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

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

Paolo Martellini
NYU Stern
paolo.martellini@nyu.edu

October 2024

Abstract

This paper studies optimal information disclosure in dynamic insurance economies with income risk in which an incumbent firm acquires more information about a consumer's persistent type than the rest of the market does. We find that if the incumbent can commit to long-term contracts but the consumer can walk away, the optimal disclosure prescribes no information revelation to maximize cross-subsidization. However, if the incumbent lacks commitment, no cross-subsidization of low-income consumers is feasible for any public information disclosure because of adverse selection. We show that partial information disclosure is typically optimal and it aims at implementing intertemporal consumption smoothing between the first period and the high-state in the second period, generating an inverse of the back-loading result in Harris and Holmstrom (1982). Lastly, we show that, without commitment, banning long-term relations can be beneficial to consumers. Our results can be used to analyze the consequences of policy proposals such as open banking and consumer data ownership.

*_First version: March 2024._ We thank Matteo Camboni, Marco Bassetto, V.V. Chari, Piero Gottardi, Natalia Kovrijnykh, Alessandro Lizzeri, Juan Pablo Nicolini, Alessandro Pavan, Marzena Rostek, Venky Venkateswaran, and Alexander Von Hafften for useful comments. Mariia Elkina provided superb research assistance.

1 Introduction

In many markets, firms engage in long-term relationships with consumers. By 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 their customers' health or driving records; credit card companies infer their customers' repayment probability by observing their spending patterns; 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. The answer to this question is important because of the pervasiveness of this type of asymmetric information and 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 the so-called *open banking*, a set of regulations that compel banks to make data on their customers' history available to competitors if the customers choose so. Another prominent example, in the opposite direction, is represented by laws that forbid employers to ask workers about their past wages. We aim 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 can 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-income consumers ("high types") and maximize ex-post cross-subsidization of low-income consumers ("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 and second periods for high-income consumers. Third, we show that when firms lack commitment and the gains from cross-subsidization are larger than the losses from adverse-selection distortions, long-term relations are harmful to consumers.

Our results have two main implications. First, making the entire history of consumers'

[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 a borrower's default probability. See Appendix A.2 for a formal treatment.

information public is never optimal. The information released should be coarsened into a rating system that bunches the high types and some low types within the highest rate. Second, transferring data ownership to consumers does not achieve the optimal amount of insurance, because high-income consumers would have a clear incentive ex-post to share their history, effectively implementing the full-information outcome.

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 income realization but only 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, because the incumbent 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, who decide whether to switch firms 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 contracts, as in the game described by Netzer and Scheuer (2014) for a Rothschild and Stiglitz (1976) economy.

We aim 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 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 the high-income consumers receive the same signal as some low-income ones, outsider firms must either distort the consumption profile of the high types not to attract low types or transfer resources to the low types. Thus, the maximal value that outsiders can offer is increasing in the share of high-income consumers with a given signal. The value of the outside option for the 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. We find that when firms can commit to long-term contracts, no information disclosure is optimal. Commitment allows incumbent firms to deliver as much intertemporal consumption smoothing and cross-subsidization of low-income consumers as permitted by the ex-post participation constraint of the high-income consumers. Information disclosure tightens such constraint, hence reducing the amount of insurance that can be sustained, consistent with the idea in Hirshleifer (1971). 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 outsiders could make. 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 types in the second period because the incumbents can deliver the same value without distorting the allocation to make it incentive compatible—as they can exclude the low types. 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; instead, they consume their expected income. Incumbents know consumers' history, hence incumbents never make negative profits on any given type. The lack of cross-subsidization for the low types is in contrast to the version of the model in Netzer and Scheuer (2014) where all firms are equally uninformed about the consumer's type. For that to happen, firms must also serve the high types and make profit on them. The presence of an informed incumbent prevents outsiders from attracting high-income consumers because the incumbent can offer them a higher value because it does not need to apply an adverse-selection distortion. Thus, in equilibrium the outsiders 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 to consumers in the first period. Thus, typically, the optimal contract is *front-loaded*.

Next, we characterize 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- and second-period consumption for the high-income consumers. We show that under mild conditions, 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: the same consumption in the first and second periods for the high types and no consumption insurance for the low types.

Our analysis sheds light on the value of long-term relationships between firms and consumers. With commitment on the firm's side, having a long-term relationship is always preferable to a sequence of spot contracts. When the firm cannot commit, there is a trade-off: long-term relationships allow for some intertemporal smoothing and dispense with the adverse-selection distortion but prevent cross-subsidization across types in the second period. We show that long-term relationships are preferable when the share of high-income consumers in the economy is low because the benefits of cross-subsidization with spot contracts are limited and the gains from intertemporal smoothing are large. Vice versa, when the share of high-income consumers is high, the lack of insurance in the second period dominates and spot contracts are preferred. Thus, when the quality of the initial pool is sufficiently high, it is optimal to enact policies such as the inability to condition insurance policies on preexisting conditions, effectively implementing the same allocation as for a sequence of spot contracts.

Our analysis offers a new perspective on the recent debate on consumer data ownership. Jones and Tonetti (2020) show that assigning data property rights to consumers improves theallocation by allowing multiple firms to use the same data. We find a complementary channel: in the presence of insurance considerations, regardless of contract duration, assigning data ownership to consumers is welfare-reducing because doing so kills all insurance within and across periods. Our paper also provides a reason for privacy concerns.

Finally, we introduce reasons for consumers to leave the incumbent by assuming that the outsider firms deliver more utility for exogenous reasons (e.g., better amenities). This changes the problem in two ways. First, public disclosure has an allocative effect: not only it affects the degree of potential competition the incumbent faces but also determines whether consumers switch in equilibrium. Second, the possibility that the high-income consumers do switch in equilibrium ameliorates the adverse-selection problem and allows for cross-subsidization of the low-income consumers in the second period. We show that if the pool quality is low, these allocative motives call for more information revelation, while if the pool quality is high, there is less information revelation relative to our baseline case.

Related literature This paper contributes to the literature on insurance contracts under lack of commitment, in the spirit of the seminal contributions by Harris and Holmstrom (1982) and Thomas and Worrall (1988). For example, this class of models has been used to study wage-tenure profiles by Burdett and Coles (2003) and, more recently, by Balke and Lamadon (2022) or life-insurance contracts by Hendel and Lizzeri (2003). The agent's outside option plays a critical role in shaping the amount of insurance attainable. We contribute to this literature by assuming that over time the incumbent firm acquires more information about the consumer's persistent type than its competitors do and show how this affects the consumer's outside option. [3] We then study how public disclosure policies affect potential competition and thereby the amount of insurance and consumption smoothing attainable in equilibrium.

In solving for the optimal disclosure policy, our paper connects to the growing literature on information design: Kamenica and Gentzkow (2011), Bergemann et al. (2015), Bergemann et al. (2017), Bergemann and Morris (2019), and Mathevet et al. (2020). Two papers related to our work, Garcia and Tsur (2021) and Immorlica et al. (2022), study optimal information disclosure in adverse-selection economies. Both papers consider a static economy and allow only for a pooling contract (this is without loss in Immorlica et al. (2022)). Calzolari and Pavan (2006) study the problem of information disclosure in a multiple-principals setting. Farinha Luz et al. (2023) characterize requirements for signals on consumer type to be welfare-improving in static adverse-selection economies. Elliott et al. (2021) and Fainmesser et al. (2023) also study how information disclosure affects the market equilibrium, but in their model both consumers and firms are short-lived. There are two main differences between these papers and ours. First, we study a dynamic relationship between the firm and the agent. By doing so, we introduce intertemporal consumption smoothing motives that are key drivers of the optimal disclosure policy. Second, firms (incumbent and outsiders) have different levels of information.

The closest papers to ours are de Garidel-Thoron (2005) and Mukherjee (2008). Both also study the role of information disclosure in dynamic insurance economies with long-term contracts where the incumbent gains an informational advantage. de Garidel-Thoron (2005)'s main analysis is under commitment on the firm's side, and he compares welfare under two polar opposites: full and no information disclosure. His main result is that the value under no information disclosure is higher. The key contribution of our paper is the result showing that some information disclosure is optimal when firms lack commitment.

Mukherjee (2008) studies an environment where outsiders in the second period are

[3] 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. Another strand of the literature allows for asymmetric information to arise between the agent and all the firms, e.g., Green (1987), or about the outside offer that the agent receives, e.g., Hopenhayn and Werning (2008).

more efficient than the incumbent (as in our analysis in Section 9) and finds conditions under which full information disclosure is optimal. Critical to his result is firms' commitment and the ability of the incumbent firm to impose a penalty[4] for agents that choose to leave. Firms can then make ex-post profits on the high-income consumers by collecting such penalty, while full information minimizes distortions from incentive provision.[5] We restrict our analysis to economies where the payment of these penalties cannot be enforced[6] and show that full information disclosure is never optimal. By departing from firms' commitment, our paper speaks to the debate on the optimality of allowing firms to use their informational advantage and it highlights the potential harm from long-term relationships.

Our paper is also related to studies of competition in frictional markets characterized by adverse selection, such as Guerrieri (2008), Guerrieri et al. (2010) and Lester et al. (2019). While in these papers information is symmetric across firms, in our work the level of potential competition ex-post is driven by the incumbent's informational advantage.

Inspired by a growing policy debate, a recent literature studies the consequences of the implementation of open banking policies: 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), and He et al. (2023). This literature considers a static environment in which information sharing stimulates entry of new fintech firms. These static analyses ignore the incentive to entry provided by the possibility of realizing profits in later periods because of informational monopoly. This aspect is analyzed by Jin and Vasserman (2021) in the context of the adoption of monitoring technology in car insurance markets. They show in a policy counterfactual that forcing firms to share data will reduce the incentives to elicit such data and lead to lower welfare in equilibrium. Although via different mechanisms, both their paper and ours show that erasing ex-post profits by fully revealing consumer information is never optimal.

The rest of the paper is organized as follows. Section 2 presents the environment. Section 3 shows how public information disclosure affects potential competition and the equilibrium in the second period. Section 4 studies the case under firm commitment. Sections 5 and 6 study the case when also the firm cannot commit. Section 7 discusses the implications of our results and some extensions. Section 8 studies when a sequence of spot contracts is superior to long-term relationships. Section 9 introduces reasons for the consumer to switch firms, and Section 10 concludes the paper.

[4]Such as a non-compete clause in a labor market application.

[5]See also Cochrane (1995) for a similar insurance mechanism.

[6]For example, such penalties are not allowed in health insurance markets, and they are legally restricted in credit markets (e.g., mortgage prepayment fees).

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. 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 increasing and strictly concave. To simplify notation, without loss of generality, we assume no discounting between period 1 and period 2 and that the price of period 2 consumption in terms of period 1 consumption is 1.

Consumers are uncertain about their income in periods 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, respectively,

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

We assume that $Y_{2H} > Y_{2L}$, so a higher income realization in period 1 forecasts a higher average income in period 2. We also define the (unconditional) expected income in period 1, $Y_1 = \sum_s \pi_1(y_s)\, y_s$,and in period 2, $Y_2 = \sum_s \pi_1(y_s)\, Y_{2s}$. We assume that $Y_{2L} \leqslant Y_1 \leqslant Y_{2H}$, or, equivalently, that $\pi_2(y_H|y_L) \leqslant \pi_1(y_H) \leqslant \pi_2(y_H|y_H)$. Last, we assume that $\pi_2(y_H|y_1) \in (0, 1)$ for all y_1 so that there is residual uncertainty in period 2.

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 consumer's type by observing the income realizations. Such an economy can be exactly mapped into our economy with $Y_1 = Y_2$, so the expected income in period 1 and the unconditional expected income in period 2 are the same. We choose to illustrate our results with a pure-exchange economy, to minimize the required notation.

This setting can be interpreted 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 this model can be reinterpreted 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 offer contracts to consumers, who choose 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 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 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 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 contracts 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 those offering the contract with the highest value, then the consumer stays with the incumbent; if not, the consumers are evenly split among the outsiders with the same offer.[7]

In setting up the interaction among firms in period 2, we follow Netzer and Scheuer (2014) and allow for a stage in which a firm observes other firms' contracts and can decide to withdraw its contract with a small cost $\varepsilon > 0$. This assumption guarantees that there exists 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 the 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).[8] The presence of a strictly positive cost of withdrawing guarantees uniqueness of continuation equilibrium outcomes by ruling out noncompetitive behavior that can arise as explained by Netzer and Scheuer (2014). [9]

[7]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. Our selection rule effectively makes the incumbent choice set closed.

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

[9]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

We will consider two cases. First, we assume, as a benchmark, 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.

If both the incumbent and the consumer can commit to the terms of the contract in period 2 and to staying 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, $c_t\left(y^t\right) = \frac{Y_1+Y_2}{2}$. 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 incumbent.

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, $c_2\left(y_1, m, y_2\right)$ and withdrawal strategy $\delta\left(x^o\right)$, where x^o is the set of contracts offered by the outsiders in period 2. This characterization does not depend on the incumbent's 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's contract to low- and high-income consumers, respectively. Denote by s the fraction of consumers that have $y_1 = y_H$:

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

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 $\left(y_H, m\right)$:

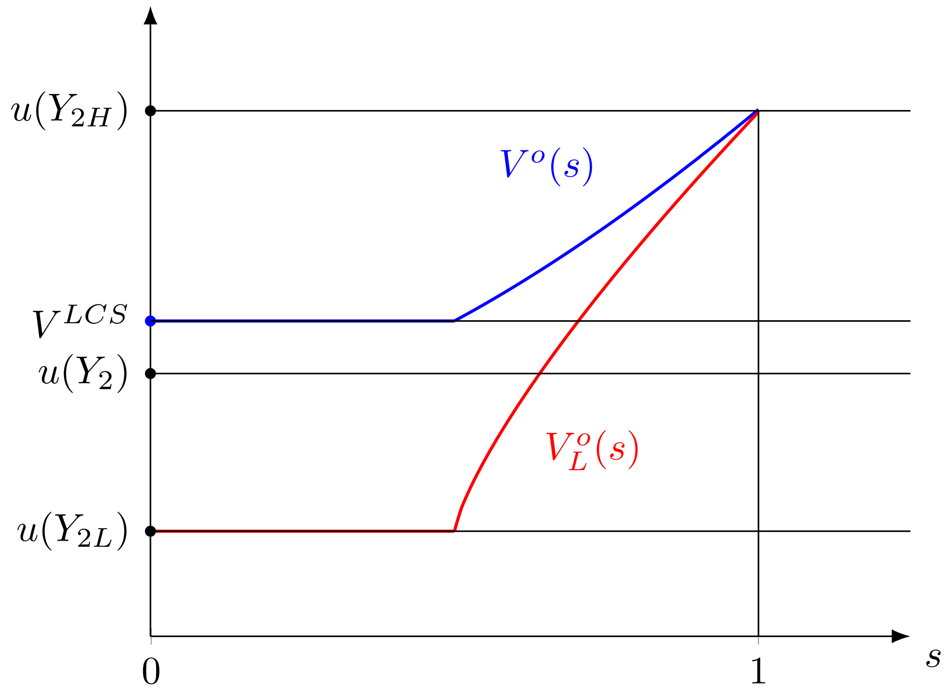
$$V_H^o\left(s\right) = \max_{c_H^o\left(y_2\right), V_L^o} \sum_{y_2} \pi_2\left(y_2|y_H\right) u\left(c_H^o\left(y_2\right)\right) \tag{2}$$

subject to the outsider's nonnegative profit condition,

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

outcome as in the presence of a positive withdrawal cost.

Figure 1: Outside options



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). Netzer and Scheuer (2014) show that without the incumbent and its informational advantage, 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 consumers (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.[10] If instead the share 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 the high-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 earlier. Moreover, it also preempts the outsiders from offering a *cream-skimming* contract that attracts only high types without attracting the low types. This cream-skimming contract delivers a maximal value $V^{cs}(V_L)$ to the high 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 nonnegative profit condition,

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

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_H^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 $u(Y_{2L})$.[11] If the incumbent withdraws its offer (to the low-income consumers), the outsiders do 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 $V_H \geqslant V_H^o(s)$, $V_L \geqslant u(Y_{2L})$, and the incumbent withdraws its offer if the outsiders offer the cream-skimming contract (4), then both the low- and high-income consumers stay with the incumbent and their value is (V_L, V_H), respectively;*

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

[10] Formally, the least-cost-separating allocation solves a restricted version of (2), where the participation constraint (3) must hold with equality.

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

3. *If* $V_H < V^{cs}(V_L)$, $V_L \geqslant u(Y_{2L})$ *and the incumbent does not withdraw its offer 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 consumers accept the cream-skimming contract.*

The proof is in the appendix. The main conclusion is that the incumbent firm retains the high-income consumers if and only if it offers them a value above $V_H^o(s)$, and it withdraws its offer if $V^{cs}(V_L)$ is offered by outsiders. Thus, the value $V_H^o(s)$ 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 the high-income consumers in order to prevent them from being offered a cream-skimming contract 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 the low-income consumers, as under one-sided commitment, but not under two-sided lack of commitment.

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 consumer 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 high 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 the consumer's expected utility subject to a zero-profit 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*

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

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

Moreover, the incumbent commits to withdraw its offer if outsiders offer 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 outsiders to offer the same value without having to distort allocations to satisfy the incentive compatibility constraint. These extra profits can then be used to increase consumption in period 1. Finally, committing to withdrawing the offer if outsiders offer the cream-skimming contract is optimal because it relaxes constraint (5), as competitors cannot poach high types without also attracting low types.

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

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

subject to the firm's nonnegative-profit condition,

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

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 income fluctuations 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 consumers (6) is slack because the incumbent wants to subsidize consumption of the low types in period 2.

Optimal disclosure policy We now turn 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. $K(s) \equiv C(V_H^o(s))$ *is convex.*

That is, the incumbent's cost of providing utility $V_H^o(s)$ in period 2, $K(s) = C(V_H^o(s))$, 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 with 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 period 1, to low-income consumers in period 2, or both, 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. 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,

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

That is,

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

subject to

$$c_1 = Y_1 + \left[Y_2 - \sum_s p(s)\left[sC(V_H(s)) + (1-s)C(V_L(s))\right]\right] \geqslant 0$$

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

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

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

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 $V_H =$

$V_H^o(\pi_1(y_H))$ increase the continuation value for high-income consumers and lower its 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 the 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

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

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 periods 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, consistent with the evidence in Hendel and Lizzeri (2003) and Balke and Lamadon (2022).

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 period 2. 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 extract 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 period 2.

The optimal disclosure policy can be described by a two-signal system, $M = \{g, b\}$. All high-income consumers receive the good signal g. Low-income consumers receive a good signal with probability $1 - \mu$ and a bad signal with probability μ. Thus, μ 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 amount of information disclosed is increasing in the share of high-income consumers.

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 incumbent as a Stackelberg leader that offers its contract to existing consumers, mimicking the timing in the case with one-sided commitment.

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 latter always consume $c_2(y_L, m, y_2) = Y_{2L}$. The consumption of the high-income consumers is*

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

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 $V_H^o(s(m))$ 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 $V_H^o(s(m))$ to the consumers with history (y_H, m). Full insurance is optimal to minimize the cost of delivering such a level of utility. Note that with full insurance the incumbent is making positive profits because $C(V_H^o(s(m))) \leqslant Y_{2H}$, with equality only if the signal is fully revealing. Offering a value $V_H^o(s(m))$ 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, because the outside options for the high-income consumers are constant in the value offered by the incumbent to the low-income consumers. Q.E.D.

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

$$C\left(V_H^o\left(s\left(m\right)\right)\right) < Y_{2H}.$$

The reason is that 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 the 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 consumers with $y_1 = y_H$, because of adverse selection, the former 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-profit condition and anticipating future contracts. Firms anticipate making profits in period 2 on the consumers with a high realization in period 1 (unless the signals fully reveal the consumer's type). Because of ex-ante competition, firms are distributing such profits to consumers in period 1. The optimal level of consumption offered in period 1 solves

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

subject to

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

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

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

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*

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

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

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

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

Summing up, the equilibrium outcome does not have cross-subsidization in period 2 between consumers with high and low income in period 1. No matter what the public disclosure policy is, a consumer with low income realization is offered a contract with full consumption insurance in period 2 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 their consumption in period 1 (for example, in the form of a teaser rate in the credit card context). 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 and moving to an outsider.

6 Optimal public disclosure policy

We now characterize the optimal public disclosure policy. Given the characterization in Lemma 4, the choice of disclosure policy does not affect the consumption of the low-income consumer that always consumes Y_{2L} in period 2, but it can affect the consumption of high-income consumers in period 2 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 options of these consumers 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_H^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

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

subject to the intertemporal zero-profit condition

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

and the share of the high type, y_H, 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:

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

that is, if under no information disclosure, consumption in period 2, $C\left(V_H^o(\pi_1(y_H))\right)$, is lower than consumption in period 1.

For low values of $\pi_1(y_H)$, condition (17) does not hold, the lower bound on high-income consumers' value is high, and the 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, $K(s) = C\left(V_H^o(s)\right)$, is convex in the share of high-income consumers with a given signal, then the optimal way to do so is by providing no information.

For higher 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, together with a fraction of low-income consumers, receive a good signal. Only low-income consumers receive a 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 period 2. This is attained for an intermediate value of signal informativeness between full information revelation and no information disclosure.

The following proposition characterizes the optimal information disclosure:

Proposition 2. *There exists a cutoffs π^* such that:*

1. *If $\pi_1(y_H) < \pi^*$, then $c_1 < c_2(y_2)$ and it is optimal to provide no information;*

2. *If $\pi_1(y_H) \geqslant \pi^*$, then consumption is equalized between periods 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.*

The formal proof is in the appendix. We first show that under the convexity of $K(s) = C(V_H^o(s))$ it is optimal to assign the same signal to all high-income consumers. This is because introducing dispersion for the high-income consumers 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 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 the 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 $[V_H^o(\pi_1(y_H)), V_H^o(1)]$.[12]

We can then write problem (16) as

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

subject to

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

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

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

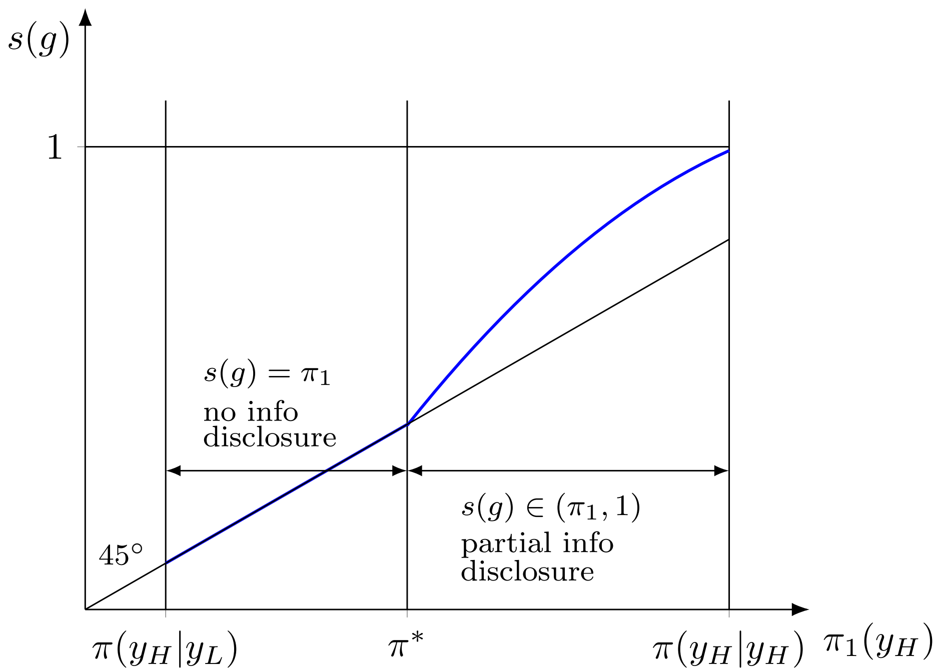
$$V_H = V_H^o(s(g)) \text{ and } s(g) = \frac{\pi_1(y_H)}{\pi_1(y_H) + \pi_1(y_L)(1 - \mu(b|y_L))}. \tag{20}$$

The optimal disclosure policy can then be represented by the share of high-income consumers who received a 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 $V_H \geqslant V_H^o(\pi_1(y_H))$ 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 higher values of $\pi_1(y_H)$, the value of expected profits under no information disclosure are then higher (and also Y_1 is higher), so condition (17) is satisfied and the V_H

[12] The size of the interval $[V_H^o(\pi_1(y_H)), V_H^o(1)]$ is hump-shaped in the degree of income persistence. If income is very persistent, e.g. $\pi_2(y_H|y_H) \to 1$, then it is easy to separate the high-income consumers, and $V_H^o(\pi_1(y_H))$ is close to $V_H^o(1)$. The same is true if income is (close to) i.i.d. over time and y_1 is not informative about y_2. Thus, the optimal public disclosure policy is more effective for intermediate values of income persistence.

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



that solves (19) is interior.[13] It follows that

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

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

Consider now the consumption profile at the optimal public disclosure policy. Consumption is *front-loaded*[14] and, for high 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 a case, consumption is back-loaded and there is perfect insurance between period 1 and period 2 after a

[13] For values of $\pi_1(y_H)$ higher than $\pi_2(y_H|y_H)$ violating our requirement that $Y_1 \leqslant Y_{2H}$, condition (17) holds but it is not possible to find a disclosure policy such that consumption in period 1 and 2 are equated. Thus, in this cases it is optimal to provide full information disclosure and consumption is $c_1 = Y_1 > Y_{2H} = c_2(y_H)$. Note that the requirement that $Y_1 \leqslant Y_{2H}$ is never violated in the type interpretation of our economy.

[14] A front-loaded consumption profile is also a feature of environments where the consumer has private information about his endowment, as in the seminal paper Green (1987).

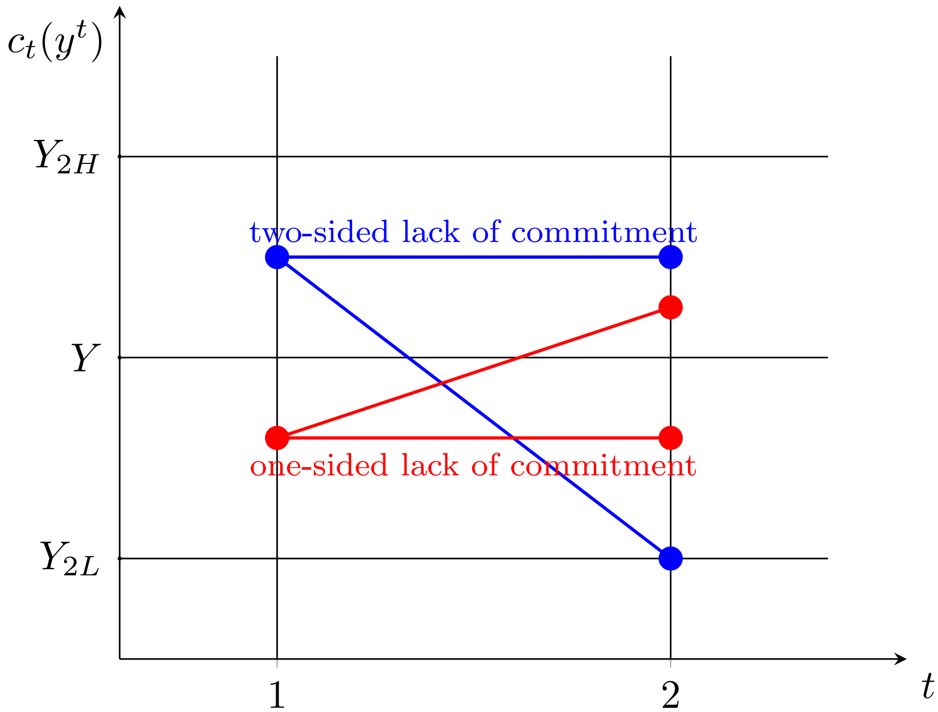
bad income realization. The different predictions for the optimal consumption profiles are illustrated in Figure 3. The critical difference is that when the incumbent can commit, it redistributes the profits it can extract from the high-income consumers in period 2 to consumption in period 1 and to low-income 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 commitment on the firm's 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.

The time series of contractual terms can then help to disentangle whether a firm can commit or not. As we showed in Proposition 1 and Proposition 2, the ability of firms to commit to a contract (because of reputation considerations, for example) is a critical factor for the optimal amount of information to disclose. If the terms are front-loaded, this is an indicator of firms' inability to commit and more information must be provided. If instead terms are back-loaded, this is an indicator that firms can commit and less (no) information should be disclosed to competitors.

We have established how optimal information disclosure varies with the share of consumers with a high-income realization in period 1, $\pi_1(y_H)$. 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 the high types, ρ, and of the conditional probability of high income realization for a given type. An increase in ρ that induces a higher probability of observing a high income in period 1, $\pi_1(y_H)$, would also induce a higher conditional probability of observing a high income in period 2, $\pi_2(y_H|y_H)$. 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 the appendix, numerical examples show that if the share of the high type ρ is small, then it is typically optimal to provide no information, while if this share is sufficiently high, it is optimal to provide partial information.

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



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-profit 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 (M, μ) 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 $V_H^o(\pi_1(y_H)) < V^{optimal}$, and the low types get the same consumption as in the original allocation. Thus, the incumbent makes more profits on the high type. Providing no information is optimal ex-post, 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.

Ownership of information We implicitly assume that firms own the information they observe and are the only ones to have access to a commitment reporting technology that allows them to truthfully communicate, according to (M, μ). Clearly, if consumers had access to such a technology in the first period, they would choose the same disclosure policy as the one in the paper. However, if consumers had access to a technology to credibly report information about y_1 at the end of period 2—but they could not commit to a reporting strategy at the beginning of period 1—then full information disclosure would be the only equilibrium outcome, following the logic in Grossman (1981) and Milgrom (1981). In fact, all high-income consumers would have an incentive to fully disclose y_H.

As shown in the previous section, while full information disclosure maximizes the value for the high-income consumers in period 2 without affecting the value of the low-income consumers (assuming firms cannot commit), it is welfare-reducing, because it destroys the valuable consumption smoothing across periods. In this sense, transferring the ownership of information to the consumers—as intended by the open banking regulation—is detrimental, and it is better to have the firm owning the consumers' information.

The optimality of assigning data ownership to firms would typically be true even if they cannot commit to a disclosure policy. As argued above, the incumbent has incentives to reveal no information ex-post. Consumers instead will reveal all the information. As shown in de Garidel-Thoron (2005), the no-information outcome is preferred to the full-information outcome when $Y_1 = Y_2$, which is the relevant assumption in a stationary environment with unobservable types.

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 incumbent's profits that are rebated to the consumer 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 p must satisfy the Bayesian plausibility constraint requiring that

$$\sum_s p(s)\, s = \pi_1(y_H); \tag{22}$$

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

Define

$$\begin{aligned} w_2(s) &\equiv sV_H^o(s) + (1-s)\, u(Y_{2L}), \\ w_1(p) &\equiv u\left(Y_1 + \sum_s p(s)\left(Y_{2H} - C\left(V_H^o(s)\right)\right) s\right) \end{aligned}$$

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

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

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

Extensions INFORMATION STRUCTURE. We assumed that income y_1 is not directly observable by outsiders and 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 held 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 lay-offs) 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 contracts to potentially separate different types of consumers. 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 agents' characteristics but also of individual effort, as in Holmström (1979). For example, a worker's human capital is determined by her intrinsic ability and 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 provides more information than the base case analyzed earlier. This is because more information allows for more spreading in continuation values that incentivize the agent to exert effort in the first period.

8 Long-term vs. short-term contracts

We next study the value of long-term relationships between firms and consumers. We can do so by comparing the equilibrium outcome in our baseline case to the case where all consumers must change firms in the second period and there are no long-term relationships. A literal interpretation of the latter arrangement is that all consumers are forced to sign contracts with a new firm (even if the incumbent can offer higher value because of better information). An alternative interpretation is that the incumbent is prohibited from offering different contracts (or menus) to consumers it has a relationship with. Thus, despite having information about the consumers, the incumbent is forced to offer them menus of contracts that satisfy the incentive compatibility constraint. One example of this no-discrimination requirement that is relevant in practice is the prescription of not allowing health insurance providers to condition their offerings on pre-existing health status.[15]

[15] Another way to attain the same outcome would require departing from the assumption that the incumbent firm is always informed about the realization of y_1 and considering situations in which the consumer has the power to control the information collected. The consumer could then opt to reveal no information to the incumbent.

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, in the second period, low-income consumers consume the expected value of their income with no transfers from the high-income consumers. This is because the presence of an informed incumbent aggravates the adverse-selection problem: new firms know that only low types are willing to switch (absent additional utility benefits) and therefore are not willing to offer anything more than actuarially fair contracts. Thus, only intertemporal smoothing is possible between the first period and the second period in the high-income state.

Consider next the environment without long-term relationships between firms and consumers. In this case all high-income consumers are forced to switch to new firms. This allows for the possibility of transfers to the low-income consumers in the second period, as shown by Netzer and Scheuer (2014). For a given public disclosure policy, consumption in the first period is Y_1, as there are no future expected profits. In the second period, the spot contract offered by the outsiders with a signal with share s is the Miyazaki-Wilson contract that solves (2). The expected value without long-term relationships is then equal to

$$u\left(Y_1\right) + \sum_s p\left(s\right) W\left(s\right),$$

where $p\left(s\right)$ is the distribution of high types' shares induced by the disclosure policy and

$$W\left(s\right) \equiv sV_H^o\left(s\right) + \left(1-s\right)V_L^o\left(s\right)$$

is the expected value associated with the Miyazaki–Wilson contract with high-income consumers' share s.

The value W is illustrated in Figure 4. To understand the shape of W, we use the following lemma that characterizes the Miyazaki–Wilson program:

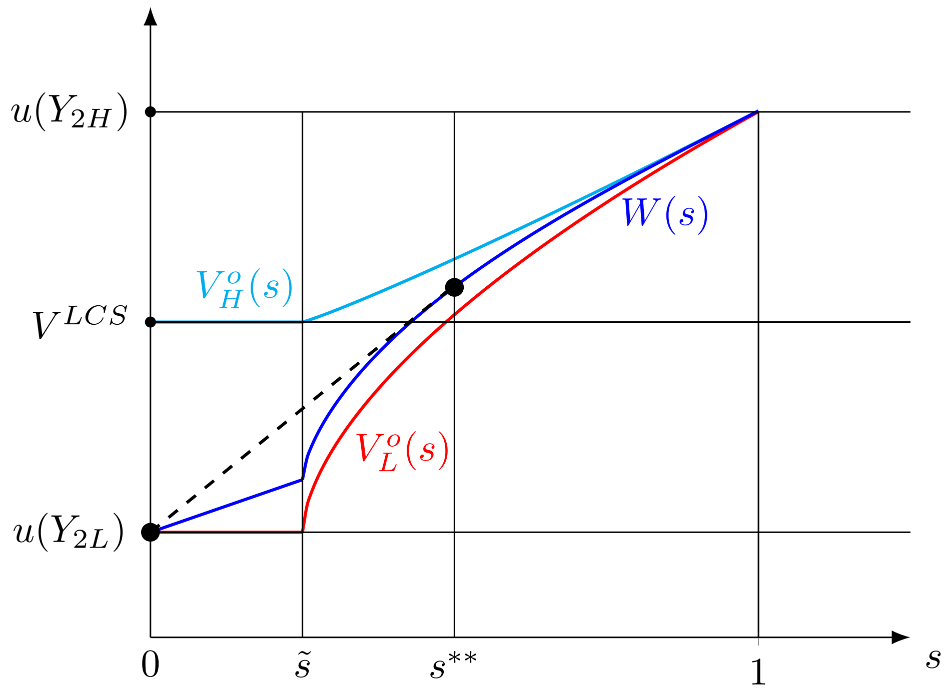
Lemma 5. *There exists a cutoff pool composition* $\tilde{s} \in (0,1)$ *such that* $V_H^o\left(s\right) > V_H^{LCS}$ *and* $V_L^o\left(s\right) > u\left(Y_{2L}\right)$ *if and only if* $s > \tilde{s}$.

For $s < \tilde{s}$ the value for the low-income consumers is $u\left(Y_{2L}\right)$ and for the high-income consumers is V^{LCS}. Both are independent of s, so W is linear in s over the interval $[0, \tilde{s}]$. For shares above $\tilde{s}$, the solution to (2) calls for subsidizing the low types to relax the incentive compatibility constraint. We make the following assumption:

Assumption 2. $W\left(s\right)$ *is concave for all* $s \geqslant \tilde{s}$.

This is automatically satisfied if $V_H^o\left(s\right)$ is concave, and it is satisfied in all our numerical examples.

Figure 4: Expected continuation value with spot contracts, $W(s)$



The optimal disclosure policy solves

$$\max_{p(s)\in\Delta([0,1])} \sum_{s} p(s)\, W(s) \tag{24}$$

subject to $\sum_s p(s)\, s = \pi_1(y_H)$. The solution of this problem can be best understood with the geometric interpretation offered in Kamenica and Gentzkow (2011). Let $\hat{W}$ be the concavification of W; that is, the smallest concave function weakly greater than W for all s. As illustrated in Figure 4, W has a convex kink at $\tilde{s}$ where the participation constraint (3) in the problem (2) starts to be slack and the outsider finds it optimal to subsidize low-income consumers. Thus, there exists a value $s^{**} \in (0,1)$ such that the concavification of W is the segment between $W(0)$ and $W(s^{**})$ for $s \in [0, s^{**}]$, and it coincides with W for $s \geqslant s^{**}$.

Lemma 6. *Under Assumptions 1 and 2, the optimal disclosure policy is such that:*

1. *If $\pi_1(y_H) < s^{**}$, then it is optimal to provide partial information disclosure with a bad-signal structure: $M = \{g, b\}$ with $s(g) = s^{**}$ and $s(b) = 0$.*

2. *If $\pi_1(y_H) \geqslant s^{**}$, then it is optimal to provide no information.*

The optimal information design problem trades off the cross-subsidization it delivers to the low-income consumers and the distortions associated with the incentive constraint for the high-income consumers. If $\pi_1(y_H) < s^{**}$, it is optimal to provide partial information disclosure with all the high-income consumers receiving the good signal and only a fraction of the low-income consumers receiving the good signal. The share of the high type among the consumers with a good signal is s^{**}. Mechanically, the partial information disclosure takes advantage of the non-concavity of W over $[0, s^{**}]$ and has two signals with shares 0 and s^{**}.

Intuitively, if $\pi_1(y_H)$ is low, the composition of the consumer pool absent any information disclosure is bad enough that it is optimal for the outsiders to offer a heavily distorted allocation to the high-income consumers with little to no subsidies (0 if $\pi_1(y_H) \leqslant \tilde{s}$) to the low-income consumers. Thus, disclosing some information is beneficial to both the high-income consumers that can attain a less distorted allocation and, in expectations, also to the low-income consumers. In fact, with some disclosure, at least a fraction of the low-income consumers get to receive a large subsidy. This result is similar to that in Goldstein and Leitner (2018).

When $\pi_1(y_H) > s^{**}$, it is optimal to provide no information, because the pool is sufficiently good and there are no gains from generating more information: the lower distortions for the high-income consumers are dominated by the reduction in insurance for the low-income consumers.

Lemma 6 implies that at least a fraction of the low-income consumers receive consumption higher than Y_{2L}.[16] Thus, the absence of an informed incumbent in the second period has the advantage of delivering higher consumption for the low-income consumers in the second period. There are also costs: no intertemporal smoothing is possible, and there are distortions in the allocation offered to the high-income consumers in the second period in order to screen out the low-income ones.[17]

It is then not obvious whether from an ex-ante perspective consumers prefer to be in an economy with long-term relationships or one without. The next proposition shows that for $\pi_1(y_H)$ close to 0, long-term relationships are beneficial, while for $\pi_1(y_H)$ close to 1, they are detrimental:

Proposition 3. *For $\pi_1(y_H) > 0$ sufficiently close to 0, the expected value with long-term relationships is higher than in the economy with spot contracts only. For $\pi_1(y_H) < 1$ but sufficiently close to 1, the expected value with long-term relationships is lower than in the economy with spot contracts only.*

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 the first period, there are

[16]The fraction of the low-income consumers that receive a good signal and consumption above Y_{2L} is $\frac{\pi_1(y_H)}{\max\{\pi_1(y_H), s^{**}\}} \frac{(1-\max\{\pi_1(y_H), s^{**}\})}{1-\pi_1(y_H)} > 0$.

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

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 $V^i(\pi)$ be the equilibrium value to consumers at the optimal information disclosure policy given $\pi = \pi_1(y_H)$ and $i = LT, ST$, where LT stands for long-term contracts and ST for the economy with spot contracts only. It is easy to see that

$$\begin{aligned}\frac{\partial V^{ST}(\pi)}{\partial \pi}\Big|_{\pi=0} &= u'(y_L)(y_H - y_L) + \left[V^{LCS} - u(Y_{2L})\right] \\ &< u'(y_L)(y_H - y_L) + \left[V^{LCS} - u(Y_{2L})\right] + u'(y_L)\left(Y_{2H} - C\left(V^{LCS}\right)\right) \\ &= \frac{\partial V^{LT}(\pi)}{\partial \pi}\Big|_{\pi=0}.\end{aligned}$$

Thus, $V^{LT}(\pi) > V^{ST}(\pi)$ if π is sufficiently close to 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 the high-income consumers and to distribute the ex-post profits $Y_{2H} - C\left(V^{LCS}\right)$ as first-period consumption, valued at the marginal utility $u'(y_L)$. It follows that when adverse selection is severe, long-term relationships are beneficial.

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

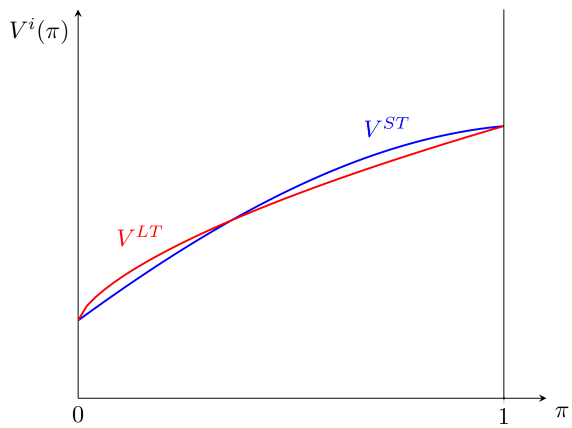
$$\begin{aligned}\frac{\partial V^{ST}(\pi)}{\partial \pi}\Big|_{\pi=1} &= u'(y_H)(y_H - y_L) + u'(Y_{2H})[Y_{2H} - Y_{2L}] \\ &< u'(y_H)(y_H - y_L) + [u(Y_{2H}) - u(Y_{2L})] = \frac{\partial V^{LT}(\pi)}{\partial \pi}\Big|_{\pi=1}\end{aligned}$$

where the inequality follows from strict concavity of u. Thus, $V^{ST}(\pi) > V^{LT}(\pi)$ if π is sufficiently close to 1. Intuitively, when the pool of consumers in the economy has sufficiently many high-income consumers, there are large gains from cross-subsidization 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 above which $V^{ST}(\pi) > V^{LT}(\pi)$, as illustrated in Figure 5.

Our result that the quality of the consumer pool modulates the optimality of inducing the reclassification of consumers extends a similar insight in Handel et al. (2015), who

Figure 5: Value of long-term relationships between firms and consumers



restrict attention to spot contracts with symmetrically uninformed firms, to a dynamic environment with asymmetrically informed firms.[18] In both our paper and theirs, prohibiting consumers' reclassification is optimal when adverse selection is not severe (high π_1). However, in our paper with dynamic relationships, a low value of π calls for *partial*—instead of full–information disclosure, since the informed incumbent is able to avoid the adverse-selection distortion and, by making positive profits ex-post, is able to increase consumption ex-ante.

9 Ex-post competition and switchers

So far we have studied the role of information revelation in affecting potential competition since there are no allocative motives for the consumer to switch to a new firm. In this section, we introduce allocative motives by increasing the utility that the outsiders offer to the consumer who decides to switch. It is then possible that high-income consumers do switch in equilibrium, and this ameliorates the adverse-selection problem and allows for cross-subsidization of low-income consumers in the second period from new firms. We ask whether the introduction of allocative motives to switch in the second period entails an increase in the optimal amount of information disclosed to the market or whether the possibility of cross-subsidization induced by said switching calls for more secrecy. We show that the answer depends on the quality of the initial pool of consumers. If the pool quality is low—i.e., low $\pi_1(y_H)$—then allocative motives call for more information being revealed to incentivize consumers to switch to a new firm, while if the pool quality is high—i.e., high $\pi_1(y_H)$—then less information is revealed.

[18] See also Finkelstein et al. (2009) for evidence on the redistributive implications of policies that restrict the dependency of contracts on observable consumer characteristics.

We modify the baseline environment by assuming that consumers who switch obtain a utility benefit $\Delta \geqslant 0$ in addition to the utility from the consumption profile prescribed by the offered contract. Thus, the utility they obtain by switching to the new firm is $u(c) + \Delta$. This additive increase in utility captures the advantage outsiders have over incumbents without affecting the Miyazaki-Wilson contract described in (2). Therefore, it is immediate to see that $V_i^o(s;\Delta) = V_i^o(s) + \Delta$ for $i = H, L$.

We first consider the equilibrium outcome in the second period for a given public disclosure policy. When $\Delta > 0$, all low-income consumers leave the incumbent, breaking the indeterminacy in the baseline case. This is because they can get at least utility $u(Y_{2L}) + \Delta$ from new firms, while the incumbent offers them at most utility $u(Y_{2L})$. However, the critical difference with the baseline case is that when $\Delta > 0$, also some high-income consumers may choose to switch to a new firm if the value offered to them is higher than the best the incumbent can offer them without making negative profits on them in the second period. That is, the incumbent chooses to retain high-income consumers only if $V_H^o(s) + \Delta \leqslant u(Y_{2H})$. If instead $V_H^o(s) + \Delta > u(Y_{2H})$, then high-income consumers switch to outsiders. For any $\Delta > 0$, we define $s^*(\Delta)$ as the maximal share of the high type with a given signal such that it is optimal for the incumbent to retain high-income consumers with that signal, implicitly defined by $V_H^o(s^*(\Delta)) + \Delta = u(Y_{2H})$. Note that if Δ is sufficiently large, then it is impossible for the incumbent to retain consumers independently of s. This is true for all $\Delta \geqslant \bar{\Delta}$ defined as $V_H^{LCS} + \bar{\Delta} = u(Y_{2H})$. [19] Throughout this section we restrict attention to values of $\Delta \in [0, \bar{\Delta}]$ because for $\Delta > \bar{\Delta}$ the allocation is the same as for $\bar{\Delta}$.

The possibility that consumers switch in the second period if $s \geqslant s^*(\Delta)$ has two important effects. First, if some high-income consumers are not retained by the incumbent, then it is not making profits on them in the second period and consequently the contractual consumption offered in the first period is going to be lower—to guarantee zero profits in present value to the firm. Second, switches by high-income consumers mitigate the adverse-selection problem and make it possible for low-income consumers to receive a payoff higher than $u(Y_{2L}) + \Delta$ in equilibrium. This is the case if the contract that maximizes the high-income consumers' value calls for subsidizing the low types, to relax the incentive compatibility constraint, that is, if $s \geqslant \tilde{s}$, as shown in Lemma 5.

One can also show that $\tilde{s} \leqslant s^*(\Delta)$. This follows from $V_H^o(s^*(\Delta)) + \Delta = u(Y_{2H}) = V_H^{LCS} + \bar{\Delta}$ and $V_H^o(s)$ being increasing. Therefore, whenever high-income types switch, that is, for $s > s^*(\Delta)$, low-income types with the same signal as high-income types receive a higher value from consumption than $u(Y_{2L})$.

The next lemma summarizes this discussion:

[19] Recall that V_H^{LCS} is the minimum value offered to high-income consumers by outsiders for any disclosure policy.

Lemma 7. *Given a public disclosure policy* (M, μ)*, the equilibrium outcome has*

$$c_1(y_1) = Y_1 + \pi_1(y_H) \sum_m \mu(m|y_H) \Pi(m) \tag{25}$$

$$c_2(y_L, m, y_2) = \begin{cases} C(V_L^o(s(m))) & \text{if } V_H^o(s(m)) + \Delta > u(Y_{2H}) \\ Y_{2L} & \text{otherwise} \end{cases} \tag{26}$$

$$c_2(y_H, m, y_2) = \begin{cases} c_H^o(s(m), y_2) & \text{if } V_H^o(s(m)) + \Delta > u(Y_{2H}) \\ C(V_H^o(s(m)) + \Delta) & \text{otherwise} \end{cases} \tag{27}$$

where $\Pi(m) = \max\{Y_{2H} - C(V_H^o(s(m)) + \Delta), 0\}$ *and* c_H^o *is the allocation from the Miyazaki–Wilson problem (2).*

The main differences with respect to the analogous characterization in 4 are the presence of cross-subsidization of low-income consumers and the lack of insurance against realizations of y_2 for high-income consumers whenever the latter sign a contract with an outsider firm.

Having described the equilibrium outcome for a given disclosure policy, we turn to characterizing the optimal disclosure policy. Such policy trades off intertemporal consumption smoothing described in the baseline case ($\Delta = 0$) with the gains from cross-subsidization and allocative efficiency in the second period induced by the utility benefit Δ. To characterize the optimum, it is convenient to first define the expected continuation value for consumers with a signal characterized by the share s:

$$W(s; \Delta) = sV_H^o(s) + (1 - s) V_L(s) + \Delta,$$

where

$$V_L(s) = \begin{cases} V_L^o(s) & \text{if } s \geqslant s^*(\Delta) \\ V_L^o(0) = u(Y_{2L}) & \text{if } s < s^*(\Delta) \end{cases}.$$

Under Assumption 2, $W(s; \Delta)$ is concave for all $s \geqslant s^*(\Delta)$, and it is without loss of generality to consider only three signals (e.g., low, intermediate, and good). The share of high-income consumers is zero for the low signal, $s \in (0, s^*(\Delta))$ for the intermediate signal, and $s \geqslant s^*(\Delta)$ for the high signal. [20]

Lemma 8. *Under Assumptions 1 and 2, the optimal disclosure policy can be implemented with* $M = \{b, i, g\}$ *(bad, intermediate, good), where* b *is a bad signal and* $s(b) = 0$, $s(i) \equiv s_{in} \in$

[20] This assumption holds in all numerical examples we have explored. This is because improvements in the pool composition trigger cross-subsidization for low values of s but reduce it for high values of s, since fewer low-income types receive a good signal.

$(0, s^*(\Delta))$, *so high-income consumers with signal* i *stay with the incumbent, and* $s(g) \equiv s_{out} \in [s^*(\Delta), 1]$, *so high-income consumers with signal* g *switch to an outsider.*

To see why it is optimal to have only one signal associated with consumers that are switching to a new firm, note that the incumbent does not make any profits on them, and because of the concavity of W, it is optimal to have only one signal with $s \geqslant s^*(\Delta)$. The logic we used in Proposition 2 to show that all good types that stay have the same signal can be replicated to show that there is at most one group of consumers with $s \in (0, s^*(\Delta))$ who are retained by the incumbent.

We can then write the problem for the optimal information disclosure policy as

$$\max_{s_{in}, s_{out}, p_0, p_{in}, p_{out}} u\left(Y_1 + p_{in} s_{in}\left[Y_{2H} - C\left(V_H^o(s_{in}) + \Delta\right)\right]\right) + p_0 W(0, \Delta) + p_{in} W(s_{in}, \Delta) + p_{out} W(s_{out}, \Delta)$$

subject to

$$p_0 + p_{in} + p_{out} = 1, \quad p_s \geqslant 0$$
$$p_{in} s_{in} + p_{out} s_{out} = \pi_1(y_H)$$
$$s_{in} \in (0, s^*(\Delta)), \quad s_{out} \in [s^*(\Delta), 1]$$

Proposition 4. *Full information is never optimal. Moreover, there exist cutoffs* $0 < \Delta_L < \Delta_H < \bar{\Delta}$ *such that for all* $\Delta \leqslant \Delta_L$*, we have* $p_{out} = 0$*; for all* $\Delta \geqslant \Delta_H$*, we have* $p_{in} = 0$.

We know that for $\Delta = 0$, it is always optimal for the incumbent to retain high-income consumers, as it is making ex-post profits on them. Thus, the information design problem aims at attaining *consumption smoothing* between period 1 and period 2 for the high types. For strictly positive but sufficiently small Δ, this is still the optimal thing to do because the gains from maximizing the continuation value in period 2 by letting the high-income consumers get the utility benefit Δ and allowing the low-income consumers to obtain some cross-subsidization are limited when Δ is small and therefore $s^*(\Delta)$ is close to 1: a negligible measure of low-income types benefits from cross-subsidization. Therefore, it is optimal to have $p_{out} = 0$ and choose s_{in} such that consumption in period 1 equates consumption in period 2 for high-income consumers. If feasible, the $s_{in}(\Delta)$ would solve

$$Y_1 + \pi_1(y_H)\left[Y_{2H} - C\left(V_H^o(s_{in}) + \Delta\right)\right] = C\left(V_H^o(s_{in}) + \Delta\right).$$

Thus, to maximize the gains from consumption smoothing, it is optimal to keep the continuation value for the high type constant as Δ changes, $V_H^o(s_{in}(\Delta)) + \Delta = V_H^o(s_{in}(0))$. So, $s_{in}(\Delta)$ must be decreasing in this range, and the signal for the high type must be less informative, to reduce the value of the outside option and offset the increase in value given by the higher Δ. Thus, for small Δ, the amount of information disclosed is decreasing in Δ.

When Δ is sufficiently high, it is then optimal to let all consumers leave the incumbent, $p_{in} = 0$. The higher utility offered by the outsiders and the subsidy offered in equilibrium to the low type in period 2 outweigh the possible gains from consumption smoothing. In this case, since the incumbent makes no profit in period 2, the consumption in period 1 is $c_1 = Y_1$ and the problem reduces to maximizing the static value in period 2:

$$\max_{p_{out}, s_{out}} (1 - p_{out}) W(0, \Delta) + p_{out} W(s_{out}, \Delta) \tag{28}$$

subject to

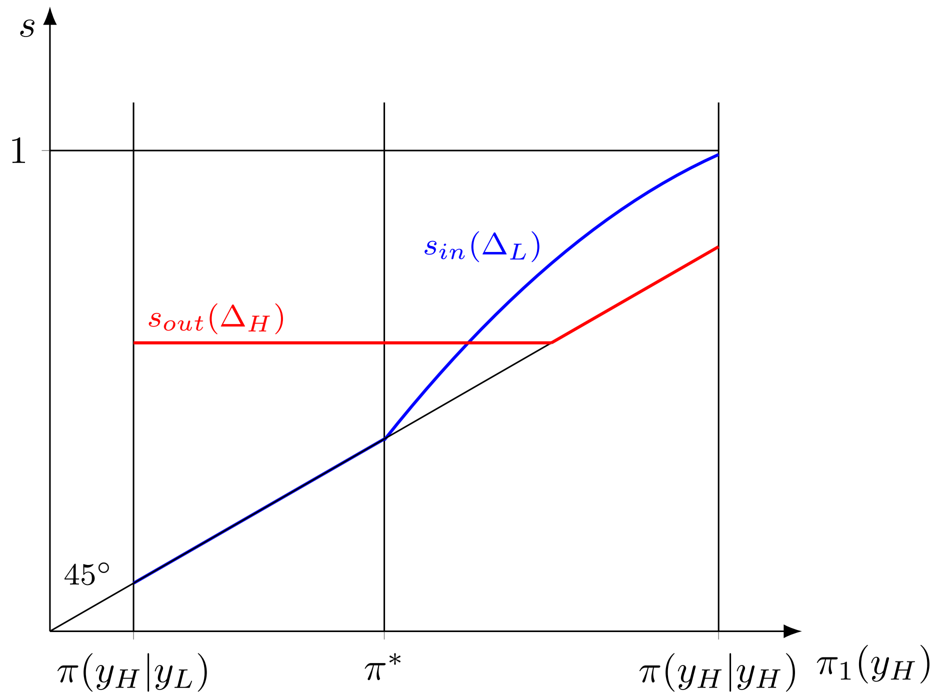
$$p_{out} s_{out} = \pi_1(y_H) \text{ and } s_{out} \geqslant s^*(\Delta).$$

The solution of this problem is the same as the one for the static problem (24) with the additional constraint $s_{out} \geqslant s^*(\Delta)$ in order to guarantee that the high-income types do in fact choose to switch in period 2. It is then convenient to define $\hat{s}(\Delta) = \max\{s^*(\Delta), s^{**}\}$. If $\pi_1(y_H) < \hat{s}(\Delta)$, then it is optimal to have a bad signal with $s = 0$ and a good signal with share $\hat{s}(\Delta)$. In this case, some information disclosure is needed to achieve both switching and cross-subsidization of low-income consumers. If instead $\pi_1(y_H) \geqslant \hat{s}(\Delta)$, then it is optimal to provide no information, since outsiders are still able to attract high-income consumers, and providing further information would only hamper cross-subsidization. Moreover, since $s^*(\Delta)$ is decreasing, stronger motives to switch (higher Δ) reduce the amount of information disclosed to the market.

For intermediate values of Δ, it might be optimal to have two signals with a positive measure of high-income consumers. The intermediate quality signal is associated with high-income consumers who do not switch, who enable intertemporal consumption smoothing; the highest-quality pool is associated with high-income consumers who switch, who enable cross-subsidization of low-income consumers, all of whom sign a contract with outsider firms.

To summarize, optimal information disclosure under allocative switching motives has three main features. First, information disclosure is locally decreasing in Δ in the sense that both s_{in} and s_{out} are decreasing in Δ for $\Delta \leqslant \Delta_L$ and $\Delta \geqslant \Delta_H$ whenever defined. Allocative motives do call for more information disclosure only at the lowest value of Δ such that $p_{out} > 0$. At that Δ, the composition of the best pool of consumers in the economy jumps up in order to induce some high-income consumers to switch. Second, whether $\Delta > 0$ increases the information content present in the optimal public disclosure policy depends on $\pi_1(y_H)$, as illustrated in Figure 6. If $\pi_1(y_H)$ is small enough,

Figure 6: How information content changes with Δ for different consumer pool compositions



then $\pi_1(y_H) < s_{in}(\Delta_L = 0) < s_{out}(\Delta_H \approx \bar{\Delta}) = s^{**}$. Intertemporal smoothing requires no information disclosure, but cross-subsidization requires a better consumer pool composition than $\pi_1(y_H)$. If $\pi_1(y_H)$ is large, then $s_{in}(\Delta_L = 0) \approx 1 > s_{out}(\Delta_H \approx \bar{\Delta}) = s^{**}$, since intertemporal smoothing calls for excessive information compared with the optimum among switchers.

Third, full information is never optimal. This is because either full information is not necessary in order to induce high-income consumers to switch—in which case insurance motives call for an interior pool composition—or full information is needed for high-income consumers to switch—but that would only be true if $\Delta = 0$, which is our benchmark case.

Note that here it is critical that firms cannot impose an exit fee agreed upon in period 1 to high-income consumers who leave the firm in period 2 (such as a non-compete clause in a labor market application or a prepayment penalty in a lending environment) and that firms lack commitment to the continuation contract. This is why our result differs from that of Mukherjee (2008), who finds that full information is optimal in an economy with $\Delta > 0$ when exit fees can be imposed and firms can commit. Under those conditions, firms can extract resources from the high-income consumers without distorting their choice about staying with the incumbent or leaving. The high-income agent can then enjoy the higher utility offered by the outside firms, and the incumbent can collect a fee and make positive profits in period 2 on the high type. Perfect insurance is then attained by paying a severance payment to the low-income consumers who leave in period 2, and possibly by giving a transfer in period 1. In this ideal world, full information maximizes the amount of resources available to the incumbent who optimally distributes them across time and space. See also Cochrane (1995) for a similar insurance mechanism.

When exit fees, severance payments, or both are not available, or firms' commitment is not possible, then full information is not optimal.[21] Without commitment, changes in information disclosure affect not only the amount of resources available to the incumbent but also the way they are distributed. Full information is not optimal because the additional resources extracted from high-income consumers cannot be redistributed to the state in which they are most valued by consumers. For sufficiently high s, reducing such misallocation of resources dominates with respect to increasing their total quantity, making full information suboptimal.

10 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 consumers 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. Full information is never optimal, as it erases the ex-post incumbent profits that are rebated in the first period.

Our analysis highlights a key trade-off between spot and long-term contracts: the former allow for cross-subsidization, while the latter remove the screening distortions due to adverse selection. We show that spot contracts are optimal only when the pool of consumers is sufficiently good to guarantee a transfer to low-income consumers absent an informed incumbent. In an extension of the model that includes utility benefits provided

[21] One can also show that with only commitment or only exit fees, full information is never optimal. In the first case, full information lowers the resources available to be distributed to low-income consumers. In the second case, full information eliminates cross-subsidization by the only firms that provide it, i.e., outsiders.

by outsider firms, information is increasing in the strength of those benefits only when necessary to induce high-income consumers to switch firms, but information is otherwise decreasing in order to preserve cross-subsidization. Yet, also in this modified environment, full information is never optimal.

We see two main avenues for further research. From a theoretical perspective, our model abstracts from endogenous firms' entry and innovation. The presence of an information monopoly and a set of agents that are locked-in with their incumbent can disincentivize entry and the introduction of innovative technologies that deliver higher utility, as in the case analyzed in Section 9. Modeling entry and studying how it interacts with the provision of insurance is a fruitful area for future research. From an empirical perspective, measuring the model fundamentals in specific settings would help the formulation of concrete information policy proposals. This is especially true because optimal disclosure depends on the consumer's incentives to switch 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

BALKE, N. AND T. LAMADON (2022): "Productivity shocks, long-term contracts, and earnings dynamics," *American Economic Review*, 112, 2139–2177. 5, 15

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

——— (2017): "First-price auctions with general information structures: Implications for bidding and revenue," *Econometrica*, 85, 107–143. 5

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

BURDETT, K. AND M. COLES (2003): "Equilibrium wage-tenure contracts," *Econometrica*, 71, 1377–1404. 5

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

COCHRANE, J. H. (1995): "Time-consistent health insurance," *Journal of Political Economy*, 103, 445–473. 6, 37

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, 24

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

ELLIOTT, M., A. GALEOTTI, A. KOH, AND W. LI (2021): "Market segmentation through information," *Available at SSRN 3432315*. 5

FAINMESSER, I. P., A. GALEOTTI, AND R. MOMOT (2023): "Consumer profiling via information design," *Available at SSRN*. 5

FARINHA LUZ, V., P. GOTTARDI, AND H. MOREIRA (2023): "Risk classification in insurance markets with risk and preference heterogeneity," *Review of Economic Studies*, 90, 3022–3082. 5

FINKELSTEIN, A., J. POTERBA, AND C. ROTHSCHILD (2009): "Redistribution by insurance market regulation: Analyzing a ban on gender-based retirement annuities," *Journal of financial economics*, 91, 38–58. 31

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

GREEN, E. J. (1987): "Lending and the smoothing of uninsurable income," *Contractual arrangements for intertemporal trade*, 1, 3–25. 5, 21

GROSSMAN, S. J. (1981): "The informational role of warranties and private disclosure about product quality," *The Journal of law and Economics*, 24, 461–483. 24

GUERRIERI, V. (2008): "Inefficient unemployment dynamics under asymmetric information," *Journal of Political Economy*, 116, 667–708. 6

GUERRIERI, V., R. SHIMER, AND R. WRIGHT (2010): "Adverse selection in competitive search equilibrium," *Econometrica*, 78, 1823–1862. 6

HANDEL, B., I. HENDEL, AND M. D. WHINSTON (2015): "Equilibria in health exchanges: Adverse selection versus reclassification risk," *Econometrica*, 83, 1261–1313. 30

HARRIS, M. AND B. HOLMSTROM (1982): "A theory of wage dynamics," *The Review of Economic Studies*, 49, 315–333. 1, 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

HENDEL, I. AND A. LIZZERI (2003): "The role of commitment in dynamic contracts: Evidence from life insurance," *The Quarterly journal of economics*, 118, 299–328. 5, 15

HIRSHLEIFER, J. (1971): "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61, 561–574. 3

HOLMSTRÖM, B. (1979): "Moral hazard and observability," *The Bell journal of economics*, 74–91. 26

HOPENHAYN, H. AND I. WERNING (2008): "Equilibrium default," *Manuscript, MIT*, 4, 21. 5

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

JONES, C. I. AND C. TONETTI (2020): "Nonrivalry and the Economics of Data," *American Economic Review*, 110, 2819–2858. 4

KAHN, L. B. (2013): "Asymmetric information between employers," *American Economic Journal: Applied Economics*, 5, 165–205. 1, 26

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615. 5, 24, 28

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. 6

LIN, X. AND C. LIU (2022): "Credible persuasion," *arXiv preprint arXiv:2205.03495*. 24

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

MILGROM, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391. 24

MIYAZAKI, H. (1977): "The rat race and internal labor markets," *The Bell Journal of Economics*, 394–418. 10

MUKHERJEE, A. (2008): "Career concerns, matching, and optimal disclosure policy," *International Economic Review*, 49, 1211–1250. 5, 36

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, 10, 27, 45

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

SHARPE, S. A. (1990): "Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships," *The journal of finance*, 45, 1069–1087. 5

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. 10

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:

$$\rho(\theta_H|y_H) = \frac{\rho(\theta_H)\pi_H(\theta_H)}{\sum_\theta \rho(\theta)\pi_H(\theta)} > \rho(\theta_H), \quad \rho(\theta_H|y_L) = \frac{\rho(\theta_H)\pi_L(\theta_H)}{\sum_\theta \rho(\theta)\pi_L(\theta)} < \rho(\theta_H). \tag{29}$$

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:

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

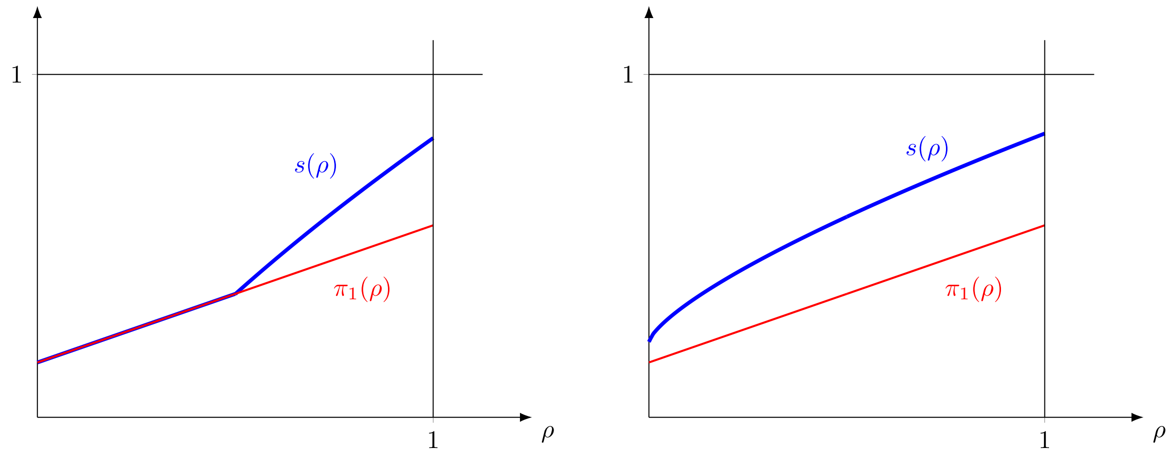
This economy is equivalent to the one considered in the main text with

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

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 7: 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 $V_H \leqslant V_H^o(1)$ is binding in problem (19) because at $s = 1$ we have

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

Thus, we are always in case i) or ii) in Proposition 2 depending on whether condition (17) holds or not.[22] 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

[22] 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.

in 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 sub-period. 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 $V_H \geqslant V_H^o(s)$, $V_L \geqslant u(Y_{2L})$, 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 $V_H^o(s)$. They could offer a cream-skimming contract if $V^{cs}(u(Y_{2L})) > V_H^o(s)$ but for that to be the case it must be that $V_L > u(Y_{2L})$ – otherwise $V^{cs}(u(Y_{2L})) \leqslant V_H^o(s)$. 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 $V_H < V_H^o(s)$ 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. Q.E.D.

B.2 Proof of Proposition 2

First we show that there exists a π^* such that condition (17) holds if and only if $\pi_1(y_H) \geqslant \pi^*$. To simplify the notation, we write $\pi_1 = \pi_1(y_H)$ when not ambiguous. Since $K(s) \equiv C(V_H^o(s))$, we can re-write condition (17) as

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

Note that

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

so f is concave in π_1 as $K'' > 0$ and $K' > 0$. Furthermore, evaluating at $\pi_1 = 0$ and $\pi_1 = \pi_2(y_H|y_H)$ we have

$$f(0) = y_L - K(0) < 0$$
$$f(\pi_2(y_H|y_H)) = 2(Y_{2H} - K(\pi_2(y_H|y_H))) > 0$$

Thus, since $f(0) < 0$, $f(1) > 0$ and f is strictly concave, f must cross 0 only once between $(0, \pi_2(y_H|y_H))$, denote such point by π^*. Then for $\pi_1 \in (0, \pi^*)$ we have $f(\pi_1) < 0$ and for $\pi \in (\pi^*, \pi_2(y_H|y_H))$ 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 $V_H^o(s(m_1)) \neq V_H^o(s(m_2))$. Let

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

Suppose that $\bar{V} \geqslant V_H^o(\pi_1(y_H))$. 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 $V_H^o(s(g)) = \bar{V}$. 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 $\bar{V} < V_H^o(\pi_1(y_H))$. Note that

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

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

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

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

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

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

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

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 $[V^o_H(\pi_1(y_H)), V^o_H(1)]$. We can then write the problem (16) as

$$\max_{c_1, V_H} u(c_1) + \pi_1(y_H) V_H + \pi_1(y_L) u(Y_{2L}) \tag{30}$$

subject to

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

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

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

$$V_H = V^o_H\left(\frac{\pi_1(y_H)}{\pi_1(y_H) + \pi_1(y_L)(1 - \mu(b|y_L))}\right). \tag{31}$$

If the last constraint in problem (30) does not bind and the optimal V_H is interior, it is clear that $u'(c_1) = u'(c_{2H})$ and so

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

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

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

Then just have to check if $u(\bar{c}) \in [V^o_H(\pi_1(y_H)), V^o_H(1)]$. Because of condition (17), $V^o_H(\pi_1(y_H)) \leqslant u(\bar{c})$. Moreover, since $\pi_1 \leqslant \pi_2(y_H|y_H)$ then $u(\bar{c}) \leqslant V^o_H(1)$. Thus, the last constraint does not bind.

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 (30) 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 $V^o_H(s(m_1)) \neq V^o_H(s(m_2))$ with

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

(Clearly, if $\bar{V} > V^o_H(\pi_1)$ 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

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

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 (30) characterizes the full problem (16). Since the last constraint is binding, one way to obtain the optimum is to provide no information so $c_2 (y_H) = C (V_H^o (\pi_1))$ and

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

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

$$
V_H^o (s) = \max_{c_H (y_2), \varepsilon} \sum_{y_2} \pi_2 (y_2|y_H) u (c_H (y_2)) \tag{32}
$$

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_H^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 $V_H^o (s_H) > V_H^o (s_L)$ 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 $V_H^o(s_H) > V_H^o(s_L)$. That $\varepsilon(s_H) > 0$ follows from the fact that $V_H^o(s_L) > V^{lcs}$ 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 $V_H^o(s)$ is close to $u(Y_{2H}) > V^{lcs}$ and so there exists a share of high-type sufficiently high such that $V_H^o(s) > V^{lcs}$.

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 (32), 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 $V_H^o \geqslant V^{lcs}$. 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.

B.4 Proof of Lemma 8

In light of Lemma 7, the optimal public disclosure policy solves

$$\max_{c_1, p(s)\in\Delta([0,1])} u(c_1) + \sum_s p(s)\, W(s;\Delta)$$

subject to

$$\sum_s p(s)\, s = \pi_1$$

$$c_1 = Y_1 + \sum_s p(s)\, s\left[Y_{2H} - \min\{C\left(V_H^o(s,\Delta)\right), Y_{2H}\}\right]$$

Suppose by way of contradiction that p^* that solves the problem above assigns positive probabilities to more than one shares above $s^*(\Delta)$. That is, there are $\{s_1, s_2, ..., s_N\}$ with $N \geqslant 2$ such that $p^*(s_n) > 0$ and $s_n \geqslant s^*(\Delta)$ for all $n = 1, ..., N$. Consider an alternative solution that only assigns probability $p(\bar{s}) = \sum_{n=1}^N p^*(s_n)$ to $\bar{s} = \sum_{n=1}^N \frac{p^*(s_n)}{p(\bar{s})} s_n$ for $s \geqslant s^*(\Delta)$ and $p(s) = p^*(s)$ for all $s < s^*(\Delta)$. Since the incumbent does not make any profits on consumers with $s \geqslant s^*(\Delta)$ then this alternative solution does not change c_1. However, by concavity of $W(s;\Delta)$ for $s \geqslant s^*(\Delta)$ we have that

$$\begin{aligned}\sum_{n=1}^N p^*(s_n)\, W(s_n;\Delta) &= p(\bar{s}) \sum_{n=1}^N \frac{p^*(s_n)}{p(\bar{s})} W(s_n;\Delta) \\ &< p(\bar{s})\, W\left(\sum_{n=1}^N \frac{p^*(s_n)}{p(\bar{s})} s_n;\Delta\right) = p(\bar{s})\, W(\bar{s};\Delta)\end{aligned}$$

so the alternative solution attains a higher value. So we have a contradiction.

The logic we used in Proposition 2 to show that all good types that stay have the same signal can be replicated to show that there is at most one group with $s \in (0, s^*(\Delta))$ that are retained by the incumbent. Q.E.D.

B.5 Proof of Proposition 4

First we show that there exists a $\Delta_L > 0$ such that $p_{out} = 0$ for $\Delta \leqslant \Delta_L$. Suppose by way of contradiction that $p_{out} > 0$. Let V be the value at the optimum

$$\begin{aligned}V = {} & u\left(Y_1 + p_{in} s_{in} \Pi(s_{in})\right) + p_{in}\left[s_{in} V_H^o(s_{in}) + (1 - s_{in}) V_L^o(0)\right] \\ & + p_{out}\left[s_{out} V_H^o(s_{out}) + (1 - s_{out}) V_L^o(s_{out})\right] + p_0 W(0) + \Delta.\end{aligned}$$

Consider a deviation to an alternative policy $(s'_{in}, p'_0, p'_{in}, p'_{out} = 0)$ that delivers the same expected value to the high-income consumers as the one with $p_{out} > 0$, $\sum_s p(s)\, s V_H(s) = \pi_1 V_H(s'_{in})$. For Δ close to 0, such policy exists and $s'_{in} < s^*(\Delta)$. Then notice that, by (strict) convexity of $K(s)$,

$$\sum_s p(s)\, s C\left(V_H^o(s) + \Delta\right) > p'_{in} s'_{in} C\left(V_H^o(s'_{in}) + \Delta\right)$$

As $\Delta \to 0$ we have that $s_{out} \to 1$ and $C(V_H(1)) = Y_{2H}$, so the expected profits in the original allocation are

$$\begin{aligned} p_{in}s_{in}\Pi(s_{in}) &= p_{in}s_{in}Y_{2H} - p_{in}s_{in}C\left(V_H^o(s_{in})\right) \\ &= \pi_1 Y_{2H} - \sum_s p(s)\, sC\left(V_H^o(s)\right) \\ &< \pi_1 Y_{2H} - \sum_s p'_{in}s'_{in}C\left(V_H^o(s'_{in})\right) \\ &= p'_{in}s'_{in}\Pi(s'_{in}) \end{aligned}$$

Then,

$$V' - V = u\left(Y_1 + p'_{in}s'_{in}\Pi'(s'_{in})\right) - u\left(Y_1 + p_{in}s_{in}\Pi(s_{in})\right) > 0$$

since, by construction, the continuation value of the high-income consumers is equal across policies, and all low-income consumers receive $u(Y_{2L})$ under both policies. As the deviation is profitable, we have reached a contradiction. Thus, for Δ close enough to 0 it is optimal to have $p_{out} = 0$.

Next, we show that there exist $\Delta_H < \bar{\Delta}$ such that $p_{in} = 0$ for $\Delta \geqslant \Delta_H$. This is the case since $p_{in} = 0$ for $\Delta \geqslant \bar{\Delta}$. By definition of $\bar{\Delta}$,$V_H^{LCS} + \bar{\Delta} = u(Y_{2H})$ so for all $s \leqslant s^*\left(\bar{\Delta}\right) = \tilde{s}$ the incumbent makes no profits on the high-income consumer it retains. Then $c_1 = Y_1$. We want to show that if Δ is sufficiently close to $\bar{\Delta}$ then it is optimal to have $p_{in} = 0$. Suppose by way of contradiction that $p_{in} > 0$. The value is

$$V = u\left(Y_1 + p_{in}s_{in}\Pi(s_{in})\right) + p_{in}W(s_{in}) + (1 - p_{in} - p_0)\, W\left(\frac{\pi_1 - s_{in}p_{in}}{1 - p_{in} - p_0}\right) + p_0 W(0)$$

Consider a perturbation that decreases p_{in} by $\varepsilon > 0$, increases p_{out} by ε, and increases s_{out} by $\left[\frac{-s_{in}(1-p_{in}-p_0)+s_{out}p_{out}}{(1-p_{in}-p_0)^2}\right]\varepsilon$ so that the perturbation is feasible. The marginal value of such perturbation is

$$\begin{aligned} dV = &-s_{in}\Pi(s_{in})\, u'\left(Y_1 + p_{in}s_{in}\Pi(s_{in})\right) - W(s_{in}) \\ &+ W(s_{out}) + (1 - p_{in} - p_0)\, W'(s_{out}) \left[\frac{-s_{in}(1 - p_{in} - p_0) + s_{out}p_{out}}{(1 - p_{in} - p_0)^2}\right] \end{aligned}$$

As $\Delta \to \bar{\Delta}$ we have $\Pi(s_{in}) \to 0$ so

$$
\begin{aligned}
dV &\to -W(s_{in}) + W(s_{out}) - W'(s_{out})\left[\frac{-s_{in}(1-p_{in}-p_0) + s_{out}p_{out}}{1-p_{in}-p_0}\right] \\
&= \left\{W(s_{out}) - W'(s_{out})\left[s_{out} - s_{in}\frac{[p_{out} + p_{in}(1-p_{out})]}{p_{out}}\right]\right\} - W(s_{in}) \\
&\geqslant \{W(s_{out}) - W'(s_{out})[s_{out} - s_{in}]\} - W(s_{in}) \\
&> 0
\end{aligned}
$$

where the first inequality follows from $\frac{[p_{out}+p_{in}(1-p_{out})]}{p_{out}} \geqslant 1$ for all $p_{in} \geqslant 0$ and the last from the fact that at an optimum it must be that $s_{out} \geqslant s^{**}$ and so

$$
W(s_{in}) > \{W(s_{out}) - W'(s_{out})[s_{out} - s_{in}]\} \quad \forall s_{in} \in (0, \tilde{s}), s_{out} \geqslant s^{**}
$$

as graphically illustrated in Figure 4. Thus, the perturbation is welfare improving for Δ sufficiently close to $\bar{\Delta}$, reaching a contradiction. Then it must be that $p_{in} = 0$ for all Δ close to $\bar{\Delta}$.

Finally, to see that full information is never optimal, simply note that a feasible solution is the solution to (28) with $s_{out} = \max\{s^*(\Delta), s^{**}, \pi_1\}$, $p_{out} = \pi_1/s_{out}$, $s_{in} = 0$. The value associated with this feasible solution is

$$
\begin{aligned}
&u(Y_1) + [(1-p_{out})W(0,\Delta) + p_{out}W(s_{out},\Delta)] \\
>&u(Y_1) + [(1-\pi_1)W(0,\Delta) + \pi_1 W(\pi_1,\Delta)]
\end{aligned}
$$

since the feasible policy maximizes the value in period 2. Thus, full information is never optimal. Q.E.D.

C Convexity of K

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

Lemma 9. *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.*

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

$$V_H^o(s) = \max_{c_H(y_2), c_L(y_2)} \sum_{y_2} \pi_2(y_2|y_H)\, u(c_H(y_2)) \tag{33}$$

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 (given $\pi_i \equiv \pi_2(y_2|y_i)$)

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

subject to

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

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

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

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

or, solving for u_L,

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

and, plugging back in the objective function and letting $x \equiv \exp(u_H)$ 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

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

so

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

and

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

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

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

where

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

Then

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

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

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 $(y_1, \tilde{y}_1)$.[23]

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

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

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:

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

$$c_1(y_1) = Y_1 + \sum_z \pi_1(y_z) p(\tilde{y}_H|y_z) \sum_m \mu(m|y_z, \tilde{y}_H) \Pi_z(m) \tag{34}$$
$$c_2(y_z, \tilde{y}_L, m, y_2) = Y_{2zL} \tag{35}$$
$$c_2(y_H, m, y_2) = Y_{2zH} - \Pi_z(m) \tag{36}$$

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):

$$\max_{c_1, V_{HH}, V_{LH}} u(c_1) + \pi_1(y_H)[p(\tilde{y}_H|y_H) V_{HH} + p(\tilde{y}_L|y_H) u(Y_{2HL})]$$
$$+ \pi_1(y_L)[p(\tilde{y}_H|y_L) V_{LH} + p(\tilde{y}_L|y_L) u(Y_{2LL})]$$

[23] 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.

subject to

$$c_1 = Y_1 + \sum_z \pi_1(y_z)\, p(\tilde{y}_H|y_z)\left[Y_{2zH} - C(V_{zH})\right]$$

$$V_{zH} \in \left[V_z^o\left(p(\tilde{y}_H|y_z)\right), V_z^o(1)\right], \quad z = H, L$$

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

$$V_{zH} = V_z^o(s_z(g)) \text{ and } s_z(g) = \frac{p(\tilde{y}_H|y_z)}{p(\tilde{y}_H|y_z) + p(\tilde{y}_H|y_z)(1 - \mu(b|y_z, \tilde{y}_L))}. \tag{37}$$

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 $c_2(y_H) < C(V_H^o(s(m)))$ to almost all agents with $y_1 = y_H$.

First consider the case $V_H^o(s(m)) = V^{lcs}$. Then the allocation coincides with that in Lemma 4,

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

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

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

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 $i \in [i^*(m), 1]$,

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

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 $\left[V^{lcs}, V_H^o(s(m))\right]$. 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_1 *sufficiently low, it is optimal to provide no information; ii) For all* π_1*, more information is disclosed under discrimination–and strictly so for* π_1 *sufficiently high.*

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) 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)

$$\max_{\mu\in[0,1]} u(Y + \pi_1(y_H)\Pi(\mu)) + \pi_1(y_H)\left[i^*(\mu)V^{lcs}(y_H) + \int_{i^*(\mu)}^1 V_H^o(s_i(\mu))\,di\right]$$

where

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

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

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

Since

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

then

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

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

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

which implies

$$c_1\left(\mu\right) = Y + \pi_1\left(y_H\right)\Pi\left(\mu\right) < Y + \pi_1\left(y_H\right)\Pi^*\left(\mu^*\right) = c_1^*\left(\mu^*\right).$$

Since $V_H^o\left(s_i\left(m\right)\right) \leqslant V_H^o\left(s_1\left(m\right)\right)\ \forall i$, 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 $\nu(e)$. 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.[24]Thus, the two economies are equivalent except for the existence of an 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

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

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

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

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

or

$$[V(h_H, m) - u(Y_{2L})] \geqslant \frac{v(e) - v(0)}{(f(h_H|e) - f(h_H|0))}. \tag{38}$$

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 (38) holds, then μ^* is optimal in the economy with effort. ii) If (38) 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 $\mu(b|h_L) \in (0,1)$ such that*

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

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 (38) and $V_H^o(s)$ being strictly increasing in s. Q.E.D.