(y_H)$.

For very high levels of $\pi_1(y_H)$ it is not possible to perfectly smooth consumption between period 1 and 2 after a high income realizations but now consumption is higher

Figure 2: Optimal share of high-income consumers conditional on good signal



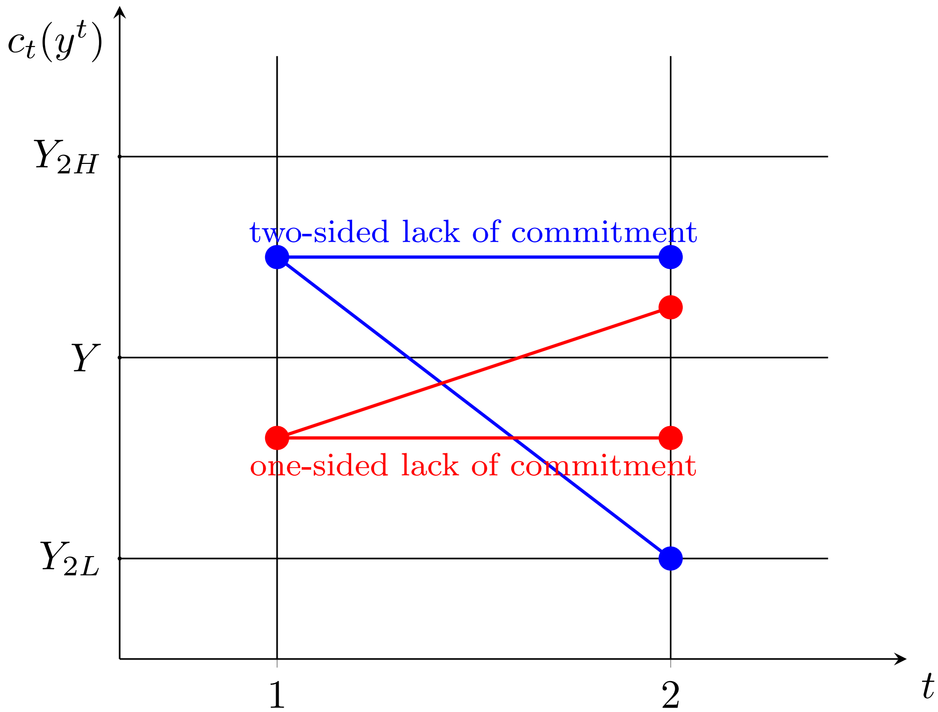
in period 1 and the optimal disclosure policy provides full information. Mechanically, the constraint $V_H \leqslant V^o(1)$ in (19) is binding for high $\pi_1(y_H)$ because increases in $\pi_1(y_H)$ keeping fixed $\pi_2(\cdot|\cdot)$ increases expected income in period 1 more than it increases Y_{2H}—except in the knife-edge case with $\pi_2(y_H|y_H) = 1$. In this case, the solution has $V_H = V^o(1)$ and $c_1 = Y_1 > c_2(y_H) = Y_{2H}$, so $\mu(b|y_L) = 1$ and $s(g) = 1$.

Consider now the consumption profile at the optimal public disclosure policy. Consumption is *front-loaded* and, for intermediate values of $\pi_1(y_H)$, there is perfect insurance between period 1 and period 2 after $y_1 = y_H$:

$$c_1 = c_2(y_H) > c_2(y_L). \tag{21}$$

This is the opposite result relative to the standard case when firms have commitment (e.g. Harris and Holmstrom (1982) and Thomas and Worrall (1988)). In such case consumption, is back-loaded and there is perfect insurance between period 1 and period 2 after a bad income realization. The different predictions for the optimal consumption profiles are illustrated in Figure 3. The critical difference is that when the incumbent can commit, it redistributes the profits it can extract from the high-income consumers in period 2 to consumption in period 1 and to low-income consumers in period 2. Consumption smoothing may not be perfect because the ex-post participation constraint for the high income consumer may be binding but consumption is constant otherwise. Without com-

Figure 3: Consumption profile under optimal disclosure policy for intermediate values of $\pi_1(y_H)$



mitment on the firm side, the incumbent does not provide any transfers to the low income consumers. (And so do the outsiders because of adverse selection since they anticipate that only low-income consumers will move to them in equilibrium.) Thus, the profits earned on high income consumers in period 2 are entirely rebated in period 1. This results in a larger increase in period 1 consumption. Thus, it is optimal to try to smooth consumption in period 2 for high-income consumers by providing some information and increasing their outside options. With commitment, it is optimal to minimize the outside option to extract as many resources as possible.

We have established how optimal information disclosure varies with the share of agents with a high-income realization in the first period, $\pi_1(y_H)$. Such income realization is central in our proposition because it determines the only information available to consumers and incumbents in the second period. Recall that our economy can be mapped to one with two unknown types that affect the distribution of income in each period. However, we cannot easily perform a similar comparative statics in the type economy. To see why, notice that, in the type economy, $\pi_1(y_H)$ is a function both of the underlying share of high types, ρ, and of the conditional probability of high income realization for given type. An increase in ρ that induces a higher probability of observing a high income in the first period, $\pi_1(y_H)$, would also induce a higher conditional probability of observing a high income in the second period, $\pi_2(y_H|y_H)$. The transition probability of income $\pi_2(\cdot|\cdot)$ affects both outsiders' profits and the cost of providing incentives, hence complicating the mapping from ρ to outside options and, ultimately, to optimal information disclosure. In Appendix, we further prove that optimal disclosure in the type economy never entails full information disclosure since in that economy $Y_1 < Y_{2H}$. Numerical examples show that if the share of high type ρ is small then it is typically optimal to provide no information while it is optimal to provide partial information if such share is sufficiently high.

7 Discussion

We now discuss some of the assumptions of our model and some extensions.

Regulation and commitment We model the choice of the optimal public disclosure as one made by a regulator. The presence of a regulator is not needed for the implementation of the optimal public disclosure policy. A firm in period 1 endowed with a commitment technology for reporting information in period 2 will choose the same disclosure policy as the planner to maximize the consumers' welfare subject to the zero profits condition otherwise another firm can adopt the optimal disclosure policy, offer the same equilibrium value and make positive profits.

A commitment technology to truthfully reporting according to (μ, M) is necessary. Consider a case in which condition (17) holds and it is optimal to provide information. The commitment technology is needed because in period 2 the incumbent would have incentives to provide no information to maximize its profits. To see this, suppose the incumbent offered the optimal disclosure policy from the ex-ante perspective. Consider a deviation where the incumbent provides no information (say it assigns signal g to everybody). In this case the value for the high type is $V^o(\pi_1) < V^{optimal}$ and the low types get same consumption as the original allocation. Thus, the incumbent makes more profits on the high type. Providing no information is optimal ex-post and a commitment disclosure technology is necessary. The same argument can be made if one assumes the notion of commitment in Lin and Liu (2022) that the distribution of signals ex-post must be conforming with μ^*. This is because ex-post the incumbent will have incentives to assign the good signal to all low-income consumers -at least as long as it is feasible and not detectable- and a bad signal to the high-income consumers.

Connection with information design literature Our environment departs from those in the traditional information design literature (see for example Kamenica and Gentzkow (2011), GK) along at least three dimensions. First, it is dynamic, in that information in the second period determines the incumbents' profits that are rebated to the agent in the first period. Second, given an information structure, payoffs are determined by the equilibrium in a game between multiple agents. Third, due to risk aversion and the dynamic nature of the problem, the objective function of the designer/sender typically depends on the whole posterior distribution and not just the posterior mean. Hence, the particular characteristics of our environment do not allow us to cast the information design problem as a variation of the well-known GK framework and its associated graphical representation.

Choosing a disclosure policy is equivalent to choosing a distribution of shares of high income consumers with a given signal. Feasible distributions f must satisfy the Bayesian plausibility constraint requiring that

$$\sum_s sf(s) = \pi_1(y_H) \tag{22}$$

That is, the total share of high income consumers is $\pi_1(y_H)$.

Define

$$\begin{aligned} w_2(s) &\equiv sV^o(s) + (1-s)\,u(Y_{2L}) \\ w_1(f) &\equiv u\left(Y_1 + q\sum_s (Y_{2H} - C(V^o(s)))\, sf(s)\right) \end{aligned}$$

where f is a probability measure over $[0,1]$. In a way similar to GK, we can reformulate the problem the problem to solve for the optimal disclosure policy, (16), as

$$\max_{f\in\Delta([0,1])} w_1(f) + \beta\sum_s w_2(s)\, f(s) \tag{23}$$

subject to (22). As it is clear from (23), the posterior mean is not a sufficient statistic for the problem. Even if $\sum_s w_2(s)\, f(s)$ is concave, it is optimal to disclose information because of $w_1(f)$.

Extensions INFORMATION STRUCTURE. We assumed that income y_1 is not directly observable by outsiders and it is (partially) revealed to them via the signal (M, μ). Here we establish that what is key for a disclosure policy to affect the equilibrium allocation is not the lack of information acquisition by outsiders, but the existence of an information advantage by the incumbent. Consider an alternative economy in which y_1 is some major outcome (e.g. a car accident, a borrower's history default, a worker's history of layoffs) that is publicly observable, while consumers and the incumbent firm also observe a private, non-redundant and non-verifiable outcome $\tilde{y}_1$ (e.g. driving style, credit usage, performance at work). In the Appendix, we show that our results naturally extend to the modified environment. In particular, optimal information disclosure is still driven by intertemporal smoothing. The incumbent firm makes ex-post profits on consumers with high realizations of $\tilde{y}_1$, no matter their observed outcome y_1, and rebates such profits as first period consumption.

CONTRACT SPACE. We assumed that firms can offer menus of contract to potentially separate different types. All our results are valid in a version of the model in which outsiders are restricted to offer one single pooling contract with consumption $c^o(s) = sY_{2H} + (1-s)Y_{2L}$ as in the labor market application in Kahn (2013).

We assume that the insider cannot discriminate among consumers with the same history. In the Appendix, we show how this assumption changes the specifics of the equilibrium allocation without affecting the main message. In particular, the optimal signal structure maintains the same features as in the restricted case.

ACTION SPACE. In many applications, outcomes are the result not only of innate agent's characteristics, but also of individual effort, as in Holmström (1979). For example, a worker's human capital is determined by her intrinsic ability and by her investment in the acquisition of skills. By affecting the spread in consumption after good and bad outcomes, information design affects the amount of effort that can be sustained. In the Appendix, we extend our analysis to this case and show that optimal disclosure policy when both sides lack commitment has the same bad-news structure but it provides more information than the base case analyzed above. This because more information allows for more spreading in continuation values that incentivize the agent to exert effort in the first period.

8 Taste shock, switchers, and ex-post competition

So far we assumed that agents obtain utility only from consumption. Due to asymmetric information, the outsiders would infer any movers to be a low type and offer them full insurance at their expected income. Hence the equilibrium features no mobility in the second period–except perhaps from low types who are indifferent between firms. In practice, consumers do move across firms and it is natural to ask how the optimal information disclosure we established in the preceding sections generalizes to an environment that allows for a richer pattern of mobility. In this section, we extend our analysis by incorporating transitions between firms that are motivated by idiosyncratic preferences and not by contractual terms. For example, a worker might have to move to a different labor because of their spouse's job, or a car driver might decide to bundle their car with their new home insurance. The presence of idiosyncratic motives weakens adverse selection by making switches less informative of the agents' types. We are particularly interested in the implications of the introduction of taste shocks for the extent of cross-subsidization in the second period–absent in our benchmark model–and for optimal information design.

We assume that, at the beginning of period 2, an exogenous share $(1-\alpha)$ of agents receives a shock that forces them to leave the incumbent firm. The realization of the shock is privately observed only by the agent. For reference, the environment we analyzed in the previous sections correspond to the case $\alpha = 1$.

One might conjecture that a decrease in α would lower the optimal amount of information disclosure as stronger idiosyncratic motives make switching to a different firm less likely to negatively affect the outsiders' beliefs about the agent's type. We show that this is not the case and, in particular, optimal information disclosure μ is not monotonically increasing in α. To understand this result, we first consider optimal information design in the extreme case in which $\alpha = 0$ so that all consumers switch in the second period. We then consider the general case with $\alpha \in (0, 1)$ in which information design responds to both intertemporal and cross-subsidization motives.

8.1 All switchers: $\alpha = 0$

Consider a version of our two-period economy in which all agents exogenously end their relationship with their incumbent firm at the beginning of the second period. Under this extreme assumption, the second period of the model is equivalent to a static economy with privately informed agents à la Rothschild and Stiglitz (1976) and Netzer and Scheuer (2014).

Since all consumers leave their firm in the second period, consumption in the first period is equal to Y and independent from the information disclosure policy. We show that full information disclosure is never optimal. Perhaps surprisingly, we show that, under some conditions, partial information disclosure is optimal as it facilitates–instead of hindering–cross-subsidization between types.

8.1.1 Environment and allocation in period 2

At the beginning of period 2, all agents receive a preference shock that makes them terminate their relationship with their incumbent firm. The outsiders are aware of such separation and offer contracts conditional on their belief about the composition of the pool of agents with a given signal. We denote the composition of switchers with signal m by $\tilde{s}(m)$ to distinguish it from the overall composition of agents $s(m)$ in the full model of Section 8.2 below in which only a fraction of agents switch firms in the second period. The equilibrium outcome in period 2 is then the solution to problem (2) with $s = \tilde{s}(m)$.

Our first result states that for the equilibrium contract to feature cross-subsidization, the minimum pool quality $\tilde{s}$ associated with a good signal must be bounded away from 0.

Lemma 5. *There exists a cutoff pool composition* $\tilde{s}^* \in (0,1)$ *such that* $V^o(\tilde{s}) > V^{LCS}$ *if and only if* $\tilde{s} > \tilde{s}^*$.

The proof is in the Appendix.

8.1.2 Equilibrium and optimal information disclosure

In light of the previous lemma, it is easy to find conditions such that it is optimal to disclose some information about y_1. Recall that the composition of the pool of agents with a good signal is $\tilde{s}(\mu) = \frac{\pi}{\pi+(1-\pi)(1-\mu)} \in [\pi, 1]$ where with some abuse of notation we let $\mu = \mu(b|y_L) \in [0,1]$ denote the fraction of low types that receive a bad signal, and $\pi = \pi_1(y_H)$. It is easy to see that optimal information disclosure maintains a bad news structure as in the benchmark model. Under no information disclosure, that is $\mu = 0$, the posterior belief of the outsiders is equal to the prior π. If $\pi < \tilde{s}^*$, under no information disclosure the equilibrium contracts offered by the outsiders is the least-cost-separating contract aimed at the high types and the full insurance contract to the low types with value $u(Y_{2L})$.

In contrast, consider a pool of quality $\tilde{s} = \tilde{s}^* + \varepsilon$ for arbitrarily small $\varepsilon > 0$ so that the equilibrium contract offered by the outsiders features cross-subsidization. Such contract provides strictly higher value than the least-cost-separating contract to the high types, and it provides a subsidy to the fraction $1-\mu = \frac{\pi(1-\tilde{s}^*)}{(1-\pi)\tilde{s}^*} > 0$ of low types that belong to the pool with a good signal. The low types associated with a bad signal are unaffected by information disclosure and still obtain a value equal to $u(Y_{2L})$.

Thus, providing no information is not optimal for $\pi < \tilde{s}^*$. Information disclosure guarantees strictly higher expected value to consumers. We emphasize that this argument holds only in a right neighborhood of $\tilde{s}^*$. Further information disclosure eventually hurts the low types by pooling too few of them with the high types and, in the limit of full information disclosure, no cross-subsidization can be sustained. We summarize these results in the following Proposition:

Proposition 3. *The optimal information disclosure has a bad news structure. In addition, i) if* $\pi < \tilde{s}^*$, *the optimal information design prescribes some information disclosure,* $\mu > 0$*; ii) for all* $\pi \in (0,1)$, *full information disclosure is never optimal,* $\mu < 1$.

Proof. i) The first part follows directly from Lemma 5. Since all agents' values are constant for $\tilde{s} < \tilde{s}^*$ and strictly increasing in $\tilde{s}$ in a right neighborhood of $\tilde{s}^*$, there exists a pool composition $\tilde{s}' = \tilde{s}^* + \delta$, with $\delta > 0$, that delivers higher welfare than no information. Since $\pi < \tilde{s}^* < \tilde{s}'$, the pool $\tilde{s}'$ is obtained by setting $\mu = \frac{\tilde{s}'-\pi}{\tilde{s}'(1-\pi)} > 0$.

ii) To prove the second part of the statement, we consider a marginal deviation from full information disclosure and show that such deviation is always optimal. As $\tilde{s} \rightarrow 1$, the solution to (25) converges to the full information outcome for high types, $c_H(y_2) \rightarrow Y_{2H}\ \forall y_2$. In addition, incentive compatibility requires $c_L(y_2) = Y_{2H}$ for all y_2. Clearly, $V^o(s) > V^{LCS}$ so that the equilibrium contract features cross-subsidization of the low types. Consider the necessary focs for $c_H(y_2)$ in problem (27), letting λ and λ_{ic} be the multipliers on the non-negative profits for the firm and on the incentive compatibility constraint respectively:

$$\pi_2(y_2|y_H)\, u'(c_H(y_2)) = \lambda \pi_2(y_2|y_H) + \lambda_{ic}\pi_2(y_2|y_L)\, u'(c_H(y_2))$$

Summing them up in the limit for $\tilde{s} \rightarrow 1$ gives

$$u'(Y_{2H}) = \lambda + \lambda_{ic} u'(Y_{2H}).$$

Next consider the foc for V_L,

$$\lambda(1-s)\, C'(V_L) = \lambda_{ic}$$

where the equality follows from $V_L > u(Y_{2L})$ since the constraint (3) is slack and there is cross-subsidization. From the latter equality we obtain $\lambda_{ic} \rightarrow 0$ which in turn implies $\lambda \rightarrow u'(Y_{2H})$. The planner's objective is given by

$$V^o(\tilde{s}(\mu))\,\pi + [\mu u(Y_{2L}) + (1-\mu)\, V_L^o(\tilde{s}(\mu))]\,(1-\pi)$$

where $V_L^o(\tilde{s}(\mu))$ is the value for the low-type. Recall that μ is the fraction of low types associated with a bad signal and it is also our measure of information disclosure. The derivative of the above expression with respect to μ is equal to

$$\frac{\partial V^o(\tilde{s})}{\partial \tilde{s}} \frac{\partial \tilde{s}}{\partial \mu}\pi + \left[(u(Y_{2L}) - V_L^o(\tilde{s}(\mu)))(1-\pi) + (1-\mu)\frac{\partial V_L^o(\tilde{s}(\mu))}{\partial \tilde{s}}\frac{\partial \tilde{s}}{\partial \mu}\right](1-\pi)$$

When evaluated at full information, $\mu = 1$, such derivative simplifies to

$$\begin{aligned} &\lambda(Y_{2H} - Y_{2L})\left(\frac{1-\pi}{\pi}\right)\pi + (u(Y_{2L}) - u(Y_{2H}))(1-\pi) \\ &= \left(u'(Y_{2H}) - \left(\frac{u(Y_{2H}) - u(Y_{2L})}{Y_{2H} - Y_{2L}}\right)\right)(Y_{2H} - Y_{2L})(1-\pi) < 0 \end{aligned}$$

where $\frac{\partial V^o(\tilde{s})}{\partial \tilde{s}} = \lambda(Y_{2H} - Y_{2L})$ is the envelope condition in problem (27) and the last inequality follows from strict concavity of u. Q.E.D.

When adverse selection is severe (low $\tilde{s}$), the equilibrium features a least-cost-separating contract in which all low income consumers consume Y_{2L}. In such cases, information provision enables–instead of hindering–cross-subsidization (as in Goldstein and Leitner (2018)). This is because if the pool composition becomes sufficiently good, firms find it optimal to transfer resources to the (few) low-types in the pool in order to relax their incentive-compatibility constraint and provide a less distorted allocation to the high-types. Once the equilibrium has switched to pooling, additional information disclosure eventually lowers agents' ex-ante welfare since the loss from reducing the measure of cross-subsidized agents outweighs the gain from the higher consumption level attained by those who receive a good signal. Therefore, as the disclosure policy approaches full information, the common intuition about the negative relationship between information and insurance prevails.

Before moving to the general model, we notice that, for any disclosure policy, the equilibrium outcome in the model with $\alpha = 0$ coincides with an equilibrium outcome in our benchmark model without taste shocks if the contract offered by the incumbent to each individual consumer were observable by outsiders in the second period. If outsiders observed the contract offered to each consumer, any unequal treatment in the second period would be fully revealing and erase the incumbent's profits. For any disclosure policy, the full information allocation is an equilibrium outcome. The only other possible equilibrium outcome is generated if the incumbent (and the outsiders) offers $(V^o(s), V_L(s))$ to all consumers in the second period. In that case, the incumbent would disregard its informational advantage and behave like any other uninformed outsider. Optimal information disclosure would then control the extent of cross-subsidization among consumers in the second period, in the same way as under a taste shock that terminates all relationships with the incumbent.

8.2 Full model analysis: Equilibrium and optimal disclosure

If $\alpha \in (0,1)$ the economy features both intertemporal consumption smoothing, enabled by ex-post profits of incumbent firms, and, possibly, cross-subsidization among switchers with different income realizations in period 1. The agents who do not receive the taste shock stay with the incumbent firm since the latter is always able to offer them as much as their outside offer without distorting their consumption. Specifically, given a good signal and the associated pool composition s in the second period, the high types who do not receive the taste shock obtain a value equal to $V^o(s)$. Such value is the maximum that outsiders are willing to offer to the high types while knowing that they would attract the entire population. [11] At the same time, the incumbent has no incentive to deliver more

[11] Given the bad news structure of optimal information disclosure, we omit m as an argument of s and $\tilde{s}$ with the understanding that we refer to the composition of agents with good signal.

than Y_{2L} to the low types so it is without loss of generality to assume that all the low types leave the incumbent whenever indifferent. It follows that the composition of the pool of agents that receive a good signal *and* leave the incumbent firm is equal to

$$\tilde{s}(s) = \frac{(1-\alpha)s}{(1-\alpha)s + (1-s)}.$$

It is easy to see that $s \leqslant \tilde{s}$ with strict inequality whenever $\alpha > 0$. That is, for a given disclosure policy, the pool of switchers has a worse composition than the overall pool of agents. Competition among outsiders guarantees that the high type switchers receive a value equal to $V^o(\tilde{s})$. Last, the low types receive $V_L^o(\tilde{s})$. [12]

In light of the previous discussion, we can decompose the total ex-ante welfare into three terms,

$$V^d(s) + \beta\pi(1-\alpha)V^o(\tilde{s}) + \beta(1-\pi_0)\mathbb{E}_\mu[V_L]$$

where

$$V^d(s) \equiv u(Y + \beta\pi_0\alpha(Y_{2H} - C(V^o(s)))) + \beta\pi_0\alpha V^o(s)$$

is the sum of the value in the first period and in the second period for (high type) stayers; $V^o(\tilde{s})$ is the value in the second period for high types who leave the incumbent; $\mathbb{E}_\mu[V_L] \equiv (1-\mu)V_L^o(\tilde{s}) + \mu u(Y_{2L})$ is the value in the second period for low types.

We begin by stating the following lemma.

Lemma 6. *Let $\tilde{s}^*$ be the largest $\tilde{s}$ such that $V^o(\tilde{s}) = V^{LCS}$. Let $s^*(\alpha)$ be implicitly defined by $\tilde{s}(s^*(\alpha);\alpha) = \tilde{s}^*$. If $\alpha < 1$, the optimal value of s is never equal to $s^*(\alpha)$.*

Proof. The values of switchers, $V^o(\tilde{s})$ and $\mathbb{E}_\mu[V_L]$, are constant for $\tilde{s} < \tilde{s}^*$–since the high types receive V^{LCS} and the low types receive $u(Y_{2L})$–and strictly increasing in $\tilde{s}$ in a right neighborhood of $\tilde{s}^*$. The value $V^d(s)$ is weakly concave in s and strictly concave whenever $V^o(s) > V^{LCS}$, which happens for all $s > s'$ with $s' < s^*$ since $s > \tilde{s}(s)$.

For s^* to be optimal, $V^d(s)$ must be maximal at s^*. If not, $V^d(s)$ is either strictly decreasing in a left neighborhood of s^*–in which case s^* cannot be optimal–or strictly increasing in a right neighborhood of s^*–in which case there exists a value of $s > s^*$ such that all three components of welfare are strictly greater than in s^*.

[12] The withdrawal stage is necessary in the environment with switchers as well in order to guarantee that an equilibrium exists. Recall that $V^o(\tilde{s})$ is the value of the Wilson-Myazaki contract. For sufficiently high $\tilde{s}$, $V_L^o(\tilde{s})$ exceeds $u(Y_{2L})$ since the outsiders prefer to cross-subsidize low-income consumers in order to reduce the distortion on the contract offered to the high-income consumers. Hence, each outsider would prefer to deviate and offer a less distorted separating contract to the high-income consumers without worrying about the low-income ones mimicking them. The withdrawal stage makes such cream-skimming deviation no longer profitable.

If $V^d(s)$ is maximal at s^*, its derivative must be 0 at that point. To prove that s^* is not optimal, we show that the right derivatives of $V^o(\tilde{s})$ and $\mathbb{E}_\mu[V_L]$ with respect to $\tilde{s}$ evaluated at $\tilde{s}^*$ are null and strictly positive, respectively, hence the optimal s is strictly higher than s^*. First, by the envelope condition,

$$\partial_{\tilde{s}} V^o(\tilde{s}) = \lambda \left[\sum_{y_2} \pi_2(y_2|y_H)(y_2 - c_H(y_2)) - \left(\sum_{y_2} \pi_2(y_2|y_L) y_2 - C(V_L^o(\tilde{s})) \right) \right]$$

where λ is the multiplier attached to the zero-profit condition. For $s \to s^{*+}$, $V^o(\tilde{s}) \to V^{LCS}$ which implies that $C(V_L^o) \to Y_{2L}$. Hence the second term in square brackets converges to 0 and, by zero-profit of outsiders, so does the first term. Last, we show that the right derivative of $\mathbb{E}_\mu[V_L]$ in $\tilde{s}$ is strictly positive at $\tilde{s}^*$. To see this, notice that for $\tilde{s} > \tilde{s}^*$, incentive-compatibility implies

$$\partial_{\tilde{s}} V_L^o(\tilde{s}) = \pi_L \partial_{\tilde{s}} u(c_H(y_H, \tilde{s})) + (1 - \pi_L) \partial_{\tilde{s}} u(c_H(y_L, \tilde{s}))$$

Since $\partial_{\tilde{s}} V^o(\tilde{s}) \to 0$,

$$\pi_H \partial_{\tilde{s}} u(c_H(y_H, \tilde{s})) + (1 - \pi_H) \partial_{\tilde{s}} u(c_H(y_L, \tilde{s})) \to 0$$

which implies

$$\begin{aligned}
\partial_{\tilde{s}} V_L^o(\tilde{s}) &= \frac{-\pi_L(1-\pi_H)\partial_{\tilde{s}} u(c_H(y_L,\tilde{s})) + (1-\pi_L)\pi_H \partial_{\tilde{s}} u(c_H(y_L,\tilde{s}))}{\pi_H} \\
&\to \left(\frac{\pi_H - \pi_L}{\pi_H} \right) \partial_{\tilde{s}} u(c_H(y_L, \tilde{s})) > 0
\end{aligned}$$

since $\partial_{\tilde{s}} u(c_H(y_L, \tilde{s})) > 0$ or otherwise the constraint $\sum_{y_2} \pi_2(y_2|y_L) u(c_H(y_2)) \geqslant u(Y_{2L})$ would be violated. Then,

$$\lim_{\tilde{s} \to \tilde{s}^{*+}} \partial_{\tilde{s}} \mathbb{E}_\mu[V_L] = (1 - \mu) \partial_s V_L^o(\tilde{s}) > 0$$

which concludes the proof. Q.E.D.

The previous lemma establishes that optimal information disclosure is always bounded away from the cutoff that separates the regions in which only the dynamic mechanism is operative and the region in which cross-subsidization among switchers occurs. In doing so, Lemma 6 extends an implication of Lemma 5 to the richer environment with both stayers and switchers. Lemma 5 established that, when $\alpha = 0$, the optimal pool quality among switchers must be strictly better than $\tilde{s}^*$. When $\alpha > 0$–and, in fact, for α sufficiently high–the intertemporal smoothing motive might dominate the cross-subsidization motive and generate an optimal pool quality that is strictly worse than $\tilde{s}^*$. However, regardless the value α, the optimal pool of switchers is never equal to$\tilde{s}^*$. We are now ready to state the main result of this section.

Proposition 4. *Let $s(\alpha)$ be the optimal share of high-income consumers among those with positive signal. If $\pi < \pi^{**}$, then $s(\alpha)$ is not strictly increasing in α.*

Proof. We prove our result using the fact that, if $\pi < \pi^{**}$, $s(\alpha)$ is on the opposite sides of s^*–as defined in Lemma 6–for $\alpha = 1$ and $\alpha = 0$ and that it is never equal to $s^*(\alpha)$First notice that $\lim_{\alpha \to 1} s^*(\alpha) = 1$ since absent idiosyncratic motives only low types leave the incumbent, the pool of switchers has arbitrarily low quality, and the outsiders offer V^{LCS} to the high types. Proposition 2 establishes that for $\alpha = 1$ full information is never optimal, hence $\lim_{\alpha \to 1} s(\alpha) - s^*(\alpha) <0$. Second, from Proposition 3, $\lim_{\alpha \to 0} s(\alpha) - s^*(\alpha) > 0$. Since $s(\alpha)$ is continuous while on the either side of $s^*(\alpha)$, and Lemma 6 established that $s(\alpha) \neq s^*(\alpha)$, it follows that there exists at least one value $\hat{\alpha}$ such that

$$\lim_{\alpha \to \hat{\alpha}^+} s(\alpha) < \lim_{\alpha \to \hat{\alpha}^-} s(\alpha)$$

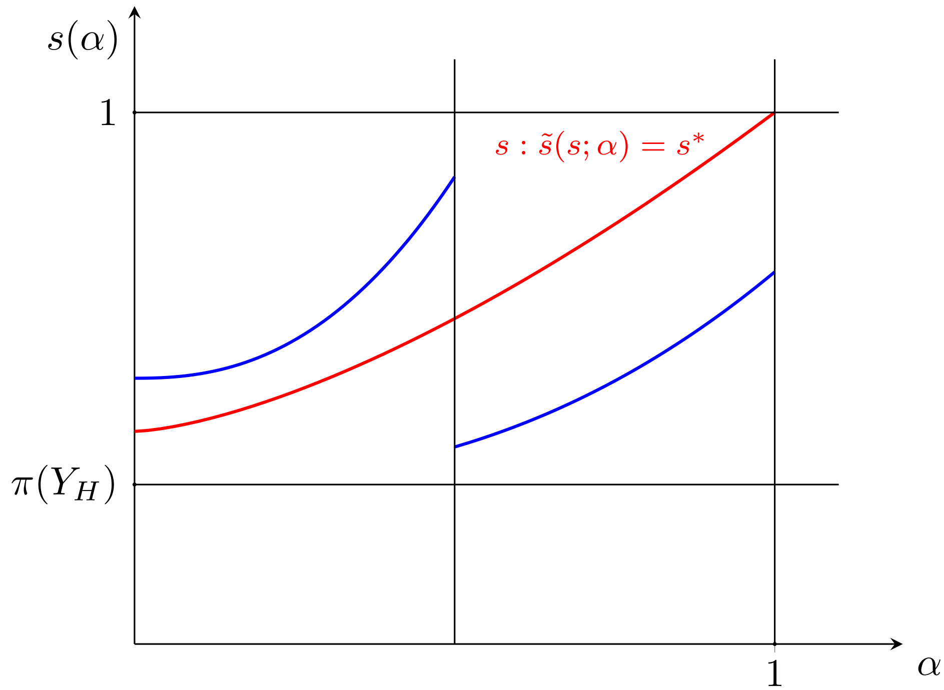
which concludes the proof. Q.E.D.

Recall that $\pi < \pi^{**}$ always holds in a stationary environment with $Y_1 = Y_2$. This restriction on π is necessary since if $\pi \geqslant \pi^{**}$ then $s(1) = s^*(1) = 1$ and $s(\alpha)$ can be strictly increasing and lie weakly above $s^*(\alpha)$ for all α.Optimal information design trades-off intertemporal consumption smoothing for stayers with cross-subsidization between switchers. Starting from $\alpha = 1$, a marginal decrease in α lowers the optimal composition s since when fewer high type consumers stick with the incumbent, the latter makes less profit, and rebates less consumption to the first period. Hence, less information is needed to reduce the amount of front-loading of consumption. Starting from $\alpha = 0$, a marginal increase in α raises the optimal composition s since a higher α implies a worse composition of switchers for given information policy, and a sufficiently good pool of consumers is needed for cross-subsidization to prevail in equilibrium. However, Proposition 4 establishes that $s(\alpha)$ is not globally increasing in α. This is because when α decreases from our benchmark value of 1, the relative importance of the two drivers of optimal disclosure becomes skewed toward cross-subsidization which requires that s is at least $s^*(\alpha)$.

8.3 Value of long-term relationship

We next study the value of long-term relationships in the model without commitment. We can do so by comparing the equilibrium outcome in our baseline case with no idiosyncratic shocks ($\alpha = 1$) to the case where all consumers change firms in the second

Figure 4: Optimal disclosure and fraction of switchers



period ($\alpha = 0$) and there are no long-term relationships. In an economy where firms can commit to the terms of the contract, long-term contracts are valuable because they allow for more insurance.

When firms cannot commit, the low-income consumers consume the expected value of their income with no transfers from the high-income ones. This is because the presence of an informed incumbent makes the adverse selection problem for the low income consumers much worse: new firms know that only low types are willing to switch (absent idiosyncratic motives) and therefore are not willing to offer anything more than actuarially fair contracts. Thus, only intertemporal smoothing is possible between period 1 and the second period conditional on having high income in period 1.

Absent long-term relationships ($\alpha = 0$), as a corollary of Proposition 3, at least a fraction of low-income consumers receive a consumption higher than Y_{2L}. Thus, the absence of an informed incumbent in the second period has the advantage of delivering higher consumption for the low income consumers in period 2. There are also costs: no intertemporal smoothing and distortions in the allocation offered to the high-income consumers in the second period to screen-out the low-income ones.[13]

It is then not obvious whether from an ex-ante perspectives consumers prefer to be

[13]It is the presence of this distortions that ensures that firms are willing to offer a transfer to the low-income consumers to reduce the efficiency costs of such distortions.

in an economy with long-term relationships or one without. To start with, notice that for $\pi_1(y_H) = \{0,1\}$ consumers obtain the same value in both economies. If all consumers have the same income in period 1, there is no asymmetric information in the second period, no scope for intertemporal smoothing, and no distortions in consumption. We can further characterize the ranking of the two economies in a neighborhood of those extreme values of $\pi_1(y_H)$. Let $V(\pi;\alpha)$ be the equilibrium value to consumers at the optimal information disclosure policy given α and $\pi = \pi_1(y_H)$. It is easy to see that

$$\frac{\partial V(\pi;0)}{\partial \pi}|_{\pi=0} = u'(y_L)(y_H - y_L) + \beta\left[V^{LCS} - u(Y_{2L})\right]$$
$$< u'(y_L)\left[(y_H - y_L) + \left(Y_{2H} - C\left(V^{LCS}\right)\right)\right] + \beta\left[V^{LCS} - u(Y_{2L})\right] = \frac{\partial V(\pi;1)}{\partial \pi}|_{\pi=0}.$$

Intuitively, when there are sufficiently few high-income consumers, the second period allocation entails little cross-subsidization across consumers regardless of the presence of an informed incumbent. However, due to its informational advantage, the incumbent is able to avoid distorting the consumption of high-income consumers, and to distribute the ex-post profits as first period consumption. It follows that, when adverse selection is severe, long-term relationships are beneficial.

Next consider the opposite extreme, $\pi = 1$. It is easy to see that,

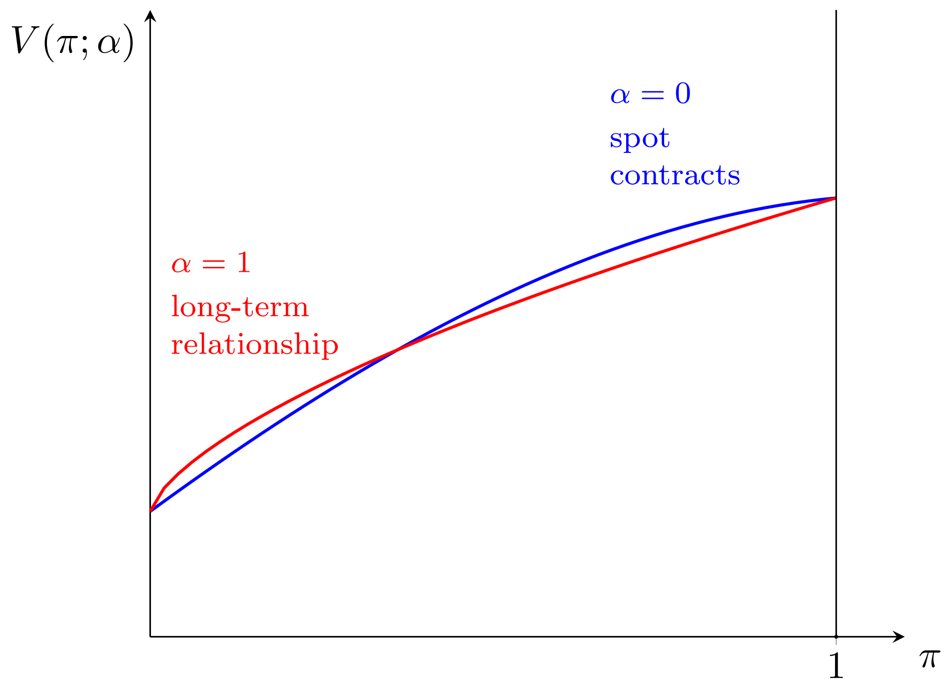
$$\frac{\partial V(\pi;0)}{\partial \pi}|_{\pi=1} = u'(y_H)(y_H - y_L) + \beta u'(Y_{2H})[Y_{2H} - Y_{2L}]$$
$$< u'(y_H)(y_H - y_L) + \beta[u(Y_{2H}) - u(Y_{2L})] = \frac{\partial V(\pi;1)}{\partial \pi}|_{\pi=1}$$

where the inequality follows from strict concavity of u. When the pool of consumers in the economy has sufficiently many high-income consumers, cross-subsidization prevails in the second period, absent the adverse selection induced by the presence of an informed firm. It follows that for a sufficiently large share of high-income consumers long-term relationships are harmful.

Characterizing the benefit of long-term relationships for interior values of π is more challenging. Yet, in all our numerical examples, we find that the intuition we provided for the ranking of values at the extremes of the support of π holds in the interior as well. Hence, there is a cutoff π_1 above which $V(\pi;1) < V(\pi;0)$, as illustrated in Figure 5.

The latest result helps understand what is the optimal number of firms a consumer would like to inform about her income realization if she had control over such information and the ability to commit to an information policy. Throughout the paper, we assumed that the incumbent firm is always informed about the realization of y_1. One could imagine situations in which the consumer has the power to control (at least some of) the information collected by the incumbent–and by outsiders, as already incorporated

Figure 5: Value of long-term relationships



in the public disclosure policy. If π is sufficiently low, consumers prefer to disclose their income to a single firm (i.e. the incumbent) in order to lower the adverse selection distortions that are necessary to screen consumers in the second period. If π is sufficiently high, consumers prefer not to disclose their income to anyone, or, equivalently, for a regulation to mandate an upper bound to the length of their contract.

9 Conclusion

We have studied optimal public disclosure of information in a dynamic insurance economy in which the incumbent firm has an informational advantage over its competitors. We showed that if the incumbent firm has commitment, it is optimal to disclose no information. Competition from outsiders limits the amounts of insurance that can be sustained, as more information allows outsiders to bid for high-income consumer more aggressively, hence limiting the transfers available to low-income consumers. If the incumbent firm has no commitment, the presence of competition in the second period disciplines the ex-post behavior of the incumbent firm. While no cross-subsidization can be sustained due to adverse selection, partial information disclosure allows intertemporal smoothing between consumption in the first period and in the second period after a high income realization. In an extension of the model that includes idiosyncratic motives to switch firms, cross-subsidization can be restored, but it is carried out exclusively by outsiders. Information disclosure controls the degree of adverse selection in the pool of switchers so that disclosing information might be necessary in order to trigger-instead of hindering-the provision of cross-subsidization.

Applied to the context of the recent adoption of open banking policies, our results suggest three main implications. First, making the entire history of consumers' information public is never optimal unless the share of high-income types is sufficiently high since doing so removes the scope for both intertemporal and intra-temporal insurance. Second, the information released should be coarsed into a rating system that bunch high-type and some low-type within the highest rate. Third, simply transferring ownership of the data to consumers does not achieve the optimal amount of insurance. Since high-income consumers would have a clear incentive to share their history, any lack of information sharing would identify low-income consumers due to adverse selection, effectively implementing the full-information outcome.

We see two main avenues along which the analysis in this paper can be pursued further. From a theoretical perspective, our model could be extended to incorporate more income realizations and longer histories. From a empirical perspective, measuring the model fundamentals in specific settings would help the formulation of concrete information policy proposals. This is especially true since optimal disclosure depends on the strength of idiosyncratic motive and on the composition of the relevant population, both of which are likely to vary systematically across markets.

References

BABINA, T., S. A. BAHAJ, G. BUCHAK, F. DE MARCO, A. K. FOULIS, W. GORNALL, F. MAZZOLA, AND T. YU (2024): "Customer data access and fintech entry: Early evidence from open banking," Tech. rep., National Bureau of Economic Research. 6

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The limits of price discrimination," *American Economic Review*, 105, 921–957. 5

BERGEMANN, D. AND S. MORRIS (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95. 5

CALZOLARI, G. AND A. PAVAN (2006): "On the optimality of privacy in sequential contracting," *Journal of Economic theory*, 130, 168–204. 5

COHEN, A. (2012): "Asymmetric learning in repeated contracting: An empirical study," *Review of Economics and Statistics*, 94, 419–432. 1

DE GARIDEL-THORON, T. (2005): "Welfare-improving asymmetric information in dynamic insurance markets," *Journal of Political Economy*, 113, 121–150. 5

DI MAGGIO, M. AND V. YAO (2021): "Fintech borrowers: Lax screening or cream-skimming?" *The Review of Financial Studies*, 34, 4565–4618. 6

GARCIA, D. AND M. TSUR (2021): "Information design in competitive insurance markets," *Journal of Economic Theory*, 191, 105160. 5

GOLDSTEIN, I. AND Y. LEITNER (2018): "Stress tests and information disclosure," *Journal of Economic Theory*, 177, 34–69. 29

GUERRIERI, V. (2008): "Inefficient unemployment dynamics under asymmetric information," *Journal of Political Economy*, 116, 667–708. 5

GUERRIERI, V., R. SHIMER, AND R. WRIGHT (2010): "Adverse selection in competitive search equilibrium," *Econometrica*, 78, 1823–1862. 5

HARRIS, M. AND B. HOLMSTROM (1982): "A theory of wage dynamics," *The Review of Economic Studies*, 49, 315–333. 3, 4, 5, 15, 21

HE, Z., J. HUANG, AND J. ZHOU (2023): "Open banking: Credit market competition when borrowers own the data," *Journal of financial economics*, 147, 449–474. 6

HELLWIG, M. (1987): "Some recent developments in the theory of competition in markets with adverse selection," *European Economic Review*, 31, 319–325. 8

HOLMSTRÖM, B. (1979): "Moral hazard and observability," *The Bell journal of economics*, 74–91. 25

HOPENHAYN, H. AND I. WERNING (2008): "Equilibrium default," *Manuscript, MIT*, 4, 21. 6

IMMORLICA, N., A. M. SZTUTMAN, AND R. M. TOWNSEND (2022): "Optimal Credit Scores Under Adverse Selection," in *Proceedings of the 23rd ACM Conference on Economics and Computation*, 737–738. 5

IOANNIDOU, V. AND S. ONGENA (2010): ""Time for a change": loan conditions and bank behavior when firms switch banks," *The Journal of Finance*, 65, 1847–1877. 1

JIN, Y. AND S. VASSERMAN (2021): "Buying data from consumers: The impact of monitoring programs in us auto insurance," Tech. rep., National Bureau of Economic Research. 6

KAHN, L. B. (2013): "Asymmetric information between employers," *American Economic Journal: Applied Economics*, 5, 165–205. 1, 25

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615. 5, 13, 23

LESTER, B., A. SHOURIDEH, V. VENKATESWARAN, AND A. ZETLIN-JONES (2019): "Screening and adverse selection in frictional markets," *Journal of Political Economy*, 127, 338–377. 5

LIN, X. AND C. LIU (2022): "Credible persuasion," *arXiv preprint arXiv:2205.03495*. 23

MATHEVET, L., J. PEREGO, AND I. TANEVA (2020): "On information design in games," *Journal of Political Economy*, 128, 1370–1404. 5

MIYAZAKI, H. (1977): "The rat race and internal labor markets," *The Bell Journal of Economics*, 394–418. 9

NETZER, N. AND F. SCHEUER (2014): "A game theoretic foundation of competitive equilibria with adverse selection," *International Economic Review*, 55, 399–422. 2, 3, 8, 9, 26, 42

ROTHSCHILD, M. AND J. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90, 629–649. 2, 8, 26

SHARPE, S. A. (1990): "Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships," *The journal of finance*, 45, 1069–1087. 6

THOMAS, J. AND T. WORRALL (1988): "Self-enforcing wage contracts," *The Review of Economic Studies*, 55, 541–554. 3, 5, 15, 21

WILSON, C. (1977): "A model of insurance markets with incomplete information," *Journal of Economic theory*, 16, 167–207. 9

Appendix

A Model

A.1 Type interpretation

Here we show that the pure exchange economy can be interpreted as one with types. Suppose that a consumers can be one of two types, θ_H and θ_L. The type affects the probability distribution of income. In particular, income in period 1 and 2 can take on two values, $y_t \in \{y_L, y_H\}$ with $y_H > y_L$. θ_H consumers are more likely to have income than θ_L consumers, $\pi_L(\theta) = \Pr(y_L|\theta)$ and $\pi_H(\theta) = \Pr(y_H|\theta)$ with $\pi_H(\theta_H) > \pi_H(\theta_L)$.

Both the consumer and the insurance companies do not know the consumer's type at the beginning of the period and learn about it through the observations of the history of income realization. Let $\rho(\theta)$ the common prior of being θ type at the beginning of period 1. After observing a high realization of income in period 1, the prior increases while it decreases after the realization of a low income realization:

$$\rho(\theta_H|y_H) = \frac{\rho(\theta_H)\pi_H(\theta_H)}{\sum_\theta \rho(\theta)\pi_H(\theta)} > \rho(\theta_H), \quad \rho(\theta_H|y_L) = \frac{\rho(\theta_H)\pi_L(\theta_H)}{\sum_\theta \rho(\theta)\pi_L(\theta)} < \rho(\theta_H). \tag{24}$$

Then, the probability that a borrower with history y_1 draws $y_2 = y_H$ is

$$\pi_H(y_1) = \rho(\theta_H|y_1)\pi_H(\theta_H) + \rho(\theta_L|y_1)\pi_H(\theta_L)$$

and the expected income is $\mathbb{E}(y_2|y_1) = \rho(\theta_H|y_1)\mathbb{E}_H y + \rho(\theta_L|y_1)\mathbb{E}_L y$. Thus, expected income in period 2 is higher after a high income realization in period 1 than after a low income realization:

$$\mathbb{E}(y_2|y_H) > \mathbb{E}(y_2|y_L).$$

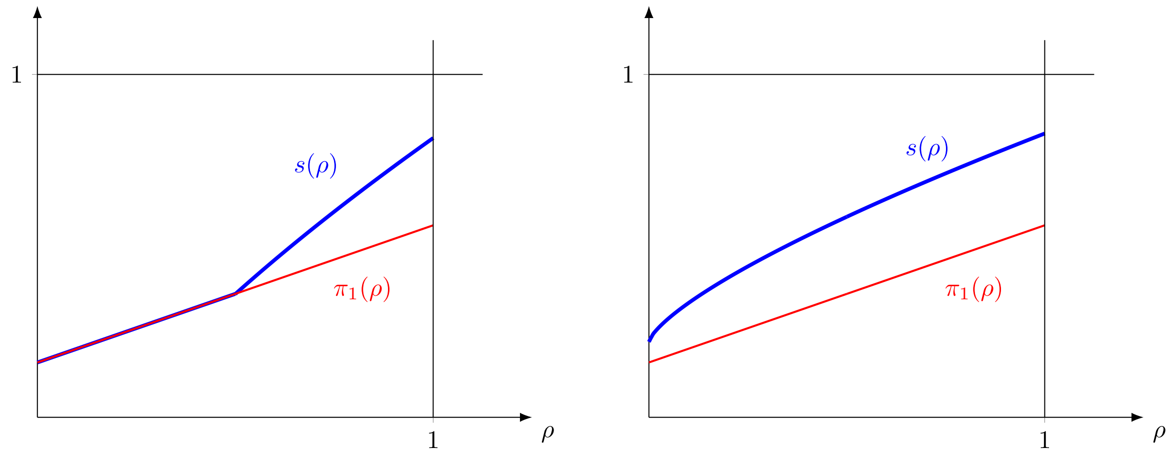
This economy is equivalent to the one considered in the main text with

$$\begin{aligned}
\pi_1(y_H) &= \rho(\theta_H)\pi_H(\theta_H) + (1-\rho(\theta_H))\pi_L(\theta_L) \\
\pi_2(y_2 = y_H|y_H) &= \rho(\theta_H|y_H)\pi_H(\theta_H) + (1-\rho(\theta_H|y_H))\pi_L(\theta_L) \\
\pi_2(y_2 = y_H|y_L) &= \rho(\theta_H|y_L)\pi_H(\theta_H) + (1-\rho(\theta_H|y_L))\pi_L(\theta_L)
\end{aligned}$$

and

$$\begin{aligned}
Y_1 &= \rho(\theta_H)\mathbb{E}_H y + (1-\rho(\theta_H))\mathbb{E}_L y \\
Y_{2H} &= \mathbb{E}(y_2|y_H) \\
Y_{2L} &= \mathbb{E}(y_2|y_L).
\end{aligned}$$

Figure 6: Optimal information disclosure in the type economy



Left panel: low risk-aversion. Right panel: high risk-aversion.

As we describe in the text when discussing Proposition 2, providing full information is never optimal. This is because in this type formulation it must be that

$$Y_1 = \pi_1 (y_H) Y_{2H} + (1 - \pi_1 (y_H)) Y_{2L} = Y_2$$

because $\rho = \pi_1 (y_H) \rho (\theta_H|y_H) + (1 - \pi_1 (y_H)) \rho (\theta_H|y_L)$. Thus, we cannot be in a situation in which constraint $V_H \leqslant V^o (1)$ is binding in problem (19) because at $s = 1$ we have

$$Y_1 < C (V^o (1)) = Y_{2H}.$$

Thus, we are always in case i) or ii) in Proposition 2 depending on whether condition (17) holds or not.[14] In most numerical examples we find a cutoff ρ^* such that for $\rho < \rho^*$ it is optimal to provide no information as condition (17) does not hold, while for $\rho > \rho^*$ it is optimal to provide partial information and have $c_1 = c_2 (y_H)$ because condition (17) holds, as illustrated in the figures below.

A.2 Credit economy

Here we sketch how we can reinterpret the model in the text as a credit economy where firms (lenders) learn about the default probability of a borrower.

Suppose there are two periods. Each period is divided into two sub-periods: AM and PM. In the AM, consumers have income $y_{AMt} = y_{AM}$ for sure and in the PM consumers can have income $y_{PMt} \in \{y_L, y_H\}$ with $y_L = 0$. The probability of drawing y_H in

[14] Another way to see the issue is that it is not possible for $\pi_1 (y_H)$ to go above the threshold π^{**} because even if $\rho \to 1$ then $\pi_1 (y_1) \to \pi_H (\theta_H)$ with constant expected income in period 1 and 2.

the first period is $\pi_1(y_H)$ and the probability in the second period is $\pi_2(y_H|y_{PM1})$ with $\pi_2(y_H|y_H) > \pi_2(y_H|y_L)$. Let $y_{AM} < y_H$ so there are motives to borrow in the first subperiod. An allocation is $\{c_{AM1}, c_{PM1}(y_{PM1}), c_{AM2}(y_{PM1}), c_{PM2}(y_{PM1}, y_{PM2})\}$. Consumer preferences are

$$u(c_{AM1}) + \pi_1 \left[u(c_{PM}(y_H)) + \beta u(c_{AM2}(y_H)) + \sum_{y_{PM2}} \pi_2(y_{PM2}|y_H) u(c_{PM2}(y_H, y_{PM2})) \right]$$
$$+ (1-\pi_1) \left[u(c_{PM}(y_L)) + \beta u(c_{AM2}(y_L)) + \sum_{y_{PM2}} \pi_2(y_{PM2}|y_L) u(c_{PM2}(y_L, y_{PM2})) \right]$$

Firms offer contracts that are amount borrowed in the AM, b, and a repayment r in the PM conditional on $y_{PM} = y_H$. If $y_{PM} = 0$ then there is a default as the firm cannot extract any payments in that state. Furthermore, assume it is not possible to save in the low-income state in the PM. Firms' period profits are then

$$-b + \Pr(y_{PM} = y_H) r$$

and period utility is

$$u(y_{AM} + b) + \Pr(y_{PM} = y_H) u(y_H - r) + (1 - \Pr(y_{PM} = y_H)) u(0).$$

This economy is equivalent to our insurance economy where consumption in the AM is consumption in the low-income state and consumption in the PM is consumption in the high-income state.

B Omitted proofs

B.1 Proof of Lemma 1

Consider the various cases:

1. If $V_H \geqslant V^o(s)$, $V_L \geqslant u(Y_{2L})$, and the incumbent withdraws its offers if the outsiders' offer the cream-skimming contract then the outsiders have no options to attract consumers. In fact, the best they can offer to the high-income consumer subject to the zero profit condition is $V^o(s)$. They could offer a cream-skimming contract if $V^{cs}(u(Y_{2L})) > V^o(s)$ but for that to be the case it must be that $V_L > u(Y_{2L})$ – otherwise $V^{cs}(u(Y_{2L})) \leqslant V^o(s)$. Thus, for the cream-skimming contract to be profitable it is required that the incumbent does not withdraw its offer as no other outsiders will offer a value higher than $u(Y_{2L})$ to the low-income consumers.

2. If $V_H < V^o(s)$ and $V_L < V_L^o(s)$ then the incumbent's offer is irrelevant and the equilibrium outcome is the one characterized in Netzer and Scheuer (2014).

3. If $V_H < V^{cs}(V_L)$, $V_L \geqslant u(Y_{2L})$ and the incumbent does not withdraw its offers, if the outsiders offer the cream-skimming contract then they will poach the high-income consumers.

B.2 Proof of Proposition 2

Proof. First we show that there exists a π^* such that condition (17) holds if and only if $\pi_1(y_H) \geqslant \pi^*$. Since $K(s) \equiv C(V^o(s))$, we can re-write condition (17) as

$$f(\pi_1) = \pi_1 y_H + (1-\pi_1) y_L + q\pi_1 (Y_{2H} - K(\pi_1)) - K(\pi_1)$$

Note that

$$f'(\pi_1) = y_H - y_L + q(Y_{2H} - K(\pi_1)) - (1 + q\pi_1) K'(\pi_1)$$
$$f''(\pi_1) = -qK''(\pi_1) - (1+q\pi_1) K''(\pi_1) - qK'(\pi_1)$$

so f is concave in π_1 as $K'' > 0$ and $K' > 0$. Furthermore, evaluating at $\pi_1 = 0$ and $\pi_1 = 1$ we have

$$f(0) = y_L - K(0) < 0$$
$$f(1) = y_H + q(Y_{2H} - K(1)) - K(1) \geqslant y_H + q(Y_{2H} - Y_{2H}) - Y_{2H} > 0$$

where we used the fact that $K(1) \leqslant C(u(Y_{2H})) = Y_{2H}$. Thus, since $f(0) < 0$, $f(1) > 0$ and f is strictly concave, f must cross 0 only once between $(0,1)$, denote such point by π^*. Then for $\pi_1 \in (0, \pi^*)$ we have $f(\pi_1) < 0$ and for $\pi \in (\pi^*, 1)$ we have $f(\pi_1) > 0$.

Suppose that $\pi_1(y_H) > \pi^*$ so condition (17) holds. First, we show that all high-income consumers receive the same signal or the signal is uninformative. Suppose by way of contradiction that the optimal disclosure policy, (M^*, μ^*), is such that there are two signals, m_1 and m_2, with $\mu^*(m_1|y_H) > 0$ and $V^o(s(m_1)) \neq V^o(s(m_2))$. Let

$$\bar{V} = \sum_m \mu^*(m|y_H) V^o(s(m)).$$

Suppose that $\bar{V} \geqslant V^o(\pi_1(y_H))$. In this case, there exists an alternative disclosure policy with $M = \{g, b\}$ with $\mu(g|y_H) = 1$ and $\mu(b|y_L) \in [0,1)$ and $V^o(s(g)) = \bar{V}$. Due to concavity of the utility function, this alternative disclosure allows the incumbent to save resources in period 2 and still deliver expected utility $\bar{V}$ to high-income consumers. Because of competition in period 1, these additional profits are rebated to the consumer in the first period. Thus, this alternative disclosure policy improves is an improvement because it delivers the same expected utility in period 2 but higher utility in period 1, reaching a contradiction.

Consider now the case with $\bar{V} < V^o(\pi_1(y_H))$. Note that

$$\begin{aligned}
&u\left(Y + q\pi_1(y_H)\sum_{m\in M}\mu(m|y_H)[Y_{2H} - C(V^o(s(m)))]\right) + \beta\pi_1(y_H)\bar{V} \\
<&u\left(Y + q\pi_1(y_H)\left[Y_{2H} - C\left(\bar{V}\right)\right]\right) + \beta\pi_1(y_H)\bar{V} \\
<&u(Y + q\pi_1(y_H)[Y_{2H} - C(V^o(\pi_1(y_H)))]) + \beta\pi_1(y_H)V^o(\pi_1(y_H))
\end{aligned}$$

where the first inequality follows from the convexity of C, and the second inequality from the concavity of u, convexity of C and condition (17). To see this last step, note that

$$F(V) \equiv u(Y + q\pi_1(y_H)[Y_{2H} - C(V)]) + \beta\pi_1(y_H)V$$

is increasing in V for $V \in [\underline{V}, V^o(\pi_1(y_H))]$ under (17). In fact,

$$\begin{aligned}
F'(V) &\equiv \beta\pi_1(y_H)\left[1 - u'(Y + \beta\pi_1(y_H)[Y_{2H} - C(V)])C'(V)\right] \\
&> \beta\pi_1(y_H)\left[1 - u'(Y + \beta\pi_1(y_H)[Y_{2H} - C(V^o(\pi_1))])C'(V^o(\pi_1))\right] \\
&= \beta\pi_1(y_H)\left[1 - \frac{u'(Y + \beta\pi_1(y_H)[Y_{2H} - C(V^o(\pi_1))])}{u'(C(V^o(\pi_1)))}\right] \\
&> 0
\end{aligned}$$

where the first inequality follows from u being concave, C increasing and convex and $V^o(\pi_1) > V$; and the last inequality from condition (17) that implies

$$u'(Y + \beta\pi_1(y_H)[Y_{2H} - C(V^o(\pi_1))]) < u'(C(V^o(\pi_1))).$$

Thus, even in this case the alternative disclosure policy improves upon the original allocation yielding a contradiction.

We established that under condition (17) all consumers with high income in period 1 receive the same signal. Without loss of generality we can consider $M = \{g, b\}$ and have $\mu(g|y_H) = 1$. We are only left to choose the fraction of consumers with low income in period 1 that receive the same "good" signal. This is not going to affect their consumption but it affects the composition of the pool of agents with a good signal and therefore the continuation value of the high income. The set of implementable continuation values for the high income consumer is $[V^o(\pi_1(y_H)), V^o(1)]$. We can then write the problem (16) as

$$\max_{c_1, V_H} u(c_1) + \beta\pi_1(y_H) V_H + \beta\pi_1(y_L) u(Y_{2L}) \tag{25}$$

subject to

$$Y_1 - c_1 + q\pi_1(y_H)[Y_{2H} - C(V_H)] \geqslant 0$$

$$V_H \in [V(\pi_1(y_H)), V^o(1)]$$

and then recover $\mu(b|y_L)$ from

$$V_H = V^o\left(\frac{\pi_1(y_H)}{\pi_1(y_H) + \pi_1(y_L)(1 - \mu(b|y_L))}\right). \tag{26}$$

If the last constraint in problem (25) does not bind and the optimal V_H is interior, it is clear that $u'(c_1) = u'(c_{2H})$ and so

$$c_1 = c_{2H} = \bar{c} \equiv \frac{Y_1 + q\pi_1(y_H) Y_{2H}}{1 + q\pi_1(y_H)} > Y_1$$

Thus, $\mu(b|y_L)$ solves

$$V^o\left(\frac{\pi_1(y_H)}{\pi_1(y_H) + \pi_1(y_L)(1 - \mu(b|y_L))}\right) = u(\bar{c})$$

Then just have to check if $u(\bar{c}) \in [V(\pi_1(y_H)), V^o(1)]$. Because of condition (17), $V^o(\pi_1(y_H)) \leqslant u(\bar{c})$. However it is possible that $u(\bar{c}) > V^o(1)$ if π_1 is sufficiently close to 1. In particular, define π^{**} such that $Y_{2H} = Y_1$ or

$$Y_{2H} = \pi^{**} y_H + (1 - \pi^{**}) y_L \iff \pi^{**} = \pi_2(y_H|y_H)$$

That is, π^{**} is the share of high-income consumers below which consumption is smoothed under full information disclosure, $\mu(b|y_L) = 1$. For $\pi_1(y_H) \in [\pi^*, \pi^{**}]$ the last constraint in problem (25) does not bind, $c_1 = c_2(y_H)$ and $\mu(b|y_L)$ solves (26) (case ii). If instead $\pi_1(y_H) > \pi^{**}$, the last constraint in problem (25) binds. Thus, the solution has $V_H = V^o(1)$ and $c_1 = Y_1 > c_2(y_H) = Y_{2H}$ and $\mu(b|y_L) = 1$ (case iii).

Finally, suppose that $\pi_1(y_H) < \pi^*$ so condition (17) does not hold. In this case, it is not possible to equalize consumption in period 1 and period 2 if $y_1 = y_H$ by assigning the same signal to all high-income consumers. This is because the last constraint in (25) binds and $u'(c_1) > u'(c_2(y_H))$. It might then be optimal to assign different signals to the high-income consumers in order to reduce their expected continuation value to economize on resources used in period 2 that can then be rebated in period 1. This is not feasible under our assumption that $K(s) = C(V(s))$ is convex. Suppose by way of contradiction that the optimal disclosure policy, (M^*, μ^*), is such that there are two signals, m_1 and m_2, with $\mu^*(m_1|y_H) > 0$ and $V^o(s(m_1)) \neq V^o(s(m_2))$ with

$$\bar{V} = \sum_m \mu^*(m|y_H) V^o(s(m)) < V^o(\pi_1(y_H)).$$

(Clearly, if $\bar{V} > V^o(\pi_1)$ an argument similar to the one in part i shows that assigning multiple signals to high income consumers is not optimal.) The period one consumption associated with this plan is

$$\begin{aligned} & Y_1 + q\pi_1(y_H)\left(Y_{2H} - \sum_m \mu^*(m|y_H) K(s(m))\right) \\ < & Y_1 + q\pi_1(y_H)\left(Y_{2H} - K\left(\sum_m \mu^*(m|y_H) s(m)\right)\right) \\ \leqslant & Y_1 + q\pi_1(y_H)(Y_{2H} - K(\pi_1(y_H))) \end{aligned}$$

where the first inequality follows from the assumed convexity of K and the second from the observation that $\sum_m \mu^*(m|y_H) s(m) \geqslant \pi_1(y_H)$ and K is increasing. Thus, disclosing no information increases both the expected continuation value in period 2 for the high-income consumers and the consumption in period 1. Thus, the original allocation cannot be optimal and it must bee optimal to assign the same signal to all high-income consumers. Therefore, problem (25) characterizes the full problem (16). Since the last constraint is binding, one way to obtain the optimum is to provide no information so $c_2(y_H) = C(V^o(\pi_1))$ and

$$c_1 = Y + q\pi_1(Y_{2H} - C(V^o(\pi_1))) < c_2(y_H).$$

Q.E.D.

B.3 Proof of Lemma 5

Since it is never optimal to distort the allocation for the low-income type, we can write (2) as

$$V^o(s) = \max_{c_H(y_2), \varepsilon} \sum_{y_2} \pi_2(y_2|y_H) u(c_H(y_2)) \tag{27}$$

subject to

$$\sum_{y_2} \pi_2(y_2|y_H)(y_2 - c_H(y_2)) - \frac{(1-s)}{s}\varepsilon \geqslant 0,$$

$$u(Y_{2L}+\varepsilon) \geqslant \sum_{y_2} \pi_2(y_2|y_L)\, u(c_H(y_2)),$$

and $\varepsilon \geqslant 0$, with $c_L(y_2) = Y_{2L} + \varepsilon$ for all y_2.

First, we argue that the value of V^o is increasing in s and strictly increasing if $\varepsilon > 0$. To see this, suppose that for some s_L it is optimal to give a subsidy to the low type to relax the incentive constraint, $\varepsilon_L = \varepsilon(s_L) > 0$. Then for all $s_H > s_L$ it $V^o(s_H) > V^o(s_L)$ and $\varepsilon(s_H) > 0$. In fact, the solution for $s = s_L$ is feasible for s_H and since $(1-s_H)/s_H < (1-s_L)/s_L$ the non-negative-profits condition is relaxed and it is possible to increase the value for the high-type in an incentive compatible way, implying that $V^o(s_H) > V^o(s_L)$. That $\varepsilon(s_H) > 0$ follows from the fact that $V^o(s_L) > V^{lcs}$ and so it cannot be that the optimum at s_H is the LCS contract with no subsidy to the low-type.

Next, note that for s sufficiently close to 1 the value of $V^o(s)$ is close to $u(Y_{2H}) > V^{lcs}$ and so there exists a share of high-type sufficiently high such that $V^o(s) > V^{lcs}$.

We are left to show that such cutoff s^* is strictly positive and for low enough s no subsidies are optimal. To see this, consider the necessary focs for problem (27), letting λ and λ_{ic} be the multipliers on the non-negative profits for the firm and on the incentive compatibility constraint respectively:

$$\pi_2(y_s|y_H)\, u'(c_s) = \lambda \pi_2(y_s|y_H) + \lambda_{ic}\pi_2(y_s|y_L)\, u'(c_s)$$

and

$$\lambda \frac{(1-s)}{s} \geqslant \lambda_{ic} u'(Y_{2L}+\varepsilon)$$

with equality if $\varepsilon > 0$. Suppose by way of contradiction that $\varepsilon(s) > 0$ for all $s > 0$. Then, for all $s > 0$,

$$\frac{(1-s)}{s} = \frac{\lambda_{ic}(s)}{\lambda(s)} u'(Y_{2L}+\varepsilon(s))$$

As $s \to 0$, $(1-s)/s \to \infty$ then it must be that also $\lambda_{ic}(s)/\lambda(s) \to \infty$ because $u'(Y_{2L}+\varepsilon(s)) \leqslant u'(Y_{2L})$. There are two possibilities then: either $\lambda_{ic}(s) \to \infty$ or $\lambda(s) \to 0$. If $\lambda_{ic}(s) \to \infty$ then from the foc

$$\frac{u'(c_s)}{\lambda} = 1 + \frac{\lambda_{ic}}{\lambda}\frac{\pi_2(y_s|y_L)}{\pi_2(y_s|y_H)} u'(c_s)$$

it must be that $c_s(s) \to 0$ which is a contradiction since $V^o \geqslant V^{lcs}$. If instead $\lambda(s) \to 0$ then foc can be written as

$$1 = \frac{\lambda}{u'(c_s)} + \lambda_{ic}\frac{\pi_2(y_s|y_L)}{\pi_2(y_s|y_H)}$$

implying that in the limit

$$\frac{\pi_2(y_H|y_H)}{\pi_2(y_H|y_L)} = \frac{\pi_2(y_L|y_H)}{\pi_2(y_L|y_L)}$$

which is not true since the two types face different output distribution. Thus, it cannot be that $\lambda(s) \to 0$. Therefore we cannot have that $\varepsilon(s) > 0$ for all $s > 0$ and there are some low s such that $\varepsilon(s) = 0$. Q.E.D.

C Convexity of K

Here we provide sufficient conditions for the function $K(s) = C(V^o(s))$ to be convex. For that to be the case, when differentiable (i.e. at all points other than at s such that $V^o(s) = V^{lcs}$) it must be that $K'' = C''(V')^2 + C'V'' \geqslant 0$.

Lemma 7. *Suppose that* $u(c) = \log c$, $\pi_2(y_H|y_L)$ *and* $Y_{2H} - Y_{2L}$ *are sufficiently small. Then* $K(s)$ *is convex.*

For s such that $V^o(s) = V^{lcs}$ then $K(s)$ is constant at $C\left(V^{lcs}\right)$. For higher s where $V^o(s) > V^{lcs}$, note that the participation constraint (3) in (2) must be slack, otherwise $V^o(s) = V^{lcs}$. Thus, for such s we can write

$$V^o(s) = \max_{c_H(y_2), c_L(y_2)} \sum_{y_2} \pi_2(y_2|y_H)\, u(c_H(y_2)) \tag{28}$$

subject to

$$s\sum_{y_2} \pi_2(y_2|y_H)(y_2 - c_H(y_2)) + (1-s)\left[\sum_{y_2} \pi_2(y_2|y_L)(y_2 - c_L(y_2))\right] \geqslant 0,$$

$$\sum_{y_2} \pi_2(y_2|y_L)\, u(c_L(y_2)) \geqslant \sum_{y_2} \pi_2(y_2|y_L)\, u(c(y_2)).$$

Further noticing that the incentive constraint must be binding, we can write $K(s)$ as

$$K(s) = \max_{u(y_2)} C\left(\sum_{y_2} \pi_2(y_2|y_H)\, u(y_2)\right)$$

subject to

$$sY_{2H} + (1-s)Y_{2L} \geqslant s\sum_{y_2} \pi_2(y_2|y_H)\, C(u(y_2)) + (1-s)\, C\left(\sum_{y_2} \pi_2(y_2|y_L)\, u(y_2)\right).$$

Assuming log utility, we can further simplify the problem as

$$K(s) = \max_{u_H, u_L} \exp(\pi_H u_H + (1-\pi_{2H})\, u_L) = \exp(u_H)^{\pi_H} \exp(u_L)^{1-\pi_H}$$

subject to

$$sY_{2H} + (1-s)\,Y_{2L} \geqslant s\left[\pi_H \exp\left(u_H\right) + (1-\pi_H)\exp\left(u_L\right)\right] + (1-s)\exp\left(\pi_L u_H + (1-\pi_L)\,u_L\right)$$

where $\pi_H = \pi_2(y_H|y_H)$ and $\pi_L = \pi_2\left(y_H|y_L\right)$.

If $\pi_L \to 0$, the zero-profit condition for the outsider reduces to

$$sY_{2H} + (1-s)\,Y_{2L} = s\left[\pi_H \exp\left(u_H\right) + (1-\pi_H)\exp\left(u_L\right)\right] + (1-s)\exp\left(u_L\right)$$

or, solving for u_L,

$$\exp\left(u_L\right) = \frac{sY_{2H} + (1-s)\,Y_{2L} - s\pi_H \exp\left(u_H\right)}{s\left(1-\pi_H\right) + (1-s)}$$

and, plugging back in the objective function and letting $x \equiv \exp\left(u_H\right)$ we can write

$$K(s) = \max_x x^{\pi_H}\left[\frac{sY_{2H} + (1-s)\,Y_{2L} - s\pi_H x}{1 - s\pi_H}\right]^{1-\pi_H}$$

The optimal x satisfies the foc

$$\left[sY_{2H} + (1-s)\,Y_{2L} - s\pi_H x\right] = x\left(1-\pi_H\right)s$$

so

$$\exp\left(u_H\right) = \frac{\left(sY_{2H} + (1-s)\,Y_{2L}\right)}{\left(1-\pi_H\right)s + s\pi_H} = \frac{\left(sY_{2H} + (1-s)\,Y_{2L}\right)}{s}$$

and

$$\exp\left(u_L\right) = \frac{\left(1-\pi_H\right)\left(sY_{2H} + (1-s)\,Y_{2L}\right)}{1 - s\pi_H}$$

Thus, letting $E\left(Y|s\right) = \left(sY_{2H} + (1-s)\,Y_{2L}\right)$ we have

$$K(s) = E\left(Y|s\right)\left(1-\pi_H\right)^{1-\pi_H} f(s)$$

where

$$f(s) \equiv \frac{1}{s^{\pi_H}\left(1 - s\pi_H\right)^{1-\pi_H}}$$

Then

$$K'(s) \propto \left(Y_{2H} - Y_{2L}\right) f(s) + E\left(Y|s\right) f'(s)$$
$$K''(s) \propto 2\left(Y_{2H} - Y_{2L}\right) f'(s) + E\left(Y|s\right) f''(s)$$

It can be easily verified that f is convex. Thus, for small $\left(Y_{2H} - Y_{2L}\right)$, K is convex for s such that $V^o(s) > V^{lcs}$. Thus, since $K(s)$ is the upper-envelope of two convex functions it is convex under our simplifying assumptions.

Numerically, we show that K is convex also when the sufficient conditions in the Lemma are not satisfied.

D Observed outcome and unobserved signals

In the main text we assume that y_1 is not observable by the outsiders. Here we sketch how our analysis extends to the case in which y_1 is observable to outsiders but the incumbent and the consumer receive a signal $\tilde{y}_1$ in period 1 that is informative about the distribution of income in period 2. Assume that $\tilde{y}_1 \in \{\tilde{y}_H, \tilde{y}_L\}$ and let $p(\tilde{y}_1|y_1)$ be the probability of getting signal $\tilde{y}_1$ given y_1. Further, let $\pi_2(y_2|y_1,\tilde{y}_1)$ be the probability of y_2 given $(y_1,\tilde{y}_1)$.[15]

The outside option for the consumer with a high-signal and observable income in period 1 y_z is

$$V_z^o(s) = \max_{c,V_L} \sum_{y_2} \pi_2(y_2|y_z,\tilde{y}_H)\, u(c(y_2))$$

subject to

$$s\left[Y_{2sH} - \sum_{y_2} \pi_2(y_2|y_z,\tilde{y}_H)\, c(y_2)\right] + (1-s)\left[Y_{2zL} - C(V_L)\right] \geqslant 0$$

$$\sum_{y_2} \pi_2(y_2|y_z,\tilde{y}_H)\, u(c(y_2)) \leqslant V_L$$

$$V_L \geqslant u(Y_{2zL})$$

where

$$Y_{2zH} \equiv \sum_{y_2} \pi_2(y_2|y_z,\tilde{y}_H)\, y_2$$
$$Y_{2zL} \equiv \sum_{y_2} \pi_2(y_2|y_z,\tilde{y}_L)\, y_2.$$

It is straightforward to show that under two-sided lack of commitment for an arbitrary (M,μ) we have:

[15] Appendix F offers another example where the incumbent and the consumer/worker see a private signal about the return on the private investment in general human capital.

Lemma. *Given a public disclosure policy* (M, μ), *the equilibrium outcome has*

$$c_1(y_1) = Y_1 + \sum_z \pi_1(y_z)\, p(\tilde{y}_H|y_z) \sum_m \mu(m|y_z, \tilde{y}_H)\, \Pi_z(m) \tag{29}$$

$$c_2(y_z, \tilde{y}_L, m, y_2) = Y_{2zL} \tag{30}$$

$$c_2(y_H, m, y_2) = Y_{2zH} - \Pi_z(m) \tag{31}$$

where $\Pi_z(m) = Y_{2zH} - C(V_z^o(s(m)))$.

Thus, the optimal disclosure policy can be found as the solution to the analog of problem (19):

$$\begin{aligned} \max u(c_1) &+ \beta\pi_1(y_H)\left[p(\tilde{y}_H|y_H)\, V_{HH} + p(\tilde{y}_L|y_H)\, u(Y_{2HL})\right] \\ &+ \beta\pi_1(y_L)\left[p(\tilde{y}_H|y_L)\, V_{LH} + p(\tilde{y}_L|y_L)\, u(Y_{2LL})\right] \end{aligned}$$

subject to

$$\begin{aligned} c_1 &= Y_1 + q\sum_z \pi_1(y_z)\, p(\tilde{y}_H|y_z)\left[Y_{2zH} - C(V_{zH})\right] \\ V_{zH} &\in \left[V_z^o(p(\tilde{y}_H|y_z)), V_z^o(1)\right] \end{aligned}$$

and then recover $s_z(g)$ and $\mu(b|y_z, \tilde{y}_L)$ from

$$V_{zH} = V_z^o(s_z(g)) \text{ and } s_z(g) = \frac{p(\tilde{y}_H|y_z)}{p(\tilde{y}_H|y_z) + p(\tilde{y}_H|y_z)(1 - \mu(b|y_z, \tilde{y}_L))}. \tag{32}$$

E Discrimination across consumers

So far we have assumed that the incumbent firm cannot discriminate among consumers with the same history. We now relax this assumption and show that the optimal signal structure maintains the same features as in the restricted case. Our main interest is in the differences in the contract structure induced by discrimination. We show that when releasing information is optimal, the incumbent firm provides $c_2(y_H) < C(V^o(s(m)))$ to almost all agents with $y_1 = y_H$.

First consider the case $V^o(s(m)) = V^{lcs}$. Then the allocation coincides with that in Lemma 3,

$$c_2(y_1, m) = C\left(V^{lcs}\right),$$

and the lack of discrimination has no impact on the allocation. The second and more interesting case is one in which $V^o(s(m)) > V^{lcs}$. In this case, discrimination allows the incumbent firm to offer less than $V^o(s(m))$ to almost all high types by exploiting the fact that lower offers from outsiders attract a worse pool of agents. First recall that $V^{lcs} = V^o(\pi_1(y_H))$. This implies that there exists a largest share of high income consumers $s^* \in [\pi_1(y_H), s(m)]$ such that

$$V^o(s^*) = V^{lcs}.$$

We then order all high income consumers in period 1 and index them by $i \in [0,1]$. We define $i^*(m)$ to be the measure of high types with signal m that generates the share s^*, for given $\mu(m|y_L)$. The value $i^*(m) \in (0,1)$ identifies two groups of agents. For $i \in [0, i^*(m)]$, the value agent i receives is

$$V(i) = V^{lcs}.$$

This is because the least cost separating contract constitutes a lower bound for all high-type agents. For $i \in [i^*(m), 1]$,

$$V(i) = V^o(s_i(m))$$

where

$$s_i(m) = \frac{\pi_1(y_H)\int_0^i d\tilde{i}}{\pi_1(y_H)\int_0^i d\tilde{i} + (1-\pi_1(y_H))(1-\mu(m|y_L))}$$

is the share of high income consumers in a pool that includes all the low types with signal m and all high types with signal m and index smaller than i. Next we solve for the optimal disclosure policy.

Optimal public disclosure

The introduction of discrimination modifies the allocation associated to a given signal structure in two ways. First, as long as the value of the outsider's contract is higher that the value of the least cost separating contract, high type agents with the same signal receive unequal consumption in the second period. In particular, their utility in the second period belongs to the interval $\left[V^{lcs}, V^o(s(m))\right]$. Second, firms are able to extract additional profits in the second period and rebate them to the agent in the first period. We show that these two features do not fundamentally change the nature of the optimal disclosure policy.

Proposition 5. *The optimal disclosure policy has a bad-signal structure i.e.* $M = \{g, b\}$ *(good or bad) and* $\mu(g|y_H) = 1$ *and* $\mu(b|y_L) \in [0,1]$. *i) For* π *sufficiently low, it is optimal to provide no information; ii) For* π *sufficiently high, full information disclosure is optimal; iii) For all* π*, more information is disclosed under discrimination–and strictly so for intermediate levels of* π.

Proof. Proposition 5 shares most predictions with Proposition 2. Part i) follows from $y_L < C\left(V^{lcs}\right)$ since $C\left(V^{lcs}\right)$ is the minimum consumption that must be guaranteed to high-income consumers. Part ii) follows from the fact that for all $\pi > \pi^{**}$ full information is optimal under no discrimination, and that, for all disclosure policies, second period consumption for high-income consumers is weakly lower under discrimination. Part iii) hinges on the fact that the incumbent firm earns more profits in the second period under discrimination, hence consumption is typically more front-loaded for all interior disclosure policies.

Formally, the optimal signal under discrimination solves (where we replace $\mu(b|y_L)$ with μ to ease notation)

$$\max_{\mu\in[0,1]} u\left(Y + q\pi\left(y_H\right)\Pi\left(\mu\right)\right) + q\pi\left(y_H\right)\left[i^*\left(\mu\right)V^{lcs}\left(y_H\right) + \int_{i^*(\mu)}^{1} V^o\left(s_i\left(\mu\right)\right)di\right]$$

where

$$\Pi\left(\mu\right) = Y_{2H} - \left[i^*\left(\mu\right)C\left(V^{lcs}\left(y_H\right)\right) + \int_{i^*(\mu)}^{1} C\left(V^o\left(s_i\left(\mu\right)\right)\right)di\right].$$

If $\mu = 1$, the result is trivial. If $\mu < 1$,

$$-u'\left(c_1\right)\int_{i^*(\mu)}^{1} C_\mu\left(V^o\left(s_i\left(\mu\right)\right)\right)di + \int_{i^*(\mu)}^{1} V^o_\mu\left(s_i\left(\mu\right)\right)di \leqslant 0.$$

Since

$$u'\left(c_1\right)C'\left(V^o\left(s_1\left(\mu\right)\right)\right)\int_{i^*(\mu)}^{1} V^o_\mu\left(s_i\left(\mu\right)\right)di \geqslant u'\left(c_1\right)\int_{i^*(\mu)}^{1} C_\mu\left(V^o\left(s_i\left(\mu\right)\right)\right)di$$

then

$$u'\left(c_1\right)C'\left(V^o\left(s_1\left(\mu\right)\right)\right) \geqslant 1.$$

Let μ^* be the optimal signal structure in the economy without discrimination. Suppose by contradiction that $\mu^* > \mu$. Then

$$u'\left(c_1\left(\mu\right)\right) \geqslant \frac{1}{C'\left(V^o\left(s_1\left(\mu\right)\right)\right)} > \frac{1}{C'\left(V^o\left(s_1\left(\mu^*\right)\right)\right)} \geqslant u'\left(c_1^*\left(\mu^*\right)\right)$$

which implies

$$c_1\left(\mu\right) = Y + q\pi\left(y_H\right)\Pi\left(\mu\right) < Y + q\pi\left(y_H\right)\Pi^*\left(\mu^*\right) = c_1^*\left(\mu^*\right).$$

Since $V^o\left(s_i\left(m\right)\right) \leqslant V^o\left(s_1\left(m\right)\right)\ \forall i$, then $\Pi\left(\mu\right) \geqslant \Pi^*\left(\mu\right)\ \forall\mu$. Since Π and Π^* are decreasing functions, the latter inequality cannot hold, which leads to a contradiction. Finally, if $\mu > 0$, then $\Pi\left(\mu\right) > \Pi^*\left(\mu\right)$ which implies that $\mu > \mu^*$. Q.E.D. □

F Effort

In this section, we extend the benchmark model to incorporate unobservable effort in the first period of the contract. In order to highlight the effect of this extension on optimal information design, we intentionally keep the environment as close as possible to our benchmark.

F.1 Environment

Consider a training model in which first period output y_1 is constant and predetermined. At the end of the first period, the worker exerts training effort e at cost $v(e)$. At the beginning of the second period, the incumbent firm and the worker jointly observe the outcome of training in the form of human capital $h \in \{h_H, h_L\}$ where $h \sim f(h|e)$. Effort is an investment-like good that does not affect first period output, but contributes to the formation of general human capital. Human capital in turn affects the distribution of second period output, $y_2 \in \{y_L, y_H\}$ with $y_2 \sim p(y_2|h)$. We assume that while effort is privately known only by the agent, human capital is observed by both the worker and the incumbent firm, but not by outsiders. Outsiders only observe a signal m about the value of human capital, $m \sim \mu(h)$.

An allocation is a contract offered by the insider,

$$x = \{c_1, e, c_2(h, m, y_2)\},$$

and a menu contract offered by the outsider $x^o(m)$.

Additional assumptions and definitions We define $E(y_2|h) = \sum_s p(y_s|h)\, y_s$, with $E(y_2|h_H) > E(y_2|h_L)$, and $E(y_2|e) = \sum_h f(h|e)\, E(y_2|h)$. To make the environment comparable to the pure exchange economy in the text, we assume that y_1 is equal to Y_1 in the benchmark model. We also assume that $E(y_2|h_s) = Y_{2s}$ and $E(y_2|e) = Y_2$ which requires us to specify a given level of effort. Hence we assume that effort can take on two values, e and 0, and in equilibrium we guess (and verify) that is optimal to exert effort e. Furthermore, we assume that $f(h_H|e) = \pi(y_H)$ and $p(y_2|h_s) = \pi_2(y_2|y_s)$ for $s = H, L$.

In this economy, human capital replaces first period income as the source of information about future income that both the agent and the incumbent have access to and that the designer conveys a signal about. The key difference with our benchmark model is that the determination of h is influenced by an action the agents performs and the incumbent cannot observe.[16]Thus, the two economies are equivalent except for the existence of an

[16]Human capital does not fully reveal the amount of effort the worker exerts. The distribution f might be

incentive compatibility constraint that guarantees that the worker exerts effort e.

Equilibrium and optimal public disclosure We solve for the equilibrium under a given signal structure. Most of the results follow directly from what we showed in the text, hence we focus on the new features that originate from the introduction of effort. We discuss the economically interesting case in which it is efficient to induce strictly positive effort. Due to adverse selection, an agent with a low realization of human capital always consumes her expected output, $E(y_2|h_L)$, which is also the only contract the outsiders are willing to offer. However, in order to induce the outsiders not to offer a contract that would attract agents with h_H, the incumbent offers them

$$V(h_H, m) = V^o(s(h_H|m; e))$$

where $s(h_H|m; e)$ is the share of agents with human capital h_H with signal m given the equilibrium effort level e,

$$s(h_H|m; e) = \frac{\mu(m|h_H) f(h_H|e)}{\sum_h \mu(m|h) f(h|e)}.$$

The key difference with the previous model is the endogeneity of human capital. For the worker to exert the effort level e, it has to be that

$$\beta(f(h_H|e) - f(h_H|0))[V(h_H, m) - u(Y_{2L})] \geqslant \nu(e) - \nu(0)$$

or

$$[V(h_H, m) - u(Y_{2L})] \geqslant \frac{\nu(e) - \nu(0)}{\beta(f(h_H|e) - f(h_H|0))}. \tag{33}$$

If this incentive compatibility constraint is satisfied, the equilibrium consumption allocation is identical to the one in the benchmark model, otherwise workers exert no effort.

Notice that the information design has no effect on the value received by workers with low human capital. The only way to induce effort is to provide additional information about workers with high human capital, hence increasing their equilibrium value. We summarize this result in the following proposition:

Proposition 6. *Let μ^* be the optimal signal structure in the economy without hidden effort. i) If (33) holds, then μ^* is optimal in the economy with effort. ii) If (33) does not hold when evaluated at μ^*, then the optimal signal structure is: $M = \{g, b\}$ (good or bad) and $\mu(g|h_H) = 1$ and*

induced by either unobserved worker type or by pure lack. What matters for our results is that the source of uncertainty underlying f is uncorrelated with that behind p.

$\mu(b|h_L) \in (0,1)$ *such that*

$$V^o\left(\frac{f(h_H|e)}{f(h_H|e)+(1-\mu(b|h_L))\,f(h_L|e)}\right) = \frac{\nu(e)-\nu(0)}{\beta\,(f(h_H|e)-f(h_H|0))} + u(Y_{2L})$$

Moreover, $\mu(b|y_L) > \mu^*(b|y_L)$. *That is, the optimal signal with hidden effort is more informative than the one without hidden effort.*

Proof. Part i) is straightforward since μ^* satisfies the IC constraint and it is optimal in its absence. Part ii) follows from concavity of (16), the binding incentive constraint (33) and $V^o(s)$ being strictly increasing in s. □