**Submitted to *Manufacturing & Service Operations Management***

# Operational Strategies for Distributing Durable Goods in the Base of the Pyramid

Andre P. Calmon
Scheller College of Business, Georgia Institute of Technology, Atlanta, GA, andre.calmon@gatech.edu

Diana Jue-Rajasingh
Essmart, Bangalore, India and Ross School of Business, University of Michigan, USA, diana@essmart-global.com

Gonzalo Romero
Rotman School of Management, University of Toronto, Toronto, Canada, gonzalo.romero@rotman.utoronto.ca

Jackie Stenson
Essmart, Bangalore, India, jackie@essmart-global.com

**Problem definition:** Novel life-improving products such as solar lanterns and energy-efficient cookstoves address essential needs of consumers in the Base of the Pyramid (BOP). However, their profitable distribution is often difficult since BOP customers are risk-averse, their ability to pay (ATP) is lower than their willingness to pay (WTP), and they face uncertainty regarding these products' value.
**Academic/Practical Relevance:** We examine two practical strategies from distributors in the BOP: (1) improving the product's affordability through a discount, and (2) increasing awareness of the product's value. Our results identify BOP-specific operational trade-offs in implementing these strategies. We also propose strategies to manage these trade-offs that can increase consumer surplus in the BOP.
**Methodology:** We introduce a supply chain model for the BOP and analyze the distributor's pricing problem with refunds, as well as the distributor's optimal budget allocation between strategies (1) and (2).
**Results:** We find that, in the BOP, the distributor's profit-maximizing budget allocation often yields the lowest consumer surplus. This misalignment between profits and consumer surplus disappears if customers' ATP is high. Moreover, the misalignment can be resolved if the distributor offers free product returns and commits to a maximum retail price. We confirm the robustness of our results through numerical simulations.
**Implications:** Best operations strategy practices in the BOP can differ significantly from developed markets. Furthermore, BOP customers' limited ATP and high risk-aversion generate a BOP-specific misalignment between profits and consumer surplus. Operational commitments such as free returns reduce this misalignment and can serve as a signal to investors of a social enterprise's focus on consumer surplus.

*Key words*: Social entrepreneurship, base of the pyramid, sustainability, supply chain management.

---

## 1. Introduction

Since the early 2000s, there has been an increasing academic and corporate interest in the creation and distribution of innovative life-improving durable goods that address the unmet needs of Base of the Pyramid (BOP) customers, i.e. the 800 million people that earn less than $2 a day in purchasing power parity (PPP) (World Bank 2016). Examples of life-improving durable goods are affordable solar lanterns, non-electric water purifiers, and smoke-reducing cooking stoves (for a sample of these technologies see Essmart 2020). The profitable distribution of these is the goal of hundreds of social enterprises[1] and billions of dollars of investment[2]. However, businesses have struggled to create and implement profitable BOP-specific distribution strategies (Simanis 2012).

Although the need for research on BOP-specific distribution strategies has been widely discussed in the management literature over the last few decades (see Prahalad 2006), there are few operations management papers that model and examine BOP supply chains. To address this gap, we model and analyze the operations strategy of a BOP distributor that allocates an investment budget to profitably distribute a new life-improving durable good. Consistent with the goal of most social enterprises, the distributor maximizes a mix of its profits and consumer surplus. We consider a model where a distributor sells a product to a profit-maximizing retailer that, in turn, sells to BOP customers, modeling the main components of most BOP markets (Bellman et al. 2018).

BOP customers' *ability to pay (ATP)* for the life-improving good is often lower than their *willingness to pay (WTP)*. Customers with ATP lower than their WTP is a context-defining feature of the BOP (Banerjee et al. 2012) and is a key feature of our model. Furthermore, we model BOP customers as being risk averse. The connection between risk aversion and poverty is well established in the literature (Haushofer and Fehr 2014). In practice, BOP customers are frequently not informed of the benefits that a new life-improving product might provide. As a result, customers might not buy the product due to uncertainty about the product's value and fear of regretting their purchase. Thus, we assume customers have two possible informational states: *uninformed*, where a customer has a high degree of uncertainty about the product's value, and *informed*, where the customer is aware of the product's value (Shugan and Xie 2000).

Recent field experiments suggest that allowing for product returns can be an effective tool for promoting product adoption in developing countries, see Dupas (2014b) and Levine et al. (2018). In practice, although a few BOP-focused organizations allow customers to return products,[3] product returns have been an underutilized option in the BOP. This is surprising, considering that many life-improving products are experience goods where customers have difficulty appraising the product's benefits before purchase. Hence, our model assumes that the retailer, incentivized by the distributor, can offer customers the option to return a product for a refund.

We use our model to study two decision-making problems. The first is the *Pricing Problem* where the distributor chooses the product's wholesale price and the refund it gives the retailer for

[1] Social enterprises are organizations whose objective combines social and economic value creation. This mixed objective differentiates social enterprises from both for-profit businesses and from charities (Miller et al. 2012).

[2] The Global Investment Impact Network estimates that the impact investment market surpassed $500 billion in 2019. See `https://thegiin.org/research/publication/impinv-survey-2019`

[3] Companies such as Burro, a solar product distributor in Ghana, EcoZoom, a cookstove manufacturer in Kenya, and Pollinate Energy, a distributor in India, are a few examples of social enterprises that offer customers the option to return products after purchase.

returned products. The retailer, in turn, decides the customer price and the refund (if any) it offers to customers that regret their purchase and choose to return the product.

The second problem we examine is the distributor's *Allocation Problem* where the distributor, based on the prices and refunds set in the Pricing Problem, allocates an investment budget to increase profits and consumer welfare. We assume the distributor allocates the budget between two common types of investments:

*1) Invest in improving product affordability through discount coupons or vouchers.* Many of the life-improving technologies are well-established, e.g. induction cooktops, water purifiers, and solar lamps. However, these technologies are rarely offered at a price point that BOP customers can afford. With the goal of managing this affordability issue, many BOP organizations address customers' low ATP with discount coupons, subsidies, or financing options.

*2) Invest in customer education to increase awareness of the product's value.* Some benefits of innovative durable goods are not immediately understood by customers.[4] To address information gaps, many social enterprises invest in educational campaigns to increase the number of BOP customers that are informed of a life-improving product's existence and benefits.[5]

We analyze (*i*) how the BOP distributor should allocate its investment budget between improving affordability and customer education (1 and 2 above), (*ii*) how the the distributor's optimal investment strategy changes depending on consumers' ATP and risk aversion, and (*iii*) the interplay between the distributor's optimal investment strategy and consumer surplus. We first analyze a setting where all customers have the same ATP. We then leverage this analysis to examine the case where customers have heterogeneous ATP.

We find that, in equilibrium, investments in affordability always improve the distributor's objective. However, investments in consumer education might worsen the distributor's objective if customers have low ATP and low to moderate risk aversion. Thus, the *value of information* (Shulman et al. 2009) can be negative in BOP contexts and depends on customer ATP and risk aversion.

Moreover, we characterize the optimal allocation and pricing strategies as a function of customers' ATP and risk aversion level. We prove that if customers' ATP is low and risk aversion is high (which is typical of a BOP context), then the distributor's profit-maximizing allocation is to increase product affordability, not allow product returns, and skim the market by targeting customers with high ATP. This pricing and allocation strategy induces the lowest average

[4] For example, smoke-reducing, fuel-efficient, clean cooking stoves could reduce the 4.3 million annual deaths that are linked to households cooking over coal, wood, and biomass stoves (World Health Organization 2016). However, BOP customers are usually uninformed of the negative long-term health repercussions associated with using inefficient stoves, or of the existence of cleaner cooking solutions. As a result, potential customers may be unwilling to pay for unfamiliar products that may not meet their immediate perceived needs (Global Alliance for Clean Cookstoves 2016).

[5] Besides Essmart, social enterprises such as Burro (BOP distributor in Ghana) and Bidhaa Sasa (BOP distributor in Kenya) rely on teams of sales executives to perform product demonstrations and promote their products

consumer surplus. Conversely, if customers' ATP is high (typical of a non-BOP context), the profit-maximizing allocation strategy is to increase customer education, accept product returns, and price the product so that all customers purchase. This pricing and allocation strategy induces the highest average consumer surplus. Hence, our analysis illustrates a *tension* between profits and consumer surplus that is BOP-specific. This result has practical implications: profit-maximizing operations strategies might have adverse implications on consumer surplus when implemented in a BOP context. Finally, we show that the distributor can mitigate this tension by committing to offer free returns to customers and by setting a maximum retail price.

Our results shed light on the trade-offs and challenges that emerge when designing an organization's operations strategy in the BOP. They provide theoretical support for novel distribution strategies and operational commitments that are being implemented by social enterprises aiming at profitably distributing life-improving products and improving consumer welfare in this context.

Several of our modeling assumptions are drawn from the experience of Essmart, an Indian social enterprise that is the main motivating case study for this paper. Essmart has several distribution centers that serve a network of over one thousand small retail shops in southern India. See Appendix A for a discussion of Essmart's operations strategy.

The rest of the paper is structured as follows. In Section 2 we present a literature review. In Section 3 we introduce our model. Section 4 examines the setting all where customers have the same ATP, while Section 5 adresses customers with heterogeneous ATP. In Section 6, we consider a continuous model and examine the robustness of our results. Finally, Section 7 concludes.

## 2. Literature review

Our results contribute to the literature on social entrepreneurship, product returns, sustainable operations, and the intersection of marketing and operations management. To the best of our knowledge, this is the first study of an equilibrium model that describes the challenges BOP distributors face when attempting to profitably reach consumers with low ATP and risk aversion. We position our paper with respect to this literature next.

### 2.1. Social entrepreneurship and development

Mair and Marti (2006) describe social entrepreneurship as "a process involving the innovative use and combination of resources to pursue opportunities to catalyze social change and/or address social needs." Social entrepreneurship emerges in contexts where goods and services are not adequately provided by public agencies or private markets (Dees 1994) and where market and government failures are perceived (McMullen 2011). Through our model, we can address the particular problems confronted by social enterprises operating in the BOP.

Recent BOP field experiments motivate the inclusion of product returns in our model. Dupas (2014b) finds through a field experiment in Kenya that short-term subsidies for bednets can boost long-term-adoptions because consumers, by trying out the bednet, learn its value. Dupas (2014b) also speculates that "free trials" could be an effective tool for promoting product adoption in the BOP. Similarly, Levine et al. (2018) observed in Uganda that giving customers the right to return a cookstove dramatically increased their adoption. Our results provide a game-theoretic foundation for these observations.

### 2.2. Socially responsible supply chains

Our paper is inscribed within the recent research trend in operations management that studies socially and environmentally responsible value chain innovations; see Lee and Tang (2018) for a motivation and literature review. Sodhi and Tang (2017) challenge the supply chain community to develop frameworks for supply chains that seek to have social impact and be financially sustainable. Given the BOP context we consider, our model is one answer to this challenge.

Recently, Uppari et al. (2017) analyze strategies for off-grid energy business models that rely on rechargeable lightbulbs in low-income markets. Similar to our paper, they assume that BOP consumers face financial distress. Zhang et al. (2017) and Gui et al. (2018) consider replenishment strategies that can help retailers in low-income markets manage costs. Since Essmart, who motivates this paper, uses a "deliver-to-order" replenishment strategy to retailers, we focus instead on the trade-off between consumer education and product discounts, modulated by product returns.

A popular research stream is the study of practical mechanisms to incentivize suppliers to comply with social and environmental standards, e.g. Plambeck and Taylor (2015), Chen and Lee (2016), and Huang et al. (2017). Our paper relates to this prior work in that we also characterize the equilibrium behavior of a three-tier supply chain, and quantify the incremental value of different strategies. However, the focus of this literature is usually on a buyer trying to incentivize a supplier that is difficult to monitor. In contrast, we focus on a distributor trying to incentivize a local retailer to carry life-improving products in a low-income market with risk-averse BOP consumers.

A paper related to ours in this area is Taylor and Xiao (2019). They compare distributing socially-desirable products through non-commercial and commercial channels, where the latter includes a for-profit intermediary, in a model that incorporates consumer awareness. Taylor and Xiao (2019) study the optimal subsidy by an international donor. Our model is different in several aspects, including the presence of product returns and locally-provided product financing.

### 2.3. Product returns, valuation uncertainty, and marketing

We build upon the literature on operations of reverse logistics systems that support product returns. In particular, Su (2009) proposes a model where consumers face valuation uncertainty before

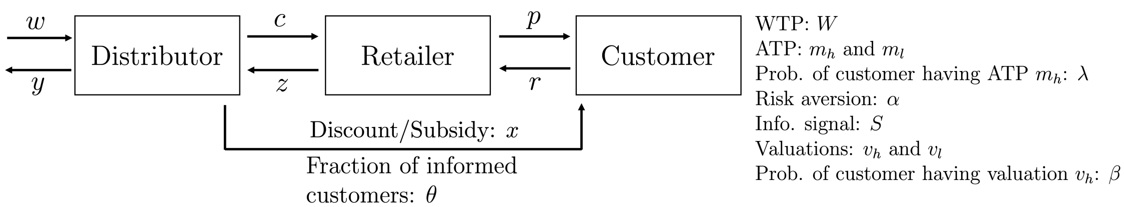



**Figure 1 Summary of model parameters and decision variables.**

purchasing and proposes contracts that coordinate the supply chain including product returns. In contrast, we incorporate information control to the consumer and increasing product affordability, while we ignore aggregate demand uncertainty and do not address the inventory management component of the problem, which were not significant concerns for Essmart, our industry partner.

Our customer education model and the Pricing Problem's outcomes are related to the marketing literature on advance-selling (Shugan and Xie 2000, Xie and Shugan 2001). However, these papers assume two pricing opportunities before the service is provided: the advance price and the spot price, where customers' valuation uncertainty is revealed before the latter. Instead, our model has one pricing opportunity and the valuation uncertainty is revealed after buying the product.

Shulman et al. (2009) consider a model with risk-neutral consumers and two horizontally differentiated products. They consider a binary decision of providing either full product fit information or no information to consumers, and identify conditions where it is optimal to provide full information. In contrast, we consider a richer information control model and incorporate consumers' risk aversion and ATP, which are key characteristics of the BOP (see Shukla and Bairiganjan 2011). While Shulman et al. (2009) shows that improving information availability might reduce a retailer's profits, we describe how the value of information changes with customer ATP and risk-aversion.

## 3. Model

We introduce a game-theoretic model of a supply chain that distributes a life-improving durable good designed for BOP customers in a low-income market. The supply chain has three echelons: the distributor, the retailer, and the customers. As is standard in the literature, we assume that the distributor anticipates the retailer and customer actions. The model parameters, which we explain in detail in the remainder of this section, are depicted schematically in Figure 1.

We assume that the market has a constant number of customers. To simplify our notation, we normalize the total demand to 1, and we work on a per-unit accounting basis. We focus on the case where there are only two possible customer valuations and two levels of customer financial distress. This setup provides novel managerial insights and is analytically tractable. In Section 6, we verify their robustness in a more sophisticated customer model through numerical simulations.

We introduce the supply chain model in three steps. First, we model customer purchasing behavior in the BOP. Next, we model the retailer's behavior as well as the distributor's Pricing and Allocation Problems. Finally, we describe the sequence of events in the game.

### 3.1. Base of the Pyramid customers

We assume customers face product uncertainty: they might not know the exact product's value before their purchase. Specifically, customers have a valuation $v_h$ or $v_l$ for the product, where $v_h > v_l$. A fraction $\beta$ of customers has a high valuation $v_h$, while a fraction $1-\beta$ has a low valuation $v_l$.

Prior to their purchasing decision, the customers may receive a marketing signal from the distributor (the signal models a marketing campaign). The random variable $S \in \{0,1\}$ indicates whether a customer received this signal and is informed ($S=1$) or uninformed ($S=0$). Thus, $\mathbb{P}(S=1)=\theta$ models the *reach* of the marketing campaign and, ultimately, the customer education level in the market. We model $\theta$ as a strategic decision made by the distributor,[6] and we assume $\theta$ to be independent of customer valuations. For simplicity, we assume that the marketing signal carries perfect information about the product. If $S=1$ the customer learns their true valuation (either $v_l$ or $v_h$), while if $S=0$ the customer receives no further information and estimates their valuation based on the distribution of valuations across the population. This customer education model is similar to the literature on advance-selling of services (Shugan and Xie 2000, Xie and Shugan 2001).

The retailer chooses the product price $p$ and might offer customers a refund $r \le p$ if they return the product. A customer returns the product if, after purchase and use, they find that their valuation for the product is less than the refund $r$. Thus, returns are a real option that guarantees customers a minimum level of satisfaction from their purchase — Essmart has found that over 30% of their retailers offer refunds, usually over a two-week return period.

The relationship between poverty and risk aversion, in particular in the BOP, is well established in the literature (see Haushofer and Fehr 2014). For tractability's sake, we model BOP customers' utility function as exhibiting constant absolute risk aversion. Namely, if the customer has a value $v$ for a product with price $p$ and is offered a refund $r$, their utility function $u$ is

$$u(v,p,r) = \frac{1-e^{-\alpha(\max(v,r)-p)}}{\alpha},$$

where $\alpha$ is the risk aversion parameter: the higher $\alpha$ the more risk averse customers are. Hence, customers explicitly take into account the option value of returning the product in their utility.

A customer's *willingness to pay* (WTP) for the product is the maximum price for which the customer's expected utility is non-negative. A customer that receives the perfectly informative

[6] In practice, companies such as Essmart, Solar Sisters (also a distributor in India), and Bidhaa Sasa (a BOP distributor in Ghana) invest in developing sales executive teams to promote and run product demonstrations in villages.

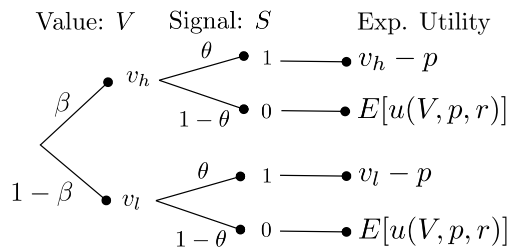



**Figure 2** **Customer utility. A fraction $\lambda$ of customers have ATP $m_h$ and a fraction $1-\lambda$ have ATP $m_l$.**

marketing signal ($S=1$) has a WTP $v_h$ or $v_l$, depending on their (known) product valuation. Conversely, a customer that is uninformed ($S=0$) has a WTP that depends on their expected utility, which we describe next. Let $V$ be a random variable that describes the valuation of a customer chosen at random. Then, the expected utility of an uninformed customer is

$$\mathbb{E}[u(V,p,r)] = \frac{\beta}{\alpha}\left(1-e^{-\alpha(\max(v_h,r)-p)}\right) + \frac{1-\beta}{\alpha}\left(1-e^{-\alpha(\max(v_l,r)-p)}\right). \tag{1}$$

As a result, the WTP of uninformed customers, $p_\alpha(r)$, solves $\mathbb{E}[u(V,p_\alpha(r),r)]=0$. Since $\mathbb{E}[u(V,p,r)]$ is strictly decreasing in $p$, $p_\alpha(r)$ exists and is unique – see Appendix B.1 for a derivation of $p_\alpha(r)$.

We model the significant credit and liquidity constraints BOP customers face by assuming that their *ability to pay* (ATP) for life-improving products is potentially lower than their WTP. This assumption is consistent with recent empirical research (see Tarozzi et al. 2014, Dupas 2014a). We assume a fraction $\lambda$ of customers have ATP $m_h$ and a fraction $1-\lambda$ have ATP $m_l$, where $m_h \geq m_l$. If $p > m_l$ then no customer with ATP $m_l$ will purchase the good, even if their WTP is larger than $p$ (if $p > m_h$ then no customer purchases the product). As a result, *a customer only purchases the product if both her WTP and ATP are larger than $p$.* We model $m_h$ and $m_l$ as exogenous parameters and we assume $v_l \leq m_l \leq m_h \leq v_h$. This assumption is commonly found in the development economics literature, e.g. Banerjee (1997) and Banerjee et al. (2012). Assuming other orderings of $v_l$, $v_h$, $m_l$ and $m_h$ does not lead to additional insights.

We now describe the probability of a customer purchasing and returning the product, as well as their expected surplus. Let $W$ and $M$ be random variables that represent, respectively, the WTP and ATP of a randomly chosen customer. The distribution of $W$ and $M$ are

$$\mathbb{P}(W=\omega) = \begin{cases} \theta\cdot\beta & \text{if } \omega = v_h, \\ \theta\cdot(1-\beta) & \text{if } \omega = v_l, \\ 1-\theta & \text{if } \omega = p_\alpha(r). \end{cases}, \quad \mathbb{P}(M=m) = \begin{cases} \lambda & \text{if } m = m_h, \\ 1-\lambda & \text{if } m = m_l. \end{cases}$$

Figure 2 depicts the distribution of customer utility. Let $\mathbb{1}_{\{e\}}$ denote the indicator function of the event $e$. Then, given a price $p$, refund $r$, and subsidy $x$, the fractions of customers that purchase and return the product are $\mathrm{B}(p,r,\theta) := E[\mathbb{1}_{\{W\geq p, M\geq p\}}|\theta]$ and $\mathrm{R}(p,r,\theta) := E[\mathbb{1}_{\{W\geq p, M\geq p, V<r\}}|\theta]$, respectively. The expected customer surplus, $CS(p,r,\theta)$, is

$$CS(p,r,\theta) := E[(V-p)\mathbb{1}_{\{W\geq p, M\geq p, V\geq r\}}|\theta] - (p-r)E[\mathbb{1}_{\{W\geq p, M\geq p, V<r\}}|\theta]. \tag{2}$$

To summarize, $\theta$ describes the customer education level; $m_h$ and $m_l$ characterize BOP customers' financial distress and the product's affordability; the probability $\beta$ and the valuations $v_l$ and $v_h$ capture characteristics of the product's design; $\alpha$ describes customers' risk aversion; and the equilibrium price $p$ and refund $r$ determine BOP customers' access to the product. Together, these parameters provide a tractable model of BOP customers' behavior.

### 3.2. Retailer

We consider a profit-maximizing retailer that buys the product from the distributor at a cost $c$ and sells it to the customers at a price $p$, while offering a refund $r$ for returns. Moreover, the distributor offers a salvage value $z$ per unit returned by customers to the retailer. The retailer's profit is

$$\Pi_R(p,r) = (p-c)\cdot \mathrm{B}(p,r,\theta) - (r-z)\cdot \mathrm{R}(p,r,\theta).$$

The retailer can choose an outside option instead of selling the distributor's product, which we normalize to zero. The outside option could be an alternative product or competing technology. Since most BOP retailers have limited shelf space, distributors often ensure frequent product replenishment. In fact, most of Essmart's retailers sell its products delivered-to-order: whenever a sale occurs, the item is delivered to the retailer, and the customer collects the product there. For these reasons, we do not explicitly model the retailer's inventory management problem.

### 3.3. Distributor

The distributor purchases products from the manufacturer at a wholesale cost $w$ and sells them to retailers in the BOP market. To rule out trivial outcomes, we assume throughout that the product has a potentially profitably market, i.e., the customers' lowest ATP is larger than the wholesale cost, $w \leq m_l$. Any product returned by the retailer is salvaged for a unit value $y$. We assume that the distributor, as a social enterprise, values consumer surplus. Specifically, the distributor's objective function is a linear mix of profits and consumer surplus.

The distributor makes four decisions. The first two are *pricing* decisions, where the distributor chooses the retailer's price $c$ and refund $z$. The distributor also makes two *strategic* decisions: a potential discount or subsidy to the customer, which we denote by $x$ per item, and the proportion of "informed" customers $\theta$. A subsidy of $x$ makes the effective price that customers face $p-x$. The subsidy $x$ is usually operationalized in the form of a discount coupon, voucher, rebate, or agreement with a microfinance institution.[7] In practice, companies like Essmart increase $\theta$ by investing in marketing campaigns and product demonstrations in BOP communities.

As discussed in Section 1, we consider two optimization problems that many BOP distributors face. The first is the *Pricing Problem* where the distributor chooses $c$ and $z$. For a given subsidy $x$ and customer education level $\theta$ this problem is

$$\begin{aligned} \Pi_D^*(x,\theta) = \max_{c,z} \ & (c-w)\cdot B(p^*-x,r^*,\theta) - (z-y)\cdot R(p^*-x,r^*,\theta) + \gamma \cdot CS(p^*-x,r^*,\theta) \\ \text{s.t. } & \{p^*,r^*\} \in \arg\max_{p,r} (p-c)\cdot B(p-x,r,\theta) - (r-z)\cdot R(p-x,r,\theta) \quad (IC) \\ & (p^*-c)\cdot B(p^*-x,r^*,\theta) - (r^*-z)\cdot R(p^*-x,r^*,\theta) \geq 0. \quad (IR) \end{aligned} \tag{Pricing Problem}$$

The first two terms in the objective function correspond to the distributor's expected profit minus the expected cost of returns, and the parameter $\gamma \geq 0$ models the distributor's relative value for consumer surplus. The first constraint corresponds to the retailer's incentive compatibility, i.e., the retailer chooses the profit-maximizing customer price and refund. Note that the subsidy $x$ changes the customer purchase and return probabilities. The second constraint corresponds to the retailer's individual rationality, i.e., the retailer's profit should be at least its outside option.

The second problem the distributor faces is the *Allocation Problem*. The distributor allocates an investment budget between (1) increasing the subsidy $x$ (a financial lever); and (2) an investment in increasing the customer education level $\theta$ (an information lever). The investment budget could represent, for example, a grant from a foundation or aid agency to improve the distributor's operations. We assume that the budget is earmarked to be spent exclusively between options (1) and (2), i.e. the distributor does not simply keep the investment budget for itself.

In more detail, the distributor allocates a budget $b$ per customer between increasing customers' ATP and improving customer education. We denote the marginal cost of the former by $c_\theta$ and assume that only customers that purchase the product receive the subsidy. Thus, given some initial customer educational level $\theta_0$ and customers' ATP $m$, the distributor chooses $\theta$ and $x$ that solve

$$\begin{aligned} \underset{x,\theta}{\text{maximize}} \quad & \Pi_D^*(x,\theta) \\ \text{subject to} \quad & c_\theta(\theta-\theta_0) + x\mathrm{B}(p^*-x,r^*,\theta) \leq b, \\ & \theta_0 \leq \theta \leq 1, 0 \leq x \leq v_h - p^*. \end{aligned} \tag{Allocation Problem}$$

The objective function of the Allocation Problem is the optimal objective value of the Pricing Problem, while $p^*$ and $r^*$ are functions of $x$ and $\theta$. We assume a natural upper bound on the subsidy $x \leq v_h - p^*$, i.e., the distributor never increases the customers' ATP above their maximum WTP. This constraint is equivalent to assuming that the maximum non-discounted price for the

[7] For various case studies on the implementation of leasing models and product financing for the distribution of life-improving good, see Clean Cooking Alliance (2015, 2019).

product will be at most the largest possible customer WTP (which is $v_h$). This constraint has a practical motivation since unbounded discounts are unrealistic. Relaxing the constraint $x \leq v_h - p^*$ does not affect our main insights, although it does slightly simplify our analysis.

In addition, we assume that the distributor's budget can at most fully maximize the ATP of all customers (i.e. set $p^* + x = v_h$ for everyone) and maximize customer education (i.e. set $\theta = 1$).

**Assumption 1** *Assume that the distributor's budget is at most enough for both the customers' ATP and the customer education to reach their upper bounds, i.e.,* $b \leq c_\theta(1-\theta_0) + \lambda\beta(v_h - m_h)$.

The term $c_\theta(1-\theta_0) + \lambda\beta(v_h - m_h)$ above comes from the fact that when $\theta = 1$ and $p^* = m_h$ only the customers with a high valuation and high ATP buy and $\mathrm{B}(m_h - x, r^*, \theta) \leq \lambda\beta$. Relaxing Assumption 1 does not change our insights.

### 3.4. Sequence of events and model discussion

We model the distributor as a Stackelberg leader. The dynamics of the game are as follows:

1. Given $\theta_0$, $m_l$, $m_h$, $\lambda$, and $\alpha$ , as well as $w$, $y$, $\beta$, $v_h$, and $v_l$, the distributor solves the Allocation Problem and chooses its strategic investments. This results in $x$ and $\theta$.
2. Given $x$ and $\theta$, the distributor then chooses $c$ and $z$ by solving the Pricing Problem, anticipating the reaction from the retailer and customers;
3. The retailer chooses $p$ and $r$ after observing $x$, $\theta$, $c$, and $z$, with knowledge of $\beta$, $v_h$, $v_l$, $m_l$, $m_h$, $\lambda$, and $\alpha$. The retailer then sells the product if its expected profit is non-negative;
4. Customers $p$, $r$, $x$, and their individual information signal $S$. Each customer purchases the product if their WTP *and* ATP are both larger than the effective price $p - x$;
5. If a purchase occurs, the product is delivered to the retailer and picked by the customer;
6. An uninformed customer ($S = 0$) that makes a purchase learns their true valuation ($v_h$ or $v_l$) from using the product, and returns it if the valuation is less than $r$;
7. The retailer salvages the returns for $z$ per unit while distributor salvages them for $y$ per unit.

Our model blends elements from the operations and development economics literature (Shulman et al. 2009, Banerjee et al. 2012). Consistent with operations management literature (Su 2009, Shulman et al. 2009), full refunds are suboptimal in our model even if ATP is lower than WTP.

We examine the Pricing and Allocation Problems in three steps. First, in Section 4, we assume that all customers have the same ATP, i.e., $m_l = m_h$. This assumption is common in the development economics literature, e.g., Banerjee (1997). This simplified setup already provides insight into the effects of limited customer ATP on the customer welfare and on the distributor's investment strategy. Second, in Section 5, we extend Section 4's results to the case where $m_l < m_h$, and show that this feature creates an additional incentive for the retailer to skim the market. Finally, in Section 6, we replicate our main insights in a model with continuous customers ATP and WTP.

# 4. Pricing and Resource Allocation Problems with Homogeneous ATP

We now solve the Pricing Problem and the Allocation Problem when $m_h = m_l = m$, and examine how the distributor's optimal strategy critically depends on customers' ATP and risk aversion. Section 4.1 examines the Pricing Problem while Section 4.2 examines the Allocation Problem. Section 4.3 considers the setup where the distributor commits to offering free returns to customers, i.e., customers can purchase the product and return it for a full refund if their valuation is low.

## 4.1. Pricing Problem with Homogeneous ATP

We first introduce notation to characterize the Pricing Problem's optimal solution when $m_h = m_l = m$. Let $r_\alpha(p)$ denote the minimum retailer's refund for uninformed customers to purchase the product at price $p$. Thus, $r_\alpha(p)$ satisfies $\mathbb{E}[u(V, p, r_\alpha(p))] = 0$ and from Equation 1 we have

$$r_\alpha(p) = \left(p - \frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - p)}}{1-\beta}\right)\right)^+, \text{ where } (\cdot)^+ = \max(\cdot, 0). \tag{3}$$

Proposition B.1 in Appendix B.1 states the Pricing Problem's optimal solution and describes the distributor's two possible non-dominated pricing strategies, which we denote strategies $(a)$ and $(b)$. In strategy $(a)$, the distributor induces the retailer to target only informed customers without accepting product returns. Thus, for a discount $x$, the product's effective retail price is $p - x = m$, $r = 0$ and a fraction $\theta\beta$ of customers purchase the product. Conversely, strategy $(b)$ is more nuanced: the distributor induces the retailer to expand sales by accepting product returns and offering refunds if necessary. In this case, $p - x = m$ and $r = r_\alpha(m)$. Under strategy $(b)$, informed customers with high valuation as well as all uninformed customers purchase the product, i.e., a fraction $\theta\beta + (1-\theta)$ of customers buy the product. Uninformed customers with low valuation return the product, i.e., a fraction $(1-\theta)(1-\beta)$ of customers purchase and then return the product. The optimal outcomes of the Pricing Problem are similar to the outcomes in Shugan and Xie (2000).

The Pricing Problem's optimal solution provides a few insights into how the equilibrium price and refund interact with model parameters. First, the function $r_\alpha(p)$ sheds light into how the customers' risk aversion level affects the optimal refund offered by the retailer under strategy $(b)$. Namely, consider the effective price, $p - x = m$, perceived by customers in their utility function. From Equation 3, if customers are risk neutral, then offering partial refunds is sufficient for uninformed customers to purchase the product, i.e., $\lim_{\alpha\to 0} r_\alpha(m) = \left(\frac{m-\beta v_h}{1-\beta}\right)^+ < m$. Conversely, if customers are extremely risk averse (in the limit max-min utility optimizers) then the retailer must refund the full (post-discount) price to induce uninformed customers to buy, i.e., $\lim_{\alpha\to\infty} r_\alpha(m) = m$.

Second, as stated in Proposition B.1, the distributor only induces the retailer to target uninformed customers when the expected value from selling to an uninformed customer is higher than the expected cost of processing a return, independent of the consumer education level $\theta$.

Third, the dependence of the distributor's optimal pricing strategy on market parameters becomes intuitive once we focus on the marginal contribution of uninformed customers to the distributor's objective. Increasing salvage value, decreasing risk aversion, or decreasing price, all make it "easier" to target uninformed customers and increases the attractiveness of pricing strategy $(b)$.

Finally, the effect of the distributor's value for consumer surplus $\gamma$ on the optimal pricing strategy is less straightforward. On the one hand, as $\gamma$ increases, strategy $(b)$ might become more attractive relative to strategy $(a)$ since it delivers positive surplus to some uninformed customers. On the other hand, if the refund that the retailer offers to customers $r_\alpha(m)$ is too low, uninformed customers that purchase and return the product might end-up with negative surplus. Hence, depending on market and product parameters, increasing $\gamma$ might make strategy $(b)$ either more or less attractive.

The consumer surplus induced by both strategies $(a)$ or $(b)$ in Proposition B.1 decreases in the customers' ATP $m$ and is independent of the distributor's subsidy for consumers $x$. These effects happens because any increase in customer ATP is recaptured by the supply chain through a price increase — these features follow from customers' homogeneous ATP, the Stackelberg game structure of the model, and are standard in the development economics literature from where we borrow the ATP formulation, e.g. Banerjee (1997). When $m_l < m_h$, the consumer surplus might increase with the average ATP as we discuss in the next section.

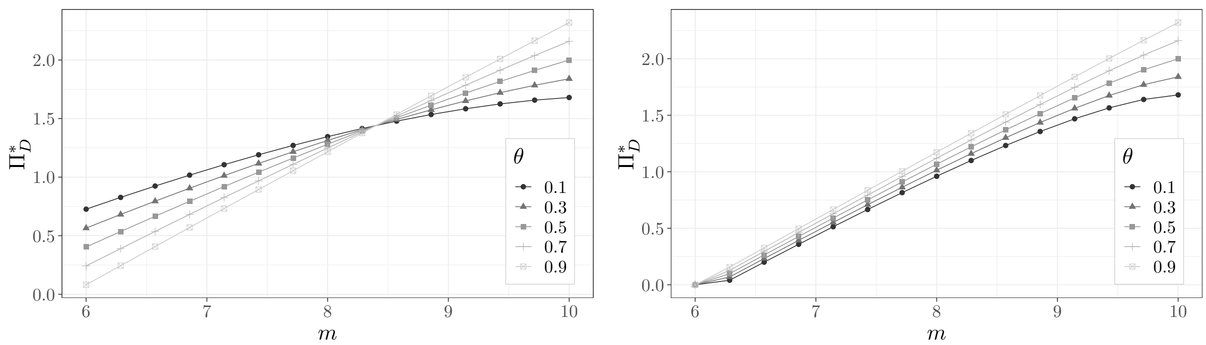
Proposition 1 provides comparative statics on the effects of $\theta$ and $x$ on the distributor's objective.

**Proposition 1** *Providing a subsidy $x \geq 0$ can only improve the distributor's objective value. Namely, let $\Pi_D^*$ be the distributor's optimal objective value, then $\frac{\partial \Pi_D^*}{\partial x} \geq 0$. In contrast, increasing the consumer education level $\theta$ might worsen the distributor's objective value. Specifically,*

$$\frac{\partial \Pi_D^*}{\partial \theta} \leq 0 \text{ if and only if } m + x - w - (r_\alpha(m) - y) - \gamma(m - r_\alpha(m)) \geq 0. \tag{4}$$

*Additionally, for the distributor, $x$ and $\theta$ are strategic complements under strategy $(a)$ in Proposition B.1, while they are strategic substitutes under strategy $(b)$ in Proposition B.1.*

Proposition 1 states that if customers' ATP is low, or if the product's salvage value is high, then increasing the consumer education level might actually be prejudicial to the distributor. Indeed, the left hand side of Equation 4 is the total contribution of a returned unit to the distributor's objective. Specifically, $m + x - w$ is the distributor's margin on each sale, $r_\alpha(m) - y$ is the distributor's cost for each returned product, and $\gamma(m - r_\alpha(m))$ is the cost perceived by the distributor due to the negative consumer surplus attained by customers who return the product. If the total contribution is positive, then returned products are valuable to the distributor and, as a result, increasing the consumer education level can reduce the distributor's objective.




(a) $\alpha = 0.1$ (b) $\alpha = 0.5$

**Figure 3** **Distributor's profits for different ATP $m$ and consumer education level $\theta$. We assume $v_h = \$10$, $v_l = 0$, $w = \$6$, $y = \$4$, $x = 0$, $\beta = 0.6$, and $\gamma = 0$.**

In particular, if the distributor's relative value for consumer surplus is moderate ($\gamma < 1$), a lower customer ATP $m$ results in the inequality in Equation 4 being violated by a larger subset of possible market and customer parameters. Thus, for a lower customer ATP, improving consumer education has a negative impact on the distributor's objective for a larger set of market and customer parameters. Figure 3(a) illustrates this result for a set of problem parameters — note how for $m > \$8.4$ increasing $\theta$ decreases the distributor's objective.

As the risk aversion parameter $\alpha$ increases, pricing strategy ($a$) becomes relatively more attractive since $r_\alpha(m)$ is increasing in $\alpha$. In strategy ($b$), as $\alpha$ increases, the marginal cost of returned products increases, and the set model parameters that violate the inequality in Equation 4 also increases. Hence, as $\alpha$ increases, increasing $\theta$ benefits the distributor for a larger set of market and customer parameters — contrasting Figure 3(a) and Figure 3(b) illustrates this observation.

Finally, Proposition 1 states that the interaction between improving consumer education and improving affordability is not the same under different pricing strategies. In strategy ($a$), they are complements. Namely, if the distributor only targets informed consumers, increasing $\theta$ increases the effective market size and the marginal value of $x$. Conversely, in strategy ($b$), they are substitutes. The cost of returned products drives this substitution effect. As $x$ increases, the cost associated to product returns decreases. Moreover, increasing $\theta$ reduces the volume of returns. Hence, if $\theta$ is large then less financing is needed to reduce the costs of returns, reducing the marginal value of $x$.

With the solution of the Pricing Problem in hand, we now examine the Allocation Problem.

### 4.2. Allocation Problem

In this section, we characterize the solution to the Allocation Problem, where the distributor allocates a budget between consumer education and increasing customers' ATP. We examine how ATP and risk aversion, as well as the distributor's relative value for consumer surplus affects the distributor's investment strategy and, ultimately, value creation in the BOP context. Theorem 1 below characterizes the four possible distributor's non-dominated investment allocation strategies: a finance-based strategy ($F$) where the distributor increases $x$; a marketing-based strategy ($M$) where the distributor increases $\theta$; a finance-based strategy including returns ($FR$) where the distributor increases $x$ and also offers customers a return option; and a marketing strategy including returns ($MR$) where the distributor increases $\theta$ and also offers customers a return option.

**Theorem 1** *The distributor has four non-dominated strategies in the Allocation Problem, which we denote by $F$, $M$, $FR$, and $MR$:*

- *$F$: invest the budget in improving product affordability by increasing $x$. If there is enough budget to set $x = v_h - m$, invest the remaining budget in increasing $\theta$. The distributor follows pricing strategy $(a)$ from Proposition B.1, targeting only informed customers.*
- *$M$: invest the budget in increasing the consumer education level $\theta$. If there is enough budget to set $\theta = 1$, invest the remaining budget in increasing $x$. The distributor follows pricing strategy $(a)$ from Proposition B.1, targeting only informed customers.*
- *$FR$: invest the budget in improving product affordability by increasing $x$. If there is enough budget to set $x = v_h - m$, the remaining budget is invested in increasing $\theta$ if $v_h - w - (r_\alpha(m) - y) - \gamma(m - r_\alpha(m)) < 0$, and not invested otherwise. The distributor follows pricing strategy $(b)$ from Proposition B.1, targeting all customers by accepting product returns.*
- *$MR$: invest the budget as in strategy $M$. However, the distributor follows pricing strategy $(b)$ from Proposition B.1, targeting all customers by accepting product returns.*

In practice, strategy $F$ could represent an investment by the distributor in discount coupons, vouchers or micro-financing for customers. Strategy $M$ could be an investment in marketing campaigns or village demonstrations. In strategies $FR$ and $MR$, the distributor complements the previous strategies by inducing the retailer to offer product returns. Offering returns, or even free trials, are relatively novel strategies to profitably distribute life-improving products in the BOP and are being used by companies like Essmart and a select number of other distributors.[8]

Strategy $FR$ is more nuanced than the other strategies. In $FR$, the distributor allocates as much budget as possible to increasing $x$. If the distributor has enough budget to set customers' ATP to $v_h$, i.e. $x = v_h - m$, then it allocates the remaining budget in one of two possible ways: if Equation 4 does not hold and, as a result, increasing $\theta$ improves the distributor's objective, then the remaining budget is allocated to increasing $\theta$; if Equation 4 holds, then the remaining budget is not invested. If we relax the constraint $x \leq v_h - m$, i.e. if we allow customers' ATP to be larger than $v_h$, then the structure of strategy $FR$ simplifies, in that all the budget is spent on $x$. However, this does not change our main results and insights, which are detailed next.

[8] For example, Burro (a distributor in Ghana), EcoZoom (Kenya), and Pollinate Energy (India).

Propositions B.2, B.3 and B.4 in the Appendix fully characterize the optimal investment strategy map for the Allocation Problem as a function of the two parameters that define customer behavior in the BOP: the (homogeneous) ATP $m$ and the risk aversion parameter $\alpha$, see Figure 4 for an illustration. The propositions specify necessary and sufficient conditions for each of the four non-dominated allocation strategies in Theorem 1 to be optimal. To facilitate discussion, we introduce a corollary to summarize the propositions' main insights.

To simplify the notation, we define $\bar{w} = \beta v_h + (1-\beta)y$ and $\bar{c}_\theta = (1-\beta)(w-y)$. We provide an interpretation of $\bar{w}$ and $\bar{c}_\theta$ in the discussion of the following corollary.

**Corollary 1** *Recall the allocation strategies from Theorem 1. Then, if $w \leq \bar{w}$ and $c_\theta \leq \bar{c}_\theta$, there exist ATPs $m^F$, $m^{MR}$, and $\bar{m}^{MR}$ where $m^F \leq m^{MR}, \bar{m}^{MR} \leq v_h$, and a function $a_F(m)$ such that:*

- *Strategy $F$ is optimal if and only if $m \leq m^F$ and $\alpha \geq a_F(m)$.*
- *Strategy $MR$ is optimal if $m \geq \max(m^{MR}, \bar{m}^{MR})$ for all $\alpha \geq 0$.*

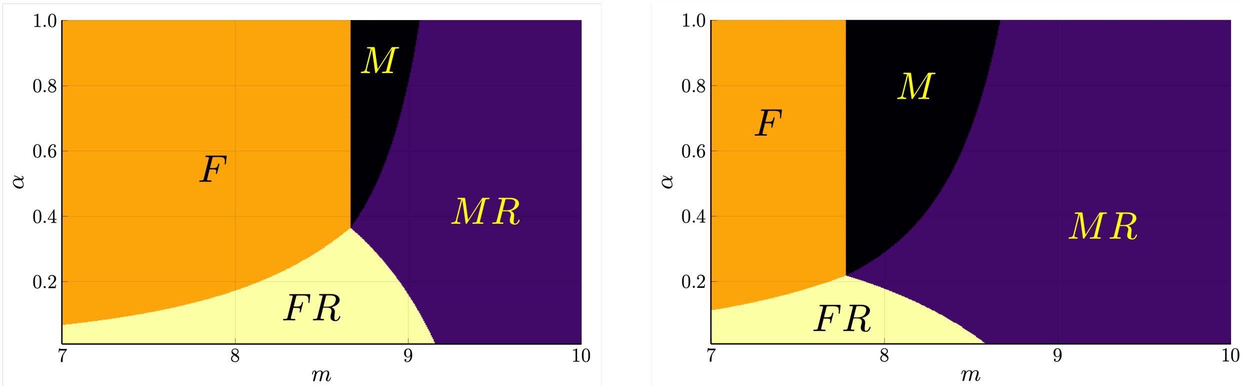
*If $w > \bar{w}$ then only strategies $F$ and $M$ can be optimal. In this case, there exists an ATP $m^F$ such that $F$ is optimal if and only if $m \leq m^F$.*

*If $c_\theta > \bar{c}_\theta$ then only strategies $F$ and $FR$ can be optimal. In this case, there exists a function $a_F(m)$ such that $F$ is optimal if and only if $\alpha \geq a_F(m)$.*

Corollary 1 differentiates between two types of products: Products that, in expectation, can be profitably distributed to uninformed customers, i.e., products where $\beta(v_h - w) + (1-\beta)(y-w) \geq 0$ (equivalently, where $w \leq \bar{w}$) and products that cannot be profitably distributed to uninformed customers (where $w > \bar{w}$). When products can be profitably distributed to uninformed customers, all four strategies can be optimal. When products cannot be profitably distributed to uninformed customers, only strategies $F$ and $M$ can be optimal.

Similarly, Corollary 1 differentiates between two possible setups. The first are setups where the cost of informing consumers about the product is not excessively large, i.e., setups where $c_\theta \leq (1-\beta)(w-y) = \bar{c}_\theta$. The second are setups where the cost of informing consumers about the product is larger than the expected cost incurred by letting uninformed consumers simply try the product out, i.e., such that $c_\theta > (1-\beta)(w-y) = \bar{c}_\theta$. When the cost of informing consumers about the product is moderate, all four strategies can be optimal. When the cost of informing consumers about the product is excessively large, then only strategies $F$ and $FR$ can be optimal.

Figure 4 depicts the strategy map for a set of problem parameters and illustrates Proposition B.2 and Corollary 1. If customers' ATP is low, then it is optimal for the distributor to invest in offering customers a discount $x$ to improve affordability (strategies $F$ and $FR$ in Figure 4). Conversely, if customers' ATP is high, the distributor should invest in customer education $\theta$ (strategies $M$ and $MR$ in Figure 4). Additionally, for any ATP, if customers' risk aversion level is low enough, then




(a) $\gamma = 0$ (b) $\gamma = 0.4$

**Figure 4** **Strategy map for the Allocation Problem as a function of $m$ and $\alpha$ for different values of $\gamma$. We assume $v_h = \$10$, $w = \$7$, $\theta = 0.2$, $b = \$0.5$, $y = \$3.25$, $\beta = 0.6$, $c_\theta = 1$.**

it is optimal for the distributor to induce the retailer to target uninformed customers, by accepting product returns and possibly offering refunds (note strategies $FR$ and $MR$ in Figure 4).

This relationship between customers' risk aversion and strategy choice occurs because, when customer risk aversion is high, the refund necessary to attract uninformed customers is overly large, making this customer segment unprofitable due to the high cost of returns. In contrast, when risk aversion is moderate or low, uninformed customers might be valuable if the distributor's budget is sufficient to increase customers' ATP above a certain threshold.

In particular, Corollary 1 and Figure 4 show that the most profitable investment strategy in a BOP context (where $\alpha$ is high and $m$ is low) might not be the same as in the non-BOP context (where $\alpha$ is low and $m$ is high). For example, in Figure 4 strategy $F$ is the most profitable in a BOP context while $MR$ is the most profitable in a non-BOP context. A natural question that follows is whether a similar difference occurs if the distributor were to maximize consumer surplus instead of profits. This question is answered in the next proposition. We denote the consumer surplus under strategies $F$, $M$, $FR$, and $MR$ by $CS^F$, $CS^M$, $CS^{FR}$, and $CS^{MR}$, respectively.

**Proposition 2** *Recall the allocation strategies from Theorem 1, and the ATP $m^F$ and the function $a_F(m)$ from Corollary 1. There exists a function $a_F^{cs}(m)$ such that, for any customers' risk aversion level $\alpha \geq 0$ and customers' ATP m $w \leq m \leq v_h$,*

- *Strategy $MR$ induces the largest consumer surplus among all strategies from Theorem 1. Moreover, $CS^{MR} \geq CS^M \geq CS^F$;*
- *Strategy $F$ induces the smallest consumer surplus among all strategies from Theorem 1 if and only if $\alpha \geq a_F^{cs}(m)$.*

The intuition behind Proposition 2 is as follows. On the one hand, strategies that involve accepting product returns ($FR$ and $MR$) can induce a higher total consumer surplus than their counterpart strategies that do not accept product returns ($F$ and $M$) since, when the distributor offers a product return option, all consumers with high valuation adopt the product and extract positive surplus. On the other hand, customers with a low valuation that buy the product (because of lack of information) will have negative surplus, since they receive a partial refund that is smaller than their payment, thus reducing total consumer surplus.

Furthermore, strategies that invest primarily in consumer education ($M$ and $MR$) can lead to higher total consumer surplus than finance-based strategies ($F$ and $FR$). This holds because the subsidies offered by the distributor to customers in strategies $F$ and $FR$ are ultimately recaptured by the supply chain through higher prices and, as a result, are ineffective in increasing consumer surplus. Conversely, the information provided through consumer education in strategies $M$ and $MR$ cannot be fully recaptured by the supply chain and hence increases consumer surplus.

Proposition 2 shows that strategy $MR$ attains the largest consumer surplus, i.e., the distributor increases the total consumer surplus by inducing the retailer to accept product returns and by educating consumers about the product's benefits. In particular, Proposition 2 implies that as the distributor's relative value for consumer surplus $\gamma$ increases, then strategy $MR$ becomes optimal for larger set of all admissible model parameters, while strategy $F$ becomes optimal for a smaller set. This insight is illustrated in Figure 4b — note how strategy $MR$ occupies a larger portion of the strategy map while strategy $F$ a smaller portion compared to Figure 4a — and is particularly relevant for social enterprises operating in the BOP context.

The combination of Proposition 2 and Corollary 1 indicates that there is tension between maximizing profits and consumer surplus in the BOP, which does not necessarily occur in non-BOP contexts. Namely, in a BOP context with high risk aversion ($\alpha \geq \max(a_F(m), a_F^{cs}(m))$) and small ATP ($m \leq m^F$), strategy $F$ is the most profitable. However, in this regime the profit maximizing allocation strategy $F$ also leads to the *lowest* total consumer surplus. Hence, *in the BOP the most profitable allocation strategy and the total consumer surplus maximizing strategy are not the same.* Conversely, in a non-BOP context with large ATP ($m \geq \max(m^{MR}, \bar{m}^{MR})$) the most profitable strategy is $MR$, which also leads to the largest consumer surplus.

We emphasize that this insight applies to the simplified setups in Corollary 1 where only two strategies can be optimal. Both simplified setups in Corollary 1 include strategy $F$ which, from Proposition 2, induces the lowest consumer surplus in the BOP. Thus the tension between maximizing profits and consumer surplus in the BOP is preserved in these setups.

In the next subsection we explore a practical strategy that the distributor can commit to when operating in the BOP, which resolves the tension between profits and consumer surplus.

### 4.3. Decoupling profits and consumer surplus by offering free returns

Corollary 1 and Proposition 2 show that, in the BOP context, the most profitable allocation strategy (strategy $F$) is the strategy that leads to lowest consumer surplus. Furthermore, even the strategy that leads to the highest consumer surplus (strategy $MR$) has the downside that uninformed customers with low valuation will have negative surplus if they purchase and then return the product. To resolve these issues we examine a *free returns* strategy (which is equivalent to offering free product trials in our model). Specifically, we assume that the distributor commits to offering customers the option of returning the product for a full refund. This commitment occurs at the start of the game, before the allocation of the investment budget. We remark that there is recent field evidence that free trials are an effective strategy for increasing adoption of life-improving products, such as cookstoves (Levine et al. 2018).

Although a free-returns commitment can only reduce the distributor's maximum profits, it has two main advantages. First, it guarantees that no customer ends up with negative surplus by trying out the product. Hence, in effect, free returns transfer the product's "fit risk" from the customer to the distributor (Che 1996). Second, we will show that the distributor's commitment to free returns ensures that the total consumer surplus is maximized independently of the allocation strategy. In other words, this commitment decouples the effects of distributor's profit-maximizing actions from worst-case consumer surplus. Thus, offering free returns can serve as a signal from the distributor to its social investors and other stakeholders that it is indeed valuing consumer surplus, satisfying an important ethical consideration when doing business in the BOP (Davidson 2009).

In this setting, we assume a wholesale contract that forces the retailer to set $r = p - x$. Recall that $p - x$ is the effective customer price after a distributor's subsidy $x$. The model details and the resulting distributor's pricing strategy are given in Proposition B.5 in the Appendix.

Proposition B.5 shows that, when the distributor commits to full refunds, the total consumer surplus is constant and equal to $(v_h - m)\beta$, independently of the values of $x$ and $\theta$. We note that this consumer surplus value is larger than the one induced by strategy $MR$ in the allocation problem and, from Proposition 2, is larger than any non-dominated distributor's budget allocation strategy.

The Allocation Problem under a free-returns commitment has only two non-dominated strategies: $FR$ and $MR$. The next proposition describes the optimal allocation strategy. The same outcomes for each strategy are achieved by assuming that customers are extremely risk averse, i.e., taking the limit as $\alpha \to \infty$ making customers max-min utility optimizers.

**Proposition 3** *Assume that the distributor commits to a free returns policy, then the distributor has two non-dominated strategies for the Allocation Problem, which we denote by $FR^f$, and $MR^f$. In both strategies the distributor follows the pricing strategy from Proposition B.5.*

- $FR^f$: *invest the budget in improving product affordability by increasing* $x$. *If there is enough budget to set* $x = v_h - m$, *the remaining budget is invested in increasing* $\theta$ *if* $v_h - m < w - y$ *and not invested otherwise.*
- $MR^f$: *invest the budget in increasing the consumer education level. If there is enough budget to set* $\theta = 1$, *the remaining budget is invested in increasing* $x$.

*Moreover, strategy* $MR^f$ *is preferred by the distributor if and only if* $(1-\beta)(w-y) \geq c_\theta$.

The allocation strategies in Proposition 3 follow the same intuition as the strategies in Theorem 1. However, in the free-returns case, the condition for choosing between investing in $\theta$ and $x$ is simpler: if the expected cost of a returned product, given by $(1-\beta)(w-y)$, is higher than the marginal cost of investing in consumer education, given by $c_\theta$, then the distributor should choose strategy $MR^f$. Conversely, if $(1-\beta)(w-y) < c_\theta$ the distributor chooses strategy $FR^f$.

## 5. Pricing and Resource Allocation Problems with Heterogeneous ATP

This section extends the previous results to the case where $m_h > m_l$. In this case, the retailer and distributor have the option to "skim" the market and choose a price such that only customers with high ATP $m_h$ are able to afford the product. While the results from the previous section still hold in this setup, the supply chain equilibrium is more nuanced due to price skimming. Specifically, more non-dominated strategies for the pricing and allocation problem exist.

When $m_h > m_l$, there are four possible non-dominated pricing strategies for the distributor:

- $A$: Target all informed customers without product returns. The equilibrium price is $p^A = m_l + x$ and $r^A = 0$. All customers that receive information signal $S = 1$ purchase the product.
- $AR$: Target all customers and offer product returns, $p^{AR} = m_l + x$ and $r^{AR} = r_\alpha(m_l)$. All customers purchase the product and a fraction $1-\beta$ (with valuation $v_l$) return it for a refund.
- $S$: Skim the market and target informed customers with ATP $m_h$ without product returns. The equilibrium price is $p^S = m_h + x$ and $r^S = 0$. The fraction $\lambda\theta$ of customers that have valuation $m_h$ and receive information signal $S = 1$ purchase the product.
- $SR$: Skim the market targeting customers with ATP $m_h$ and offer product returns, $p^{SR} = m_h + x$ and $r^{SR} = r_\alpha(m_h)$. The fraction $\lambda$ of customers with ATP $m_h$ purchase the product and a fraction $1-\beta$ of those who purchased with valuation $v_l$ return it for a refund.

Proposition C.1 in the Appendix fully characterizes these equilibrium pricing strategies, including the profits of the distributor and retailer. Strategies $S$ and $SR$ are "skimming" strategies where customers with low ATP $m_l$ cannot afford the product.

In the remainder of our analysis we make the following two assumptions that allow for an analytical treatment of the Allocation Problem with heterogeneous ATP.

**Assumption 2** *Assume that:*

1. *When there is no risk aversion ($\alpha = 0$), the purchasing decision of uninformed customers with high ATP $m_h$ only depends on their expected valuation, i.e., $m_h > \beta v_h$;*
2. *For any risk aversion level, offering a refund of $r_\alpha(m_l)$ to all customers is more expensive than offering a refund $r_\alpha(m_h)$ only to customers with high ATP $m_h$, i.e., $r_\alpha(m_l) \geq \lambda r_\alpha(m_h)$ for all $\alpha$. Since $\lim_{\alpha\to 0} r_\alpha(p) = \left(\frac{p-\beta v_h}{1-\beta}\right)^+$, this assumption is equivalent to $(m_l - \beta v_h)^+ \geq \lambda(m_h - \beta v_h)^+$.*

The first part of the assumption above states that when customers are risk neutral at least some of the uninformed customers — those with high ATP $m_h$ — base their purchasing decision on their expected valuation instead of their ATP. When this condition is not met, risk neutral uninformed customers with high ATP $m_h$ always purchase the product as long as the price is lower than $m_h$.

The second part of the assumption is consistent with the context in which BOP distributors operate. The life-improving products these distributors sell are designed for low-income customers with limited ATP such that, in practice, $\lambda$ is small and $m_h$ is close to $m_l$. Thus, the refund cost in strategy $AR$ is likely higher than in strategy $SR$. This assumption also simplifies the description of the distributor's strategy map as a function of $\alpha$, $m_h$, and $m_l$ since it eliminates the edge case where small changes in the values of $\alpha$, $m_h$ and $m_l$ can make the distributor's optimal pricing strategy shift directly from $A$ to $SR$, without first becoming $AR$ or $S$.

With Assumption 2 in hand, we extend the result from Proposition 1 to heterogeneous ATPs.

**Proposition 4** *Under Assumption 2, providing a subsidy $x \geq 0$ can only improve the distributor's objective value, i.e., $\frac{\partial \Pi_D^*}{\partial x} \geq 0$. In contrast, increasing the consumer education level $\theta$ might worsen the distributor's objective value, i.e., $\frac{\partial \Pi_D^*}{\partial \theta} \leq 0$ if*

$$m_h + x - w - (r_\alpha(m_h) - y) - \gamma(m_h - r_\alpha(m_h)) \geq 0. \tag{5}$$

*Additionally, for the distributor, $x$ and $\theta$ are strategic complements under strategies $A$ and $S$, while they are strategic substitutes under strategies $AR$ and $SR$.*

Proposition 4 states a sufficient condition for an increase in information availability to reduce the distributor's objective. The condition is analogous to the one in Proposition 1: when the contribution of a returned unit priced at $m_h$ to the distributor's objective is high enough, increasing information level's might be prejudicial for the distributor. Proposition 4 also extends Proposition 1 by stating that strategies $x$ and $\theta$ are strategic complements under pricing strategies that do not offer refunds ($A$ and $S$), while being substitutes when refunds are offered ($AR$ and $SR$).

### 5.1. Allocation Problem and Free Returns with Heterogeneous ATP

When $m_l < m_h$, there are eight possible non-dominated allocation strategies in the Allocation Problem which extend the strategies in Theorem 1 and can be split into two sets. In the first set of strategies, the distributor uses pricing strategies $A$ or $AR$, which target all customers. We denote these allocation strategies by $FA$, $MA$, $FAR$, and $MAR$ — the notation mirrors $F, M, FR$, and $MR$ from Theorem 1. In the second set of strategies, the distributor uses pricing strategies $S$ or $SR$, which skim the market and target only customers with high ATP $m_h$. We denote these allocation strategies by $FS$, $MS$, $FSR$, and $MSR$.

**Theorem 2** *Under Assumption 2, the distributor has eight non-dominated strategies in the Allocation Problem:*

- $FS$*: invest the budget in improving product affordability by increasing* $x$*. If there is enough budget to set* $x = v_h - m_h$*, invest the remaining budget in increasing* $\theta$*. The distributor follows pricing strategy* $S$*, targeting only informed customers with high ATP.*
- $MS$*: invest the budget in increasing the consumer education level* $\theta$*. If there is enough budget to set* $\theta = 1$*, invest the remaining budget in increasing* $x$*. The distributor follows pricing strategy* $S$*, targeting only informed customers with high ATP.*
- $FSR$*: invest the budget in improving product affordability by increasing* $x$*. If there is enough budget to set* $x = v_h - m_h$*, the remaining budget is invested in increasing* $\theta$ *if* $v_h - w - (r_\alpha(m_h) - y) - \gamma(m_h - r_\alpha(m_h)) < 0$*, and not invested otherwise. The distributor follows pricing strategy* $SR$*, targeting all customers with high ATP by accepting product returns.*
- $MSR$*: invest the budget as in strategy* $MS$*. However, the distributor follows pricing strategy* $SR$*, targeting all customers with high ATP by accepting product returns.*
- $FA$*: invest the budget in improving product affordability by increasing* $x$*. If there is enough budget to set* $x = v_h - m_l$*, the remaining budget is invested in increasing* $\theta$ *if* $m_l - (1-\gamma)(v_h - m_l) - \lambda/(1-\lambda)(m_h - m_l) - w > 0$*, and not invested otherwise. The distributor follows pricing strategy* $A$*, targeting all informed customers.*
- $MA$*: invest the budget in increasing the consumer education level* $\theta$*. If there is enough budget to set* $\theta = 1$*, invest the remaining budget in increasing* $x$*. The distributor follows pricing strategy* $A$*, targeting all informed customers.*
- $FAR$*: invest the budget in improving product affordability by increasing* $x$*. If there is enough budget to set* $x = v_h - m_l$*, the remaining budget is invested in increasing* $\theta$ *if* $v_h - w - (r_\alpha(m_l) - y) - \gamma(m_l - r_\alpha(m_l)) < 0$*, and not invested otherwise. The distributor follows pricing strategy* $AR$*, targeting all customers by accepting product returns.*
- $MAR$*: invest the budget as in strategy* $MA$*. However, the distributor follows pricing strategy* $AR$*, targeting all customers by accepting product returns.*

If the distributor "constrains" itself to follow either only skimming allocation strategies ($FS$, $MS$, $FSR$ or $MSR$) or non-skimming allocation strategies ($FA$, $MA$, $FAR$ or $MAR$) then the restricted strategy map as a function of $m_h$ and $\alpha$ (alternatively, $m_l$ and $\alpha$) has an analogous structure to the homogeneous ATP case where $m_h = m_l$.

We follow the analysis for the homogeneous ATP case and we define $\bar{w}^S = \beta v_h + (1-\beta)y$, $\bar{w}^A = \beta v_h + (1-\beta)y - \lambda/(1-\lambda)\left((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)\right)$, $\bar{c}_\theta^S = (1-\beta)\lambda(w-y)$, and $\bar{c}_\theta^A = (1-\beta)(w-y) - \lambda/(1-\lambda)(1-\beta)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))$.

The strategy maps for the case where the distributor constrains itself to either skimming allocation strategies or non-skimming strategies are derived in Proposition C.2 in the Appendix. Proposition C.2 assumes $w \le \bar{w}^S$ and $c_\theta \le \bar{c}_\theta^S$ (resp. $w \le \bar{w}^A$ and $c_\theta \le \bar{c}_\theta^A$), when the distributor constrains itself to skimming (resp. non-skimming) allocation strategies. Analogous to Propositions B.3 and B.4, fewer strategies can be optimal when $w > \bar{w}^S$ or $c_\theta > \bar{c}_\theta^S$ (resp. $w > \bar{w}^A$ or $c_\theta > \bar{c}_\theta^A$) and we omit the analysis of these cases since they are a special case of the analysis for Proposition C.2. Figure 7 in the Appendix illustrates Proposition C.2 by depicting the "constrained" strategy maps that underlie the strategy map of Figure 5.

When $w \le \min(\bar{w}^S, \bar{w}^A)$ and $c_\theta \le \min(\bar{c}_\theta^S, \bar{c}_\theta^A)$ the distributor's full strategy map is a combination of the constrained strategy maps from Proposition C.2, and shares many similar features with the strategy map when $m_h = m_l$, recall Proposition B.2 and Figure 4. Namely, if $\alpha$ is high and $m_h$ and $m_l$ are low, then either strategy $FA$ or strategy $FS$ dominate. Conversely, if $\alpha$ is low and $m_h$ and $m_l$ are high, then either strategy $MAR$ or $MSR$ dominate.

A sample strategy map with heterogeneous ATP is depicted in Figure 5. The horizontal axis of Figure 5 increases both $m_l$ and $m_h$ by setting the ATPs to $m_l + \Delta m$ and $m_h + \Delta m$. Note the similar disposition of the strategies in Figure 5 compared to the strategy map in Figure 4: an increase in $\alpha$ can induce the distributor to choose strategies without refunds, while an increase in $\Delta m$ induces the distributor to invest in increasing information instead of subsidies.

We now build on Proposition C.2 to extend Corollary 1 to the case where $m_l < m_h$ and $\gamma$ is small, i.e., the distributor prioritizes profits (if $\gamma = 0$ the distributor is profit-maximizing).

**Corollary 2** *Under Assumption 2, if* $w \le \min(\bar{w}^S, \bar{w}^A)$, $c_\theta \le \min(\bar{c}_\theta^S, \bar{c}_\theta^A)$ *and* $\gamma c_\theta \le (\lambda/(1-\lambda))^2 \beta(m_h - m_l)$, *then there exists ATPs* $m^{FAS}$, $m^{FS}$, $\hat{m}^{MAR}$, *and* $\hat{m}^{MSR}$ *where* $m^{FS}, m^{FAS} \le \hat{m}^{MAR}, \hat{m}^{MSR} \le v_h$, *and functions* $a_{FS}(m_l, m_h)$ *and* $a_{MAR}(m_l)$ *such that:*

- *Strategy FS is optimal if and only if* $m_l \le m^{FAS}$, $m_h \le m^{FS}$ *and* $\alpha \ge a_{FS}(m_l, m_h)$.
- *Strategy MAR is optimal if* $m_l \ge \hat{m}^{MAR}$, $m_h \ge \hat{m}^{MSR}$ *and* $\alpha \le a_{MAR}(m_l)$.

Corollary 2 strengthens the intuition on how the optimal investment and pricing strategy in a BOP context (where $m_l$ and $m_h$ are low and $\alpha$ is high) might differ from a non-BOP context (where

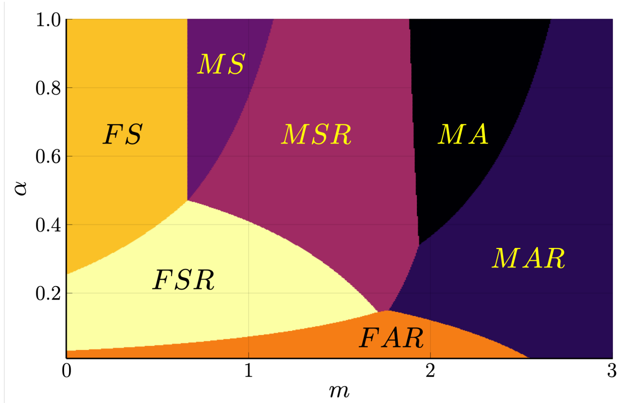



**Figure 5** **Strategy map for the Allocation Problem as a function of $\Delta m$ and $\alpha$. We assume $v_h = \$10$, $w = \$6$, $\theta_0 = 0.2$, $m_h = \$7$, $m_l = \$6$, $\lambda = 0.4$, $b = \$0.2$, $y = \$3.25$, $\beta = 0.5$, $c_\theta = 0.33$, and $\gamma = 0$. The horizontal axis sets the ATPs to $m_l + \Delta m$ and $m_h + \Delta m_h$. Strategy $FA$ is dominated by other strategies and does not appear.**

$m_l$ and $m_h$ are higher and $\alpha$ is low). In the BOP context, a profit-maximizing distributor invests in improving affordability and skims the market (strategy $FS$). Conversely, in a non-BOP context, a profit-maximizing distributor invests in increasing customer education level, offers returns, and prices the product to target all customers (strategy $MAR$). Thus, in the BOP context, the heterogeneous ATP of customers might create an additional incentive for the distributor to skim the market which does not occur in the non-BOP context.

We now characterize the consumer surplus attained in a BOP and non-BOP setting.

**Proposition 5** *Under Assumption 2, there exists a function $a^{cs}_{FS}(m_h, m_l)$ such that, for any customers' risk aversion level $\alpha \geq 0$ and customers' ATPs $m_l$, $m_h$, $w \leq m_l \leq m_h \leq v_h$,*

- *Strategy $MAR$ induces the largest consumer surplus among all strategies from Theorem 2;*
- *Strategy $FS$ induces the smallest consumer surplus among all strategies from Theorem 2 if $\alpha \geq a^{cs}_{FS}(m_h, m_l)$ and strategy $FA$ is individually rational for the distributor.*

Proposition 5 highlights that the misalignment between consumer surplus and the distributor's profits in the BOP persists under heterogeneous ATPs. Given the market characteristics in the BOP, where customers' ATP is low and risk aversion is high, the most profitable allocation strategy is $FS$, where the distributor must skim the market and invest in increasing $x$ to have a positive objective function. This strategy is also the strategy that leads to the lowest consumer surplus if $\alpha$ is high enough since only a small fraction of the market purchases the product (the market size is $\lambda\beta\theta^{FS}$, where $\theta^{FS}$ is the information level under strategy $FS$) and receive surplus $v_h - m_h$.

The total misalignment between the customer-surplus-maximizing strategy and the distributor's profit-maximizing strategy does not occur in a non-BOP context. From Corollary 2, when $\alpha$ is low and customer ATPs are high then strategy $MAR$, which from Proposition 5 induces the largest consumer surplus, is the most profitable (despite the fact that some customers that return the product might have negative surplus). Thus, Corollary 2 and Proposition 5 strengthen and extend the main insight from Section 4.

Corollary 2 and Proposition 5 further indicate that, unlike the case where $m_h = m_l$, consumer surplus can increase as ATP increases. If both $m_h$ and $m_l$ increase, the distributor can shift from a skimming pricing strategy to a pricing strategy that targets all customers (for example from $FS$ to $MAR$) which can lead to a price reduction for customers and a net increase in consumer surplus.

Finally, we revisit Section 4.3 and examine the value of an operational commitment by the distributor to offer free returns in resolving the misalignment between customer surplus and distributor profits in a BOP setting with heterogeneous ATPs. Recall that free returns eliminate the downside of pricing strategies $AR$ and $SR$ where customers that purchase and return the product for a partial refund have negative consumer surplus. The proposition below extends Proposition 3.

**Proposition 6** *Assume that the distributor commits to a free-returns policy, then the distributor has four non-dominated strategies in the Allocation Problem:*

- $FSRf$*: invest the budget in improving product affordability by increasing* $x$*. If there is enough budget to set* $x = v_h - m_h$*, the remaining budget is invested in increasing* $\theta$ *if* $v_h - m_h < w - y$*, and not invested otherwise. The distributor follows pricing strategy* $FSR$*.*
- $MSRf$*: invest the budget in increasing the consumer education level. If there is enough budget to set* $\theta = 1$*, the remaining budget increases* $x$*. The distributor follows pricing strategy* $MSR$*.*
- $FARf$*: invest the budget in improving product affordability by increasing* $x$*. If there is enough budget to set* $x = v_h - m_l$*, the remaining budget is invested in increasing* $\theta$ *if* $v_h - m_l < w - y$*, and not invested otherwise. The distributor follows pricing strategy* $FAR$*.*
- $MARf$*: invest the budget as in strategy* $MSRf$*. The distributor follows pricing strategy* $MAR$*.*

*Moreover, the consumer surplus induced by strategies* $FSRf$ *and* $MSRf$ *is* $\lambda\beta(v_h - m_h)$*, which is larger than the consumer surplus induced by the allocation strategies* $FS$*,* $FSR$*,* $MS$ *and* $MSR$ *from Theorem 2. The consumer surplus induced by strategies* $FARf$ *and* $MARf$ *is* $\beta(v_h - m_l)$*, which is larger than the consumer surplus induced by any allocation strategy from Theorem 2.*

Thus, a commitment to free returns guarantees that no customers that purchase the product have negative surplus in the BOP. Free returns also eliminate strategies $FS$ and $FA$, which induce a low consumer surplus. However, free returns do not remove the incentive for the distributor and retailer to skim the market and target only customers with high $ATP$.

To mitigate price skimming, the distributor can commit to a maximum retail price (MRP). A MRP is required in a few developing countries such as India — the Indian parliament passed the Consumer Goods Act in 2006 (The Indian Express 2017), which requires consumer goods distributed and sold in India to have a MRP printed on good's package. A distributor can use such policy to print a price on the box that mitigates skimming, ensuring that low ATP customers can afford the product. Such MRP commitment coupled with free returns can ensure that the optimal strategies for the distributor are either $FARf$ or $MARf$, and that there is no misalignment between consumer surplus and the optimal allocation and pricing strategies.

## 6. Robustness Check: Model with Continuous Customer Types

We now illustrate that our main results and insights hold for a more sophisticated customer model with continuous types. This model is analytically intractable and we study its equilibrium outcomes through numerical simulations. We briefly describe each component of the continuous model.

There is a continuum of customer types defined by their product valuation. We assume valuations are normally distributed as $V \sim N(\mu, \sigma^2)$. As in the discrete model from Section 3, without perfect information customers only learn their valuation if they purchase the product. Customers' ATP are also normally distributed as $M \sim N(\mu_m, \sigma_m^2)$. Customers' ATP and valuations can be correlated, where $\kappa$ is the correlation between $V$ and $M$. While we could assume other distributions of $V$ and $M$, a normal distribution allows for a tractable model of information disclosure. For a discussion on the microfoundations of information disclosure models with normally-distributed customer valuations see Chu and Zhang (2011) and Johnson and Myatt (2006).

Customers do not know their true valuation before purchasing the product and form an estimate of the product's utility. Customers of type $v$ receive a noisy signal $v + \eta$, where $\eta \sim N(0, \sigma_\eta^2)$ is random noise that models the quality of the information customers have about the product. The distribution of signals is $S = V + \eta$, and we assume $V$ and $\eta$ to be independent. Customers form a valuation estimate which has the distribution of the random variable $\widehat{V} = E[V|S]$. Then, as in Chu and Zhang (2011), it follows that

$$\widehat{V} = E[V|S] = \rho^2 S + (1 - \rho^2)\mu,$$

where $\rho^2 = \frac{\sigma^2}{\sigma^2 + \sigma_\eta^2} \in [0, 1]$ is the correlation coefficient of $V$ and $S$, and models the quality of the information a customer receives. Note that $\widehat{V}$ is normally distributed and $\widehat{V} \sim N(\mu, \rho^2\sigma^2)$. Moreover, $V - \widehat{V} \sim N(0, (1 - \rho^2)\sigma^2)$. Hence, if $\rho = 0$ customers have no information about their own valuation, while if $\rho = 1$ they have perfect information.

All customers receive a baseline "low-quality" information signal $S_l$ where the correlation coefficient of $V$ and $S_l$ is $\rho_l^2$. The distributor can then invest in customer education campaigns that provide a "high-quality" information signal $S_h$ with correlation coefficient $\rho_h^2$ (where $\rho_h \geq \rho_l$) to a fraction of the potential customers. The probability of a customer being reached by a customer education campaign is $\theta$. Thus, a customer receives signal $S_h$ with probability $\theta$ and signal $S_l$ with probability $1-\theta$. In short, as in the model of Section 3.1, $\theta$ models the campaign's *reach*, while $\rho_l$ and $\rho_h$ model the *quality* of the education campaign. We emphasize that allowing $0<\rho_l<\rho_h<1$, generalizes the consumer education model from Section 3.1.

As in the discrete-valuation model, customers are risk averse and have the utility function described in Section 3.1. We denote the utility estimate of a customer chosen at random by $U$. Note that the expected utility and ATP are correlated if $V$ and $M$ are correlated. Given $p$, $r$, and $\theta$, the fractions of customers that purchase and return the product are

$$\mathrm{B}(p,r,\theta,\rho_h,\rho_l) := E[\mathbb{1}_{\{U\geq 0, M\geq p\}}|\theta,\rho_h,\rho_l] \qquad \text{(Prob. of buying the product)},$$
$$\mathrm{R}(p,r,\theta,\rho_h,\rho_l) := E[\mathbb{1}_{\{U\geq 0, M\geq p, V<r\}}|\theta,\rho_h,\rho_l] \qquad \text{(Prob. of returning the product)},$$

and the consumer surplus is

$$CS(p,r,\theta,\rho_h,\rho_l) := E[(V-p)\mathbb{1}_{\{U\geq 0, M\geq p, V\geq r\}}|\theta,\rho_h,\rho_l] - (p-r)E[\mathbb{1}_{\{U\geq 0, M\geq p, V<r\}}|\theta,\rho_h,\rho_l]. \qquad (6)$$

The distributor's Pricing Problem in the continuous model is the same as in Section 3. The Allocation Problem in the continuous model is nearly identical to the problem stated in Section 3. We again interpret $x$ as an investment in product affordability, while an investment in increasing $\theta$ represents an increase in the fraction of customers with a high-quality information signal. However, we assume an upper bound $x\leq\bar{x}$ where $\bar{x}$ is an exogenous parameter (mirroring $x\leq v_h-p^*$, which does not directly translate to the continuous model due to the customers' continuous valuation). The bound on the subsidy $x$ rules out edge cases where, in equilibrium, the distributor offers a very steep discount and a very small fraction of customers purchase the product.

### 6.1. Numerical analysis

We now show that our main analytical results from Section 4 hold in the continuous model. In the latter, $\mu_m$, $\sigma_m$, and $\kappa$ determine the distribution of customers' ATP. We assume $\kappa\leq 0$ since we model products that would most benefit BOP customers, who might be under financial distress and, as a result, have low ATP. We evaluate the probabilities of buying and returning the product explicitly through numerical integration.

**Pricing Problem.** Figure 8 in Appendix D replicates the insight that increasing information might decrease distributor profits if consumers have low ATP and low risk aversion (a similar effect was observed in Figure 3). As before, when customers have low average ATP and risk aversion ($\alpha=0.01$ in Figure 8a) increasing $\theta$ decreases distributor profits. Conversely, when customers are very risk averse ($\alpha=1$ in Figure 8b), increasing $\theta$ always increases distributor profits. This behavior is robust to different model parameters.

| Parameter | Values | Parameter | Values |
|---|---|---|---|
| $w$ | 1 | $\rho_l$ | Uniform(0.1, 0.45) |
| $y$ | Uniform(0,1) | $\rho_h$ | Uniform(0.7, 0.95) |
| $\mu$ | Uniform(0.7,1) | $\theta_0$ | Uniform(0.01, 0.5) |
| $\sigma$ | Uniform(0.5,1) | $c_\theta$ | Uniform(0,2) |
| $\sigma_m$ | Uniform(0.1,1) | $b$ | Uniform(0,1) |
| $\kappa$ | Uniform(-0.8,-0.2) | $\bar{x}$ | $\mu - \mu_m + \sigma + \sigma_m$ |

**Table 1 Distribution of parameters used for the Allocation Problem.**

| | Avg. % of budget allocated to $\theta$ | | |
|---|---|---|---|
| | $\mu_m = 0.75$ | $\mu_m = 1.25$ | $\mu_m = 1.75$ |
| $\alpha = 0.01$ | 0.5% | 4.5% | 26% |
| $\alpha = 1$ | 1% | 4.7% | 28% |
| $\alpha = 2$ | 1.8% | 4.9% | 30% |

**Table 2 Avg. % of budget allocated to $\theta$**

**Allocation Problem.** We now examine how the optimal distributor's budget allocation strategy changes as a function of the customers' average ATP $\mu_m$ and risk aversion level $\alpha$, and discuss their connection to the main results of Section 4.2, which are illustrated in Figure 4.

The simulation sets $w = \$1$ and samples a set of initial market, product, and information parameters $(\mu, \sigma, \sigma_m, \kappa, y, c_\theta, \rho_l, \rho_h)$ drawn from the distributions described in Table 1. The values in Table 1 were chosen to generate problem instances that draw from a wide range of problem parameter values relative to the value of $w$ and from a wide variety of valuation and ATP distributions. We assume that $\gamma = 0$, i.e. a profit-maximizing distributor. For each sample of parameters we solve the allocation problem for different values of $\mu_m$ and $\alpha$. When solving the allocation problem, we calculate the fraction of consumers that purchase and return the product through Monte Carlo integration. We sampled 10,000 sets of problem parameters and, for each set, we considered nine combinations of $\mu_m$ and $\alpha$. We solved a total of 90,000 instances of the allocation problem.

We first discuss the relative investment in $\theta$ and $x$. Table 2 describes the average fraction of the budget invested in $\theta$ (the remainder budget is allocated to $x$) for different values of $\alpha$ and $\mu_m$. The results are consistent with Section 4.2: when comparing Figure 4 and Table 2 along their horizontal axis, the relative investment in $\theta$ (resp. $x$) decreases (resp. increases) with the average ATP in both of them. The distributor invests more in consumer education $\theta$ if the customers' ATP is high.

Second, we discuss the prevalence of strategies that target uninformed customers via product returns. Tables 3 and 4 describe, respectively, the average fraction of customers that receive the low-information or high-information signal and purchase the product for different values of $\alpha$ and $\mu_m$. As before, the results are consistent with Section 4.2. When comparing Figure 4 and Tables 3 and 4 along their vertical axis, the fraction of purchases from customers that receive the low-information (resp. high-information) signal decreases (resp. increases) with $\alpha$. That is, the distributor should induce the retailer to target uninformed consumers if the customers' risk aversion level is low.

| | Avg. % of low-info customers that buy | | |
|---|---|---|---|
| | $\mu_m = 0.75$ | $\mu_m = 1.25$ | $\mu_m = 1.75$ |
| $\alpha = 0.01$ | 11% | 16.5% | 20.3% |
| $\alpha = 1$ | 10.6% | 14.7% | 17.8% |
| $\alpha = 2$ | 9.7% | 13.5% | 16.5% |

**Table 3 Avg. % of purchases from customers that received the low-information signal**

| | Avg. % of high-info customers that buy | | |
|---|---|---|---|
| | $\mu_m = 0.75$ | $\mu_m = 1.25$ | $\mu_m = 1.75$ |
| $\alpha = 0.01$ | 5.4% | 11.4% | 18.3% |
| $\alpha = 1$ | 6.2% | 13.4% | 18.6% |
| $\alpha = 2$ | 7.1% | 15.1% | 18.8% |

**Table 4 Avg. % of purchases from customers that received the high-information signal**

| | % of instances where $\theta^{CS} \geq \theta^*$ | | |
|---|---|---|---|
| | $\mu_m = 0.75$ | $\mu_m = 1.25$ | $\mu_m = 1.75$ |
| $\alpha = 0.01$ | 98% | 92% | 75% |
| $\alpha = 1$ | 96% | 86% | 68% |
| $\alpha = 2$ | 93% | 82% | 61% |

**Table 5 % of sampled instances where maximizing consumer surplus requires a higher $\theta$**

| | % of simulations where $x^{CS} \geq x^*$ | | |
|---|---|---|---|
| | $\mu_m = 0.75$ | $\mu_m = 1.25$ | $\mu_m = 1.75$ |
| $\alpha = 0.01$ | 0.6% | 2.4% | 21% |
| $\alpha = 1$ | 0.5% | 0.8% | 17% |
| $\alpha = 2$ | 0.5% | 0.6% | 16% |

**Table 6 % of sampled instances where maximizing consumer surplus requires a higher $x$**

We now examine the resource allocation strategy that maximizes consumer surplus. For each sample of instance parameters from the previous simulations, we solve the allocation problem with consumer surplus as the distributor's objective for different values of $\mu_m$ and $\alpha$. Tables 5 and 6 summarize the fraction of consumer-surplus-maximizing allocations where the budget allocation to $x$ and $\theta$ is *larger* than in the profit-maximizing allocation. Specifically, let $x^{CS}$ and $\theta^{CS}$ be the optimal solution to the allocation problem when the distributor maximizes consumer surplus, and let $x^*$ and $\theta^*$ be the profit-maximizing optimal allocation. Table 5 lists the fraction of instances where $\theta^{CS} \geq \theta^*$, while Table 6 lists the fraction of instances where $x^{CS} \geq x^*$.

Consistent with the discrete model, Tables 5 and 6 illustrate that strategies that invest more in consumer education generally lead to a higher consumer surplus than finance-based strategies. Additionally, Tables 5 and 6 confirm a strong tension between consumer surplus and profits in the BOP, where $\alpha$ is high and $\mu_m$ is medium or low. That is, in the BOP the consumer-surplus maximizing allocation almost always involves a higher investment in $\theta$ and a lower investment in $x$ than in the profit-maximizing allocation. Moreover, the tension between consumer surplus and profits dissipates in non-BOP contexts with a higher average ATP.

## 7. Conclusions

We introduce a game-theoretic model to analyze the operations strategy of distributors of innovative, life-improving, durable goods in supply chains that serve BOP customers. Our model incorporates two key features of BOP customers: risk aversion and customers with a lower ATP than WTP. We use our model to analyze two operational optimization problems faced by BOP distributors. The first is a Pricing Problem, where the distributor decides the price and refund for returned

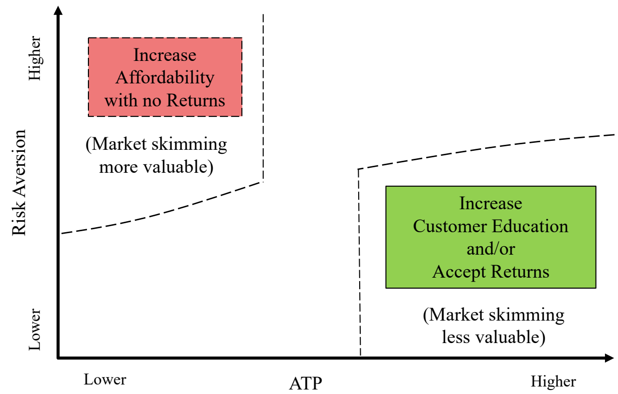



**Figure 6** **Depiction of the distributor's profit-maximizing allocation strategy as a function of customer ATP and risk aversion. The green solid box highlights the strategy with highest consumer surplus and the red dotted box highlights the strategy with lowest consumer surplus.**

products it offers a retailer which, in turn, chooses a customer price and refund. The second is a resource Allocation Problem, where the distributor allocates a given budget to (1) improve product affordability (a financial lever), and (2) improve consumer education (an informational lever).

The distinction between risk-averse customers' ATP and WTP has profound effects on the solutions of these two problems and on supply chain equilibrium behavior. Figure 6 summarizes these effects. The green solid box in Figure 6 highlights the equilibrium allocation strategy that induces the highest consumer surplus, while the red dotted box highlights the strategy where consumer surplus is lowest. Specifically, when customer risk aversion is high and customers' ATPs are low (which is common in the BOP), the distributor's profit-maximizing allocation leads to the *lowest* consumer surplus among non-dominated strategies. Conversely, when customers' ATPs are high, this tension between profits and consumer surplus disappears. Namely, we find a BOP-specific misalignment between consumer surplus and profits.

We propose an operational commitment from the distributor to offering free returns as a potential solution to this misalignment. Such commitment ensures that no customer obtains a negative surplus by trying out the product. We show that, when all customers have the same ATP, the resulting consumer surplus is independent of profit maximization decisions and larger than any equilibrium strategy without free returns, even in a BOP context. When customers have heterogeneous ATP, a commitment to free returns increases customer surplus but does not address the incentives for the retailer to skim the market. Nevertheless, market skimming can be addressed by a distributor's commitment to charging a maximum retail price (which has become a legal requirement in India, for example). More generally, a commitment to free returns and a maximum retail price can signal to socially-conscious investors that the distributor cares about consumer surplus.

Many social enterprises spend significant resources measuring the social impact of their operations to provide "proof" that they are achieving their social goals. Our results indicate that, beyond measuring impact, certain operational commitments (such as guaranteeing free returns and a maximum retail price) can also serve as "proof" of social-mindedness. These commitments constrain the distributor's decisions in a way that forces an alignment between social and financial objectives throughout the supply chain. Exploring other operational commitments that force such alignment can be valuable future research direction. Finally, there is a pressing need for more empirical research and field experiments to validate stylized models of BOP supply chains. For example, a field experiment to test the value of offering free returns on the adoption of durable goods in the BOP could lead to new managerial practices and research questions.

## References

Banerjee A, Hanna R, Mullainathan S (2012) Gibbons R, Roberts J, eds., *The Handbook of Organizational Economics*, chapter Corruption (Princeton: Princeton University Press), ISBN 978-0-691-13279-2.

Banerjee AV (1997) A Theory of Misgovernance. *The Quarterly Journal of Economics* 112(4):1289–1332.

Bellman E, Agarwal V, Sharma S (2018) India's Biggest Competitors to Walmart and Amazon? Mom and Pop. *Wall Street Journal* ISSN 0099-9660, URL `https://www.wsj.com/articles/indias-biggest-competitors-to-walmart-and-amazon-mom-and-pop-1527512400`.

Che YK (1996) Customer return policies for experience goods. *The Journal of Industrial Economics* 17–24.

Chen L, Lee HL (2016) Sourcing under supplier responsibility risk: The effects of certification, audit, and contingency payment. *Management Science* 63(9):2795–2812.

Chu LY, Zhang H (2011) Optimal preorder strategy with endogenous information control. *Management Science* 57(6):1055–1077.

Clean Cooking Alliance (2015) Consumer finance models for clean cookstoves: Global mapping. Technical report, Clean Cooking Alliance.

Clean Cooking Alliance (2019) 2019 clean cooking industry snapshot. Technical report, Clean Cooking Alliance, URL `https://www.cleancookingalliance.org/resources/566.html`.

Costello A (2010) Troubled Water. URL `http://www.pbs.org/frontlineworld/stories/southernafrica904/video_index.html`.

Davidson K (2009) Ethical concerns at the bottom of the pyramid: Where csr meets bop. *Journal of International Business Ethics* 2(1).

Dees JG (1994) *Social enterprise: Private initiatives for the common good* (Harvard Business School).

Dupas P (2014a) Getting essential health products to their end users: Subsidize, but how much? *Science* 345(6202):1279–1281.

Dupas P (2014b) Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica* 82(1):197–228.

Essmart (2020) Life improving technologies catalogue URL `http://www.essmart-global.com/`.

Garrette B, Karnani A (2010) Challenges in marketing socially useful goods to the poor. *California Management Review* 52(4):29–47.

Global Alliance for Clean Cookstoves (2016) Demand creation. URL `http://cleancookstoves.org/market-development/demand-creation/`.

Gui L, Tang CS, Yin S (2018) Improving Microretailer and Consumer Welfare in Developing Economies: Replenishment Strategies and Market Entries. *Manufacturing & Service Operations Management* .

Haushofer J, Fehr E (2014) On the psychology of poverty. *Science* 344(6186):862–867.

Huang L, Song JSJ, Swinney R (2017) Managing Social Responsibility in Multitier Supply Chains. SSRN Scholarly Paper ID 2837332, Social Science Research Network.

Johnson JP, Myatt DP (2006) On the Simple Economics of Advertising, Marketing, and Product Design. *American Economic Review* 96(3):756–784.

Jue DM (2012) *From the lab to the land: social impact technology dissemination in rural southern India.* Thesis, Massachusetts Institute of Technology.

Kohli R, Bhagwati J (2011) Organized retailing in india: Issues and outlook. *CBS Research Paper* (12/25).

Lee HL, Tang CS (2018) Socially and environmentally responsible value chain innovations: New operations management research opportunities. *Management Science* 64(3):983–996.

Levine DI, Beltramo T, Blalock G, Cotterman C, Simons AM (2018) What impedes efficient adoption of products? evidence from randomized sales offers for fuel-efficient cookstoves in uganda. *Journal of the European Economic Association* 16(6):1850–1880.

Mair J, Marti I (2006) Social entrepreneurship research: A source of explanation, prediction, and delight. *Journal of world business* 41(1):36–44.

McMullen JS (2011) Delineating the domain of development entrepreneurship: a market-based approach to facilitating inclusive economic growth. *Entrepreneurship Theory and Practice* 35(1):185–193.

Miller TL, Grimes MG, McMullen JS, Vogus TJ (2012) Venturing for others with heart and head: How compassion encourages social entrepreneurship. *Academy of management review* 37(4):616–640.

Plambeck EL, Taylor TA (2015) Supplier evasion of a buyer's audit: Implications for motivating supplier social and environmental responsibility. *M&SOM* 18(2):184–197.

Prahalad CK (2006) *The Fortune at the Bottom of the Pyramid* (Pearson Prentice Hall).

Shugan SM, Xie J (2000) Advance pricing of services and other implications of separating purchase and consumption. *Journal of Service Research* 2(3):227–239.

Shukla S, Bairiganjan S (2011) The base of pyramid distribution challenge: evaluating alternate distribution models of energy products for rural base of pyramid in india. *XIFMR Research, Centre for Dev. Fin.* .

Shulman JD, Coughlan AT, Savaskan RC (2009) Optimal restocking fees and information provision in an integrated demand-supply model of product returns. *M&SOM* 11(4):577–594.

Simanis E (2012) Reality Check at the Bottom of the Pyramid. *Harvard Business Review* ISSN 0017-8012.

Sodhi MS, Tang CS (2017) Social responsibility in supply chains. *Sustainable Supply Chains*, 465–483 (Springer International Publishing).

Su X (2009) Consumer returns policies and supply chain performance. *M&SOM* 11(4):595–612.

Swaminathan JM (2007) Managing supply chain operations in india. *Building supply chain excellence in emerging economies*, 137–154 (Springer).

Tarozzi A, Mahajan A, Blackburn B, Kopf D, Krishnan L, Yoong J (2014) Micro-loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India. *American Economic Review* 104(7):1909–1941.

Taylor TA, Xiao W (2019) Donor product-subsidies to increase consumption: Implications of consumer awareness and profit-maximizing intermediaries. *Prod. and Operations Management* 28(7):1757–1772.

The Indian Express (2017) From 2018, mrp labels on products sold online mandatory. URL `https://indianexpress.com/article/business/from-2018-mrp-labels-on-products-sold-online-mandatory-4729215/`.

Uppari BS, Popescu I, Netessine S (2017) SSRN Scholarly Paper ID 2782302, Social Science Research Network, Rochester, NY, URL `https://papers.ssrn.com/abstract=2782302`.

Vidal J (2013) India's villagers reap visible benefits from solar electricity scheme. *The Guardian* .

World Bank (2016) *Poverty and shared prosperity 2016: taking on inequality* (World Bank Publications).

World Health Organization (2016) Burning opportunity: clean household energy for health, sustainable development, and well-being of women and children. Technical report, World Health Organization.

Xie J, Shugan SM (2001) Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Science* 20(3):219–243.

Zhang K, Tang CS, Zhou S (2017) Replenishment strategies for micro-retailers in developing countries. *Prod. and Operations Management* 26(12):2207–2225.

## Appendix A: Challenges faced by BOP distributors: the case of Essmart

Profitably scaling the distribution of life-improving durable goods designed for BOP consumers is challenging (Swaminathan 2007, Garrette and Karnani 2010). Although non-profit organizations, multinational companies, and start-ups have attempted to distribute these technologies at scale, no strategy has been completely successful (Jue 2012). Non-profit organizations that have traditionally distributed technologies on a project-by-project basis are limited in scale and financial independence. Multinational companies and start-ups alike have tried increasing consumer awareness and access through massive door-to-door campaigns that combine education and subsidized direct-sales to consumers (for an example, see Vidal 2013). However, managing these massive campaigns with their large networks of sales agents is expensive and labor-intensive, and scaling them is difficult due to high personnel turnover and the limited reach of the sales agents. Finally, all types of distributors struggle with inappropriate product design, lack of long-term maintenance, and lack of proper incentives, leading to commercial failures (see examples in Costello 2010, Simanis 2012).

Given this context, the motivation for our research and the main case study for this paper comes from our collaboration with Essmart, a social enterprise that distributes innovative life-improving technologies to BOP consumers in India. For a distributor like Essmart, the mixed objective includes both consumer surplus and profitability.

Essmart was founded in 2012 and operates a hub-and-spoke distribution model that leverages India's extensive network of local retail shops. Each modular Essmart office, located in tier 3 towns in India, has four Sales Executives who drive on routes of up to 100 kilometers per day via motorbike to build relationships with peri-urban and rural retail shops, turning these shop owners into extensions of Essmart's sales force but with an established local presence and existing trusted buying relationships with customers. Essmart currently operates 18 offices in Tamil Nadu, Karnataka, and Andhra Pradesh, India that have collectively built a network of over 2,700 local retail shops.

Essmart's innovative operations strategy, which allowed the company to achieve the unit economic profitability of their offices in 2016, has three main components:

1. *Distribution*: Essmart partners with small "mom-and-pop" retail shops as points of sale and offers them expedited product delivery, leveraging the fact that 85% of the annual retail spending in India occurs through more than 12 million local retailers (Kohli and Bhagwati 2011). Essmart gives retailers a few sample items and a catalogue that lists all of its products. With Essmart's help, retailers select the products from Essmart's catalogue of 350+ SKUs that best fit their needs and shop profile, enabling a diverse set of shop types to start selling Essmart products. When the retailer has a sale opportunity, Essmart delivers the product within a few days. This "deliver-to-order" strategy effectively removes the inventory risk and shelf space requirement from the retailer.

2. *Consumer Education (marketing)*: Essmart Sales Executives run product demonstrations at local shops and markets to educate consumers on product features and the needs they address, as well as refer consumers to local retailers if they are interested in purchasing a product. Although the demonstrations have less reach than door-to-door campaigns, they are significantly less labour intensive. The demonstrations create awareness of Essmart's products and seek to build trust between Essmart and consumers, as well as between Essmart and their retailers, who are involved in the demonstrations.

3. *Consumer Guarantees*: Essmart offers consumers a range of guarantees, including the opportunity to return products and facilitating the service of faulty products that have manufacturer warranties. This is the most innovative component of their operations strategy, since most companies and organizations that distribute life-improving technologies in the BOP do not offer customers or retailers any form of protection or recourse.

## Appendix B: Proofs and Additional Results from Section 4

### B.1. Additional Results from Section 4.1

Recall that $r_\alpha(p)$ is the minimum refund the retailer must offer to uninformed consumers such that they purchase the product at an effective price $p$, and $p_\alpha(r)$ is the WTP of uninformed consumers given a refund $r$ offered by the retailer.

We now argue that $r_\alpha(p)$, given in Equation 3, is invertible for any $p$ such that $r_\alpha(p) > 0$, thus $p_\alpha(r) = r_\alpha^{-1}(r)$ is well-defined. Indeed, the derivative $r'_\alpha(p)$ for any $p$ such that $r_\alpha(p) > 0$ is

$$r'_\alpha(p) = 1 + \frac{\beta}{e^{\alpha(v_h - p + x)} - \beta} > 0, \tag{7}$$

where the inequality follows since $e^{\alpha(v_h - p + x)} \geq 1$ for $p \in [0, v_h + x]$ and $0 < \beta < 1$. Then, $r_\alpha(p)$ is invertible.

We are now ready to characterize the optimal solution to the Pricing Problem.

**Proposition B.1** *Consider any consumer education level $\theta \in [0,1)$, ATP $m \in [w, v_h]$, and subsidy $x \in [0, v_h - m]$. Then, the distributor's optimal objective is*

$$\Pi^*_D(x,\theta) = \max\left\{\Pi^a_D(x,\theta), \Pi^b_D(x,\theta)\right\}. \tag{8}$$

*Where $\Pi^a_D(x,\theta)$ and $\Pi^b_D(x,\theta)$ each correspond to the distributor's profits in a non-dominated strategy. Specifically, strategies (a) and (b) are characterized by:*

*(a) Target only informed customers without product returns. The customer price is $p^a = m + x$ and refund is $r^a = 0$. The retailer's price and refund are $c^a = m + x$ and $z^a = 0$, respectively. The customer surplus is $CS^a = (v_h - m)\theta\beta$. The retailer's profit is $\Pi^a_R = 0$ while the distributor's profit is $\Pi^a_D(x,\theta) = (m + x - w)\theta\beta + \gamma CS^a$.*

*(b) Target both informed and uninformed customers with product returns. The customer price is $p^b = m + x$ and the refund is $r^b = r_\alpha(m) = \max\left(0, m - \frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right)\right)$. The equilibrium retailer refund is $z^b = r_\alpha(m)$ and the retailer price $c^b$ is*

$$c^b = m + x + \frac{(1-\theta)(1-\beta)\theta\beta}{\theta\beta + (1-\theta)}(z^b - \bar{z}). \tag{9}$$

*The consumer surplus is*

$$CS^b = (v_h - m)\beta - (m - r_\alpha(m))(1-\theta)(1-\beta),$$

*the retailer's profit is $\Pi^b_R = 0$, and the distributor's profit is*

$$\Pi^b_D(x,\theta) = (m + x - w)\beta + (1-\beta)(m + x - r_\alpha(m) - w + y)(1-\theta) + \gamma CS^b.$$

*Strategy (b) is preferred by the distributor over strategy (a) if and only if*

$$m+x-w-(1-\beta)(r_\alpha(m)-y)+\gamma(\beta(v_h-m)-(1-\beta)(m-r_\alpha(m)))\geq 0. \tag{10}$$

*Proof.* First, we analyze the retailer's pricing problem and show that the optimal nominal price to consumers is $p^*=m+x$ and the optimal refund is such that $r^*\in\{0,r_\alpha(m)\}$. Note that the expected fraction of customers who purchase the product is $\mathrm{B}(p,r,\theta)=\Big(\theta\beta+(1-\theta)\mathbb{1}_{\{r\geq r_\alpha(p-x)\}}\Big)\,\mathbb{1}_{\{p-x\leq m\}}$. Since both indicator functions are decreasing in $p$ (cf. Equation 7), it follows that, for a given $r$, the optimal retailer's price $p^*$ is such that $p^*\in\{m+x,p_\alpha(r)+x\}$.

Then, if $p^*=m+x$ the retailer's profit function is

$$\Pi_R\,(m+x,r)=(m+x-c)\left(\theta\beta+(1-\theta)\mathbb{1}_{\{r\geq r_\alpha(m)\}}\right)-(r-z)(1-\theta)(1-\beta)\mathbb{1}_{\{r\geq r_\alpha(m)\}}.$$

The profit function above is constant for $r<r_\alpha(m)$, has an increasing or decreasing step at $r=r_\alpha(m)$, and is linear decreasing for $r>r_\alpha(m)$. Therefore, $r^*\in\{0,r_\alpha(m)\}$ when $p^*=m+x$.

Conversely, if $p^*=p_\alpha(r)+x$ the retailer's profit function is

$$\Pi_R\,(p_\alpha(r)+x,r)=(p_\alpha(r)+x-c)\,(\theta\beta+(1-\theta))\,\mathbb{1}_{\{p_\alpha(r)\leq m\}}-(r-z)(1-\theta)(1-\beta)\mathbb{1}_{\{p_\alpha(r)\leq m\}}.$$

The profit function above is zero if $p_\alpha(r)>m$ since customers cannot afford the product. When $p_\alpha(r)\leq m$ (equivalently when $r\leq r_\alpha(m)$) the profit function is increasing in $r$. To show this, first note that when $r\leq r_\alpha(m)$ we have

$$\frac{\partial\Pi_R\,(p_\alpha(r)+x,r)}{\partial r}=p'_\alpha(r)\,(\theta\beta+(1-\theta))-(1-\theta)(1-\beta)=\frac{\theta\beta+(1-\theta)}{r'_\alpha(p_\alpha(r))}-(1-\theta)(1-\beta).$$

where the second equality comes from the fact that $p_\alpha=r_\alpha^{-1}$. Then, from Equation 7, for any $p\in[0,v_h+x]$, we have that

$$r'_\alpha(p)\leq 1+\frac{\beta}{1-\beta}\leq 1+\frac{\beta}{(1-\theta)(1-\beta)}=\frac{\theta\beta+(1-\theta)}{(1-\theta)(1-\beta)}.$$

The first inequality comes from noting that $r'_\alpha(p)$, given in Equation 7, is decreasing in $p$ for $p\in[0,v_h+x]$ and that $r'_\alpha(v_h+x)=1+\frac{\beta}{1-\beta}$. It follows that $\frac{\partial\Pi_R(p_\alpha(r)+x,r)}{\partial r}\geq 0$, thus $r^*=r_\alpha(m)$ when $p^*=p_\alpha(r)+x$, and $p^*=m+x$ in this case as well.

Hence, we conclude that $p^*=m+x$ and $r^*\in\{0,r_\alpha(m)\}$, fully characterizing the potential equilibrium behaviors of the retailer. With the retailer's equilibrium behavior in hand, we now characterize the distributor's equilibrium pricing and refund strategies.

Assume first that the distributor is interested in inducing the retailer to target informed consumers with high valuation, i.e. set $p^*=m+x$ and $r^*=0$. In this case, the distributor's problem can be written as

$$\begin{aligned}
&\max_{c,z\geq 0}\;(c-w)\theta\beta+\gamma(v_h-m)\theta\beta\\
&\text{s.t. }(m+x-c)\theta\beta\geq(m+x-c)\,(\theta\beta+(1-\theta))-(r_\alpha(m)-z)(1-\theta)(1-\beta) && (IC)\\
&\qquad(m+x-c)\theta\beta\geq 0. && (IR)
\end{aligned}$$

Note that the $(IC)$ constraint above implies that the retailer sets $p^*=m+x$ and $r^*=0$. The objective is increasing in $c$ and independent of $z$, with an upper bound $c\leq m+x$ given by the $(IR)$ constraint as long as $z \le r_\alpha(m)$, otherwise the problem is infeasible. Hence, in particular $c^* = m + x$ and $z^* = 0$, leading to strategy $(a)$ in the statement of the proposition.

Now assume that the distributor is interested in inducing the retailer to target uninformed consumers too, i.e. set $p^* = m + x$ and $r^* = r_\alpha(m)$. In this case, the distributor's problem can be written as

$$\begin{aligned}
&\max_{c,z\ge 0}\ (c-w)\left(\theta\beta+(1-\theta)\right)-(z-y)(1-\theta)(1-\beta)+\gamma((v_h-m)\beta-(m-r_\alpha(m))(1-\theta)(1-\beta))\\
&\text{ s.t. } (m+x-c)\left(\theta\beta+(1-\theta)\right)-(r_\alpha(m)-z)(1-\theta)(1-\beta)\ge(m+x-c)\theta\beta && (IC)\\
&\qquad (m+x-c)\left(\theta\beta+(1-\theta)\right)-(r_\alpha(m)-z)(1-\theta)(1-\beta)\ge 0. && (IR)
\end{aligned}$$

The objective is increasing in $c$ and decreasing in $z$, leading to $c^* = m + x$ and $z^* = r_\alpha(m)$ where both the $(IR)$ and $(IC)$ constraints are tight. This corresponds to strategy $(b)$ in the statement of the proposition. A direct comparison between $\Pi_D^b$ and $\Pi_D^a$ shows that $\Pi_D^b(x,\theta) \ge \Pi_D^a(x,\theta)$ if and only if

$$m+x-w-(1-\beta)(r_\alpha(m)-y)+\gamma(\beta(v_h-m)-(1-\beta)(m-r_\alpha(m)))\ge 0,$$

completing the proof. □

Note that in Equation 10 the term $m + x - w$ is the expected distributor's margin on each sale, $(1-\beta)(r_\alpha(m)-y)$ is the expected cost of a product return, and $\beta(v_h - m) - (1-\beta)(m - r_\alpha(m))$ is the uninformed consumer's expected surplus.

### B.2. Proof of Proposition 1

*Proof.* We first show that the distributor's objective function is increasing in $x$. Note that

$$\frac{\partial \Pi_D^a}{\partial x} = \theta\beta > 0, \text{ and } \frac{\partial \Pi_D^b}{\partial x} = \beta + (1-\theta)(1-\beta) > 0.$$

Hence, $\frac{\partial \Pi_D^*}{\partial x} > 0$.

In contrast, the distributor's objective function is not always increasing in $\theta$. To prove this, first note that $\frac{\partial \Pi_D^a}{\partial \theta} = (\gamma v_h + (1-\gamma)m + x - w)\beta > 0$ and

$$\frac{\partial \Pi_D^b}{\partial \theta} = -(1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) + (1-\beta)(w-y-x).$$

Hence, $\frac{\partial \Pi_D^b}{\partial \theta} \le 0$ if and only if

$$y + (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) + x - w \ge 0.$$

Moreover, recall from Proposition B.1 that $\Pi_D^* = \Pi_D^b$ if and only if

$$m+x-w-(1-\beta)(r_\alpha(m)-y)+\gamma(\beta(v_h-m)-(1-\beta)(m-r_\alpha(m)))\ge 0,$$

or equivalently

$$y + (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) + x - w + \frac{\beta}{1-\beta}(\gamma(v_h-m)+m+x-w) \ge 0. \tag{11}$$

By noting that $(\gamma(v_h - m) + m + x - w) \ge 0$ we conclude that $\frac{\partial \Pi_D^b}{\partial \theta} \le 0$ implies $\Pi_D^* = \Pi_D^b$. Therefore, $\frac{\partial \Pi_D^*}{\partial \theta} < 0$ if and only if the inequality (11) holds.

Finally, $\frac{\partial^2 \Pi_D^a}{\partial \theta \partial x} = \beta > 0$, and $\frac{\partial^2 \Pi_D^b}{\partial \theta \partial x} = -(1-\beta) < 0$. This concludes the proof. □

### B.3. Proof of Theorem 1

*Proof.* Let $(x^*, \theta^*)$ be the optimal solution of the distributor's allocation problem. We analyze strategies $(a)$ and $(b)$ from Proposition B.1 separately.

$\underline{F \text{ and } M}$: Assume first that $\Pi_D^*(x^*, \theta^*) = \Pi_D^a(x^*, \theta^*)$.

Since $\frac{\partial \Pi_D^a}{\partial x} = \theta\beta > 0$ and $\frac{\partial \Pi_D^a}{\partial \theta} = (\gamma v_h + (1-\gamma)m + x - w)\beta > 0$ then from Assumption 1 it follows that the optimal solution to the resource allocation problem exhausts the budget, i.e., $c_\theta(\theta^* - \theta_0) + \theta^*\beta x^* = b$, or equivalently $x^*(\theta) = \frac{b - c_\theta(\theta - \theta_0)}{\theta\beta}$. By replacing $x^*(\theta)$ in the resource allocation problem, it simplifies to the following one-variable optimization problem

$$\begin{aligned} \max_{\theta} \quad & \Pi_D^a(\theta) = (\gamma v_h + (1-\gamma)m - w)\theta\beta + b - c_\theta(\theta - \theta_0) \\ \text{s.t.} \quad & \theta \in [\theta_0, 1], \ \frac{b - c_\theta(\theta - \theta_0)}{\theta\beta} \in [0, v_h - m]. \end{aligned} \tag{12}$$

Let $\theta^a$ be the optimal solution of problem (12), and $x^a = \frac{b - c_\theta(\theta^a - \theta_0)}{\theta^a \beta}$. The objective function of problem (12), $\Pi_D^a(\theta)$, is linear. Moreover,

$$\frac{d\Pi_D^a(\theta)}{d\theta} = (\gamma v_h + (1-\gamma)m - w)\beta - c_\theta. \tag{13}$$

Hence, $\theta^a$ must be equal to one of its upper or lower bound, i.e.,

$$\theta^a \in \left\{ \theta_0 + \frac{(b - \theta_0\beta(v_h - m))^+}{c_\theta + \beta(v_h - m)}, \min\left(1, \theta_0 + \frac{b}{c_\theta}\right) \right\},$$

and thus

$$x^a \in \left\{ \min\left(v_h - m, \frac{b}{\theta_0\beta}\right), \frac{(b - c_\theta(1 - \theta_0))^+}{\min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\beta} \right\}.$$

Namely, when following strategy $(a)$ in Proposition B.1 the distributor either invests the budget in increasing the consumers' maximum ability to pay first, and then the consumer education level only if there is budget available (strategy $F$), or invests the budget in increasing the consumer education level first, and then the consumers' maximum ability to pay only if there is budget available (strategy $M$).

Specifically,

$$\theta^F = \theta_0 + \frac{(b - \theta_0\beta(v_h - m))^+}{c_\theta + \beta(v_h - m)}, \ x^F = \min\left(v_h - m, \frac{b}{\theta_0\beta}\right),$$

and

$$\Pi_D^F = \left(\gamma v_h + (1-\gamma)m + \min\left(v_h - m, \frac{b}{\theta_0\beta}\right) - w\right)\left(\theta_0 + \frac{(b - \theta_0\beta(v_h - m))^+}{c_\theta + \beta(v_h - m)}\right)\beta$$

Similarly,

$$\theta^M = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right), \ x^M = \frac{(b - c_\theta(1 - \theta_0))^+}{\beta},$$

and

$$\Pi_D^M = (\gamma v_h + (1-\gamma)m - w)\min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\beta + (b - c_\theta(1 - \theta_0))^+$$

$\underline{FR \text{ and } MR}$: Assume now that $\Pi_D^*(x^*, \theta^*) = \Pi_D^b(x^*, \theta^*)$.

Since $\frac{\partial \Pi_D^b}{\partial x} = \beta + (1-\theta)(1-\beta) > 0$, then from Assumption 1 it follows that the optimal subsidy must either be equal to its upper bound or the budget constraint must be tight, i.e., $x^*(\theta) = \min\left(v_h - m, \frac{b - c_\theta(\theta - \theta_0)}{\beta + (1-\theta)(1-\beta)}\right)$.

By replacing $x^*(\theta)$ in the resource allocation problem, it simplifies to the following one-variable optimization problem

$$\begin{aligned}\max_{\theta} \quad & \Pi_D^b(\theta) = (\gamma v_h + (1-\gamma)m - w)\beta + \min\big((v_h - m)(\beta + (1-\theta)(1-\beta)), b - c_\theta(\theta - \theta_0)\big) \\ & + \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right) - (1-\beta)(w-y)\right)(1-\theta) \\ \text{s.t.} \quad & \theta \in [\theta_0, 1],\ c_\theta(\theta - \theta_0) \le b.\end{aligned} \tag{14}$$

Let $\theta^b$ be the optimal solution of problem (14) and $x^b = x^*(\theta^b)$. Since the minimum of two linear functions is concave, then the objective function of problem (14), $\Pi_D^b(\theta)$, is piece-wise linear concave with at most two pieces. Moreover, note that

$$\begin{aligned}\frac{d\Pi_D^b(\theta)}{d\theta} = & -\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right) - (1-\beta)(w-y)\right) \\ & - c_\theta - ((v_h - m)(1-\beta) - c_\theta)\mathbb{1}_{\{c_\theta(\theta-\theta_0)+(v_h-m)(\beta+(1-\theta)(1-\beta))\le b\}}.\end{aligned} \tag{15}$$

Hence, we conclude that either $\theta^b$ is equal to its upper bound, i.e., $\theta^b = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)$, or alternatively $\theta^b$ must be equal to one of its lower bound or the kink between the linear pieces of $\Pi_D^b(\theta)$, i.e., $\theta^b \in \left\{\theta_0, \frac{b + c_\theta\theta_0 - (v_h - m)}{c_\theta - (v_h - m)(1-\beta)}\right\}$. Namely, when following strategy $(b)$ in Proposition B.1 the distributor either invests the budget in increasing the consumer education level first, and then the consumers' maximum ability to pay only if there is budget available (strategy $MR$), or invests the budget in increasing the consumers' maximum ability to pay first, and then in increasing the consumer education level only if it is beneficial and there is budget available (strategy $FR$).

Specifically,

$$\theta^{MR} = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right),\ x^{MR} = \frac{(b - c_\theta(1-\theta_0))^+}{\beta},$$

and

$$\begin{aligned}\Pi_D^{MR} = \quad & (\gamma v_h + (1-\gamma)m - w)\beta + (b - c_\theta(1-\theta_0))^+ \\ & + \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right) - (1-\beta)(w-y)\right)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+.\end{aligned} \tag{16}$$

In order to write $\theta^{FR}$ in closed form there are two possible cases depending on whether $\Pi_D^b(\theta)$ has a kink in the feasible interval of problem (14), $\left[\theta_0, \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\right]$. We analyze these cases next.

First assume $c_\theta \le (v_h - m)(1-\beta)$, then from Assumption 1 it follows that $\Pi_D^b(\theta)$ does not have a kink in the feasible interval of problem (14). Specifically, if $c_\theta < (v_h - m)(1-\beta)$ then $\frac{b + c_\theta\theta_0 - (v_h - m)}{c_\theta - (v_h - m)(1-\beta)} \ge 1$, and if $c_\theta = (v_h - m)(1-\beta)$ then $b - c_\theta(\theta - \theta_0) \le (v_h - m)(\beta + (1-\theta)(1-\beta))$ for all $\theta$. Namely, at $\theta = \theta_0$ there is no leftover budget after investing in $x$ and $\theta^{FR} = \theta_0$ in this sub-case.

Now assume $c_\theta > (v_h - m)(1-\beta)$, then from Assumption 1 it follows that $\Pi_D^b(\theta)$ has a kink in the feasible interval of problem (14) if and only if $\frac{b + c_\theta\theta_0 - (v_h - m)}{c_\theta - (v_h - m)(1-\beta)} \ge \theta_0$, or equivalently $(v_h - m)(\beta + (1-\theta_0)(1-\beta)) \le b$. Moreover, from Equation 15 it follows that the kink will attain an objective value at least as large as the solution $\theta = \theta_0$ if and only if $y \le y_\theta$, where

$$y_\theta \equiv w - (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right) - (v_h - m).$$

To simplify the notation, we define the function $\frac{1}{x^{++}} := \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{1}{x} & \text{if } x > 0. \end{cases}$ Then,

$$\theta^{FR} = \theta_0 + \frac{\big(b-(v_h-m)(\beta+(1-\theta_0)(1-\beta))\big)^+}{\big(c_\theta-(v_h-m)(1-\beta)\big)^{++}} \mathbb{1}_{\{y \leq y_\theta\}},$$

$$x^{FR} = \min\left(v_h - m, \frac{b}{\beta+(1-\theta_0)(1-\beta)}\right),$$

and

$$\begin{aligned}\Pi_D^{FR} = \quad & (\gamma v_h + (1-\gamma)m - w)\,\beta + \min\left((v_h-m)(\beta+(1-\theta_0)(1-\beta)), b\right) \\ & + \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)\,(w-y)\right)(1-\theta_0) \\ & + \left(w - y - \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (v_h-m)\right)^+ (1-\beta)\frac{\big(b-(v_h-m)(\beta+(1-\theta_0)(1-\beta))\big)^+}{\big(c_\theta-(v_h-m)(1-\beta)\big)^{++}}.\end{aligned} \tag{17}$$

To conclude, note that the optimal objective value of the resource allocation problem is

$$\Pi_D^{OPT} = \max(\Pi_D^F, \Pi_D^M, \Pi_D^{FR}, \Pi_D^{MR}).$$

□

### B.4. Strategy Map with Homogeneous ATP

To characterize the strategy map of the Allocation Problem in Proposition B.2 we use three auxiliary functions $a_1(m)$, $a_2(m)$, and $a_3(m)$, as well as three auxiliary thresholds on the consumers ATP $m^F$, $m^{MR}$ and $\bar{m}^{MR}$, which are defined in its proof.

**Proposition B.2** *Assume $w \leq \beta v_h + (1-\beta)y$ and $c_\theta \leq (1-\beta)(w-y)$. Let $F$, $M$, $FR$, and $MR$ be the investment strategies from Theorem 1. Then, there exist ATPs $m^F$, $m^{MR}$, and $\bar{m}^{MR}$ where $w \leq m^F \leq m^{MR}, \bar{m}^{MR} \leq v_h$ and functions $a_1(m)$, $a_2(m)$, and $a_3(m)$, such that for a given customer risk aversion parameter $\alpha$ and base ATP $m$,*

- *Strategy $F$ is optimal if $\alpha$ is large enough and $m$ is smaller than $m^F$. Specifically,*

$$\Pi_D^F \geq \max(\Pi_D^M, \Pi_D^{FR}, \Pi_D^{MR}) \text{ if and only if } m \leq m^F \text{ and } \alpha \geq a_3(m).$$

- *Strategy $M$ is optimal if $\alpha$ is large enough and $m$ is larger than $m^F$ and smaller than $\bar{m}^{MR}$. Specifically,*

$$\Pi_D^M \geq \max(\Pi_D^F, \Pi_D^{FR}, \Pi_D^{MR}) \text{ if and only if } m^F \leq m \leq \bar{m}^{MR} \text{ and } \alpha \geq a_2(m).$$

- *Strategy $FR$ is optimal if $\alpha$ and $m$ are small enough. Specifically,*

$$\Pi_D^{FR} \geq \max(\Pi_D^F, \Pi_D^M, \Pi_D^{MR}) \text{ if and only if } m \leq m^{MR} \text{ and } \alpha \leq \min(a_1(m), a_3(m)).$$

- *Strategy $MR$ is optimal if $m$ is larger than $m^F$ and $\alpha$ is larger than $a_1(m)$ and smaller than $a_2(m)$. Specifically,*

$$\Pi_D^{MR} \geq \max(\Pi_D^F, \Pi_D^{MF}, \Pi_D^{FR}) \text{ if and only if } m^F \leq m \text{ and } a_1(m) \leq \alpha \leq a_2(m),$$

*where $a_1(m) \leq a_2(m)$ if and only if $m \geq m^F$. Also, $a_1(m) = 0$ for all $m^{MR} \leq m$ and $a_2(m) = \infty$ for all $\bar{m}^{MR} \leq m$.*

Before stating the proof of Proposition B.2, we provide intuition on the proof's structure and define auxiliary functions.

Our goal is to describe the optimal allocation strategy as a function of $(\alpha, m)$. For such, we write the profit of each strategy as a function of $(\alpha, m)$. The profits under strategies $F, M, FR$, and $MR$ are, respectively $\Pi_D^F(\alpha,m), \Pi_D^M(\alpha,m), \Pi_D^{FR}(\alpha,m)$, and $\Pi_D^{MR}(\alpha,m)$.

We will define a threshold, $m^F$, and functions $a_1(m), a_2(m)$, and $a_3(m)$ where $a_1(m^F) = a_2(m^F) = a_3(m^F)$. We will then establish the following properties:

- Property 1: $\Pi_D^F(\alpha,m) \geq \Pi_D^M(\alpha,m)$ if and only if $m \leq m^F$;
- Property 2: $\Pi_D^{MR}(\alpha,m) \geq \Pi_D^{FR}(\alpha,m)$ if and only if $\alpha \geq a_1(m)$;
- Property 3: $\Pi_D^M(\alpha,m) \geq \Pi_D^{MR}(\alpha,m)$ if and only if $\alpha \geq a_2(m)$;
- Property 4: $\Pi_D^F(\alpha,m) \geq \Pi_D^{FR}(\alpha,m)$ if and only if $\alpha \geq a_3(m)$;
- Property 5: $a_1(m) \leq a_2(m)$ if and only if $m \geq m^F$. Moreover, $a_2(m) = \infty$ for all $m \geq \bar{m}^{MR}$ and $a_1(m) = 0$ for all $m \geq m^{MR}$, where $w \leq m^F \leq m^{MR}, \bar{m}^{MR} \leq v_h$

Note that Property 5 combined with the first four properties establishes the following additional properties:

- Property 6: when $m \leq m^F$ then $\Pi_D^{MR}(\alpha,m) \leq \max(\Pi_D^{FR}(\alpha,m), \Pi_D^M(\alpha,m))$;
- Property 7: when $m \geq m^F$ then $\alpha \geq a_2(m)$ implies $\Pi_D^{FR}(\alpha,m) \leq \Pi_D^M(\alpha,m)$;
- Property 8: when $m \leq m^F$ then $\alpha \leq \alpha_3(m)$ implies $\Pi_D^{FR}(\alpha,m) \geq \Pi_D^M(\alpha,m)$;
- Property 9: when $m \geq m^F$ then $\alpha \leq \alpha_1(m)$ implies $\Pi_D^{FR}(\alpha,m) \geq \Pi_D^M(\alpha,m)$.

The nine properties above completely characterize the strategy map in the statement of Proposition B.2. Namely,

- $F$ is optimal if and only if $m \leq m^F$ and $\alpha \geq a_3(m)$ (combine Properties 1, 4, and 6);
- $M$ is optimal if and only if $m^F \leq m \leq \bar{m}^{MR}$ and $\alpha \geq \alpha_2(m)$ (combine Properties 1, 3, 5, and 7);
- $FR$ is optimal if and only if $m \leq m^{MR}$ and $\alpha \leq \min(a_1(m), a_3(m))$ (combine Properties 2, 4, 5, 8, 9);
- $MR$ is optimal if and only if $m^F \leq m$ and $a_1(m) \leq \alpha \leq a_2(m)$, where $a_1(m) \leq a_2(m)$ if and only if $m \geq m^F$, $a_1(m) = 0$ for all $m^{MR} \leq m$, and $a_1(m) = \infty$ for all $\bar{m}^{MR} \leq m$. (combine Properties 1, 2, 3, 5).

We now prove each property.

*Proof of Property 1.* Recall the Linear Program in (12) for deciding between strategies $M$ and $F$. From Equation 13, we have that $\Pi_D^F \geq \Pi_D^M$ if and only if $(\gamma v_h + (1-\gamma)m - w)\beta \leq c_\theta$. Thus, $m^F = \frac{\frac{c_\theta}{\beta} + w - \gamma v_h}{(1-\gamma)}$ .

*Proof of Property 2.* Recall the optimization problem in (14). From the concavity of $\Pi_D^b(\theta)$ it follows that it is sufficient to check the derivative $\frac{d\Pi_D^b(\theta)}{d\theta}$ at $\theta = 1$, i.e.,

$$\Pi_D^{MR} \geq \Pi_D^{FR} \iff \frac{d\Pi_D^b(1)}{d\theta} \geq 0.$$

We have

$$\begin{aligned}
\frac{d\Pi_D^b(1)}{d\theta} &= -\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y)\right) \\
&\quad - c_\theta - ((v_h-m)(1-\beta) - c_\theta)\mathbb{1}_{\left\{c_\theta(1-\theta_0)+(v_h-m)\beta \leq b\right\}} \\
&= -(1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) + (1-\beta)(w-y) - c_\theta,
\end{aligned}$$

where the second equality follows from Assumption 1. Thus,

$$\frac{d\Pi_D^b(1)}{d\theta} \geq 0 \iff \frac{1-\gamma}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) \leq g_1(m),$$

where $g_1(m) = w - y - \frac{c_\theta}{(1-\beta)}$. First assume that $\gamma \leq 1$, note that then the term $\frac{1-\gamma}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $\left[0, \frac{(1-\gamma)\beta(v_h-m)}{1-\beta}\right]$. If $0 \leq g_1(m) \leq \frac{(1-\gamma)\beta(v_h-m)}{1-\beta}$, then there exists $\hat{\alpha}_1$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_1} \ln\left(\frac{1-\beta e^{-\hat{\alpha}_1(v_h-m)}}{1-\beta}\right) = g_1(m).$$

Then, we define $a_1(m)$ as

$$a_1(m) = \begin{cases} \infty, & \text{if } g_1(m) < 0 \\ \hat{\alpha}_1, & \text{if } 0 \leq g_1(m) \leq \frac{(1-\gamma)\beta(v_h-m)}{1-\beta} \\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m)}{1-\beta} < g_1(m). \end{cases}$$

Since $c_\theta \leq (1-\beta)(w-y)$ by assumption, then $g_1(m) \geq 0$ thus $a_1(m) < \infty$. The definition of $a_1(m)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity. This establishes Property 2.

*Proof of Property 3.* To compare strategy $M$ and $MR$, we directly compare their distributor's value. From the proof of Theorem 1:

$$\begin{aligned}\Pi_D^{MR}(\alpha, m) =& (\gamma v_h + (1-\gamma)m - w)\beta + (b - c_\theta(1-\theta_0))^+ \\ &+ \left((1-\gamma)\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y)\right)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+,\end{aligned}$$

and

$$\Pi_D^M(\alpha, m) = (\gamma v_h + (1-\gamma)m - w)\min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\beta + (b - c_\theta(1-\theta_0))^+.$$

Thus, $\Pi_D^M(\alpha, m) \geq \Pi_D^{MR}(\alpha, m)$ if and only if

$$-(\gamma v_h + (1-\gamma)m - w)\beta\left(1-\theta_0-\frac{b}{c_\theta}\right)^+ \geq \left((1-\gamma)\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y)\right)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+.$$

If $\left(1-\theta_0-\frac{b}{c_\theta}\right)^+ = 0$, then the budget is sufficient to set $\theta = 1$ and the profits of policy $M$ and $MR$ are the same since all customers are informed and there are no returns. Thus we consider the case where $\left(1-\theta_0-\frac{b}{c_\theta}\right)^+ > 0$. In such case, the inequality above becomes

$$-(\gamma v_h + (1-\gamma)m - w)\beta \geq (1-\gamma)\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y).$$

Rearranging the terms yields the relationship,

$$\Pi_D^M(\alpha, m) \geq \Pi_D^{MR}(\alpha, m) \iff \frac{1-\gamma}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) \leq g_2(m).$$

where $g_2(m) = w - y - (\gamma v_h + (1-\gamma)m - w)\frac{\beta}{1-\beta}$. As in Property 2, first assume that $\gamma \leq 1$, note that then the term $\frac{1-\gamma}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $\left[0, \frac{(1-\gamma)\beta(v_h-m)}{1-\beta}\right]$. If $0 \leq g_2(m) \leq \frac{(1-\gamma)\beta(v_h-m)}{1-\beta}$, then there exists a $\hat{\alpha}_2$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_2} \ln\left(\frac{1-\beta e^{-\hat{\alpha}_2(v_h-m)}}{1-\beta}\right) = g_2(m).$$

Then, we define $a_2(m)$ as

$$a_2(m) = \begin{cases} \infty, & \text{if } g_2(m) < 0 \\ \hat{\alpha}_2, & \text{if } 0 \le g_2(m) \le \frac{(1-\gamma)\beta(v_h - m)}{1-\beta} \\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m)}{1-\beta} < g_2(m). \end{cases}$$

Since $w \le \beta v_h + (1-\beta)y$ by assumption, then $g_2(m) \le (1-\gamma)\beta(v_h - m)/(1-\beta)$ thus $a_2(m) > 0$. As before, the definition of $a_2(m)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity. This establishes Property 3.

*Proof of Property 4.* The proof of this property is analogous to the proof of Properties 2 and 3. We denote the information level in strategy $F$ and $FR$ by $\theta^F$ and $\theta^{FR}$, respectively. We directly compare the distributor's value in $F$ and $FR$. Hence,

$$\begin{aligned}
&\Pi_D^F(\alpha, m) \ge \Pi_D^{FR}(\alpha, m) \\
\iff &(\gamma v_h + (1-\gamma)m - w)\theta^F\beta + b - c_\theta(\theta^F - \theta_0) \ge \\
&(\gamma v_h + (1-\gamma)m - w)\beta + \min\left((v_h - m)(\beta + (1-\theta^{FR})(1-\beta)), b - c_\theta(\theta^{FR} - \theta_0)\right) \\
&+ \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y)\right)(1-\theta^{FR}) \\
\iff &(\gamma v_h + (1-\gamma)m - w)\theta^F\beta + b - c_\theta(\theta^F - \theta_0) \ge \\
&(\gamma v_h + (1-\gamma)m - w)\theta^{FR}\beta + b - c_\theta(\theta^{FR} - \theta_0) - \left(b - c_\theta(\theta^{FR} - \theta_0) - (v_h - m)(\beta + (1-\theta^{FR})(1-\beta))\right)^+ \\
&+ \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y) + (\gamma v_h + (1-\gamma)m - w)\beta\right)(1-\theta^{FR}) \\
\iff &((\gamma v_h + (1-\gamma)m - w)\beta - c_\theta)(\theta^F - \theta^{FR}) \\
&+ \left(b - (v_h - m)(\beta + (1-\theta_0)(1-\beta)) - (c_\theta - (v_h - m)(1-\beta))(\theta^{FR} - \theta_0)\right)^+ \ge \\
&\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y) + (\gamma v_h + (1-\gamma)m - w)\beta\right)(1-\theta^{FR}) \\
\iff &((\gamma v_h + (1-\gamma)m - w)\beta - c_\theta)(\theta^F - \theta^{FR}) \\
&+ \left(b - (v_h - m)(\beta + (1-\theta_0)(1-\beta)) - \left(b - (v_h - m)(\beta + (1-\theta_0)(1-\beta))\right)^+ \mathbb{1}_{\left\{c_\theta \ge (v_h - m)(1-\beta)\right\}} \mathbb{1}_{\{y \le y_\theta\}}\right)^+ \ge \\
&\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (1-\beta)(w-y) + (\gamma v_h + (1-\gamma)m - w)\beta\right)(1-\theta^{FR}) \\
\iff &\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) \le g_3(m),
\end{aligned}$$

where the first equivalence is by definition of $\Pi_D^F$ and $\Pi_D^{FR}$, the second and third equivalences follow by rearranging terms, the fourth equivalence follows from the definition of $\theta^{FR}$, and the last equivalence follows by defining $g_3(m)$ as

$$\begin{aligned}
g_3(m) = &\frac{1}{(1-\beta)}\Bigg[(1-\beta)(w-y) - (\gamma v_h + (1-\gamma)m - w)\beta + \left((\gamma v_h + (1-\gamma)m - w)\beta - c_\theta\right)\left(\frac{\theta^F - \theta^{FR}}{1-\theta^{FR}}\right) \\
&+ \left(b - (v_h - m)(\beta + (1-\theta_0)(1-\beta))\right)^+ \frac{1 - \mathbb{1}_{\left\{y \le y_\theta, c_\theta \ge (v_h - m)(1-\beta)\right\}}}{1-\theta^{FR}}\Bigg],
\end{aligned}$$

where $y_\theta = w - \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) - (v_h - m)$. Similar to the previous two properties, assume first that $\gamma \le 1$, if $0 \le g_3(m) \le \frac{(1-\gamma)\beta(v_h-m)}{1-\beta}$, then there exists a $\hat{\alpha}_3$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_3}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_3(v_h-m)}}{1-\beta}\right) = g_3(m).$$

We then define $a_3$ as

$$a_3(m) = \begin{cases} \infty, & \text{if } g_3(m) \le 0 \\ \hat{\alpha}_3, & \text{if } 0 \le g_3(m) \le \frac{(1-\gamma)\beta(v_h - m)}{1-\beta} \\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h - m)}{1-\beta} \le g_3(m). \end{cases}$$

Again, the definition of $a_3(m)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity. This establishes Property 4.

*Proof of Property 5.* We first show that $a_1(m^F) = a_2(m^F)$. By definition $m^F$ satisfies $(\gamma v_h + (1-\gamma)m^F - w)\beta = c_\theta$. Therefore,

$$g_2(m^F) = w - y - (\gamma v_h + (1-\gamma)m^F - w)\frac{\beta}{1-\beta} = w - y - \frac{c_\theta}{(1-\beta)} = g_1(m^F).$$

We now show that $a_1(m) \ge a_2(m)$ if and only if $m \le m^F$ . Since $(\gamma v_h + (1-\gamma)m - w)\beta \ge c_\theta$ if and only if $m \ge m^F$, then $g_1(m) \ge g_2(m)$ if and only if $m \ge m^F$, hence $a_1(m) \le a_2(m)$ if and only if $m \ge m^F$.

We now prove the second statement in Property 5. Let $m^{MR} = v_h - \frac{(1-\beta)(w-y)-c_\theta}{(1-\gamma)\beta}$. Then, from the definition of $a_1(m)$ in Property 2 we have that if $m \ge m^{MR}$ then $a_1(m) = 0$, i.e., strategy MR dominates strategy FR for all $\alpha \ge a_1(m) = 0$. Moreover, the first statement in Property 5 then implies $m^{MR} \ge m^F$ while $c_\theta \le (1-\beta)(w-y)$ implies $m^{MR} \le v_h$. Furthermore, Property 1 then also implies that strategy MR dominates strategy F for all $\alpha \ge a_1(m) = 0$ and $m \ge m^F$.

Similarly, let $\bar{m}^{MR} = v_h - \frac{\beta v_h + (1-\beta)y - w}{(1-\gamma)\beta}$. Then, from the definition of $a_2(m)$ in Property 3 we have that if $m \ge \bar{m}^{MR}$ then $a_2(m) = \infty$, i.e., strategy MR dominates strategy M for all $\alpha \le a_2(m) = \infty$. Moreover, the first statement in Property 5 implies $\bar{m}^{MR} \ge m^F$, while $w \le \beta v_h + (1-\beta)y$ implies $\bar{m}^{MR} \le v_h$, establishing Property 5.

Hence, we have shown Properties 1 to 5, completing the proof of Proposition B.2. □

### B.5. Strategy Map when $w > \beta v_h + (1-\beta)y$ with Homogeneous ATP

The strategy map when $\beta(v_h - w) + (1-\beta)(y - w) < 0$ is simpler since only strategies $F$ and $M$ can be optimal, as shown in the following proposition.

**Proposition B.3** *Assume $\beta(v_h - w) + (1-\beta)(y - w) < 0$ then the distributor's optimal strategy in the resource allocation problem can be characterized as follows. Strategies $MR$ and $FR$ in Theorem 1 are always dominated. Moreover, consider the threshold $m^F$ on the customers' base ability to pay from Proposition B.2, then:*

- *Strategy $F$ in Theorem 1 is optimal if and only if the customers' base ability to pay $m$ is smaller than the threshold $m^F$. Specifically,*

$$\max(\Pi_D^M, \Pi_D^{FR}, \Pi_D^{MR}) \le \Pi_D^F \text{ if and only if } m \le m^F.$$

- *Strategy $M$ in Theorem 1 is optimal if the customers' base ability to pay $m$ is larger than the threshold $m^F$. Specifically,*

$$\max(\Pi_D^F, \Pi_D^{FR}, \Pi_D^{MR}) \le \Pi_D^M \text{ if and only if } m \ge m^F.$$

*Proof.* We build on the proof of Proposition B.2. Specifically, recall the threshold $m^F$ and the functions $a_1(m), a_2(m)$, and $a_3(m)$. Moreover, recall Properties 1 to 5.

We show that $\beta(v_h - w) + (1-\beta)(y-w) < 0$, or equivalently $w > \beta v_h + (1-\beta)y$, implies that strategies $MR$ and $FR$ are dominated for any $m$ and $\alpha \geq 0$.

First, we show that $w > \beta v_h + (1-\beta)y$ implies $\Pi_D^M(\alpha, m) \geq \Pi_D^{MR}(\alpha, m)$ for any $m$ and $\alpha \geq 0$. We have

$$\begin{aligned}
0 &> \beta v_h + (1-\beta)y - w \\
&= (\gamma v_h + (1-\gamma)m - w)\beta + (1-\gamma)\beta(v_h - m) - (1-\beta)(w-y) \\
&= (1-\gamma)\beta(v_h - m) - (1-\beta)g_2(m).
\end{aligned}$$

Hence, $g_2(m) > (1-\gamma)\beta(v_h - m)/(1-\beta)$. Thus, by definition $a_2(m) = 0$ for any $m$ and Property 3 implies $\Pi_D^M(\alpha, m) \geq \Pi_D^{MR}(\alpha, m)$ for any $m$ and $\alpha \geq 0$.

Second, we show that $w > \beta v_h + (1-\beta)y$ implies $\Pi_D^{MR}(\alpha, m) \geq \Pi_D^{FR}(\alpha, m)$ for any $m \geq m^F$ and $\alpha \geq 0$. Indeed, we have already shown that $w > \beta v_h + (1-\beta)y$ implies $a_2(m) = 0$ for any $m$, then Property 5 implies $a_1(m) = 0$ for any $m \geq m^F$, and from Property 2 we conclude $\Pi_D^{MR}(\alpha, m) \geq \Pi_D^{FR}(\alpha, m)$ for any $m \geq m^F$ and $\alpha \geq 0$.

Third, we show that $w > \beta v_h + (1-\beta)y$ implies $\Pi_D^F(\alpha, m) \geq \Pi_D^{FR}(\alpha, m)$ for any $m \leq m^F$ and $\alpha \geq 0$. Indeed,

$$\begin{aligned}
\Pi_D^{FR}(\alpha, m) =& (\gamma v_h + (1-\gamma)m - w)\beta + \min\left((v_h - m)(\beta + (1-\theta^{FR})(1-\beta)), b - c_\theta(\theta^{FR} - \theta_0)\right) \\
&+ \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right) - (1-\beta)(w-y)\right)(1-\theta^{FR}) \\
=& (\gamma v_h + (1-\gamma)m - w)\theta^{FR}\beta + \min\left((v_h - m)(\beta + (1-\theta^{FR})(1-\beta)), b - c_\theta(\theta^{FR} - \theta_0)\right) \\
&+ \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right) - (1-\beta)(w-y) + \beta(\gamma v_h + (1-\gamma)m - w)\right)(1-\theta^{FR}) \\
\leq& (\gamma v_h + (1-\gamma)m - w)\theta^{FR}\beta + b - c_\theta(\theta^{FR} - \theta_0) + \left(\beta(v_h - w) - (1-\beta)(w-y)\right)(1-\theta^{FR}) \\
\leq& (\gamma v_h + (1-\gamma)m - w)\theta^{FR}\beta + b - c_\theta(\theta^{FR} - \theta_0) \\
\leq& (\gamma v_h + (1-\gamma)m - w)\theta^{F}\beta + b - c_\theta(\theta^{F} - \theta_0) \\
=& \Pi_D^F(\alpha, m),
\end{aligned}$$

where the first inequality follows from $\min\left((v_h - m)(\beta + (1-\theta^{FR})(1-\beta)), b - c_\theta(\theta^{FR} - \theta_0)\right) \leq b - c_\theta(\theta^{FR} - \theta_0)$, and since $\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right)$ is decreasing in $\alpha > 0$ and taking the limit $\alpha \to 0$. The second inequality follows from the assumption $w > \beta v_h + (1-\beta)y$. The last inequality follows since $\theta^F \geq \theta^{FR}$ from Lemma B.1 below, and $(\gamma v_h + (1-\gamma)m - w)\beta - c_\theta \geq 0$ for any $m \leq m^F$. Hence, we conclude that $w > \beta v_h + (1-\beta)y$ implies $\Pi_D^F(\alpha, m) \geq \Pi_D^{FR}(\alpha, m)$ for any $m \leq m^F$ and $\alpha \geq 0$.

Last, Properties 2, 3, and 5 imply $\Pi_D^{MR} \leq \max(\Pi_D^F, \Pi_D^{FR})$ for any $m \leq m^F$ and $\alpha \geq 0$.

Putting all these observations together, we conclude that $w > \beta v_h + (1-\beta)y$ implies $\max(\Pi_D^F(\alpha, m), \Pi_D^M(\alpha, m)) \geq \max(\Pi_D^{FR}(\alpha, m), \Pi_D^{MR}(\alpha, m))$ for any $m$ and $\alpha \geq 0$.

Namely, strategies $MR$ and $FR$ are dominated for any $m$ and $\alpha \geq 0$, and it is enough to compare between strategies $F$ and $M$. Finally, Property 1 states $\Pi_D^F(\alpha, m) \geq \Pi_D^M(\alpha, m)$ if and only if $m \geq m^F$, concluding the proof. □

Proposition B.3 shows that for products that are socially efficient to try by high valuation consumers only the distributor generally implements strategy $M$ or $F$ in Theorem 1, depending on whether the consumers' base ability to pay $m$ is high or low, respectively.

### B.6. Strategy Map when $c_\theta > (1-\beta)(w-y)$ with Homogeneous ATP

The strategy map when $c_\theta > (1-\beta)(w-y)$ is simpler since only strategies $F$ and $FR$ can be optimal, as shown in the following proposition.

**Proposition B.4** *Assume $c_\theta > (1-\beta)(w-y)$ then the distributor's optimal strategy in the resource allocation problem can be characterized as follows. Strategies $M$ and $MR$ in Theorem 1 are always dominated for any $m \le v_h$. Moreover, consider the function $a_3(m)$ from Proposition B.2, then:*

- *Strategy $F$ in Theorem 1 is optimal if and only if the customers' risk aversion parameter is large enough. Specifically,*
$$\max(\Pi_D^M, \Pi_D^{FR}, \Pi_D^{MR}) \le \Pi_D^F \text{ if and only if } \alpha \ge a_3(m).$$
- *Strategy $FR$ in Theorem 1 is optimal if the customers' the customers' risk aversion parameter is small enough. Specifically,*
$$\max(\Pi_D^F, \Pi_D^M, \Pi_D^{MR}) \le \Pi_D^{FR} \text{ if and only if } \alpha \le a_3(m).$$

*Proof.* We build on the proof of Proposition B.2. Specifically, recall the threshold $m^F$ and the functions $a_1(m), a_2(m)$, and $a_3(m)$. Moreover, recall Properties 1 to 5.

We show that $c_\theta > (1-\beta)(w-y)$ implies that $M$ and $MR$ are dominated for any $m \le v_h$ and $\alpha \ge 0$.

First, we show that $c_\theta > (1-\beta)(w-y)$ implies $\Pi_D^{FR}(\alpha,m) \ge \Pi_D^{MR}(\alpha,m)$ for any $m$ and $\alpha \ge 0$. Indeed, $c_\theta > (1-\beta)(w-y)$ implies $g_1(m) < 0$. Thus, by definition $a_1(m) = 0$ for any $m$ and Property 2 implies $\Pi_D^{FR}(\alpha,m) \ge \Pi_D^{MR}(\alpha,m)$ for any $m$ and $\alpha \ge 0$.

Second, we show that $c_\theta > (1-\beta)(w-y)$ implies $\max(\Pi_D^F(\alpha,m), \Pi_D^{FR}(\alpha,m)) \ge \Pi_D^M(\alpha,m)$ for any $m \le v_h$ and $\alpha \ge 0$. We consider two cases. Indeed, first assume $\beta(v_h - w) \le c_\theta$ which is equivalent to $m^F \ge v_h$. Thus, Property 1 implies $\Pi_D^F(\alpha,m) \ge \Pi_D^M(\alpha,m)$ for any $m \le v_h$, completing the proof in the first case. Now assume $\beta(v_h - w) > c_\theta$, since by assumption $c_\theta > (1-\beta)(w-y)$ we conclude $(1-\beta)(w-y) < \beta(v_h - w)$, which implies $g_2(m) < 0$ for all $m \le v_h$. Thus, by definition $a_2(m) = \infty$ for any $m \le v_h$ and Property 3 implies $\Pi_D^{MR}(\alpha,m) \ge \Pi_D^M(\alpha,m)$ for any $m \le v_h$ and $\alpha \le \infty$. Since we have already shown $\Pi_D^{FR}(\alpha,m) \ge \Pi_D^{MR}(\alpha,m)$ for any $m$ and $\alpha \ge 0$, we conclude that $\beta(v_h - w) > c_\theta$ and $c_\theta > (1-\beta)(w-y)$ imply $\Pi_D^{FR}(\alpha,m) \ge \Pi_D^M(\alpha,m)$ for any $m \le v_h$ and $\alpha \ge 0$, completing the proof of the second case.

Putting all these observations together, we conclude that $c_\theta > (1-\beta)(w-y)$ implies $\max(\Pi_D^F(\alpha,m), \Pi_D^{FR}(\alpha,m)) \ge \max(\Pi_D^M(\alpha,m), \Pi_D^{MR}(\alpha,m))$ for any $m \le v_h$ and $\alpha \ge 0$.

Namely, strategies $M$ and $MR$ are dominated for any $m$ and $\alpha \ge 0$, and it is enough to compare between strategies $F$ and $FR$. Finally, Property 4 states $\Pi_D^F(\alpha,m) \ge \Pi_D^{FR}(\alpha,m)$ if and only if $\alpha \ge a_3(m)$, concluding the proof. □

### B.7. Proof of Corollary 1

*Proof.* The first statement in the corollary is the same as the first statement in Proposition B.2 taking $a_F(m) = a_3(m)$, while the second statement in the corollary follows directly from the last statement in Proposition B.2.

The third statement in the corollary is the same as Proposition B.3, while the last statement in the corollary is the same as Proposition B.4, completing the proof. □

### B.8. Consumer Surplus Analysis with Homogeneous ATP

We now state and prove a useful auxiliary lemma.

**Lemma B.1** *Let $\theta^i$, $i \in \{F, M, FR, MR\}$ be the education level of each strategy from Theorem 1. Then, under Assumption 1,*

$$\theta^M = \theta^{MR} \geq \theta^F \geq \theta^{FR}. \tag{18}$$

*Proof.* From their definition, in the proof of Theorem 1, we have $\theta^M = \theta^{MR}$, where they are equal to the natural upper bound on $\theta$. Moreover, from Assumption 1 it follows that $\theta^M = \theta^{MR} \geq \theta^F$. We now argue that $\theta^F \geq \theta^{FR}$. In fact, if $\theta^{FR} = \theta_0$ then $\theta^F \geq \theta^{FR}$ follows trivially from their definition, in the proof of Theorem 1. If $\theta^{FR} > \theta_0$ then by definition again we must have $c_\theta > (v_h - m)(1-\beta) > 0$, $c_\theta(\theta^{FR} - \theta_0) + (v_h - m)(\beta + (1-\theta^{FR})(1-\beta)) = b$, and $c_\theta(\theta^F - \theta_0) + (v_h - m)\theta^F(1-\beta) = b$, hence we conclude $\theta^F \geq \theta^{FR}$ in this case as well, completing the proof. □

### B.9. Proof of Proposition 2

*Proof.* Recall, from Proposition B.1, that

$$CS^i = (v_h - m)\theta^i\beta, \quad i \in \{F, M\},$$

and

$$CS^i = (v_h - m)\beta - \frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m)}}{1-\beta}\right)\left(1-\theta^i\right), \quad i \in \{FR, MR\}.$$

We now show that $CS^{MR}$ is the largest consumer surplus of the non-dominated distributor's strategies from Theorem 1. In particular, since $\theta^{MR} \geq \theta^{FR}$ from Lemma B.1 then $CS^{MR} \geq CS^{FR}$.

Moreover,

$$CS^{MR} \geq (v_h - m)\theta^{MR}\beta = (v_h - m)\theta^M\beta = CS^M \geq (v_h - m)\theta^F\beta = CS^F,$$

where the first inequality follows since $\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right)$ is decreasing in $\alpha > 0$ and taking the limit $\alpha \to 0$. The second inequality follows since, from Lemma B.1, $\theta^M = \theta^{MR} \geq \theta^F$. Hence, we conclude $CS^{MR} \geq \max(CS^F, CS^M, CS^{FR})$.

Finally, we show that there exists $a_F^{cs}(m)$ such that $CS^F$ is the smallest consumer surplus of the non-dominated distributor's strategies from Theorem 1 if and only if $\alpha \geq a_F^{cs}(m)$. We have already shown $CS^{MR} \geq CS^M \geq CS^F$. We now show $CS^{FR} \geq CS^F$ if and only if $\alpha \geq a_4(m)$. Then, together with the first statement in Proposition B.2, we will conclude $\Pi_D^F \geq \max(\Pi_D^M, \Pi_D^{FR}, \Pi_D^{MR})$ and $CS^F \leq \min(CS^{MR}, CS^M, CS^{FR})$ if and only if $m \leq \bar{m}$ and $\alpha \geq \max(a_F(m), a_F^{cs}(m))$, where $a_F(m) = a_3(m)$. In fact, we have

$$CS^{FR} \geq CS^F \iff \frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) \leq \beta(v_h - m)\frac{(1-\theta^F)}{(1-\theta^{FR})},$$

where the term $\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $[0, \beta(v_h - m)]$. Since from Lemma B.1 we have $\theta^F \geq \theta^{FR}$, then $0 \leq \beta(v_h - m)\frac{(1-\theta^F)}{(1-\theta^{FR})} \leq \beta(v_h - m)$. Moreover, since $\theta^F$ is independent of $\alpha$ and $\theta^{FR}$ is non-decreasing in $\alpha$ it follows that there exists $a_F^{cs}(m)$ such that

$$\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right) \leq \beta(v_h - m)\frac{(1-\theta^F)}{(1-\theta^{FR})} \iff \alpha \geq a_F^{cs}(m),$$

completing the proof. □

### B.10. Results from Section 4.3 on Free Returns with Homogeneous ATP

We first specify the model under the assumption that the distributor commits to a free return policy.

With free returns, the distributor's pricing problem then becomes

$$\begin{aligned}\Pi_D^*(x,\theta) = \max_{c,z}\ & (c-w)\cdot B(p^*-x,p^*-x,\theta)-(z-y)\cdot R(p^*-x,p^*-x,\theta)+\gamma\cdot CS(p^*-x,p^*-x,\theta)\\ \text{s.t. } & \{p^*\}\in\arg\max_p (p-c)\cdot \mathrm{B}(p-x,p-x,\theta)-(p-x-z)\cdot\mathrm{R}(p-x,p-x,\theta)\quad (IC) \\ & \mathrm{B}(p^*-x,p^*-x,\theta)-(p^*-x-z)\cdot\mathrm{R}(p^*-x,p^*-x,\theta)\geq 0.\quad (IR)\end{aligned} \tag{19}$$

and the distributor's resource allocation problem remains unchanged.

Note that the same outcomes can be achieved in the original model by assuming that customers are extremely risk averse, by taking the limit as $\alpha\to\infty$, making customers max-min utility optimizers.

With the formulation at hand, we adapt Proposition B.1 to characterize the distributor's pricing strategy in this restricted setup, in the next proposition.

**Proposition B.5** *Consider any customer education level $\theta\in[0,1)$, ATP $m\in[w,v_h]$, and subsidy $x\in[0,v_h-m]$. Assume the distributor commits to a free returns policy. Then, the equilibrium customer price is $p^f=m+x$ and the refund is $r^f=m$. The equilibrium retailer refund is $z^f=m$ and the retailer price $c^f$ is*

$$c^f = m + x + \frac{(1-\theta)(1-\beta)\theta\beta}{\theta\beta+(1-\theta)}(z^f-\bar{z}^f).$$

*The consumer surplus is $CS^f=(v_h-m)\beta$, retailer's profit is 0, and the distributor's profit is*

$$\Pi_D^f(x,\theta)=(m+x-w)\beta+(1-\beta)(x-w+y)(1-\theta)+\gamma CS^f.$$

Proposition B.5 is a special case of Proposition B.1 when $\alpha\to\infty$. Therefore, we omit the proof.

*Proof of Proposition 3.* The proof of the first part of the proposition is the same as the proof of Theorem 1 for the special case when $\alpha\to\infty$. Therefore, we only focus on the expressions that change. Specifically,

$$\Pi_D^{MR^f} = (\gamma v_h+(1-\gamma)m-w)\beta+(b-c_\theta(1-\theta_0))^+-(1-\beta)(w-y)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+. \tag{20}$$

To simplify the notation, recall the function $\frac{1}{x^{++}} := \begin{cases}0 & \text{if } x\leq 0\\ \frac{1}{x} & \text{if } x>0.\end{cases}$ Then,

$$\theta^{FR^f}=\theta_0+\frac{\left(b-(v_h-m)(\beta+(1-\theta_0)(1-\beta))\right)^+}{\left(c_\theta-(v_h-m)(1-\beta)\right)^{++}}\mathbb{1}_{\{(v_h-m)\leq(w-y)\}},$$

and

$$\begin{aligned}\Pi_D^{FR^f} = & \left(\gamma v_h+(1-\gamma)m-w\right)\beta+\min\left((v_h-m)(\beta+(1-\theta_0)(1-\beta)),b\right)\\ & -(1-\beta)\left(w-y\right)\left(1-\theta_0\right)+\left(w-y-(v_h-m)\right)^+(1-\beta)\frac{\left(b-(v_h-m)(\beta+(1-\theta_0)(1-\beta))\right)^+}{\left(c_\theta-(v_h-m)(1-\beta)\right)^{++}}.\end{aligned} \tag{21}$$

The proof of the second part of the proposition is the same as the proof of Property 2 in the proof of Proposition B.2 for the special case when $\alpha\to\infty$. Namely,

$$\Pi_D^{MR^f}\geq\Pi_D^{FR^f} \iff \frac{d\Pi_D^f(1)}{d\theta}=(1-\beta)(w-y)-c_\theta\geq 0,$$

completing the proof. □

## Appendix C: Results from Section 5 on Heterogeneous ATP

**Proposition C.1** *Consider any consumer education level $\theta \in [0,1)$, ATPs $m_l, m_h \in [w, v_h]$, $m_l < m_h$, and subsidy $x \in [0, v_h - m_l]$. Then, the distributor's optimal objective is*

$$\Pi_D^*(x,\theta) = \max\left\{\Pi_D^A(x,\theta), \Pi_D^{AR}(x,\theta), \Pi_D^S(x,\theta), \Pi_D^{SR}(x,\theta)\right\}. \tag{22}$$

*Where $\Pi_D^i(x,\theta)$ $i \in \{A, AR, S, SR\}$ each correspond to the distributor's profits in a non-dominated strategy. Specifically, these strategies are characterized by:*

*(A) Target all informed customers without product returns. The customer price is $p^A = m_l + x$ and refund is $r^A = 0$. The retailer's price and refund are $c^A = m_l + x - \frac{\hat{O}_R^A}{\theta\beta}$ and $z^A = 0$, respectively. The customer surplus is $CS^A = (v_h - m_l)\theta\beta$. The retailer's profit is $\Pi_R^A = \hat{O}_R^A$ while the distributor's profit is*

$$\Pi_D^A(x,\theta) = (m_l + x - w)\theta\beta - \hat{O}_R^A + \gamma CS^A,$$

*where*

$$\hat{O}_R^A = \begin{cases} \frac{\lambda\theta\beta(m_h - m_l)}{1-\lambda} & \text{if } \theta\beta \le \lambda(\theta\beta + (1-\theta)) \\ \max\left\{\frac{\lambda\theta\beta(m_h-m_l)}{1-\lambda}, \lambda\theta\beta\frac{r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h - m_l)(\theta\beta + (1-\theta))}{\lambda(\theta\beta+(1-\theta)) - \theta\beta}\right\} & \text{if } \theta\beta > \lambda(\theta\beta + (1-\theta)). \end{cases}$$

*This strategy can be sustained in equilibrium if and only if $\hat{O}_R^A \le UB^A$, where*

$$UB^A = \begin{cases} \min\left\{r_\alpha(m_l)\theta\beta(1-\beta), \lambda\theta\beta\frac{(r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h-m_l)(\theta\beta+(1-\theta)))}{\lambda(\theta\beta+(1-\theta))-\theta\beta}\right\} & \text{if } \theta\beta \le \lambda(\theta\beta + (1-\theta)) \\ r_\alpha(m_l)\theta\beta(1-\beta) & \text{if } \theta\beta > \lambda(\theta\beta+(1-\theta)). \end{cases}$$

*(AR) Target all informed and uninformed customers with product returns. The customer price is $p^{AR} = m_l + x$ and the refund is $r^{AR} = r_\alpha(m_l) = \max\left(0, m_l - \frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right)\right)$. Let $\bar{z}^{AR} = r_\alpha(m_l) - \frac{\hat{O}_R^{AR}}{(1-\beta)\theta\beta}$. Then, without loss of generality, the equilibrium retailer refund is $z^{AR} = \max(0, \bar{z}^{AR})$ and the retailer price is $c^{AR} = m_l + x - \frac{\hat{O}_R^{AR}}{\theta\beta} + \frac{(1-\theta)(1-\beta)}{\theta\beta+(1-\theta)}(z^{AR} - \bar{z}^{AR})$. The consumer surplus is $CS^{AR} = (v_h - m_l)\beta - (m_l - r_\alpha(m_l))(1-\theta)(1-\beta)$, the retailer's profit is $\Pi_R^{AR} = \hat{O}_R^{AR}$, and the distributor's profit is*

$$\Pi_D^{AR}(x,\theta) = (m_l + x - w)\beta + (m_l + x - r_\alpha(m_l) - w + y)(1-\beta)(1-\theta) - \hat{O}_R^{AR} + \gamma CS^{AR},$$

*where*

$$\hat{O}_R^{AR} = \frac{\lambda}{1-\lambda}((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\theta)(1-\beta))^+.$$

*This strategy can be sustained in equilibrium if and only if*

$$\hat{O}_R^{AR} \le (m_l + x)\beta + (m_l + x - r_\alpha(m_l))(1-\theta)(1-\beta).$$

*(S) Target only informed customers with high ATP without product returns. The customer price is $p^S = m_h + x$ and refund is $r^S = 0$. The retailer's price and refund are $c^S = m_h + x$ and $z^S = 0$, respectively. The customer surplus is $CS^S = (v_h - m_h)\lambda\theta\beta$. The retailer attains no profit, $\Pi_R^S = 0$, while the distributor's profit is*

$$\Pi_D^S(x,\theta) = (m_h + x - w)\lambda\theta\beta + \gamma CS^S.$$

*This strategy can always be sustained in equilibrium.*

*(SR) Target both informed and uninformed customers with product returns. The customer price is* $p^{SR} = m_h + x$ *and the refund is* $r^{SR} = r_\alpha(m_h) = \max\left(0, m_h - \frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right)\right)$. *The equilibrium retailer price is* $c^{SR} = m_h + x$, *and the retailer refund is* $z^{SR} = r_\alpha(m_h)$. *The consumer surplus is* $CS^{SR} = (v_h - m_h)\lambda\beta - (m_h - r_\alpha(m_h))\lambda(1-\theta)(1-\beta)$, *the retailer attains no profit,* $\Pi_R^{SR} = 0$, *and the distributor's profit is*

$$\Pi_D^{SR}(x,\theta) = (m_h + x - w)\lambda\beta + (m_h + x - r_\alpha(m_h) - w + y)\lambda(1-\beta)(1-\theta) + \gamma CS^{SR}.$$

*This strategy can always be sustained in equilibrium.*

*Proof.* First, we analyze the retailer's pricing problem and show that the optimal nominal price to consumers and refund are such that $(p^*, r^*) \in \{(m_l + x, 0), (m_l + x, r_\alpha(m_l)), (m_h + x, 0)(m_h + x, r_\alpha(m_h))\}$. Note that the expected fraction of customers who purchase the product is $\mathrm{B}(p,r,\theta) = (\lambda \mathbb{1}_{\{p-x\le m_h\}} + (1-\lambda)\mathbb{1}_{\{p-x\le m_l\}})\left(\theta\beta + (1-\theta)\mathbb{1}_{\left\{r \ge r_\alpha(p-x)\right\}}\right)$. Since all indicator functions are decreasing in $p$ (cf. Equation 7), it follows that, for a given $r$, the optimal retailer's price $p^*$ is such that $p^* \in \{m_l + x, m_h + x, p_\alpha(r) + x\}$.

Then, if $p^* = m_l + x$ the retailer's profit function is

$$\Pi_R\left(m_l + x, r\right) = (m_l + x - c)\left(\theta\beta + (1-\theta)\mathbb{1}_{\left\{r \ge r_\alpha(m_l)\right\}}\right) - (r - z)(1-\theta)(1-\beta)\mathbb{1}_{\left\{r\ge r_\alpha(m_l)\right\}}.$$

The profit function above is constant for $r < r_\alpha(m_l)$, has an increasing or decreasing step at $r = r_\alpha(m_l)$, and is linear decreasing for $r > r_\alpha(m_l)$. Therefore, $r^* \in \{0, r_\alpha(m_l)\}$ when $p^* = m_l + x$.

Similarly, if $p^* = m_h + x$ the retailer's profit function is

$$\Pi_R\left(m_h + x, r\right) = (m_h + x - c)\lambda\left(\theta\beta + (1-\theta)\mathbb{1}_{\left\{r \ge r_\alpha(m_h)\right\}}\right) - (r - z)\lambda(1-\theta)(1-\beta)\mathbb{1}_{\left\{r\ge r_\alpha(m_h)\right\}}.$$

As before, the profit function above is constant for $r < r_\alpha(m_h)$, has an increasing or decreasing step at $r = r_\alpha(m_h)$, and is linear decreasing for $r > r_\alpha(m_h)$. Therefore, $r^* \in \{0, r_\alpha(m_h)\}$ when $p^* = m_h + x$.

Conversely, if $p^* = p_\alpha(r) + x$ the retailer's profit function is

$$\begin{aligned}\Pi_R\left(p_\alpha(r) + x, r\right) = (p_\alpha(r) + x - c)&\left(\lambda\mathbb{1}_{\left\{p_\alpha(r)\le m_h\right\}} + (1-\lambda)\mathbb{1}_{\left\{p_\alpha(r)\le m_l\right\}}\right)(\theta\beta + (1-\theta))\\ - (r - z)&\left(\lambda\mathbb{1}_{\left\{p_\alpha(r)\le m_h\right\}} + (1-\lambda)\mathbb{1}_{\left\{p_\alpha(r)\le m_l\right\}}\right)(1-\theta)(1-\beta).\end{aligned}$$

The profit function above is zero if $p_\alpha(r) > m_h$ since customers cannot afford the product. When $p_\alpha(r) \le m_l$ or $m_l < p_\alpha(r) \le m_h$ (equivalently when $r \le r_\alpha(m_l)$ or $r_\alpha(m_l) < r \le r_\alpha(m_h)$) the profit function is increasing in $r$. To show this, first note that when $r \le r_\alpha(m_l)$ we have

$$\frac{\partial \Pi_R\left(p_\alpha(r) + x, r\right)}{\partial r} = p'_\alpha(r)\left(\theta\beta + (1-\theta)\right) - (1-\theta)(1-\beta) = \frac{\theta\beta + (1-\theta)}{r'_\alpha(p_\alpha(r))} - (1-\theta)(1-\beta),$$

where the second equality comes from the fact that $p_\alpha = r_\alpha^{-1}$. Then, from Equation 7, for any $p \in [0, v_h + x]$, we have that

$$r'_\alpha(p) \le 1 + \frac{\beta}{1-\beta} \le 1 + \frac{\beta}{(1-\theta)(1-\beta)} = \frac{\theta\beta + (1-\theta)}{(1-\theta)(1-\beta)}.$$

The first inequality comes from noting that $r'_\alpha(p)$, given in Equation 7, is increasing in $p$ for $p \in [0, v_h + x]$ and that $r'_\alpha(v_h + x) = 1 + \frac{\beta}{1-\beta}$. The analysis for the case when $r_\alpha(m_l) < r \le r_\alpha(m_h)$ is analogous. It follows that

$\frac{\partial \Pi_R(p_\alpha(r)+x,r)}{\partial r} \geq 0$ when $r \leq r_\alpha(m_l)$ or $r_\alpha(m_l) < r \leq r_\alpha(m_h)$, thus $r^* \in \{r_\alpha(m_l), r_\alpha(m_h)\}$ when $p^* = p_\alpha(r)+x$, hence $p^* \in \{m_l+x, m_h+x\}$ in this case as well.

Hence, we conclude that $(p^*, r^*) \in \{(m_l+x, 0), (m_l+x, r_\alpha(m_l)), (m_h+x, 0), (m_h+x, r_\alpha(m_h))\}$, fully characterizing the equilibrium behavior of the retailer. With the retailer's equilibrium behavior in hand, we now characterize the distributor's equilibrium pricing and refund strategies.

Strategy $(S)$: First assume that the distributor is interested in inducing the retailer to target informed consumers with high valuation and high ATP, i.e. set $p^* = m_h + x$ and $r^* = 0$. In this case, the distributor's problem can be written as

$$\begin{aligned}
\max_{c,z\geq 0}\ & (c-w)\lambda\theta\beta + \gamma(v_h - m_h)\lambda\theta\beta \\
\text{s.t. } & (m_h + x - c)\lambda\theta\beta \geq (m_l + x - c)\,\theta\beta && (IC_{SA}) \\
& (m_h + x - c)\lambda\theta\beta \geq (m_l + x - c)\,(\theta\beta + (1-\theta)) - (r_\alpha(m_l) - z)(1-\theta)(1-\beta) && (IC_{SAR}) \\
& (m_h + x - c)\lambda\theta\beta \geq (m_h + x - c)\,\lambda\,(\theta\beta + (1-\theta)) - (r_\alpha(m_h) - z)(1-\theta)\lambda(1-\beta) && (IC_{SSR}) \\
& (m_h + x - c)\lambda\theta\beta \geq 0. && (IR_S)
\end{aligned}$$

The $(IC_{SA})$, $(IC_{SAR})$, $(IC_{SSR})$ constraints above imply that the retailer sets $p^* = m_h + x$ and $r^* = 0$. Moreover, note that $(IC_{SAR})$ is redundant since it is implied by $(IC_{SA})$ and $(IC_{SSR})$ combined. Indeed, $(IC_{SSR})$ implies $(IC_{AAR})$, the latter combined with $(IC_{SA})$ implies $(IC_{SAR})$.

The objective function is increasing in $c$ and independent of $z$, with an upper bound $c \leq m_h + x$ given by the $(IR_S)$ constraint (note that the $(IC_{SA})$ and $(IC_{SSR})$ are trivially satisfied for any $m_l \leq m_h$). Hence, in particular $c^* = m_h + x$ and $z^* = 0$, and strategy $(S)$ can always be induced by the distributor.

Strategy $(SR)$: Now assume that the distributor is interested in inducing the retailer to target all informed and uninformed consumers with high ATP, i.e. set $p^* = m_h + x$ and $r^* = r_\alpha(m_h)$. In this case, the distributor's problem can be written as

$$\begin{aligned}
\max_{c,z\geq 0}\ & (c-w)\lambda\,(\theta\beta + (1-\theta)) - (z-y)\lambda(1-\theta)(1-\beta) + \gamma((v_h - m_h)\beta - (m_h - r_\alpha(m_h))\lambda(1-\theta)(1-\beta)) \\
\text{s.t. } & (m_h + x - c)\,\lambda\,(\theta\beta + (1-\theta)) - (r_\alpha(m_h) - z)\lambda(1-\theta)(1-\beta) \geq (m_l + x - c)\theta\beta && (IC_{SRA}) \\
& (m_h + x - c)\,\lambda\,(\theta\beta + (1-\theta)) - (r_\alpha(m_h) - z)\lambda(1-\theta)(1-\beta) \geq \\
& (m_l + x - c)\,(\theta\beta + (1-\theta)) - (r_\alpha(m_l) - z)(1-\theta)(1-\beta) && (IC_{SRAR}) \\
& (m_h + x - c)\,\lambda\,(\theta\beta + (1-\theta)) - (r_\alpha(m_h) - z)\lambda(1-\theta)(1-\beta) \geq (m_h + x - c)\lambda\theta\beta && (IC_{SRS}) \\
& (m_h + x - c)\,\lambda\,(\theta\beta + (1-\theta)) - (r_\alpha(m_h) - z)\lambda(1-\theta)(1-\beta) \geq 0. && (IR_{SR})
\end{aligned}$$

The $(IC_{SRA})$, $(IC_{SRAR})$, $(IC_{SRS})$ constraints above imply that the retailer sets $p^* = m_h + x$ and $r^* = r_\alpha(m_h)$.

Note that the objective is increasing in $c$ and decreasing in $z$, leading to $c^* = m_h + x$ and $z^* = r_\alpha(m_h)$, i.e., the constraints $(IC_{SRS})$ and $(IR_{SR})$ are tight. Note that constraint $(IC_{SRA})$ is trivially satisfied for any $m_l \leq m_h$, while constraint $(IC_{SRAR})$ becomes redundant in this solution since it is implied by the combination of $(IC_{SRA})$, and $(IC_{SRS})$ being tight. Specifically, since $(IC_{SRS})$ is tight it is equivalent to $(IC_{SSR})$, while $(IC_{SRA})$ is equivalent to $(IC_{SA})$ and $(IC_{SRAR})$ is equivalent to $(IC_{SAR})$. Thus, since $(IC_{SSR})$ implies

$(IC_{AAR})$, and the latter combined with $(IC_{SA})$ imply $(IC_{SAR})$ we conclude that $(IC_{SRAR})$ is guaranteed to be satisfied in this solution. Thus, strategy $(SR)$ can always be induced by the distributor.

Strategy $(AR)$: Assume that the distributor is interested in inducing the retailer to target all informed and uninformed consumers, i.e. set $p^* = m_l + x$ and $r^* = r_\alpha(m_l)$. In this case, the distributor's problem can be written as

$$\begin{aligned}
\max_{c,z\geq 0}\ & (c-w)(\theta\beta+(1-\theta)) - (z-y)(1-\theta)(1-\beta) + \gamma((v_h-m_l)\beta - (m_l - r_\alpha(m_l))(1-\theta)(1-\beta)) \\
\text{s.t. }\ & (m_l+x-c)(\theta\beta+(1-\theta)) - (r_\alpha(m_l)-z)(1-\theta)(1-\beta) \geq (m_l+x-c)\theta\beta \qquad (IC_{ARA}) \\
& (m_l+x-c)(\theta\beta+(1-\theta)) - (r_\alpha(m_l)-z)(1-\theta)(1-\beta) \geq (m_h+x-c)\lambda\theta\beta \qquad (IC_{ARS}) \\
& (m_l+x-c)(\theta\beta+(1-\theta)) - (r_\alpha(m_l)-z)(1-\theta)(1-\beta) \geq \\
& (m_h+x-c)\lambda(\theta\beta+(1-\theta)) - (r_\alpha(m_h)-z)\lambda(1-\theta)(1-\beta) \qquad (IC_{ARSR}) \\
& (m_l+x-c)(\theta\beta+(1-\theta)) - (r_\alpha(m_l)-z)(1-\theta)(1-\beta) \geq 0. \qquad (IR_{AR})
\end{aligned}$$

The $(IC_{ARA})$, $(IC_{ARS})$, $(IC_{ARSR})$ constraints above imply that the retailer sets $p^* = m_l + x$ and $r^* = r_\alpha(m_l)$. Moreover, note that $(IC_{ARS})$ is redundant, implied by $(IC_{ARA})$ and $(IC_{ARSR})$ combined. Indeed, $(IC_{ARA})$ implies $(IC_{SRS})$, the latter combined with $(IC_{ARSR})$ implies $(IC_{ARS})$.

Let $\hat{O}_R^{AR} = \lambda/(1-\lambda)((m_h-m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\theta)(1-\beta))^+$. Note that the objective is increasing in $c$ and decreasing in $z$, leading to $c^* = m_l + x - \frac{\hat{O}_R^{AR}}{\theta\beta} + \frac{(1-\theta)(1-\beta)}{\theta\beta+(1-\theta)}(z^* - \bar{z}^{AR})$ and $z^* = \max(0, \bar{z}^{AR})$, where $\bar{z}^{AR} = r_\alpha(m_l) - \frac{\hat{O}_R^{AR}}{(1-\beta)\theta\beta}$, i.e., the constraints $(IC_{ARA})$ and one of $(IR_{AR})$ or $(IC_{ARSR})$ are tight when $\bar{z}^{AR} \geq 0$, or the constraint $(IC_{ARA})$ is redundant and one of $(IR_{AR})$ or $(IC_{ARSR})$ is tight when $\bar{z}^{AR} < 0$, as long as

$$\hat{O}_R^{AR} \leq (m_l+x)\beta + (m_l + x - r_\alpha(m_l))(1-\theta)(1-\beta), \tag{23}$$

that is, as long as $c^* \geq 0$. Alternatively, if Equation 23 is not satisfied then the problem is infeasible, i.e., strategy $(AR)$ cannot be induced by the distributor.

Strategy $(A)$: Finally, assume that the distributor is interested in inducing the retailer to target all informed consumers, i.e. set $p^* = m_l + x$ and $r^* = 0$. In this case, the distributor's problem can be written as

$$\begin{aligned}
\max_{c,z\geq 0}\ & (c-w)\theta\beta + \gamma(v_h - m_l)\theta\beta \\
\text{s.t. }\ & (m_l+x-c)\theta\beta \geq (m_l+x-c)(\theta\beta+(1-\theta)) - (r_\alpha(m_l)-z)(1-\theta)(1-\beta) \qquad (IC_{AAR}) \\
& (m_l+x-c)\theta\beta \geq (m_h+x-c)\lambda\theta\beta \qquad (IC_{AS}) \\
& (m_l+x-c)\theta\beta \geq (m_h+x-c)\lambda(\theta\beta+(1-\theta)) - (r_\alpha(m_h)-z)\lambda(1-\theta)(1-\beta) \qquad (IC_{ASR}) \\
& (m_l+x-c)\theta\beta \geq 0. \qquad (IR_A)
\end{aligned}$$

The $(IC_{ASR})$, $(IC_{AS})$, $(IC_{AAR})$ constraints above imply that the retailer sets $p^* = m_l + x$ and $r^* = 0$.

Assume first that $\theta\beta \leq \lambda(\theta\beta + (1-\theta))$. Let $\hat{O}_R^{A1} = \frac{\lambda\theta\beta}{1-\lambda}(m_h - m_l)$. The objective is increasing in $c$ and independent of $z$, with an upper bound $c \leq m_l + x - \frac{\hat{O}_R^{A1}}{\theta\beta}$ given by the $(IR_A)$ and $(IC_{AS})$ constraints as long as

$$\hat{O}_R^{A1} \leq \min\left\{ r_\alpha(m_l)\theta\beta(1-\beta), \lambda\theta\beta\frac{(r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h-m_l)(\theta\beta+(1-\theta)))}{\lambda(\theta\beta+(1-\theta)) - \theta\beta} \right\}, \tag{24}$$

i.e., as long as constraints $(IC_{AAR})$ and $(IC_{ASR})$ are satisfied. Hence, in particular $c^* = m_l + x - \frac{\hat{O}_R^{A1}}{\theta\beta}$ and $z^* = 0$, leading to strategy $(A)$ in the statement of the proposition. Alternatively, if Equation 24 is not satisfied then the problem is infeasible, i.e., strategy $(A)$ cannot be induced by the distributor when $\theta\beta \leq \lambda(\theta\beta + (1-\theta))$.

Now assume that $\theta\beta > \lambda(\theta\beta + (1-\theta))$. Let

$$\hat{O}_R^{A2} = \max\left\{\frac{\lambda\theta\beta}{1-\lambda}(m_h - m_l), \lambda\theta\beta\frac{r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h - m_l)(\theta\beta + (1-\theta))}{\lambda(\theta\beta + (1-\theta)) - \theta\beta}\right\}.$$

The objective is increasing in $c$ and independent of $z$, with an upper bound $c \leq m_l + x - \frac{\hat{O}_R^{A2}}{\theta\beta}$ given by the $(IR_A)$, $(IC_{AS})$, and $(IC_{ASR})$ constraints as long as

$$\hat{O}_R^{A2} \leq r_\alpha(m_l)\theta\beta(1-\beta), \tag{25}$$

i.e., as long as constraint $(IC_{AAR})$ is satisfied. Hence, in particular $c^* = m_h + x - \frac{\hat{O}_R^{A2}}{\theta\beta}$ and $z^* = 0$. Alternatively, if Equation 25 is not satisfied then the problem is infeasible, i.e., strategy $(A)$ cannot be induced by the distributor when $\theta\beta > \lambda(\theta\beta + (1-\theta))$. □

### C.1. Proof of Proposition 4

*Proof.* We first show that the distributor's objective function is increasing in $x$. Note that

$$\frac{\partial \Pi_D^A}{\partial x} = \theta\beta > 0, \frac{\partial \Pi_D^{AR}}{\partial x} = \beta + (1-\theta)(1-\beta) > 0, \frac{\partial \Pi_D^S}{\partial x} = \lambda\theta\beta > 0, \text{ and } \frac{\partial \Pi_D^{SR}}{\partial x} = \lambda(\beta + (1-\theta)(1-\beta)) > 0.$$

Hence, $\frac{\partial \Pi_D^*}{\partial x} > 0$.

In contrast, the distributor's objective function is not always increasing in $\theta$. To prove this, we show that Equation 5 in the statement of the proposition is equivalent to $\frac{\partial \Pi_D^{SR}}{\partial \theta} \leq 0$, and it implies $\Pi_D^S \leq \Pi_D^{SR}$, $\frac{\partial \Pi_D^{AR}}{\partial \theta} \leq 0$, and $\frac{\partial \Pi_D^A}{\partial \theta} \leq 0$ or $\Pi_D^A \leq \Pi_D^{AR}$. Hence, Equation 5 implies $\frac{\partial \Pi_D^*}{\partial \theta} \leq 0$.

First, we show that Equation 5 is equivalent to $\frac{\partial \Pi_D^{SR}}{\partial \theta} \leq 0$. Note that,

$$\frac{\partial \Pi_D^{SR}}{\partial \theta} = -(1-\gamma)\lambda\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) + \lambda(1-\beta)(w - y - x).$$

Then,

$$\begin{aligned}\frac{\partial \Pi_D^{SR}}{\partial \theta} \leq 0 &\iff y + (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) + x - w \geq 0 \\ &\iff m_h + x - w - (r_\alpha(m_h) - y) - \gamma(m - r_\alpha(m_h)) \geq 0.\end{aligned}$$

Second, we now show that Equation 5 implies $\Pi_D^S \leq \Pi_D^{SR}$. Indeed, from Proposition C.1 it follows that

$$\Pi_D^S \leq \Pi_D^{SR} \iff m_h + x - w - (1-\beta)(r_\alpha(m_h) - y) + \gamma(\beta(v_h - m_h) - (1-\beta)(m_h - r_\alpha(m_h))) \geq 0,$$

or equivalently

$$\Pi_D^S \leq \Pi_D^{SR} \iff y + (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) + x - w + \frac{\beta}{1-\beta}(\gamma v_h + (1-\gamma)m_h + x - w) \geq 0.$$

By noting that $(\gamma v_h + (1-\gamma)m_h + x - w) \geq 0$ we conclude that Equation 5 implies $\Pi_D^S \leq \Pi_D^{SR}$.

Third, we show that Equation 5 also implies $\frac{\partial \Pi_D^{AR}}{\partial \theta} \le 0$. Note that,

$$\frac{\partial \Pi_D^{AR}}{\partial \theta} = -(1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) + (1-\beta)(w-y-x) - \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta),$$

Then,

$$\frac{\partial \Pi_D^{AR}}{\partial \theta} \le 0 \iff y + (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) + x - w + \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h))) \ge 0.$$

Since $m - r_\alpha(m)$ is decreasing in $m$, then Equation 5 also implies $\frac{\partial \Pi_D^{AR}}{\partial \theta} \le 0$.

Fourth, we show that Equation 5 implies $\frac{\partial \Pi_D^{A}}{\partial \theta} \le 0$ or $\Pi_D^A \le \Pi_D^{AR}$. On the one hand, from Proposition C.1 it follows that

$$\Pi_D^A \le \Pi_D^{AR} \iff m_l + x - w - (1-\beta)(r_\alpha(m_l) - y) + \gamma(\beta(v_h - m_l) - (1-\beta)(m_l - r_\alpha(m_l))) - \frac{\lambda}{1-\lambda}(\beta(m_h - m_l) - (1-\beta)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))) \ge 0,$$

or equivalently

$$\Pi_D^A \le \Pi_D^{AR} \iff y + (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) + x - w + \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h))) + \frac{\beta}{1-\beta}\left(\gamma v_h + (1-\gamma)m_l + x - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right) \ge 0.$$

On the other hand, Lemma C.1 below implies

$$\frac{\partial \Pi_D^A}{\partial \theta} = (\gamma v_h + (1-\gamma)m_l + x - w)\beta - \frac{\lambda\beta}{1-\lambda}(m_h - m_l).$$

Hence, we conclude that if Equation 5 holds, then $\Pi_D^A > \Pi_D^{AR}$ implies $\frac{\partial \Pi_D^{A}}{\partial \theta} \le 0$, i.e., $\Pi_D^A \le \Pi_D^{AR}$ or $\frac{\partial \Pi_D^{A}}{\partial \theta} \le 0$ must hold.

Thus, we have shown Equation 5 is equivalent to $\frac{\partial \Pi_D^{SR}}{\partial \theta} \le 0$, and it implies $\Pi_D^S \le \Pi_D^{SR}$, $\frac{\partial \Pi_D^{AR}}{\partial \theta} \le 0$, and $\frac{\partial \Pi_D^{A}}{\partial \theta} \le 0$ or $\Pi_D^A \le \Pi_D^{AR}$. Hence, Equation 5 implies $\frac{\partial \Pi_D^{*}}{\partial \theta} \le 0$.

Finally, $\frac{\partial^2 \Pi_D^A}{\partial \theta \partial x} = \beta > 0$, $\frac{\partial^2 \Pi_D^{AR}}{\partial \theta \partial x} = -(1-\beta) < 0$, $\frac{\partial^2 \Pi_D^S}{\partial \theta \partial x} = \lambda\beta > 0$, and $\frac{\partial^2 \Pi_D^{SR}}{\partial \theta \partial x} = -\lambda(1-\beta) < 0$. This concludes the proof. □

**Lemma C.1** *Under Assumption 2, if $\theta\beta > \lambda(\theta\beta + (1-\theta))$ then*

$$\lambda\theta\beta\frac{r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h - m_l)(\theta\beta + (1-\theta))}{\lambda(\theta\beta + (1-\theta)) - \theta\beta} \le \frac{\lambda\theta\beta(m_h - m_l)}{1-\lambda}.$$

*Thus, by its definition in Proposition C.1 it follows that under Assumption 2 we have $\hat{O}_R^a = \frac{\lambda\theta\beta(m_h - m_l)}{1-\lambda}$.*

*Proof.* The assumptions in the lemma imply $r_\alpha(m_h)(1-\beta)(1-\lambda) \ge (m_h - \beta v_h)(1-\lambda) \ge (m_h - m_l)$, since the first inequality follows from $r_\alpha(m_h)$ being increasing in $\alpha > 0$ and taking the limit as $\alpha \to 0$, and the second inequality is equivalent to the assumption $\lambda \le \frac{m_l - \beta v_h}{m_h - \beta v_h}$ when $m_h > \beta v_h$.

This inequality implies the result in the lemma. Indeed,

$$\begin{aligned}
&\lambda\theta\beta\frac{r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h - m_l)(\theta\beta + (1-\theta))}{\lambda(\theta\beta + (1-\theta)) - \theta\beta} \le \frac{\lambda\theta\beta(m_h - m_l)}{1-\lambda} \\
\iff &(r_\alpha(m_h)(1-\theta)(1-\beta) - (m_h - m_l)(\theta\beta + (1-\theta)))(1-\lambda) \ge (m_h - m_l)(\lambda(\theta\beta + (1-\theta)) - \theta\beta) \\
\iff &r_\alpha(m_h)(1-\theta)(1-\beta)(1-\lambda) \ge (m_h - m_l)(1-\theta) \\
\iff &r_\alpha(m_h)(1-\beta)(1-\lambda) \ge (m_h - m_l),
\end{aligned}$$

where the first equivalence follows from the assumption $\theta\beta > \lambda(\theta\beta + (1-\theta))$, completing the proof. □

### C.2. Proof of Theorem 2

*Proof.* Let $(x^*, \theta^*)$ be the optimal solution of the distributor's allocation problem with heterogeneous ATPs. We analyze strategies $A$, $AR$, $S$, and $SR$ from Proposition C.1 separately.

<u>$FS$ and $MS$</u>: Assume first that $\Pi_D^*(x^*, \theta^*) = \Pi_D^S(x^*, \theta^*)$.

Since $\frac{\partial \Pi_D^S}{\partial x} = \lambda\theta\beta > 0$ and $\frac{\partial \Pi_D^S}{\partial \theta} = \lambda(\gamma v_h + (1-\gamma)m + x - w)\beta > 0$ then from Assumption 1 it follows that the optimal solution to the resource allocation problem exhausts the budget, i.e., $c_\theta(\theta^* - \theta_0) + \lambda\theta^*\beta x^* = b$, or equivalently $x^*(\theta) = \frac{b - c_\theta(\theta - \theta_0)}{\lambda\theta\beta}$. By replacing $x^*(\theta)$ in the resource allocation problem, it simplifies to the following one-variable optimization problem

$$\begin{aligned} \max_{\theta} \quad & \Pi_D^S(\theta) = (\gamma v_h + (1-\gamma)m_h - w)\lambda\theta\beta + b - c_\theta(\theta - \theta_0) \\ \text{s.t.} \quad & \theta \in [\theta_0, 1], \; \frac{b - c_\theta(\theta - \theta_0)}{\lambda\theta\beta} \in [0, v_h - m_h]. \end{aligned} \tag{26}$$

Let $\theta^S$ be the optimal solution of problem (26), and $x^S = \frac{b - c_\theta(\theta^S - \theta_0)}{\lambda\theta^S\beta}$. The objective function of problem (26), $\Pi_D^S(\theta)$, is linear. Moreover,

$$\frac{d\Pi_D^S(\theta)}{d\theta} = (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta - c_\theta. \tag{27}$$

Hence, $\theta^S$ must be equal to one of its upper or lower bound, i.e.,

$$\theta^S \in \left\{ \theta_0 + \frac{(b - \lambda\theta_0\beta(v_h - m_h))^+}{c_\theta + \lambda\beta(v_h - m_h)}, \min\left(1, \theta_0 + \frac{b}{c_\theta}\right) \right\},$$

and thus

$$x^S \in \left\{ \min\left(v_h - m_h, \frac{b}{\lambda\theta_0\beta}\right), \frac{(b - c_\theta(1-\theta_0))^+}{\lambda\beta} \right\}.$$

Namely, when following strategy $S$ in Proposition C.1 the distributor either invests the budget in increasing the consumers' maximum ability to pay first, and then the consumer education level only if there is budget available (strategy $FS$), or invests the budget in increasing the consumer education level first, and then the consumers' maximum ability to pay only if there is budget available (strategy $MS$). Specifically,

$$\theta^{FS} = \theta_0 + \frac{(b - \lambda\theta_0\beta(v_h - m_h))^+}{c_\theta + \lambda\beta(v_h - m_h)}, \; x^{FS} = \min\left(v_h - m_h, \frac{b}{\lambda\theta_0\beta}\right),$$

and

$$\Pi_D^{FS} = \left(\gamma v_h + (1-\gamma)m_h + \min\left(v_h - m_h, \frac{b}{\lambda\theta_0\beta}\right) - w\right)\left(\theta_0 + \frac{(b - \lambda\theta_0\beta(v_h - m_h))^+}{c_\theta + \lambda\beta(v_h - m_h)}\right)\lambda\beta.$$

Similarly,

$$\theta^{MS} = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right), \; x^{MS} = \frac{(b - c_\theta(1-\theta_0))^+}{\lambda\beta},$$

and

$$\Pi_D^{MS} = (\gamma v_h + (1-\gamma)m_h - w)\min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\lambda\beta + (b - c_\theta(1-\theta_0))^+.$$

<u>$FSR$ and $MSR$</u>: Assume now that $\Pi_D^*(x^*, \theta^*) = \Pi_D^{SR}(x^*, \theta^*)$.

Since $\frac{\partial \Pi_D^{SR}}{\partial x} = \lambda(\beta + (1-\theta)(1-\beta)) > 0$, then from Assumption 1 it follows that the optimal subsidy must either be equal to its upper bound or the budget constraint must be tight, i.e., $x^*(\theta) =$

$\min\left(v_h - m_h, \frac{b-c_\theta(\theta-\theta_0)}{\lambda(\beta+(1-\theta)(1-\beta))}\right)$. By replacing $x^*(\theta)$ in the resource allocation problem, it simplifies to the following one-variable optimization problem

$$\begin{aligned}\max_{\theta} \quad & \Pi_D^{SR}(\theta) = (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta + \min\left((v_h - m_h)\lambda(\beta + (1-\theta)(1-\beta)), b - c_\theta(\theta-\theta_0)\right) \\ & + \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (w-y)\right)\lambda(1-\beta)(1-\theta) \\ \text{s.t.} \quad & \theta \in [\theta_0, 1],\ c_\theta(\theta-\theta_0) \le b.\end{aligned} \tag{28}$$

Let $\theta^{SR}$ be the optimal solution of problem (28) and $x^{SR} = x^*(\theta^{SR})$. Since the minimum of two linear functions is concave, then the objective function of problem (28), $\Pi_D^{SR}(\theta)$, is piece-wise linear concave with at most two pieces. Moreover, note that

$$\begin{aligned}\frac{d\Pi_D^{SR}(\theta)}{d\theta} = & -\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (w-y)\right)\lambda(1-\beta) \\ & - c_\theta - ((v_h - m_h)\lambda(1-\beta) - c_\theta)\mathbb{1}_{\left\{c_\theta(\theta-\theta_0)+(v_h-m_h)\lambda(\beta+(1-\theta)(1-\beta))\le b\right\}}.\end{aligned} \tag{29}$$

Hence, we conclude that either $\theta^{SR}$ is equal to its upper bound, i.e., $\theta^{SR} = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)$, or alternatively $\theta^{SR}$ must be equal to one of its lower bound or the kink between the linear pieces of $\Pi_D^{SR}(\theta)$, i.e., $\theta^{SR} \in \left\{\theta_0, \frac{b+c_\theta\theta_0-\lambda(v_h-m)}{c_\theta-(v_h-m)\lambda(1-\beta)}\right\}$. Namely, when following strategy $SR$ in Proposition C.1 the distributor either invests the budget in increasing the consumer education level first, and then the consumers' maximum ability to pay only if there is budget available (strategy $MSR$), or invests the budget in increasing the consumers' maximum ability to pay first, and then in increasing the consumer education level only if it is beneficial and there is budget available (strategy $FSR$). Specifically,

$$\theta^{MSR} = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right),\ x^{MSR} = \frac{(b - c_\theta(1-\theta_0))^+}{\lambda\beta},$$

and

$$\begin{aligned}\Pi_D^{MSR} = \quad & (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta + (b - c_\theta(1-\theta_0))^+ \\ & + \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (w-y)\right)(1-\beta)\lambda\left(1-\theta_0-\frac{b}{c_\theta}\right)^+.\end{aligned} \tag{30}$$

In order to write $\theta^{FSR}$ in closed form there are two possible cases depending on whether $\Pi_D^{SR}(\theta)$ has a kink in the feasible interval of problem (28), $\left[\theta_0, \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\right]$. We analyze these cases next.

First assume $c_\theta \le (v_h - m_h)\lambda(1-\beta)$, then from Assumption 1 it follows that $\Pi_D^{SR}(\theta)$ does not have a kink in the feasible interval of problem (28). Specifically, if $c_\theta < (v_h - m_h)\lambda(1-\beta)$ then $\frac{b+c_\theta\theta_0-\lambda(v_h-m_h)}{c_\theta-(v_h-m_h)\lambda(1-\beta)} \ge 1$, and if $c_\theta = (v_h - m_h)\lambda(1-\beta)$ then $b - c_\theta(\theta-\theta_0) \le (v_h - m_h)\lambda(\beta + (1-\theta)(1-\beta))$ for all $\theta$. Namely, at $\theta = \theta_0$ there is no leftover budget after investing in $x$ and $\theta^{FSR} = \theta_0$ in this sub-case.

Now assume $c_\theta > (v_h - m_h)\lambda(1-\beta)$, then from Assumption 1 it follows that $\Pi_D^{SR}(\theta)$ has a kink in the feasible interval of problem (28) if and only if $\frac{b+c_\theta\theta_0-\lambda(v_h-m_h)}{c_\theta-(v_h-m_h)\lambda(1-\beta)} \ge \theta_0$, or equivalently $(v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta)) \le b$. Moreover, from Equation 29 it follows that the kink will attain an objective value at least as large as the solution $\theta = \theta_0$ if and only if $y \le y_\theta^{SR}$, where

$$y_\theta^{SR} \equiv w - (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (v_h - m_h).$$

To simplify the notation, we define the function $\frac{1}{x^{++}} := \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{1}{x} & \text{if } x > 0. \end{cases}$ Then,

$$\theta^{FSR} = \theta_0 + \frac{(b - (v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta)))^+}{(c_\theta - (v_h - m_h)\lambda(1-\beta))^{++}} \mathbb{1}_{\{y \leq y_\theta^{SR}\}}, \; x^{FSR} = \min\left(v_h - m_h, \frac{b}{\lambda(\beta + (1-\theta_0)(1-\beta))}\right),$$

and

$$\begin{aligned} \Pi_D^{FSR} = \quad & (\gamma v_h + (1-\gamma)m_h - w)\,\lambda\beta + \min\left((v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta)), b\right) \\ & + \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) - (w - y)\right)(1-\beta)\lambda(1-\theta_0) \\ & + \left(w - y - \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (v_h - m_h)\right)^+ (1-\beta)\lambda\frac{(b - (v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta)))^+}{(c_\theta - (v_h - m_h)\lambda(1-\beta))^{++}}. \end{aligned} \tag{31}$$

<u>$FA$ and $MA$</u>: Assume now that $\Pi_D^*(x^*, \theta^*) = \Pi_D^A(x^*, \theta^*)$.

Since $\frac{\partial \Pi_D^A}{\partial x} = \theta\beta > 0$ then from Assumption 1 it follows that the optimal subsidy must either be equal to its upper bound or the budget constraint must be tight, i.e., $x^*(\theta) = \min\left(v_h - m_l, \frac{b - c_\theta(\theta - \theta_0)}{\theta\beta}\right)$. By replacing $x^*(\theta)$ in the resource allocation problem, it simplifies to the following one-variable optimization problem

$$\begin{aligned} \max_\theta \quad & \Pi_D^A(\theta) = (\gamma v_h + (1-\gamma)m_l - w)\theta\beta + \min((v_h - m_l)\theta\beta, b - c_\theta(\theta - \theta_0)) - \frac{\lambda\theta\beta}{1-\lambda}(m_h - m_l) \\ \text{s.t.} \quad & \theta \in [\theta_0, 1], \; c_\theta(\theta - \theta_0) \leq b, \end{aligned} \tag{32}$$

where $\hat{O}_R^A = \frac{\lambda\theta\beta}{1-\lambda}(m_h - m_l)$ follows from Assumption 2 and Lemma C.1.

Let $\theta^A$ be the optimal solution of problem (32), and $x^A = x^*(\theta^A)$. Since the minimum of two linear functions is concave, then the objective function of problem (32), $\Pi_D^A(\theta)$, is piece-wise linear concave with at most two pieces. Moreover, note that

$$\frac{d\Pi_D^A(\theta)}{d\theta} = (\gamma v_h + (1-\gamma)m_l - w)\beta - c_\theta + ((v_h - m_l)\beta + c_\theta)\mathbb{1}_{\{c_\theta(\theta-\theta_0) + (v_h - m_l)\theta\beta \leq b\}} - \frac{\lambda\beta}{1-\lambda}(m_h - m_l). \tag{33}$$

Hence, we conclude that either $\theta^a$ is equal to its upper bound, i.e., $\theta^a = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)$, or alternatively $\theta^a$ must be equal to one of its lower bound or the kink between the linear pieces of $\Pi_D^A(\theta)$, i.e., $\theta^a \in \left\{\theta_0, \frac{b + c_\theta\theta_0}{c_\theta + (v_h - m_l)\beta}\right\}$.

Namely, when following strategy $A$ in Proposition C.1 the distributor either invests the budget in increasing the consumer education level first, and then the consumers' maximum ability to pay only if there is budget available (strategy $MA$), or invests the budget in increasing the consumers' maximum ability to pay first, and then in increasing the consumer education level only if it is beneficial and there is budget available (strategy $FA$). Specifically,

$$\theta^{MA} = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right), \; x^{MA} = \frac{(b - c_\theta(1-\theta_0))^+}{\beta},$$

and

$$\Pi_D^{MA} = \left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\beta + (b - c_\theta(1-\theta_0))^+.$$

In order to write $\theta^{FA}$ in closed form there are two possible cases depending on whether $\Pi_D^A(\theta)$ has a kink in the feasible interval of problem (32), $\left[\theta_0, \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\right]$. From Assumption 1, $\Pi_D(\theta)$ has a kink in the feasible interval of problem (32) if and only if $\frac{b + c_\theta\theta_0}{c_\theta + (v_h - m_l)\beta} \geq \theta_0$, or equivalently $(v_h - m)\theta_0\beta \leq b$. Moreover, from Equation 33 it follows that the kink will attain an objective value at least as large as the solution $\theta = \theta_0$ if and only if $w \le w_\theta^A$ where

$$w_\theta^A \equiv v_h + \gamma(v_h - m_l) - \frac{\lambda}{1-\lambda}(m_h - m_l).$$

Then,

$$\theta^{FA} = \theta_0 + \frac{\big(b - \theta_0\beta(v_h - m_l)\big)^+}{c_\theta + (v_h - m_l)\beta}\mathbb{1}_{\left\{w \le w_\theta^A\right\}},\ x^{FA} = \min\left(v_h - m_l, \frac{b}{\theta_0\beta}\right),$$

and

$$\Pi_D^{FA} = \left(\gamma v_h + (1-\gamma)m_l + \min\left(v_h - m_l, \frac{b}{\theta_0\beta}\right) - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\left(\theta_0 + \frac{(b - \theta_0\beta(v_h - m_l))^+}{c_\theta + \beta(v_h - m_l)}\mathbb{1}_{\left\{w \le w_\theta^A\right\}}\right)\beta.$$

<u>*FAR* and *MAR*:</u> Finally, assume that $\Pi_D^*(x^*, \theta^*) = \Pi_D^{AR}(x^*, \theta^*)$.

Since $\frac{\partial \Pi_D^{AR}}{\partial x} = \beta + (1-\theta)(1-\beta) > 0$, then from Assumption 1 it follows that the optimal subsidy must either be equal to its upper bound or the budget constraint must be tight, i.e., $x^*(\theta) = \min\left(v_h - m_l, \frac{b - c_\theta(\theta - \theta_0)}{\beta + (1-\theta)(1-\beta)}\right)$. By replacing $x^*(\theta)$ in the resource allocation problem, it simplifies to the following one-variable optimization problem

$$\begin{aligned}
\max_\theta \quad & \Pi_D^{AR}(\theta) = (\gamma v_h + (1-\gamma)m_l - w)\beta + \min\big((v_h - m_l)(\beta + (1-\theta)(1-\beta)), b - c_\theta(\theta - \theta_0)\big) \\
& - \frac{\lambda}{1-\lambda}((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\theta)(1-\beta)) \\
& + \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1 - \beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right) - (w - y)\right)(1-\beta)(1-\theta) \\
\text{s.t.} \quad & \theta \in [\theta_0, 1],\ c_\theta(\theta - \theta_0) \le b.
\end{aligned} \tag{34}$$

Let $\theta^{AR}$ be the optimal solution of problem (34) and $x^{AR} = x^*(\theta^{AR})$. Since the minimum of two linear functions is concave, then the objective function of problem (34), $\Pi_D^{AR}(\theta)$, is piece-wise linear concave with at most two pieces. Moreover, note that

$$\begin{aligned}
\frac{d\Pi_D^{AR}(\theta)}{d\theta} = & -\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1 - \beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right) - (w - y)\right)(1-\beta) \\
& - c_\theta - ((v_h - m_l)(1-\beta) - c_\theta)\mathbb{1}_{\left\{c_\theta(\theta - \theta_0) + (v_h - m_l)(\beta + (1-\theta)(1-\beta)) \le b\right\}} \\
& - \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta).
\end{aligned} \tag{35}$$

Hence, we conclude that either $\theta^b$ is equal to its upper bound, i.e., $\theta^b = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)$, or alternatively $\theta^b$ must be equal to one of its lower bound or the kink between the linear pieces of $\Pi_D^{AR}(\theta)$, i.e., $\theta^b \in \left\{\theta_0, \frac{b + c_\theta\theta_0 - (v_h - m_l)}{c_\theta - (v_h - m_l)(1-\beta)}\right\}$. Namely, when following strategy $(b)$ in Proposition C.1 the distributor either invests the budget in increasing the consumer education level first, and then the consumers' maximum ability to pay only if there is budget available (strategy *MAR*), or invests the budget in increasing the consumers' maximum ability to pay first, and then in increasing the consumer education level only if it is beneficial and there is budget available (strategy *FAR*).

Specifically,

$$\theta^{MAR} = \min\left(1, \theta_0 + \frac{b}{c_\theta}\right),\ x^{MAR} = \frac{(b - c_\theta(1-\theta_0))^+}{\beta},$$

and

$$\begin{aligned}\Pi_D^{MAR} = \quad & (\gamma v_h + (1-\gamma)m_l - w)\beta + (b - c_\theta(1-\theta_0))^+ \\ & - \frac{\lambda}{1-\lambda}\left((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+\right) \\ & + \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) - (w-y)\right)(1-\beta)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+ .\end{aligned} \tag{36}$$

In order to write $\theta^{FAR}$ in closed form there are two possible cases depending on whether $\Pi_D^{AR}(\theta)$ has a kink in the feasible interval of problem (34), $\left[\theta_0, \min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\right]$. We analyze these cases next.

First assume $c_\theta \le (v_h - m_l)(1-\beta)$, then from Assumption 1 it follows that $\Pi_D^{AR}(\theta)$ does not have a kink in the feasible interval of problem (34). Specifically, if $c_\theta < (v_h - m_l)(1-\beta)$ then $\frac{b + c_\theta\theta_0 - (v_h - m_l)}{c_\theta - (v_h-m_l)(1-\beta)} \ge 1$, and if $c_\theta = (v_h - m_l)(1-\beta)$ then $b - c_\theta(\theta - \theta_0) \le (v_h - m_l)(\beta + (1-\theta)(1-\beta))$ for all $\theta$. Namely, at $\theta = \theta_0$ there is no leftover budget after investing in $x$ and $\theta^{FAR} = \theta_0$ in this sub-case.

Now assume $c_\theta > (v_h - m_l)(1-\beta)$, then from Assumption 1 it follows that $\Pi_D^{AR}(\theta)$ has a kink in the feasible interval of problem (34) if and only if $\frac{b + c_\theta\theta_0 - (v_h - m_l)}{c_\theta - (v_h-m_l)(1-\beta)} \ge \theta_0$, or equivalently $(v_h - m_l)(\beta + (1-\theta_0)(1-\beta)) \le b$. Moreover, from Equation 35 it follows that the kink will attain an objective value at least as large as the solution $\theta = \theta_0$ if and only if $y \le y_\theta^{AR}$, where

$$y_\theta^{AR} \equiv w - (1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) - (v_h - m_l) - \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h))).$$

Then,

$$\theta^{FAR} = \theta_0 + \frac{\left(b - (v_h - m_l)(\beta + (1-\theta_0)(1-\beta))\right)^+}{\left(c_\theta - (v_h - m_l)(1-\beta)\right)^{++}}\mathbb{1}_{\{y \le y_\theta^{AR}\}}, \; x^{FAR} = \min\left(v_h - m_l, \frac{b}{\beta + (1-\theta_0)(1-\beta)}\right),$$

and

$$\begin{aligned}\Pi_D^{FAR} = \quad & (\gamma v_h + (1-\gamma)m_l - w)\,\beta + \min\left((v_h - m_l)(\beta + (1-\theta_0)(1-\beta)), b\right) \\ & - \frac{\lambda}{1-\lambda}((m_h - m_l)\beta - (m_l - r_\alpha(m_l) + (m_h - r_\alpha(m_h)))(1-\theta_0)(1-\beta)) \\ & + \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) - (w-y)\right)(1-\beta)(1-\theta_0) \\ & + \left(w - y - \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) - (v_h - m_l) - \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))\right)^+ \\ & \cdot (1-\beta)\frac{\left(b - (v_h - m_l)(\beta + (1-\theta_0)(1-\beta))\right)^+}{\left(c_\theta - (v_h - m_l)(1-\beta)\right)^{++}}.\end{aligned} \tag{37}$$

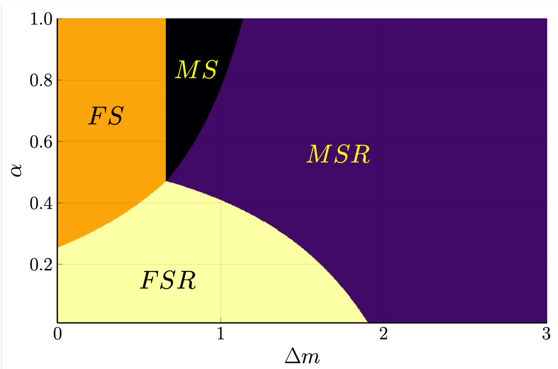
To conclude, note that the optimal objective value of the resource allocation problem is

$$\Pi_D^{OPT} = \max(\Pi_D^{FS}, \Pi_D^{MS}, \Pi_D^{FSR}, \Pi_D^{MSR}, \Pi_D^{FA}, \Pi_D^{MA}, \Pi_D^{FAR}, \Pi_D^{MAR}).$$

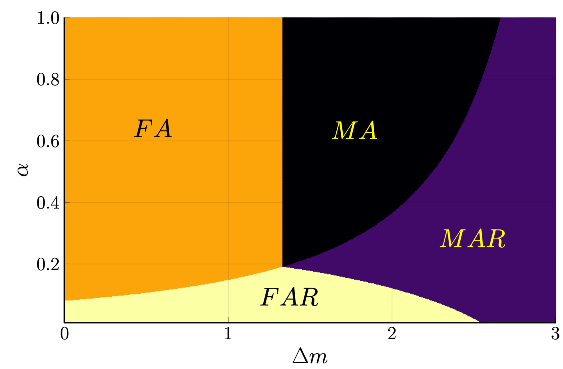
□

### C.3. Restricted Strategy Maps with Heterogeneous ATPs

In this section, we provide a generalization of Proposition B.2 that holds for heterogeneous ATPs. Proposition C.2 below shows that the structure of the strategy map from Proposition B.2 is preserved when the distributor constrains itself to using either only skimming or non-skimming strategies from Theorem 2. Figure 7 illustrates the motivating intuition for Proposition C.2 by depicting the "constrained" strategy maps




(a) Strategy map constrained to strategies $FS, MS, FSR$ and $MSR$.




(b) Strategy map constrained to strategies $FA, MA, FAR$ and $MAR$.

**Figure 7** **Strategy map for the Allocation Problem as a function of $\Delta m$ and $\alpha$ when the distributor is constrained to either (a) skimming strategies or (b) strategies that target all customers. We assume $v_h = \$10$, $w = \$6$, $\theta_0 = 0.2$, $m_h = \$7$, $m_l = \$6$, $\lambda = 0.4$, $b = \$0.2$, $y = \$3.25$, $\beta = 0.5$, $c_\theta = 0.33$, and $\gamma = 0$. The horizontal axis sets the ATPs to $m_l + \Delta m$ and $m_h + \Delta m_h$.**

that form the strategy map of Figure 5. Figure 7a depicts the strategy map when the distributor constrains itself to "skimming" strategies $FS$, $MS$, $FSR$ and $MSR$, while Figure 7b depicts the strategy map when the distributor constrains itself to strategies that target all customers $FA$, $MA$, $FAR$ and $MAR$.

Generalizations of Propositions B.3 and B.4 –which describe simpler settings where a smaller number of strategies can be optimal– follow exactly the same logic, and are thus omitted for the sake of brevity.

**Proposition C.2** *Under Assumption 2, assume $w \le \beta v_h + (1-\beta)y$ and $c_\theta \le (1-\beta)\lambda(w-y)$. Let $FS$, $MS$, $FSR$, and $MSR$ be the investment strategies from Theorem 2. Then, there exist ATPs $m^{FS}$, $m^{MSR}$, and $\bar{m}^{MSR}$ where $w \le m^{FS} \le m^{MSR}, \bar{m}^{MSR}$ and functions $a_{S1}(m)$, $a_{S2}(m)$, and $a_{S3}(m)$ such that for a given customer risk aversion parameter $\alpha$ and base ATPs $m_l$, $m_h$, $m_l \le m_h$,*

- $\Pi_D^{FS} \ge \max(\Pi_D^{MS}, \Pi_D^{FSR}, \Pi_D^{MSR})$ *if and only if* $m_h \le m^{FS}$ *and* $\alpha \ge a_{S3}(m_h)$
- $\Pi_D^{MS} \ge \max(\Pi_D^{FS}, \Pi_D^{FSR}, \Pi_D^{MSR})$ *if and only if* $m^{FS} \le m_h \le \bar{m}^{MSR}$ *and* $\alpha \ge \alpha_{S2}(m_h)$
- $\Pi_D^{FSR} \ge \max(\Pi_D^{FS}, \Pi_D^{MS}, \Pi_D^{MSR})$ *if and only if* $m_h \le m^{MSR}$ *and* $\alpha \le \min(a_{S1}(m_h), a_{S3}(m_h))$
- $\Pi_D^{MSR} \ge \max(\Pi_D^{FS}, \Pi_D^{MS}, \Pi_D^{FSR})$ *if and only if* $m^{FS} \le m_h$ *and* $a_{S1}(m_h) \le \alpha \le a_{S2}(m_h)$*, where* $a_{S1}(m_h) \le a_{S2}(m_h)$ *if and only if* $m_h \ge m^{FS}$*. Also,* $a_{S1}(m_h) = 0$ *for all* $m^{MSR} \le m_h$ *and* $a_{S1}(m_h) = \infty$ *for all* $\bar{m}^{MSR} \le m_h$*.*

*Further, assume $w \le \beta v_h + (1-\beta)y - \lambda/(1-\lambda)\left((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)\right)$ and $c_\theta \le (1-\beta)(w-y) - \lambda/(1-\lambda)(1-\beta)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))$. Let $FA$, $MA$, $FAR$, and $MAR$ be the investment strategies from Theorem 2. Then, there exist ATPs $m^{FA}$, $m^{MAR}$, and $\bar{m}^{MAR}$ where $w \le m^{FA} \le m^{MAR}, \bar{m}^{MAR}$, and functions $a_{A1}(m)$, $a_{A2}(m)$, and $a_{A3}(m)$ such that for a given customer risk aversion parameter $\alpha$ and base ATPs $m_l$, $m_h$, $m_l \le m_h$,*

- $\Pi_D^{FA} \ge \max(\Pi_D^{MA}, \Pi_D^{FAR}, \Pi_D^{MAR})$ *if and only if* $m_l \le m^{FA}$ *and* $\alpha \ge a_{A3}(m_l)$
- $\Pi_D^{MA} \ge \max(\Pi_D^{FA}, \Pi_D^{FAR}, \Pi_D^{MAR})$ *if and only if* $m^{FA} \le m_l \le \bar{m}^{MAR}$ *and* $\alpha \ge \alpha_{A2}(m_l)$

- $\Pi_D^{FAR} \geq \max(\Pi_D^{FA}, \Pi_D^{MA}, \Pi_D^{MAR})$ *if and only if* $m_l \leq m^{MAR}$ *and* $\alpha \leq \min(a_{A1}(m_l), a_{A3}(m_l))$
- $\Pi_D^{MAR} \geq \max(\Pi_D^{FA}, \Pi_D^{MA}, \Pi_D^{FAR})$ *if and only if* $m^{FA} \leq m_l$ *and* $a_{A1}(m_l) \leq \alpha \leq a_{A2}(m_l)$*, where* $a_{A1}(m_l) \leq a_{A2}(m_l)$ *if and only if* $m_l \geq m^{FA}$*. Also,* $a_{A1}(m_l) = 0$ *for all* $m^{MAR} \leq m_l$ *and* $a_{A1}(m_l) = \infty$ *for all* $\bar{m}^{MAR} \leq m_l$*.*

The proof of Proposition C.2 follows the same logic as the proof of Proposition B.2. Namely, we establish the following properties:

- Property 1': $\Pi_D^{FS} \geq \Pi_D^{MS}$ if and only if $m_h \leq m^{FS}$, and $\Pi_D^{FA} \geq \Pi_D^{MA}$ if and only if $m_l \leq m^{FA}$;
- Property 2': $\Pi_D^{MSR} \geq \Pi_D^{FSR}$ if and only if $\alpha \geq a_{S1}(m_h)$, and $\Pi_D^{MAR} \geq \Pi_D^{FAR}$ if and only if $\alpha \geq a_{A1}(m_l)$;
- Property 3': $\Pi_D^{MS} \geq \Pi_D^{MSR}$ if and only if $\alpha \geq a_{S2}(m_h)$, and $\Pi_D^{MA} \geq \Pi_D^{MAR}$ if and only if $\alpha \geq a_{A2}(m_l)$;
- Property 4': $\Pi_D^{FS} \geq \Pi_D^{FSR}$ if and only if $\alpha \geq a_{S3}(m_h)$, and $\Pi_D^{FA} \geq \Pi_D^{FAR}$ if and only if $\alpha \geq a_{A3}(m_l)$;
- Property 5': $a_{S1}(m_h) \leq a_{S2}(m_h)$ if and only if $m_h \geq m^{FS}$, $a_{S2}(m_h) = \infty$ for all $m_h \geq \bar{m}^{MSR}$ and $a_{S1}(m_h) = 0$ for all $m_h \geq m^{MSR}$, where $w \leq m^{FS} \leq m^{MSR}, \bar{m}^{MSR}$, and $a_{A1}(m_l) \leq a_{A2}(m_l)$ if and only if $m_l \geq m^{FA}$, $a_{A2}(m_l) = \infty$ for all $m_l \geq \bar{m}^{MAR}$ and $a_{A1}(m_l) = 0$ for all $m_l \geq m^{MAR}$, where $w \leq m^{FA} \leq m^{MAR}, \bar{m}^{MAR}$.

Note that the first five properties combined establish the following additional properties:

- Property 6': when $m \leq m^{FS}$ then $\Pi_D^{MSR} \leq \max(\Pi_D^{FSR}, \Pi_D^{MS})$; and when $m \leq m^{FA}$ then $\Pi_D^{MAR} \leq \max(\Pi_D^{FAR}, \Pi_D^{MA})$;
- Property 7': when $m \geq m^{FS}$ then $\alpha \geq a_{S2}(m_h)$ implies $\Pi_D^{FSR} \leq \Pi_D^{MS}$; and when $m \geq m^{FA}$ then $\alpha \geq a_{A2}(m_l)$ implies $\Pi_D^{FAR} \leq \Pi_D^{MA}$;
- Property 8': when $m \leq m^{FS}$ then $\alpha \leq \alpha_{S3}(m_h)$ implies $\Pi_D^{FSR} \geq \Pi_D^{MS}$; and when $m \leq m^{FA}$ then $\alpha \leq \alpha_{A3}(m_l)$ implies $\Pi_D^{FAR} \geq \Pi_D^{MA}$;
- Property 9': when $m \geq m^{FS}$ then $\alpha \leq \alpha_{S1}(m_h)$ implies $\Pi_D^{FSR} \geq \Pi_D^{MS}$; and when $m \geq m^{FA}$ then $\alpha \leq \alpha_{A1}(m_l)$ implies $\Pi_D^{FAR} \geq \Pi_D^{MA}$.

Properties 1'-9' above completely characterize the restricted strategy maps given in Proposition C.2. Namely,

- $\Pi_D^{FS} \geq \max(\Pi_D^{MS}, \Pi_D^{FSR}, \Pi_D^{MSR})$ if and only if $m \leq m^{FS}$ and $\alpha \geq a_{S3}(m_h)$; and $\Pi_D^{FA} \geq \max(\Pi_D^{MA}, \Pi_D^{FAR}, \Pi_D^{MAR})$ if and only if $m \leq m^{FA}$ and $\alpha \geq a_{A3}(m_l)$ (combine Properties 1', 4', and 6');
- $\Pi_D^{MS} \geq \max(\Pi_D^{FS}, \Pi_D^{FSR}, \Pi_D^{MSR})$ if and only if $m \geq m^{FS}$ and $\alpha \geq \alpha_{S2}(m_h)$; and $\Pi_D^{MA} \geq \max(\Pi_D^{FA}, \Pi_D^{FAR}, \Pi_D^{MAR})$ if and only if $m \geq m^{FA}$ and $\alpha \geq \alpha_{A2}(m_l)$ (combine Properties 1', 3', and 7');
- $\Pi_D^{FSR} \geq \max(\Pi_D^{FS}, \Pi_D^{MS}, \Pi_D^{MSR})$ if and only if $\alpha \leq \min(a_{S1}(m_h), a_{S3}(m_h))$; and $\Pi_D^{FAR} \geq \max(\Pi_D^{FA}, \Pi_D^{MA}, \Pi_D^{MAR})$ if and only if $\alpha \leq \min(a_{A1}(m_l), a_{A3}(m_l))$ (combine Properties 2', 4', 8' and 9');
- $\Pi_D^{MSR} \geq \max(\Pi_D^{FS}, \Pi_D^{MS}, \Pi_D^{FSR})$ if and only if $m^{FS} \leq m_h$ and $a_{S1}(m_h) \leq \alpha \leq a_{S2}(m_h)$, where $a_{S1}(m_h) \leq a_{S2}(m_h)$ if and only if $m_h \geq m^{FS}$, $a_{S1}(m_h) = 0$ for all $m^{MSR} \leq m_h$, and $a_{S1}(m_h) = \infty$ for all $\bar{m}^{MSR} \leq m_h$; and $\Pi_D^{MAR} \geq \max(\Pi_D^{FA}, \Pi_D^{MA}, \Pi_D^{FAR})$ if and only if $m^{FA} \leq m_l$ and $a_{A1}(m_l) \leq \alpha \leq a_{A2}(m_l)$, where $a_{A1}(m_l) \leq a_{A2}(m_l)$ if and only if $m_l \geq m^{FA}$, $a_{A1}(m_l) = 0$ for all $m^{MAR} \leq m_l$, and $a_{A1}(m_l) = \infty$ for all $\bar{m}^{MAR} \leq m_l$ (combine Properties 1', 2', 3', and 5').

We now prove each property.

*Proof of Property 1'.* Recall the Linear Program in (26) for deciding between strategies $MS$ and $FS$. From Equation 27, we have that $\Pi_D^{FS} \geq \Pi_D^{MS}$ if and only if $(\gamma v_h + (1-\gamma)m_h - w)\lambda\beta \leq c_\theta$. Thus, $m^{FS} = \frac{\frac{c_\theta}{\lambda\beta} + w - \gamma v_h}{(1-\gamma)} = \frac{\frac{c_\theta}{\beta} + w - \gamma v_h}{(1-\gamma)} + \frac{1-\lambda}{\lambda}\frac{c_\theta}{(1-\gamma)\beta}$.

Similarly, recall the optimization problem in (32). From the concavity of $\Pi_D^A(\theta)$ it follows that it is sufficient to check the derivative $\frac{d\Pi_D^A(\theta)}{d\theta}$ at $\theta = 1$, i.e.,

$$\Pi_D^{MA} \geq \Pi_D^{FA} \iff \frac{d\Pi_D^A(1)}{d\theta} \geq 0.$$

We have

$$\begin{aligned}\frac{d\Pi_D^A(1)}{d\theta} &= (\gamma v_h + (1-\gamma)m_l - w)\beta - c_\theta + ((v_h - m_l)\beta + c_\theta)\mathbb{1}_{\left\{c_\theta(1-\theta_0)+(v_h-m_l)\beta \leq b\right\}} - \frac{\lambda\beta}{1-\lambda}(m_h - m_l)\\ &= (\gamma v_h + (1-\gamma)m_l - w)\beta - c_\theta - \frac{\lambda\beta}{1-\lambda}(m_h - m_l).\end{aligned}$$

where the second equality follows from Assumption 1. Thus, $\Pi_D^{FA} \geq \Pi_D^{MA}$ if and only if $(\gamma v_h + (1-\gamma)m_l - w)\beta \leq c_\theta + \frac{\lambda\beta}{1-\lambda}(m_h - m_l)$. Thus, $m^{FA} = \frac{\frac{c_\theta}{\beta} + w - \gamma v_h}{(1-\gamma)} + \frac{\lambda}{1-\lambda}\left(\frac{m_h - m^{FA}}{1-\gamma}\right)$, completing the proof of Property 1'.

*Proof of Property 2'.* Recall the optimization problem in (28). From the concavity of $\Pi_D^{SR}(\theta)$ it follows that it is sufficient to check the derivative $\frac{d\Pi_D^{SR}(\theta)}{d\theta}$ at $\theta = 1$, i.e.,

$$\Pi_D^{MSR} \geq \Pi_D^{FSR} \iff \frac{d\Pi_D^{SR}(1)}{d\theta} \geq 0.$$

We have

$$\begin{aligned}\frac{d\Pi_D^{SR}(1)}{d\theta} =& -\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (1-\beta)(w-y)\right)\lambda\\ &- c_\theta - ((v_h - m_h)\lambda(1-\beta) - c_\theta)\mathbb{1}_{\left\{c_\theta(1-\theta_0)+(v_h-m_h)\lambda\beta \leq b\right\}}\\ =& -(1-\gamma)\frac{(1-\beta)\lambda}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) + (1-\beta)\lambda(w-y) - c_\theta,\end{aligned}$$

where the second equality follows from Assumption 1. Thus,

$$\frac{d\Pi_D^{SR}(1)}{d\theta} \geq 0 \iff \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) \leq w - y - \frac{c_\theta}{\lambda(1-\beta)},$$

First assume that $\gamma \leq 1$, note that then the term $\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $\left[0, \frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta}\right]$. If $0 \leq w - y - \frac{c_\theta}{\lambda(1-\beta)} \leq \frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta}$, then there exists $\hat{\alpha}_{S1}$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_{S1}}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_{S1}(v_h-m_h)}}{1-\beta}\right) = w - y - \frac{c_\theta}{\lambda(1-\beta)}.$$

Then, we define $a_{S1}(m_h)$ as

$$a_{S1}(m_h) = \begin{cases}\infty, & \text{if } w - y - \frac{c_\theta}{\lambda(1-\beta)} < 0\\ \hat{\alpha}_{S1}, & \text{if } 0 \leq w - y - \frac{c_\theta}{\lambda(1-\beta)} \leq \frac{(1-\gamma)\beta(v_h-m)}{1-\beta}\\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta} < w - y - \frac{c_\theta}{\lambda(1-\beta)}.\end{cases}$$

Since $c_\theta \leq (1-\beta)\lambda(w-y)$ by assumption, then $a_{S1}(m_h) < \infty$. The definition of $a_{S1}(m_h)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity.

Similarly, recall the optimization problem in (34). From the concavity of $\Pi_D^{AR}(\theta)$ it follows that it is sufficient to check the derivative $\frac{d\Pi_D^{AR}(\theta)}{d\theta}$ at $\theta=1$, i.e.,

$$\Pi_D^{MAR} \geq \Pi_D^{FAR} \iff \frac{d\Pi_D^{AR}(1)}{d\theta} \geq 0.$$

We have

$$\begin{aligned}\frac{d\Pi_D^{AR}(1)}{d\theta} =& -\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)-(1-\beta)(w-y)\right)\\ &-c_\theta-((v_h-m_l)(1-\beta)-c_\theta)\mathbb{1}_{\left\{c_\theta(1-\theta_0)+(v_h-m_l)\beta\leq b\right\}}\\ &-\frac{\lambda}{1-\lambda}(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))(1-\beta)\\ =& -(1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)+(1-\beta)(w-y)-c_\theta\\ &-\frac{\lambda}{1-\lambda}(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))(1-\beta)\end{aligned}$$

where the second equality follows from Assumption 1. Thus,

$$\frac{d\Pi_D^{AR}(1)}{d\theta} \geq 0 \iff \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) \leq g_{A1}(m_l),$$

where $g_{A1}(m_l) = w - y - \frac{c_\theta}{1-\beta} - \frac{\lambda}{1-\lambda}(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))$. First assume that $\gamma \leq 1$, note that then the term $\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $\left[0, \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}\right]$. If $0 \leq g_{A1}(m_l) \leq \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}$, then there exists $\hat{\alpha}_{A1}$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_{A1}}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_{A1}(v_h-m_l)}}{1-\beta}\right) = g_{A1}(m_l).$$

Then, we define $a_{A1}(m_l)$ as

$$a_{A1}(m_l) = \begin{cases}\infty, & \text{if } g_{A1}(m_l) < 0\\ \hat{\alpha}_{A1}, & \text{if } 0 \leq g_{A1}(m_l) \leq \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}\\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta} < g_{A1}(m_l).\end{cases}$$

Since $c_\theta \leq (1-\beta)(w-y) - \lambda/(1-\lambda)(1-\beta)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))$ by assumption, then $g_{A1}(m_l) \geq 0$ thus $a_{A1}(m_l) < \infty$. The definition of $a_{A1}(m_l)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity. This establishes Property 2'.

*Proof of Property 3'.* To compare strategy $MS$ and $MSR$, we directly compare their distributor's value. From the proof of Theorem 2:

$$\begin{aligned}\Pi_D^{MSR} =& (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta + (b - c_\theta(1-\theta_0))^+\\ &+\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)-(w-y)\right)(1-\beta)\lambda\left(1-\theta_0-\frac{b}{c_\theta}\right)^+.\end{aligned}$$

and

$$\Pi_D^{MS} = (\gamma v_h + (1-\gamma)m_h - w)\min\left(1, \theta_0 + \frac{b}{c_\theta}\right)\lambda\beta + (b - c_\theta(1-\theta_0))^+$$

Thus, $\Pi_D^{MS} \geq \Pi_D^{MSR}$ if and only if

$$-(\gamma v_h + (1-\gamma)m_h - w)\beta\left(1-\theta_0-\frac{b}{c_\theta}\right)^+ \geq \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)-(w-y)\right)(1-\beta)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+.$$

If $\left(1-\theta_0-\frac{b}{c_\theta}\right)^+=0$, then the budget is sufficient to set $\theta=1$ and the profits of policy $M$ and $MR$ are the same since all customers are informed and there are no returns. Thus we consider the case where $\left(1-\theta_0-\frac{b}{c_\theta}\right)^+>0$. In such case, the inequality above becomes

$$-(\gamma v_h+(1-\gamma)m_h-w)\beta \geq \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)-(w-y)\right)(1-\beta).$$

Rearranging the terms yields the relationship,

$$\Pi_D^{MS}\geq \Pi_D^{MSR} \iff \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)\leq g_{S2}(m_h).$$

where $g_{S2}(m_h)=w-y-(\gamma v_h+(1-\gamma)m_h-w)\frac{\beta}{1-\beta}$.

As in the proof of Property 2', first assume that $\gamma\leq 1$, note that then the term $\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $\left[0,\frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta}\right]$. If $0\leq g_{S2}(m_h)\leq \frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta}$, then there exists a $\hat{\alpha}_{S2}$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_{S2}}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_{S2}(v_h-m_h)}}{1-\beta}\right)=g_{S2}(m_h).$$

Then, we define $a_{S2}(m_h)$ as

$$a_{S2}(m_h)=\begin{cases}\infty, & \text{if } g_{S2}(m_h)<0\\ \hat{\alpha}_{S2}, & \text{if } 0\leq g_{S2}(m_h)\leq \frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta}\\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m_h)}{1-\beta}<g_{S2}(m_h).\end{cases}$$

Since $w\leq \beta v_h+(1-\beta)y$ by assumption, then $g_{S2}(m_h)\leq(1-\gamma)\beta(v_h-m_h)/(1-\beta)$ thus $a_{S2}(m_h)>0$. As before, the definition of $a_{S2}(m_h)$ when $\gamma>1$ is analogous and is skipped for the sake of brevity.

Similarly, to compare strategy $MA$ and $MAR$, we directly compare their distributor's value. From the proof of Theorem 1:

$$\begin{aligned}\Pi_D^{MAR}=&(\gamma v_h+(1-\gamma)m_l-w)\beta+(b-c_\theta(1-\theta_0))^+\\ &-\frac{\lambda}{1-\lambda}\left((m_h-m_l)\beta-(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))(1-\beta)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+\right)\\ &+\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)-(w-y)\right)(1-\beta)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+.\end{aligned}$$

and

$$\Pi_D^{MA}=(\gamma v_h+(1-\gamma)m_l-w)\min\left(1,\theta_0+\frac{b}{c_\theta}\right)\beta+(b-c_\theta(1-\theta_0))^+-\frac{\lambda\beta}{1-\lambda}\min\left(1,\theta_0+\frac{b}{c_\theta}\right)(m_h-m_l)$$

Thus, $\Pi_D^{MA}\geq\Pi_D^{MAR}(\alpha,m_l)$ if and only if

$$\begin{aligned}-(\gamma v_h+(1-\gamma)m_l-w)\beta\left(1-\theta_0-\frac{b}{c_\theta}\right)^+\geq&\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m)}}{1-\beta}\right)-(w-y)\right)(1-\beta)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+.\\ &-\frac{\lambda}{1-\lambda}\left((m_h-m_l)-(r_\alpha(m_h)-r_\alpha(m_l))(1-\beta)\right)\left(1-\theta_0-\frac{b}{c_\theta}\right)^+\end{aligned}$$

If $\left(1-\theta_0-\frac{b}{c_\theta}\right)^+=0$, then the budget is sufficient to set $\theta=1$ and the profits of policy $MA$ and $MAR$ are the same since all customers are informed and there are no returns. Thus we consider the case where $\left(1-\theta_0-\frac{b}{c_\theta}\right)^+>0$. In such case, the inequality above becomes

$$\begin{aligned}&\frac{\lambda}{1-\lambda}\left((m_h-m_l)-(r_\alpha(m_h)-r_\alpha(m_l))(1-\beta)\right)-(\gamma v_h+(1-\gamma)m_l-w)\beta\\ &\qquad\geq\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)-(w-y)\right)(1-\beta).\end{aligned}$$

Rearranging the terms yields the relationship,

$$\Pi_D^{MA}(\alpha,m) \geq \Pi_D^{MAR}(\alpha,m) \iff \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right) \leq g_{A2}(m_l).$$

where $g_{A2}(m_l) = w - y + \frac{\lambda}{1-\lambda}\left(\frac{m_h-m_l}{1-\beta} - (r_\alpha(m_h) - r_\alpha(m_l))\right) - (\gamma v_h + (1-\gamma)m_l - w)\frac{\beta}{1-\beta}$.

As before, first assume that $\gamma \leq 1$, note that then the term $\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $\left[0, \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}\right]$. If $0 \leq g_{A2}(m_l) \leq \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}$, then there exists a $\hat{\alpha}_{A2}$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_{A2}}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_{A2}(v_h-m_l)}}{1-\beta}\right) = g_{A2}(m_l).$$

Then, we define $a_{A2}(m_l)$ as

$$a_{A2}(m_l) = \begin{cases} \infty, & \text{if } g_{A2}(m_l) < 0 \\ \hat{\alpha}_{A2}, & \text{if } 0 \leq g_{A2}(m_l) \leq \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta} \\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta} < g_{A2}(m_l). \end{cases}$$

Since $c_\theta \leq (1-\beta)(w-y) - \lambda/(1-\lambda)(1-\beta)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))$ by assumption, then $g_{A2}(m_l) \leq (1-\gamma)\beta(v_h-m_l)/(1-\beta)$ thus $a_{A2}(m_l) > 0$. As before, the definition of $a_{A2}(m_l)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity. This establishes Property 3'.

*Proof of Property 4'.* The proof of this property is analogous to the proof of Properties 2' and 3'. We denote the information level in strategy $FS$ and $FSR$ by $\theta^{FS}$ and $\theta^{FSR}$, respectively. We directly compare the distributor's value in $FS$ and $FSR$. Hence,

$$\begin{aligned}
&\Pi_D^{FS} \geq \Pi_D^{FSR} \\
\iff &(\gamma v_h + (1-\gamma)m_h - w)\theta^{FS}\lambda\beta + b - c_\theta(\theta^{FS}-\theta_0) \geq \\
&(\gamma v_h + (1-\gamma)m_h - w)\lambda\beta + \min\left((v_h-m_h)\lambda(\beta + (1-\theta^{FSR})(1-\beta)), b - c_\theta(\theta^{FSR}-\theta_0)\right) \\
&+ \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (w-y)\right)(1-\beta)\lambda(1-\theta^{FSR}) \\
\iff &(\gamma v_h + (1-\gamma)m_h - w)\theta^{FS}\lambda\beta + b - c_\theta(\theta^{FS}-\theta_0) \geq \\
&(\gamma v_h + (1-\gamma)m_h - w)\theta^{FSR}\lambda\beta + b - c_\theta(\theta^{FSR}-\theta_0) - \left(b - c_\theta(\theta^{FSR}-\theta_0) - (v_h-m_h)\lambda(\beta + (1-\theta^{FSR})(1-\beta))\right)^+ \\
&+ \left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (1-\beta)(w-y) + (\gamma v_h + (1-\gamma)m_h - w)\beta\right)\lambda(1-\theta^{FSR}) \\
\iff &((\gamma v_h + (1-\gamma)m_h - w)\lambda\beta - c_\theta)(\theta^{FS}-\theta^{FSR}) \\
&+ \left(b - (v_h-m_h)\lambda(\beta + (1-\theta_0)(1-\beta)) - (c_\theta - (v_h-m_h)\lambda(1-\beta))(\theta^{FSR}-\theta_0)\right)^+ \geq \\
&\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (1-\beta)(w-y) + (\gamma v_h + (1-\gamma)m_h - w)\beta\right)\lambda(1-\theta^{FSR}) \\
\iff &((\gamma v_h + (1-\gamma)m_h - w)\lambda\beta - c_\theta)(\theta^{FS}-\theta^{FSR}) \\
&+ \left(b - (v_h-m_h)\lambda(\beta + (1-\theta_0)(1-\beta)) - \left(b - (v_h-m_h)\lambda(\beta + (1-\theta_0)(1-\beta))\right)^+ \mathbb{1}_{\left\{c_\theta \geq (v_h-m_h)\lambda(1-\beta)\right\}}\mathbb{1}_{\left\{y \leq y_\theta^S\right\}}\right)^+ \geq \\
&\left((1-\gamma)\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - (1-\beta)(w-y) + (\gamma v_h + (1-\gamma)m_h - w)\beta\right)\lambda(1-\theta^{FSR}) \\
\iff &\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) \leq g_{S3}(m_h),
\end{aligned}$$

where the first equivalence is by definition of $\Pi_D^{FS}$ and $\Pi_D^{FSR}$, the second and third equivalences follow by rearranging terms, the fourth equivalence follows from the definition of $\theta^{FSR}$, and the last equivalence follows by defining $g_{S3}(m_h)$ as

$$g_{S3}(m_h) = \frac{1}{(1-\beta)}\left[(1-\beta)(w-y) - (\gamma v_h + (1-\gamma)m_h - w)\beta + \big((\gamma v_h + (1-\gamma)m_h - w)\lambda\beta - c_\theta\big)\left(\frac{\theta^{FS} - \theta^{FSR}}{\lambda(1-\theta^{FSR})}\right)\right.$$
$$\left. + \big(b - (v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta))\big)^+ \frac{1 - \mathbb{1}_{\{y \le y_\theta^d, c_\theta \ge (v_h - m_h)\lambda(1-\beta)\}}}{\lambda(1-\theta^{FSR})}\right],$$

where $y_\theta^S = w - \frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) - (v_h - m_h)$. Similar to the previous two properties, assume first that $\gamma \le 1$, if $0 \le g_{S3}(m_h) \le \frac{(1-\gamma)\beta(v_h - m_h)}{1-\beta}$, then there exists a $\hat{\alpha}_{S3}$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_{S3}}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_{S3}(v_h - m)}}{1-\beta}\right) = g_{S3}(m_h).$$

We then define $a_{S3}(m_h)$ as

$$a_{S3}(m_h) = \begin{cases} \infty, & \text{if } g_{S3}(m_h) \le 0 \\ \hat{\alpha}_{S3}, & \text{if } 0 \le g_{S3}(m) \le \frac{(1-\gamma)\beta(v_h - m_h)}{1-\beta} \\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h - m_h)}{1-\beta} \le g_{S3}(m_h). \end{cases}$$

Again, the definition of $a_{S3}(m_h)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity.

Similarly, we denote the information level in strategy $FA$ and $FAR$ by $\theta^{FA}$ and $\theta^{FAR}$, respectively. We directly compare the distributor's value in $FA$ and $FAR$. Hence,

$$\Pi_D^A(\theta) = (\gamma v_h + (1-\gamma)m_l - w)\theta\beta + \min((v_h - m_l)\theta\beta, b - c_\theta(\theta - \theta_0)) - \frac{\lambda\theta\beta}{1-\lambda}(m_h - m_l)$$

$$\begin{aligned}
&\Pi_D^{FA} \ge \Pi_D^{FAR}(\alpha, m_l) \\
\iff &\left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\theta^{FA}\beta + \min((v_h - m_l)\theta^{FA}\beta, b - c_\theta(\theta^{FA} - \theta_0)) \ge \\
&(\gamma v_h + (1-\gamma)m_l - w)\beta + \min\big((v_h - m_l)(\beta + (1-\beta)(1-\theta^{FAR})), b - c_\theta(\theta^{FAR} - \theta_0)\big) \\
&- \frac{\lambda}{1-\lambda}((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)(1-\theta^{FAR})) \\
&+ \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right) - (w-y)\right)(1-\beta)(1-\theta^{FAR}) \\
\iff &\left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\theta^{FA}\beta + b - c_\theta(\theta^{FA} - \theta_0) - (b - c_\theta(\theta^{FA} - \theta_0) - (v_h - m_l)\theta^{FA}\beta)^+ \ge \\
&\left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\theta^{FAR}\beta + b - c_\theta(\theta^{FAR} - \theta_0) \\
&- \big(b - c_\theta(\theta^{FAR} - \theta_0) - (v_h - m_l)(\beta + (1-\theta^{FAR})(1-\beta))\big)^+ \\
&- \frac{\lambda}{1-\lambda}\left((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)\right)(1-\theta^{FAR}) \\
&+ \left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right) - (w-y) + (\gamma v_h + (1-\gamma)m_l - w)\frac{\beta}{1-\beta}\right)(1-\beta)(1-\theta^{FAR}) \\
\iff &\left(\left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\beta - c_\theta\right)(\theta^{FA} - \theta^{FAR}) \\
&- (b - (v_h - m_l)\theta_0\beta - (c_\theta + (v_h - m_l)\beta)(\theta^{FA} - \theta_0))^+ \\
&+ \big(b - (v_h - m_l)(\beta + (1-\theta_0)(1-\beta)) - (c_\theta - (v_h - m_l)(1-\beta))(\theta^{FAR} - \theta_0)\big)^+
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\lambda}{1-\lambda}\left((m_h-m_l)\beta-(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))(1-\beta)\right)(1-\theta^{FAR})\geq\\
&\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)-(w-y)+(\gamma v_h+(1-\gamma)m_l-w)\frac{\beta}{1-\beta}\right)(1-\beta)(1-\theta^{FAR})\\
\iff&\left(\left(\gamma v_h+(1-\gamma)m_l-w-\frac{\lambda}{1-\lambda}(m_h-m_l)\right)\beta-c_\theta\right)(\theta^{FA}-\theta^{FAR})\\
&-\left(b-(v_h-m_l)\theta_0\beta-(b-\theta_0\beta(v_h-m_l))^{+}\mathbb{1}_{\{w\leq w_\theta^A\}}\right)^{+}\\
&+\left(b-(v_h-m_l)(\beta+(1-\theta_0)(1-\beta))-\left(b-(v_h-m_l)(\beta+(1-\theta_0)(1-\beta))\right)^{+}\mathbb{1}_{\{c_\theta\geq(v_h-m_l)(1-\beta)\}}\mathbb{1}_{\{y\leq y_\theta^b\}}\right)^{+}\\
&+\frac{\lambda}{1-\lambda}\left((m_h-m_l)\beta-(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))(1-\beta)\right)(1-\theta^{FAR})\geq\\
&\left(\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)-(w-y)+(\gamma v_h+(1-\gamma)m_l-w)\frac{\beta}{1-\beta}\right)(1-\beta)(1-\theta^{FAR})\\
\iff&\frac{1-\gamma}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)\leq g_{A3}(m_l),
\end{aligned}
$$

where the first equivalence is by definition of $\Pi_D^{FA}$ and $\Pi_D^{FAR}$, the second and third equivalences follow by rearranging terms, the fourth equivalence follows from the definition of $\theta^{FA}$ and $\theta^{FAR}$, and the last equivalence follows by defining $g_{A3}(m_l)$ as

$$
\begin{aligned}
g_{A3}(m_l)=&\frac{1}{1-\beta}\Big[(w-y)(1-\beta)-(\gamma v_h+(1-\gamma)m_l-w)\beta\\
&+\left(\left(\gamma v_h+(1-\gamma)m_l-w-\frac{\lambda}{1-\lambda}(m_h-m_l)\right)\beta-c_\theta\right)\left(\frac{\theta^{FA}-\theta^{FAR}}{1-\theta^{FAR}}\right)\\
&-(b-\theta_0\beta(v_h-m_l)^{+}\frac{\mathbb{1}_{\{w>w_\theta^A\}}}{1-\theta^{FAR}}\\
&+\left(b-(v_h-m_l)(\beta+(1-\theta_0)(1-\beta))\right)^{+}\frac{1-\mathbb{1}_{\{y\leq y_\theta^b,c_\theta\geq(v_h-m_l)(1-\beta)\}}}{1-\theta^{FAR}}\\
&+\frac{\lambda}{1-\lambda}\left((m_h-m_l)\beta-(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))(1-\beta)\right)\Big],
\end{aligned}
$$

where $w_\theta^A=v_h+\gamma(v_h-m_l)-\frac{\lambda}{1-\lambda}(m_h-m_l)$ and $y_\theta^b=w-(1-\gamma)\frac{1}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)-(v_h-m_l)-\frac{\lambda}{1-\lambda}(m_l-r_\alpha(m_l)-(m_h-r_\alpha(m_h)))$. Similar to the previous two properties, assume first that $\gamma\leq 1$, if $0\leq g_{A3}(m_l)\leq\frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}$, then there exists a $\hat{\alpha}_{A3}$ that satisfies

$$\frac{1-\gamma}{\hat{\alpha}_{A3}}\ln\left(\frac{1-\beta e^{-\hat{\alpha}_{A3}(v_h-m_l)}}{1-\beta}\right)=g_{A3}(m_l).$$

We then define $a_{A3}(m_l)$ as

$$a_{A3}(m_l)=\begin{cases}\infty, & \text{if } g_{A3}(m_l)\leq 0\\ \hat{\alpha}_{A3}, & \text{if } 0\leq g_{A3}(m_l)\leq\frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}\\ 0, & \text{if } \frac{(1-\gamma)\beta(v_h-m_l)}{1-\beta}\leq g_{A3}(m_l).\end{cases}$$

Again, the definition of $a_{A3}(m_l)$ when $\gamma>1$ is analogous and is skipped for the sake of brevity. This establishes Property 4'.

*Proof of Property 5'.* We show the first sentence in the first statement in Property 5'. We first show that $a_{S1}(m^{FS})=a_{S2}(m^{FS})$. By definition $m^{FS}$ satisfies $(\gamma v_h+(1-\gamma)m^{FS}-w)\lambda\beta=c_\theta$. Therefore,

$$g_{S2}(m^{FS})=w-y-(\gamma v_h+(1-\gamma)m^{FS}-w)\frac{\beta}{1-\beta}=w-y-\frac{c_\theta}{\lambda(1-\beta)}.$$

We now show that $a_{S1}(m_h) \geq a_{S2}(m_h)$ if and only if $m_h \leq m^{FS}$ . Since $(\gamma v_h + (1-\gamma)m_h - w)\lambda\beta \geq c_\theta$ if and only if $m_h \geq m^{FS}$, then $w - y - \frac{c_\theta}{\lambda(1-\beta)} \geq g_{S2}(m_h)$ if and only if $m_h \geq m^{FS}$, hence $a_{S1}(m_h) \leq a_{S2}(m_h)$ if and only if $m_h \geq m^{FS}$.

We now prove the second sentence in the first statement in Property 5. Let $m^{MSR} = v_h - \frac{(1-\beta)\lambda(w-y)-c_\theta}{(1-\gamma)\lambda\beta}$. Then, from the definition of $a_{S1}(m_h)$ in Property 2' we have that if $m_h \geq m^{MSR}$ then $a_{S1}(m_h) = 0$, i.e., strategy MSR dominates strategy FSR for all $\alpha \geq a_{S1}(m_h) = 0$. Moreover, the first sentence in the first statement in Property 5' then implies $m^{MSR} \geq m^{FS}$ while $c_\theta \leq (1-\beta)\lambda(w-y)$ implies $m^{MSR} \leq v_h$. Furthermore, Property 1' then also implies that strategy MSR dominates strategy FS for all $\alpha \geq a_{S1}(m_h) = 0$ and $m_h \geq m^{FS}$.

Analogously, let $\bar{m}^{MSR} = v_h - \frac{\beta v_h + (1-\beta)y - w}{(1-\gamma)\beta}$. Then, from the definition of $a_{S2}(m_h)$ in Property 3' we have that if $m_h \geq \bar{m}^{MSR}$ then $a_{S2}(m_h) = \infty$, i.e., strategy MSR dominates strategy MS for all $\alpha \leq a_{S2}(m_h) = \infty$. Moreover, the first sentence in the first statement in Property 5' implies $\bar{m}^{MSR} \geq m^{FS}$, while $w \leq \beta v_h + (1-\beta)y$ implies $\bar{m}^{MSR} \leq v_h$, establishing the first statement in Property 5'.

Similarly, we now show the first sentence in the second statement in Property 5'. Namely, we first show that $a_{A1}(m^{FA}) = a_{A2}(m^{FA})$. By definition $m^{FA}$ satisfies $(\gamma v_h + (1-\gamma)m^{FA} - w)\beta = c_\theta + \frac{\lambda\beta}{1-\lambda}(m_h - m^{FA})$. Therefore,

$$\begin{aligned} g_{A2}(m^{FA}) &= w - y + \frac{\lambda}{1-\lambda}\left(\frac{m_h - m^{FA}}{1-\beta} - (r_\alpha(m_h) - r_\alpha(m^{FA}))\right) - (\gamma v_h + (1-\gamma)m^{FA} - w)\frac{\beta}{1-\beta} \\ &= w - y - \frac{\lambda}{1-\lambda}\left(m^{FA} - r_\alpha(m^{FA}) - (m_h - r_\alpha(m_h))\right) - \frac{c_\theta}{1-\beta} = g_{A1}(m^{FA}) \end{aligned}$$

We now show that $a_{A1}(m_l) \geq a_{A2}(m_l)$ if and only if $m \leq m^{FA}$ . Since $(\gamma v_h + (1-\gamma)m_l - w)\beta \geq c_\theta + \frac{\lambda\beta}{1-\lambda}(m_h - m_l)$ if and only if $m_l \geq m^{FA}$, then $g_{A1}(m_l) \geq g_{A2}(m_l)$ if and only if $m_l \geq m^{FA}$, hence $a_{A1}(m_l) \leq a_{A2}(m_l)$ if and only if $m_l \geq \bar{m}_l$.

To conclude, we now prove the second sentence in the second statement in Property 5. Let $m^{MAR} = v_h - \frac{(1-\beta)\lambda(w-y)-c_\theta}{(1-\gamma)\lambda\beta} + \frac{\lambda}{1-\lambda}(1-\beta)(m^{MAR} - r_\alpha(m^{MAR}) - (m_h - r_\alpha(m_h)))$. Then, from the definition of $a_{A1}(m_l)$ in Property 2' we have that if $m_l \geq m^{MAR}$ then $a_{A1}(m_l) = 0$, i.e., strategy MAR dominates strategy FAR for all $\alpha \geq a_{A1}(m_l) = 0$. Moreover, the first sentence in the first statement in Property 5' then implies $m^{MAR} \geq m^{FA}$ while $c_\theta \leq (1-\beta)(w-y) - \lambda/(1-\lambda)(1-\beta)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))$ implies $m^{MAR} \leq v_h$. Furthermore, Property 1' then also implies that strategy MAR dominates strategy FA for all $\alpha \geq a_{A1}(m_l) = 0$ and $m_l \geq m^{FA}$.

Analogously, let $\bar{m}^{MAR} = v_h - \frac{\beta v_h + (1-\beta)y - w}{(1-\gamma)\beta} + \frac{\lambda}{1-\lambda}\frac{(m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)}{(1-\gamma)\beta}$. Then, from the definition of $a_{A2}(m_l)$ in Property 3' we have that if $m_l \geq \bar{m}^{MAR}$ then $a_{A2}(m_l) = \infty$, i.e., strategy MAR dominates strategy MA for all $\alpha \leq a_{A2}(m_l) = \infty$. Moreover, the first sentence in the first statement in Property 5' implies $\bar{m}^{MAR} \geq m^{FA}$, while $w \leq \beta v_h + (1-\beta)y - \lambda/(1-\lambda)\left((m_h - m_l)\beta - (m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))(1-\beta)\right)$ implies $\bar{m}^{MAR} \leq v_h$, establishing Property 5'.

Hence, we have shown Properties 1' to 5', completing the proof of Proposition C.2. □

### C.4. Proof of Corollary 2

*Proof.* To prove the corollary we build on Properties 1'-9' from the proof of Proposition C.2, and show the following additional properties at the end of this proof.

- Property 10': if $\gamma c_\theta \le (\lambda/(1-\lambda))^2\beta(m_h - m_l)$ and $\Pi_D^{FS} \ge \Pi_D^{MS}$ then $\Pi_D^{FS} \ge \Pi_D^{FA}$ if and only if $m_l \le m^S$;
- Property 11': $\Pi_D^{MAR} \ge \Pi_D^{MSR}$ if and only if $\alpha \le a_{MAR}(m_l)$.

Properties 1'-11' prove the corollary. Namely,

- If $\gamma c_\theta \le (\lambda/(1-\lambda))^2\beta(m_h - m_l)$ then

$$\Pi_D^{FS} \ge \max(\Pi_D^{MS}, \Pi_D^{FSR}, \Pi_D^{MSR}, \Pi_D^{FA}, \Pi_D^{MA}, \Pi_D^{FAR}, \Pi_D^{MAR})$$

  if and only if $m_l \le m^{FAS} = \min(m^{FA}, m^S)$, $m_h \le m^{FS}$ and $\alpha \ge a_{FS}(m_l, m_h) = \max(a_{A3}(m_l), a_{S3}(m_h))$ (combine Properties 1', 4', 5', 6' and 10');

- If $m_l \ge \hat{m}^{MAR} = \max(m^{MAR}, \bar{m}^{MAR})$, $m_h \ge \hat{m}^{MSR} = \max(m^{MSR}, \bar{m}^{MSR})$ and $\alpha \le a_{MAR}(m_l)$, then

$$\Pi_D^{MAR} \ge \max(\Pi_D^{MS}, \Pi_D^{FSR}, \Pi_D^{MSR}, \Pi_D^{MA}, \Pi_D^{FAR}, \Pi_D^{FS}),$$

  (combine Properties 1', 2', 3', 5', and 11').

We now complete the proof of the corollary by proving Properties 10' and 11'.

*Proof of Property 10.* Let $m^S$ be the unique value of $m_l$ that satisfies

$$m_l = \frac{w - \gamma v_h}{1-\gamma} + \frac{\lambda}{1-\lambda}(m_h - m_l)\left(1 + \frac{1}{(1-\lambda)(1-\gamma)}\right) - \frac{(c_\theta - (\gamma v_h + (1-\gamma)m_h - w)}{(1-\gamma)(1-\lambda)\beta}\frac{\theta^{FS} - \theta^{FA}}{\theta^{FA}}. \qquad (38)$$

We analyze three exhaustive cases.

First, assume $b < \lambda\theta_0\beta(v_h - m_h) \le \theta_0\beta(v_h - m_l)$ then $\theta^{FS} = \theta^{FS} = \theta_0$ and after some algebra we conclude that in this case

$$\Pi_D^{FS} \ge \Pi_D^{FA} \iff m_l \le \frac{w - \gamma v_h}{1-\gamma} + \frac{\lambda}{1-\lambda}(m_h - m_l)\left(1 + \frac{1}{(1-\lambda)(1-\gamma)}\right) \equiv m^S,$$

concluding the proof of Property 10 in this case.

Second, if $\lambda\theta_0\beta(v_h - m_h) \le b < \theta_0\beta(v_h - m_l)$ then $\theta^{FS} > \theta^{FA} = \theta_0$ and after some algebra we conclude

$$\begin{aligned}\Pi_D^{FS} \ge \Pi_D^{FA} \iff & (c_\theta - (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta)(c_\theta\theta_0 + b) \\ & \le \left(c_\theta - \left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\beta\right)\theta_0(c_\theta + \lambda\beta(v_h - m_h)). \end{aligned} \qquad (39)$$

We show that if $\gamma c_\theta \le (\lambda/(1-\lambda))^2\beta(m_h - m_l)$ and $\Pi_D^{FS} \ge \Pi_D^{SR}(\theta)$ then the left hand side of (39) increases strictly faster in $m_h$ than the right hand side in $m_h$, $m_l$. Indeed, after some algebra the latter statement holds if and only if

$$(1-\gamma)\lambda(\theta_0\beta(v_h - m_l) - b) + \theta_0(1-\lambda)\left(\left(\frac{\lambda}{1-\lambda}\right)^2\beta(m_h - m_l) - \gamma c_\theta\right) + \theta_0(c_\theta - (\gamma v_h + (1-\gamma)m_h - w) > 0,$$

where the first term is strictly positive in this case, and the second and third term are positive by assumption (the third term being positive is equivalent to $\Pi_D^{FS} \ge \Pi_D^{SR}(\theta)$). Hence, (38) follows directly from (39).

Finally, if $\lambda\theta_0\beta(v_h - m_h) \le \theta_0\beta(v_h - m_l) \ge b$ then $\theta^{FS} \ge \theta^{FA} > \theta_0$ then

$$\Pi_D^{FS} \ge \Pi_D^{FA} \iff (c_\theta - (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta)(c_\theta + \beta(v_h - m_l)) \\ \le \left(c_\theta - \left(\gamma v_h + (1-\gamma)m_l - w - \frac{\lambda}{1-\lambda}(m_h - m_l)\right)\beta\right)(c_\theta + \lambda\beta(v_h - m_h)). \tag{40}$$

After some algebra, it is not hard to see that the left hand side of (40) increases strictly faster in $m_h$, $m_l$, than the right hand side in $m_h$, $m_l$ if and only if $\gamma c_\theta \le (\lambda/(1-\lambda))^2\beta(m_h - m_l)$, which holds by assumption. Hence, (38) follows directly from (39) completing the proof of Property 10'.

*Proof of Property 11'.* We now show that $\Pi_D^{MAR} \ge \Pi_D^{MSR}$ if and only if $\alpha \le a_{MAR}(m_l)$. From equations (30) and (36) we conclude, after some algebra, that

$$\Pi_D^{MAR} \ge \Pi_D^{MSR} \\ \iff \left((1-\gamma)(m_l - r_\alpha(m_l)) + \frac{\lambda}{1-\lambda}\left((1-\gamma) + \frac{1}{1-\lambda}\right)(m_l - r_\alpha(m_l) - (m_h - r_\alpha(m_h)))\right)(1-\beta)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+ \\ \ge \frac{\lambda\beta}{1-\lambda}(m_h - m_l)\left((1-\gamma) + \frac{1}{1-\lambda}\right) + (w-y)(1-\beta)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+ - (\gamma v_h + (1-\gamma)m_l - w)\beta \tag{41}$$

To simplify the notation, let lhs(41) and rhs(41) denote the left and right hand side of (41), respectively. As before, first assume that $\gamma \le 1$, note that then lhs(41) is monotonically decreasing in $\alpha$ and has range $\left[0, \left((1-\gamma)\beta(v_h - m_l) + \frac{\lambda}{1-\lambda}\left((1-\gamma) + \frac{1}{1-\lambda}\right)\beta(m_h - m_l)\right)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+\right]$. If rhs(41) falls in this range then there exists $\hat{\alpha}_{MAR}$ that satisfies (41) with equality. Then, we define $a_{MAR}(m_l)$ as

$$a_{MAR}(m_l) = \begin{cases} \infty, & \text{if lhs}(41) \le 0 \\ \hat{\alpha}_{bd}, & \text{if } 0 \le \text{lhs}(41) \le \left((1-\gamma)\beta(v_h - m_l) + \frac{\lambda}{1-\lambda}\left((1-\gamma) + \frac{1}{1-\lambda}\right)\beta(m_h - m_l)\right)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+ \\ 0, & \text{if } \left((1-\gamma)\beta(v_h - m_l) + \frac{\lambda}{1-\lambda}\left((1-\gamma) + \frac{1}{1-\lambda}\right)\beta(m_h - m_l)\right)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+ \le \text{lhs}(41). \end{cases}$$

The definition of $a_{MAR}(m_l)$ when $\gamma > 1$ is analogous and is skipped for the sake of brevity. This completes the proof of Property 11', thus the proof of the corollary. □

### C.5. Consumer Surplus Analysis with Heterogeneous ATPs

We start by extending Lemma B.1 to hold for heterogeneous ATPs.

**Lemma C.2** *Let $\theta^i$, $i \in \{FA, MA, FAR, MAR, FS, MS, FSR, MSR\}$ be the education level of each strategy from Theorem 2. Then, under Assumption 1,*

$$\theta^{MS} = \theta^{MSR} \ge \theta^{FS} \ge \theta^{FSR}, \quad \theta^{MA} = \theta^{MAR} \ge \theta^{FA} \ge \theta^{FAR}. \tag{42}$$

*Proof.* From their definition, in the proof of Theorem 2, we have $\theta^{MA} = \theta^{MS} = \theta^{MAR} = \theta^{MSR}$, where they are equal to the natural upper bound on $\theta$. Moreover, from Assumption 1 it follows that $\theta^{MS} = \theta^{MSR} \ge \theta^{FS}$ and $\theta^{MA} = \theta^{MAR} \ge \theta^{FA}$.

We now argue that $\theta^{FS} \ge \theta^{FSR}$. If $\theta^{FSR} = \theta_0$ then $\theta^{FS} \ge \theta^{FSR}$ follows trivially from their definition, in the proof of Theorem 2. If $\theta^{FSR} > \theta_0$ then by definition again we must have $c_\theta > (v_h - m_h)\lambda(1-\beta) > 0$, $c_\theta(\theta^{FSR} - \theta_0) + (v_h - m_h)\lambda(\theta^{FSR}\beta + (1 - \theta^{FSR})) = b$, and $c_\theta(\theta^{FS} - \theta_0) + (v_h - m_h)\theta^{FS}\lambda\beta = b$, hence we conclude $\theta^{FS} \ge \theta^{FSR}$.

We now argue that $\theta^{FA} \ge \theta^{FAR}$. If $\theta^{FAR} = \theta_0$ then $\theta^{FA} \ge \theta^{FAR}$ follows trivially from their definition, in the proof of Theorem 2. If $\theta^{FAR} > \theta_0$ then by definition again we must have $c_\theta > (v_h - m_l)(1-\beta) > 0$, $c_\theta(\theta^{FAR} - \theta_0) + (v_h - m_l)(\theta^{FAR}\beta + (1 - \theta^{FAR})) = b$, and $c_\theta(\theta^{FA} - \theta_0) + (v_h - m_l)\theta^{FA}\beta = b$, hence we conclude $\theta^{FA} \ge \theta^{FAR}$, completing the proof. □

### C.6. Proof of Proposition 5

*Proof.* Recall, from Proposition C.1, that

$$CS^i = (v_h - m_l)\theta^i\beta, \quad i \in \{FA, MA\},$$
$$CS^i = (v_h - m_h)\lambda\theta^i\beta, \quad i \in \{FS, MS\},$$

and

$$CS^i = (v_h - m_l)\beta - \frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)\left(1-\theta^i\right), \quad i \in \{FAR, MAR\},$$
$$CS^i = (v_h - m_h)\lambda\beta - \frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)\lambda\left(1-\theta^i\right), \quad i \in \{FSR, MSR\}.$$

We prove the proposition by showing the following results,

(i) $CS^{MSR} \geq \max(CS^{MS}, CS^{FS}, CS^{FSR})$; and $CS^{MAR} \geq \max(CS^{MA}, CS^{FA}, CS^{FAR})$.

(ii) $CS^{MSR} \leq CS^{MAR}$;

(iii) $CS^{FS} \leq \min(CS^{MS}, CS^{FSR}, CS^{MSR})$ if and only if $\alpha \geq a_{S4}(m_h)$; and $CS^{FA} \leq \min(CS^{MA}, CS^{FAR}, CS^{MAR})$ if and only if $\alpha \geq a_{A4}(m_l)$;

(iv) If $\Pi_D^{FA} \geq 0$ then $CS^{FS} \leq CS^{FA}$.

Then, the first statement in the proposition follows from $(i)$ and $(ii)$. While the second statement in the proposition follows from $(iii)$ and $(iv)$ by taking $a_{FS}^{cs}(m_h, m_l) = \max(a_{S4}(m_h), a_{A4}(m_l))$.

We first prove $(i)$. We start by showing that $CS^{MSR} \geq \max\left(CS^{MS}, CS^{FSR}, CS^{FS}\right)$. In particular, since $\theta^{MSR} \geq \theta^{FSR}$ from Lemma C.2 then $CS^{MSR} \geq CS^{FSR}$. Moreover,

$$CS^{MSR} \geq (v_h - m_h)\lambda\theta^{MSR}\beta = (v_h - m_h)\lambda\theta^{MS}\beta = CS^{MS} \geq (v_h - m_h)\lambda\theta^{FS}\beta = CS^{FS},$$

where the first inequality follows since $\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)$ is decreasing in $\alpha > 0$ and taking the limit $\alpha \to 0$. The second inequality follows since, from Lemma C.2, $\theta^{MS} = \theta^{MSR} \geq \theta^{FS}$. Hence, we conclude $CS^{MSR} \geq \max\left(CS^{FS}, CS^{MS}, CS^{FSR}\right)$.

Similarly, we now show that $CS^{MAR} \geq \max\left(CS^{MA}, CS^{FAR}, CS^{FA}\right)$. In particular, since $\theta^{MAR} \geq \theta^{FAR}$ from Lemma C.2 then $CS^{MAR} \geq CS^{FAR}$. Moreover,

$$CS^{MAR} \geq (v_h - m_l)\theta^{MAR}\beta = (v_h - m_l)\theta^{MA}\beta = CS^{MA} \geq (v_h - m_l)\theta^{FA}\beta = CS^{FA},$$

where the first inequality follows since $\frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)$ is decreasing in $\alpha > 0$ and taking the limit $\alpha \to 0$. The second inequality follows since, from Lemma C.2, $\theta^{MA} = \theta^{MAR} \geq \theta^{FA}$. Hence, we conclude $CS^{MAR} \geq \max\left(CS^{FA}, CS^{MA}, CS^{FAR}\right)$, completing the proof of $(i)$.

We now show $(ii)$, i.e., $CS^{MSR} \leq CS^{MAR}$. Indeed, note that

$$\beta(m_h - m_l) + \frac{1-\beta}{\alpha}\left(\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right) - \ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)\right)\left(1-\theta^{MSR}\right) \geq \theta^{MSR}\beta(m_h - m_l) \geq 0, \tag{43}$$

where the first inequality follows from noticing that the right hand side is monotonically increasing in $\alpha$ and taking the limit as $\alpha \to 0$. Thus, we conclude that the first term in (43) is non-negative, which is equivalent to the second inequality in the following chain of inequalities

$$\lambda \leq 1 \leq \frac{(v_h - m_l)\beta - \frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_l)}}{1-\beta}\right)(1-\theta^{MAR})}{(v_h - m_h)\beta - \frac{1-\beta}{\alpha}\ln\left(\frac{1-\beta e^{-\alpha(v_h-m_h)}}{1-\beta}\right)(1-\theta^{MSR})}, \tag{44}$$

which completes the proof by noticing that (44) is equivalent to $CS^{MSR} \leq CS^{MAR}$.

We now prove *(iii)*. We start by showing that there exists $a_{S4}(m_h)$ such that $\min(CS^{MS}, CS^{FSR}, CS^{MSR}) \geq CS^{FS}$ if and only if $\alpha \geq a_{S4}(m_h)$. We have already shown $CS^{MSR} \geq CS^{MS} \geq CS^{FS}$. We now show that $CS^{FSR} \geq CS^{FS}$ if and only if $\alpha \geq a_{S4}(m_h)$. In fact, we have

$$CS^{FSR} \geq CS^{FS} \iff \frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) \leq \beta(v_h - m_h)\frac{(1-\theta^{FS})}{(1-\theta^{FSR})},$$

where the term $\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $[0, \beta(v_h - m_h)]$. Since from Lemma C.2 we have $\theta^{FS} \geq \theta^{FSR}$, then $0 \leq \beta(v_h - m_h)\frac{(1-\theta^{FS})}{(1-\theta^{FSR})} \leq \beta(v_h - m_h)$. Moreover, since $\theta^{FS}$ is independent of $\alpha$ and $\theta^{FSR}$ is non-decreasing in $\alpha$ it follows that there exists $a_{S4}(m_h)$ such that

$$\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h - m_h)}}{1-\beta}\right) \leq \beta(v_h - m_h)\frac{(1-\theta^{FS})}{(1-\theta^{FSR})} \iff \alpha \geq a_{S4}(m_h).$$

Hence, we conclude $\min(CS^{MS}, CS^{FSR}, CS^{MSR}) \geq CS^{FS}$ if and only if $\alpha \geq a_{S4}(m_h)$.

Similarly, we now show that there exists $a_{A4}(m_l)$ such that $\min(CS^{MA}, CS^{FAR}, CS^{MAR}) \geq CS^{FA}$ if and only if $\alpha \geq a_{A4}(m_l)$. We have already shown $CS^{MAR} \geq CS^{MA} \geq CS^{FA}$. We now show that $CS^{FAR} \geq CS^{FA}$ if and only if $\alpha \geq a_{A4}(m_l)$. In fact, we have

$$CS^{FAR} \geq CS^{FA} \iff \frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right) \leq \beta(v_h - m_l)\frac{(1-\theta^{FA})}{(1-\theta^{FAR})},$$

where the term $\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right)$ is monotonically decreasing in $\alpha$ and has range $[0, \beta(v_h - m_l)]$. Hence, if $\theta^{FA} \leq \theta^{FAR}$ we conclude $CS^{FAR} \geq CS^{FA}$ and thus we set $a_{S4}(m_h) = 0$. Alternatively, if $\theta^{FA} > \theta^{FAR}$ then $0 \leq \beta(v_h - m_l)\frac{(1-\theta^{FA})}{(1-\theta^{FAR})} \leq \beta(v_h - m_l)$. Moreover, since $\theta^{FA}$ is independent of $\alpha$ and $\theta^{FAR}$ is non-decreasing in $\alpha$ it follows that there exists $a_{A4}(m_l)$ such that

$$\frac{1-\beta}{\alpha} \ln\left(\frac{1-\beta e^{-\alpha(v_h - m_l)}}{1-\beta}\right) \leq \beta(v_h - m_l)\frac{(1-\theta^{FA})}{(1-\theta^{FAR})} \iff \alpha \geq a_{A4}(m_l).$$

Hence, we conclude $\min(CS^{MS}, CS^{FSR}, CS^{MSR}) \geq CS^{FS}$ if and only if $\alpha \geq a_{A4}(m_h)$, completing the proof of *(iii)*.

Finally, we conclude by proving *(iv)*, i.e., that $\Pi_D^{FA} \geq 0$ implies $CS^{FS} \leq CS^{FA}$. Note that $\Pi_D^{FA} \geq 0$ implies

$$\lambda \leq \frac{(v_h - m_l)(c_\theta + \lambda\beta(v_h - m_h))}{(v_h - m_h)(c_\theta + \beta(v_h - m_l))} \leq \frac{(v_h - m_l)\theta^{FA}}{(v_h - m_h)\theta^{FS}}, \tag{45}$$

where the first inequality is equivalent to $\lambda \leq (v_h - m_l)/(v_h - m_h)$, which holds for any $\lambda$ satisfying the assumption $\lambda \leq m_l/m_h$. We now verify the second inequality holds in three exhaustive cases. First, if $b \leq \lambda\theta_0\beta(v_h - m_h) \leq \theta_0\beta(v_h - m_l)$ then $\theta^{FS} = \theta^{FA} = \theta_0$ and the second inequality holds since $(c_\theta + \lambda\beta(v_h - m_h))/(c_\theta + \beta(v_h - m_l)) \leq 1 = \theta^{FA}/\theta^{FS}$. Second, if $\lambda\theta_0\beta(v_h - m_h) \leq b \leq \theta_0\beta(v_h - m_l)$ then $\theta^{FS} > \theta^{FA} = \theta_0$ and the second inequality holds since $(c_\theta + \lambda\beta(v_h - m_h))/(c_\theta + \beta(v_h - m_l)) \leq \theta_0(c_\theta + \lambda\beta(v_h - m_h))/(c_\theta\theta_0 + b) = \theta^{FA}/\theta^{FS}$, where the latter inequality follows from $b \leq \theta_0\beta(v_h - m_l)$. Third, if $\lambda\theta_0\beta(v_h - m_h) \leq \theta_0\beta(v_h - m_l) \leq b$ then $\Pi_D^{FA} \geq 0$ implies $\theta^{FS} \geq \theta^{FA} > \theta_0$ and the second inequality becomes tight. To conclude, note that the extremes of the chain of inequalities (45) are equivalent to $CS^{FS} \leq CS^{FA}$, completing the proof. $\square$

### C.7. Free Refunds Analysis with Heterogeneous ATP

The problem formulation is the same, see Problem (19), and the analysis is analogous to the case with homogeneous ATP in Section B.10. Namely, we first adapt Proposition C.1 to characterize the distributor's pricing strategy in this restricted setup, in the next proposition.

**Proposition C.3** *Consider any consumer education level $\theta \in [0,1)$, ATPs $m_l$, $m_h$, $w \le m_l < m_h \le v_h$, and subsidy $x \in [0, v_h - m_h]$. Then, the distributor's optimal objective is*

$$\Pi_D^*(x,\theta) = \max\left\{\Pi_D^{ARf}(x,\theta), \Pi_D^{SRf}(x,\theta)\right\}. \tag{46}$$

*Where $\Pi_D^i(x,\theta)$ $i \in \{ARf, SRf\}$ each correspond to the distributor's profits in a non-dominated strategy. Specifically, they are characterized by:*

*(ARf) Target all informed and uninformed customers with product returns. The customer price is $p^{ARf} = m_l + x$ and the refund is $r^{ARf} = m_l$. Let $\bar{z}^{ARf} = m_l - \frac{\hat{O}_R^{ARf}}{(1-\beta)\theta\beta}$. Then, without loss of generality, the equilibrium retailer refund is $z^{ARf} = \max(0, \bar{z})$ and the retailer price is $c^{ARf} = m_l + x - \frac{\hat{O}_R^{ARf}}{\theta\beta} + \frac{(1-\theta)(1-\beta)}{\theta\beta+(1-\theta)}(z^{ARf} - \bar{z}^{ARf})$. The consumer surplus is $CS^{ARf} = (v_h - m_l)\beta$, the retailer's profit is $\Pi_R^{ARf} = \hat{O}_R^{ARf}$, and the distributor's profit is*

$$\Pi_D^{ARf}(x,\theta) = (m_l + x - w)\beta + (x - w + y)(1-\beta)(1-\theta) - \hat{O}_R^{ARf} + \gamma CS^{ARf},$$

*where*

$$\hat{O}_R^{ARf} = \frac{\lambda\beta(m_h - m_l)}{1-\lambda}.$$

*Moreover, this strategy can be sustained in equilibrium if and only if*

$$\hat{O}_R^{ARf} \le m_l\beta + x(\beta + (1-\theta)(1-\beta)).$$

*(SRf) Target both informed and uninformed customers with product returns. The customer price is $p^{SRf} = m_h + x$ and the refund is $r^{SRf} = m_h$. The equilibrium retailer price is $c^{SRf} = m_h + x$, and the retailer refund is $z^{SRf} = m_h$. The consumer surplus is $CS^{SRf} = (v_h - m_h)\lambda\beta$, the retailer attains no profit, $\Pi_R^{SRf} = 0$, and the distributor's profit is*

$$\Pi_D^{SRf}(x,\theta) = (m_h + x - w)\lambda\beta + (x - w + y)\lambda(1-\beta)(1-\theta) + \gamma CS^{SRf}.$$

*Moreover, this strategy can always be sustained in equilibrium.*

Proposition C.3 is a special case of Proposition C.1 when $\alpha \to \infty$. Therefore, we omit the proof.

### C.8. Proof of Proposition 6

*Proof.* The proof of the first part of the proposition is the same as the proof of Theorem 2 for the special case when $\alpha \to \infty$. Therefore, we only focus on the expressions that change. Specifically,

$$\Pi_D^{MSRf} = \quad (\gamma v_h + (1-\gamma)m_h - w)\lambda\beta + (b - c_\theta(1-\theta_0))^+ - (1-\beta)(w-y)\lambda\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+. \tag{47}$$

To simplify the notation, recall the function $\frac{1}{x^{++}} := \begin{cases} 0 & \text{if } x \le 0 \\ \frac{1}{x} & \text{if } x > 0. \end{cases}$ Then,

$$\theta^{FSRf} = \theta_0 + \frac{\big(b - (v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta))\big)^+}{\big(c_\theta - (v_h - m_h)\lambda(1-\beta)\big)^{++}} \mathbb{1}_{\{v_h - m_h \le w - y\}},$$

and

$$\begin{aligned} \Pi_D^{FSRf} = \quad & (\gamma v_h + (1-\gamma)m_h - w)\,\lambda\beta + \min\left((v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta)), b\right) \\ & - (w-y)\,(1-\beta)\lambda\,(1-\theta_0) + (w - y - (v_h - m_h))^+\,(1-\beta)\lambda \frac{\big(b - (v_h - m_h)\lambda(\beta + (1-\theta_0)(1-\beta))\big)^+}{\big(c_\theta - (v_h - m_h)\lambda(1-\beta)\big)^{++}}. \end{aligned} \tag{48}$$

Similarly,

$$\Pi_D^{MARf} = \quad (\gamma v_h + (1-\gamma)m_l - w)\beta + (b - c_\theta(1-\theta_0))^+ - (1-\beta)(w-y)\left(1 - \theta_0 - \frac{b}{c_\theta}\right)^+ - \frac{\lambda\beta(m_h - m_l)}{1-\lambda}, \tag{49}$$

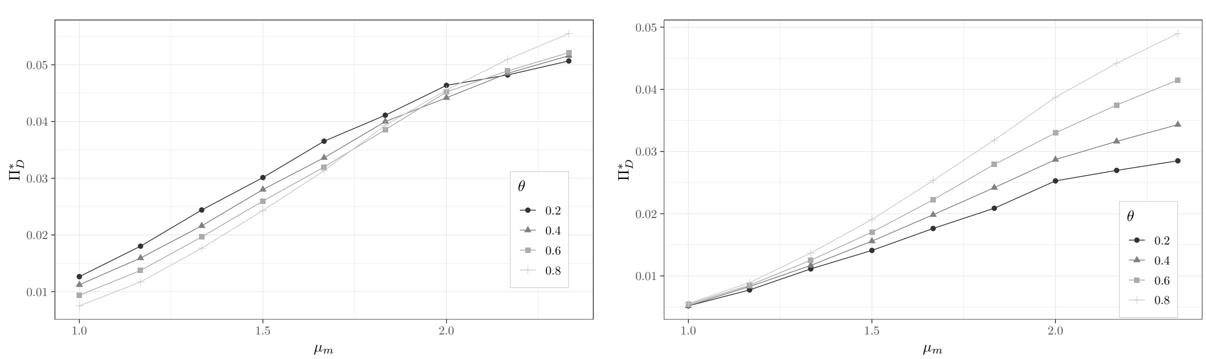
$$\theta^{FARf} = \theta_0 + \frac{\big(b - (v_h - m_l)(\beta + (1-\theta_0)(1-\beta))\big)^+}{\big(c_\theta - (v_h - m_l)(1-\beta)\big)^{++}} \mathbb{1}_{\{v_h - m_l \le w - y\}},$$

and

$$\begin{aligned} \Pi_D^{FARf} = \quad & (\gamma v_h + (1-\gamma)m_l - w)\,\beta + \min\left((v_h - m_l)(\beta + (1-\theta_0)(1-\beta)), b\right) - \frac{\lambda\beta(m_h - m_l)}{1-\lambda} \\ & - (w-y)\,(1-\beta)\,(1-\theta_0) + (w - y - (v_h - m_l))^+\,(1-\beta) \frac{\big(b - (v_h - m_l)(\beta + (1-\theta_0)(1-\beta))\big)^+}{\big(c_\theta - (v_h - m_l)(1-\beta)\big)^{++}}. \end{aligned} \tag{50}$$

The proof of the second part in the Proposition follows directly from Propositions B.5 and C.3 □

## Appendix D: Figure from Section 6




(a) $\alpha = 0.01$ (b) $\alpha = 1.0$

**Figure 8 Distributor's profits for different average ATP $\mu_m$ and consumer education level $\theta$. We assume $w = \$1.0$, $y = \$0.7$, $\mu = \$0.95$, $\sigma = \$1.0$, $\sigma_m = \$0.5$, $\kappa = -0.5$, $\rho_l = 0.1$, $\rho_h = 0.9$, $x = 0$, and $\gamma = 0$.**