

# Screening with Frames\*

Franz Ostrizek<sup>†</sup>      Denis Shishkin<sup>‡</sup>

April 18, 2019

## Abstract

We analyze screening with frame-dependent valuations. A monopolist principal designs an extensive-form decision problem with frames at each stage. This allows the firm to induce dynamic inconsistency and thereby reduce information rents. We show that the optimal extensive form has a simple three-stage structure and uses only the two highest frames (high-low-high). Some types buy in the first stage, while others continue the interaction and buy at the last stage. The principal offers unchosen decoy contracts. Sophisticated consumers correctly anticipate that if they deviated, they would choose a decoy, which they want to avoid in a lower frame. This eliminates incentive compatibility constraints into types who don't buy in the first stage. With naive consumers, the principal can perfectly screen by cognitive type and extract full surplus from naifs.

**JEL Codes:** D42, D82, D90, L12

**Keywords:** Screening, Framing, Extensive-Form Decision Problems, Dynamic Inconsistency, Sophistication, Naivete

---

\*We are grateful for valuable discussions and suggestions to Roland Bénabou, Elliot Lipnowski, Stephen Morris, Pietro Ortoleva, Wolfgang Pesendorfer, Doron Ravid, and Leeat Yariv; and for helpful comments to seminar participants at Princeton University, the Econometric Society Summer School 2018 in Singapore, Young Economists Symposium 2018, ESWM 2018, the NOeG Winter Workshop 2018, and participants of the RSF Summer Institute 2018. Ostrizek gratefully acknowledges financial support from the Philip Bennett \*79 Public Policy and Finance Fund.

<sup>†</sup>Department of Economics, Princeton University, [ostrizek@princeton.edu](mailto:ostrizek@princeton.edu)

<sup>‡</sup>Department of Economics, Princeton University, [shishkin@princeton.edu](mailto:shishkin@princeton.edu)

# 1 Introduction

Ample evidence, casual empiricism and introspection suggest that framing effects are common in choice.<sup>1</sup> In particular, the way a product is presented and the setting of the sales interaction can have a strong impact on consumer valuations.<sup>2</sup> Concordantly, many firms go to great expenses to improve the presentation of their product in largely non-informative and payoff-irrelevant ways through packaging, in-store presentation, and the emotions invoked by the sales pitch.

Most of the literature focuses on framing in static decision situations. However, many economic interactions including sales unfold in several stages. For instance, when buying a car, a consumer is first exposed to a manufacturer's marketing material, contemplates his purchase decision at home, and is then affected by the way the product is presented by the dealer. Even the sales pitch itself unfolds sequentially. As a result, firms have the opportunity to frame the options offered to the consumers differently at different stages of the decision and to use such changes of framing strategically. What is the optimal structure of a sales interaction? In particular, is it always best to present a product in the most favorable light? In general, how can a principal leverage the power to affect agents' preferences throughout a sequential interaction?

We investigate these questions by adding framing and extensive forms to a classic screening problem. The interaction of framing and sequential mechanisms allows the principal to exploit dynamic inconsistency to reduce information rents, whether consumers are sophisticated or not. While a growing literature analyzes this possibility assuming that the pattern of taste changes is given by the consumer's preferences (e.g. temptation or  $\beta$ - $\delta$ ), in our model this pattern is *endogenous*. The monopolist induces changing tastes by varying framing throughout the interaction. She designs not only the contracts, but also the structure of the sequential decision problem along with a frame at each stage.

---

<sup>1</sup>For example, decision makers overvalue the impact of certain product attributes if they vary strongly in the choice set (see [Bordalo et al., 2013](#), and references therein) and tend to be risk averse in decisions framed as gains and risk seeking for losses ([Tversky and Kahneman, 1981](#)).

<sup>2</sup>Consumer decisions are affected by the framing of insurance coverage ([Johnson et al., 1993](#)), the description of a surcharge ([Hardisty et al., 2010](#)), whether discounts are presented in relative or absolute terms ([DelVecchio et al., 2007](#)), prices as totals or on a per-diem basis ([Gourville, 1998](#)), and by background music ([Areni and Kim, 1993](#); [North et al., 1997, 2003, 2016](#)). Large effects of framing on consumer valuation are also found in incentivized lab experiments and across policy discontinuities ([Bushong et al., 2010](#); [Schmitz and Ziebarth, 2017](#)).

Our main result characterizes the structure of an optimal extensive-form decision problem (EDP) in a setting with quasi-linear single crossing utility and any finite number of types and frames under regularity conditions. The optimum is achieved in *three stages* using *two frames* (Theorem 1). If the consumers are sophisticated, the optimal EDP has the following key features:

1. *Short Interaction.* All types make at most three choices, and some types make only one choice.
2. *Natural Structure: approach–“cool-off”–close.* First, the agent is presented with a range of choices under a “hard sell” condition (highest valuation frame) and either buys now or expresses interest in one of the contracts, but is given time to consider. Then he is allowed to “cool-off” (second highest valuation frame) and decides whether to continue or take the outside option. Finally, again in the “hard sell” frame, he is presented with the contract he expressed interest in and a range of decoy contracts designed to throw off agents that misrepresented their type initially.
3. *Gains from Framing vs. Rent Extraction.* The principal can either “reveal” or “conceal” each type. Therefore, she faces a trade-off between maximizing the valuation by using only the highest frame (revealed types) and reducing information rents by using frames in a high-low-high pattern to induce dynamic inconsistency (concealed types).

We illustrate these features and the main construction for sophisticated consumers in the following example.

**Example 1.** There are two equally likely types  $\theta^1$  (low) and  $\theta^2$  (high) and two frames, low and high. Preferences of type  $\theta^i$  in frame  $f \in \{l, h\}$  are represented by  $u_f^i(p, q) = \theta_f^i q - p$ , where the marginal utility  $\theta_f^i$  depends both on the type and the current frame (see Fig. 1a). The monopolist principal (she) produces a good of quality  $q$  at cost  $\frac{1}{2}q^2$  and maximizes profits.

If the principal offers a menu, this is a standard screening problem with an additional choice of a frame. It is easy to see that it is optimal to pick the high frame  $h$  and offer contracts so that  $\theta_h^1$ 's participation constraint and  $\theta_h^2$ 's incentive compatibility constraint bind, which yields a profit of 20.<sup>3</sup>

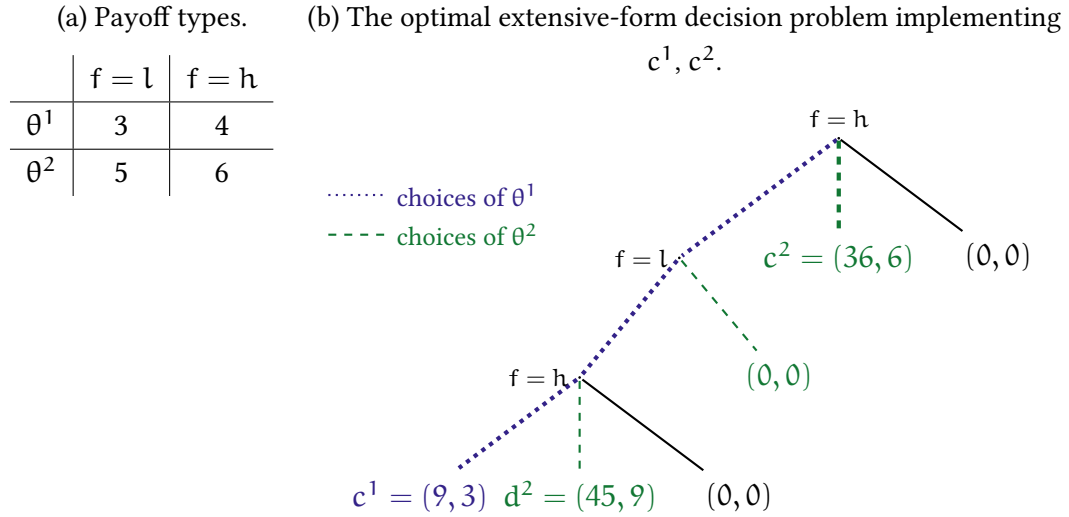
The principal can do better. Consider the allocation that would arise if the

---

<sup>3</sup>In particular, the optimal contracts are  $(p^1, q^1) = (8, 2)$  and  $(p^2, q^2) = (32, 6)$ . Note that with these functional forms,  $q^i = \theta_f^i$  is efficient for frame  $f$  and the quality of type 2 is distorted downward compared to the efficient quality for *both* frames.

principal could make types observable at the cost of always putting the low type in the low frame. Then, she could implement the efficient full-extraction contract for  $\theta_l^1, c^1 = (9, 3)$ , and  $\theta_h^2, c^2 = (36, 6)$ , obtaining a profit of  $\Pi_{\text{extensive}} = 22.5 > 20$ . We show that this is indeed possible in an EDP by varying the frames:  $h \rightarrow l \rightarrow h$ .

**Figure 1:** Example 1



To see how the principal achieves this, consider Fig. 1b. It is easy to check that the low type prefers  $c^1$  to any other contract in the EDP in both frames and therefore proceeds through the tree to  $c^1$ . What about the high type? Because  $c^1$  is preferable to  $c^2$  for him in both frames, we need to show that such a deviation is *infeasible* in this extensive form. To deviate to  $c^1$ , at the root the high type needs to choose the continuation problem leading to this contract. As he is sophisticated, he correctly anticipates his future choices but cannot commit. That is, at the second stage he anticipates that at the final stage he would pick the decoy  $d^2$  (in the high frame). But according to his taste at the second stage (in the low frame), the decoy is very unappealing, so he would choose the outside option. Hence, at the root the choice of the continuation problem is effectively equivalent to the outside option, thus, making the deviation to  $c^1$  impossible.

By placing a decoy contract as a “tempting poison pill” in the extensive form, the principal effectively removes the incentive compatibility constraint. Hence, the high type doesn’t obtain any information rent as the low type is concealed. This comes at the cost of adding an additional participation constraint, namely for the low type in the low frame, who has to pass through the low frame on the path to

his contract. Consequently, the maximal surplus that can be extracted from the low type is lower than in the static menu. There is a trade-off between concealing the contract intended for the low type in the continuation problem and thereby eliminating information rents and extracting surplus from this type. ▲

In general, the profit maximization problem is an optimization over the set of all extensive-form decision problems. However, based on the structure of the optimal EDP established in Theorem 1, we identify an equivalent simple optimization problem in contract space (Theorem 2). The principal partitions the set of types into revealed and concealed. This partition determines the participation and incentive constraints: Concealing a type eliminates incoming IC constraints at the cost of a tighter participation constraint. In contrast to the classic setting, it is never optimal to exclude any type, as it is strictly better to sell a strictly positive quality to every type and conceal some of them instead (Proposition 3).<sup>4</sup>

For the main sections, we assume that consumers are sophisticated. They correctly anticipate their choices, but cannot commit to a course of action.<sup>5</sup> As the optimal sales interaction has a simple 3-stage structure, correctly anticipating behavior in this extensive form is relatively easy. Moreover, consumers are exposed to sales pitches on a daily basis, they are experienced and understand the flow of the interaction. Sophistication reflects the idea that consumers understand that they are more prone to choose premium option when under pressure from the salesperson (high frame), and (in a low frame) avoid putting themselves in such situations that lead to excessive purchases. In addition, sophistication serves as a benchmark, by making it difficult for the principal to extract surplus. Even if consumers are fully strategically sophisticated and can opt out of the sales interaction at any point, framing in extensive forms affects the sales interaction and its outcomes. Indeed, the principal turns consumers' sophistication against them.

We also consider naive consumers. They fail to anticipate that their tastes may change and choose a continuation problem as if their choice from this problem would be made according to their current tastes. For naive consumers, the principal

---

<sup>4</sup>This is in line with Corollary 2 in [Salant and Siegel \(2018\)](#), which states that there is no exclusion with two types, when the principal offers a framed menu under a participation constraint in a neutral frame. A related result is in [Eliaz and Spiegler \(2006\)](#). They show that there is no exclusion when the principal screens by the degree of sophistication. We show that no-exclusion holds when the principal screens by payoff type.

<sup>5</sup>Sophistication is a common modeling choice in the domain of time preference following the seminal work of [Strotz \(1955\)](#); [Laibson \(1997\)](#).

can implement the efficient quantities in the highest frame and extract all surplus with a three-stage decision problem. She does so using decoy contracts in a bait-and-switch: Naive consumers expect to choose a decoy option tailored to them and reveal their type by choosing the continuation problem containing it at the root (bait), but end up signing a different contract due to the preference reversals induced by a change of frame (switch). When both naive and sophisticated consumers are present in arbitrary proportions and this cognitive type is not observable to the firm, our results generalize (Theorem 3)<sup>6</sup>. The optimal extensive-form still has three stages and implements the same contracts as if the cognitive type were observable. There are no cross-subsidies from naive to sophisticated consumers.

Many jurisdictions mandate a right to return goods and cancel contracts, especially when the sale happened under pressure (e.g. door to door). This gives consumers the option to reconsider their purchase in a calm state of mind, unaffected by the immediate presence of the salesperson. We analyze such regulation and find that, while the principal can no longer use framing to exaggerate surplus, she can still use the resulting dynamic inconsistency to fully eliminate the information rent of all types. Sophisticated consumers do not require protection by a right to return if they can decide to avoid the seller, e.g. by not visiting the store, but naive consumers would benefit even in this case.

Beyond the setting of framing in screening problems, we view our results as steps towards understanding the impact of behavioral choice patterns (both framing and choice set dependence) when a principal (or mechanism designer) can offer extensive-form decision problems in order to exploit the resulting violations of dynamic consistency and demand for commitment. We return to this discussion in the conclusion.

We set up the model in Section 2. In Section 3, we show that the optimal extensive-form decision problem is of a simple three-stage structure. We find a relaxed problem in price-quality space that characterizes the optimal vector of contracts. In Section 4, we construct the optimal extensive-form decision problem if some consumers are naive about the effect of framing. We also consider the case when the principal's choice of extensive form is restricted to account for a participation decision (e.g. a right to return the product) in an exogenous "neutral" frame. We conclude with discussions. Proofs are collected in the Appendix.

---

<sup>6</sup>Spiegler (2011) notes that the principal can costlessly screen by cognitive type in a setting without taste heterogeneity.

## Related Literature

A growing literature studies the manipulation of framing by firms. [Piccione and Spiegler \(2012\)](#) and [Spiegler \(2014\)](#) focus on the impact of framing on the comparability of different products. [Salant and Siegel \(2018\)](#) study screening when framing affects the taste for quality, as in our setting. In this paper, the principal chooses a framed menu, while we study the optimal design of an extensive-form decision problem to exploit the dynamic inconsistency generated by choice with frames and make predictions about the structure of interactions. In addition, our model makes different predictions for the use of framing and efficiency in the setting where the two are closely comparable:<sup>7</sup> Using extensive forms, it is *always* more profitable to use framing (not only when it is sufficiently weak) and framing removes *all* distortions created by second-degree price discrimination (not only some) in our setting.

Our article is also related to behavioral contract theory more generally (for a recent survey, see [Kőszegi 2014](#), for a textbook treatment, see [Spiegler 2011](#)), in particular to screening problems with dynamically inconsistent agents ([Eliaz and Spiegler, 2006, 2008](#); [Esteban et al., 2007](#); [Esteban and Miyagawa, 2006a,b](#); [Zhang, 2012](#); [Galperti, 2015](#); [Heidhues and Kőszegi, 2010, 2017](#); [Yu, 2018](#); [Moser and Olea de Souza e Silva, 2019](#)).<sup>8</sup> These papers consider situations when the taste changes are *given* by the preferences of the agents (e.g. [Gul and Pesendorfer \(2001\)](#) or  $\beta$ - $\delta$ ) and consequently design a 2-stage decision problem as induced by the natural time structure of the problem. We study how a principal *chooses* the sequence of frames and an extensive form of arbitrary (finite) length to induce dynamic inconsistency and we show that a 3-stage mechanism is optimal.

---

<sup>7</sup>That is, comparing their Section 3 with our Section 4.2, where we impose a right to return the product in an exogenously given "neutral" frame. They also consider a model without returns but with a "basic" product that has to be offered and an insurance problem in which the monopolist can highlight one of the options, turning it into a reference point relative to which consumers experience regret.

<sup>8</sup>[Eliaz and Spiegler \(2006, 2008\)](#) screen dynamically inconsistent agents by their degree of sophistication and optimistic agents by their degree of optimism, respectively. [Esteban et al. \(2007\)](#); [Esteban and Miyagawa \(2006a,b\)](#) study screening when agents are tempted to over- or underconsume. [Zhang \(2012\)](#) studies screening by sophistication when consumption is habit inducing. [Galperti \(2015\)](#) studies screening in the provision of commitment contracts to agents with private information on their degree of time inconsistency, [Heidhues and Kőszegi \(2017\)](#) study selling credit contracts in this setting. [Yu \(2018\)](#) and [Moser and Olea de Souza e Silva \(2019\)](#) study optimal taxation problem, where agents are also heterogeneous in the degree of present-bias.



Given the optimal sequence of frames, this mechanism employs techniques similar to those in the literature. In particular, it involves off-path options that remain unchosen by every type ("decoys"). In [Esteban and Miyagawa \(2006a\)](#) and [Galperti \(2015\)](#) such decoys make deviations less attractive and are thus analogous to the decoy contracts we introduce in the optimal extensive form for sophisticated agents. [Heidhues and Köszegi \(2010\)](#) show that credit contracts for partially sophisticated quasi-hyperbolic discounters feature costly delay of the payment which the consumer fails to expect when signing the credit contract. Immediate repayment is hence an unused option analogous to the "bait" decoys we introduce to screen naive consumers.

[Glazer and Rubinstein \(2012, 2014\)](#) consider models where the principal designs a procedure such that misrepresenting their type is beyond the boundedly rational agents' capabilities. While their decision problems are based on hypothetical questions about the agent's type, we show that it is possible to structure a choice problem with framing to make it impossible to imitate certain types.

There is a large literature on endogenous context effects, e.g. through focusing the attention of the decision maker on attributes that vary strongly or are exceptional within the choice set ([Bordalo et al., 2013](#); [Köszegi and Szeidl, 2013](#)). We consider the case of framing through features of the choice situation, such as the sales pitch or the presentation format. Thus, consumers in our model fit into the choice with frames framework of [Salant and Rubinstein \(2008\)](#).

The presence of different frames and extensive forms places our screening setting close to implementation. If we reinterpret our decision maker as a group of individuals with common knowledge of their type but different tastes, one individual corresponding to each frame, the principal applies implementation in backward induction ([Herrero and Srivastava, 1992](#)). While they give abstract conditions for implementability in a very general setting, we characterize the structure of the optimal decision problem for our screening model and derive properties of the optimal contracts.

## 2 Screening with Frames and Extensive Forms

We build on the classic model of price discrimination ([Mussa and Rosen, 1978](#); [Maskin and Riley, 1984](#)), extending the framework in two ways. Instead of simple menus, firms design an extensive-form decision problem. Furthermore, for every



decision node the firm picks a frame affecting the valuation of consumers. Our results are driven by the interaction of both ingredients.

## 2.1 Contracts and Frames

The firm produces a good with a one-dimensional characteristic  $q \geq 0$ , interpreted as quantity or quality. Throughout the exposition, we maintain the latter interpretation. A contract  $c$  is a pair of a price  $p$  and a quality  $q$ , the space of contracts is  $\mathcal{C} = \mathbb{R} \times \mathbb{R}_+$ .

There is a finite set of frames  $F$  with  $|F| \geq 2$  and a finite type space  $\Theta$  endowed with a full support prior  $\mu$ . Each type is a function  $\theta : F \rightarrow \mathbb{R}$  that maps frames into payoff types, denoted as  $\theta_f \equiv \theta(f)$ . For a given payoff type  $\theta_f$  the consumer is maximizing the utility function

$$u_{\theta_f}(p, q) = v_{\theta_f}(q) - p.$$

where  $v : \mathbb{R} \times \mathbb{R}_+$  is a thrice differentiable function, that satisfies

$$\frac{\partial v}{\partial q} > 0, \quad \frac{\partial v}{\partial \theta} \geq 0, \quad \frac{\partial^2 v}{\partial \theta q} > 0, \quad \frac{\partial^2 v}{\partial q^2} < 0, \quad \frac{\partial^3 v}{\partial q^2 \partial \theta} > 0.$$

For convenience, we normalize  $\forall \theta_f, v_{\theta_f}(0) = 0$ . Note that we assumed that utility is quasi-linear in money and frames affect a consumer's taste for quality. This is consistent with framing effects on price perception, as long as these effects are multiplicatively separable.

For a given vector of contracts  $\mathbf{c} = (c_\theta)_{\theta \in \Theta}$ , we refer to the constraints

$$u_{\theta_f}(\mathbf{c}_\theta) \geq 0, \text{ and} \tag{P_\theta^f}$$

$$u_{\theta_f}(\mathbf{c}_\theta) \geq u_{\theta_f}(\mathbf{c}_{\theta'}) \tag{IC_{\theta\theta'}^f}$$

as the participation constraint for  $\theta$  and the incentive compatibility constraint (IC) from  $\theta$  to  $\theta'$  in frame  $f$ .

First, we require a non-triviality condition.

**Assumption 1** (Relevant Frames). *For any  $f, f' \in F$  there exists a type  $\theta \in \Theta$  such that  $\theta_f \neq \theta_{f'}$ , i.e. the two frames induce a different valuation for this type.*

Second, we also assume additional structure on payoff types across frames in order to ensure that the problem remains one-dimensional despite the addition of frames.

**Assumption 2** (Comonotonic Environment). *For any  $\theta, \theta' \in \Theta, f, f' \in F$ :*

$$\theta_f > \theta_{f'} \implies \theta'_f > \theta'_{f'} \quad \text{and} \quad \theta_f > \theta'_f \implies \theta_{f'} > \theta'_{f'}.$$

The first part of the assumption implies that frames can be ordered by their impact on the valuation. There is a lowest frame, i.e. a frame inducing the lowest valuation for every type and a highest frame, i.e. a frame inducing the highest valuation for every type. The second part implies that types can also be ordered by their valuation independently of the frame. With slight abuse of notation, we denote the order on frames and types using regular inequality signs.

In many cases, a frame has a similar impact on different consumer types. The more effectively a seller emphasizes quality, for instance, the higher a consumer values quality irrespective of their type. The first part of our assumption is satisfied as long as the *direction* of the impact of a given frame is the same for all types. The second part is satisfied as long as the *size* of the effect is not too different between types relative to their initial difference in valuation. In particular, suppose there is a neutral frame  $f_n$ . The assumption is satisfied when the absolute impact of an enthusiastic frame  $f_e > f_n$  is greater for high valuation consumers and the absolute impact of a pessimistic frame  $f_p < f_n$  is greater for low valuation consumers: The firm can amplify the initial feelings of consumers and make all types more or less interested in quality, but it cannot manipulate them to the degree that the order is reversed.

Assumption 2 precludes any frame from impacting the valuations of different types in a different direction. For example, focusing a car buyers attention on emissions may increase the valuation of a “green” car for some buyers while reducing the valuation of all cars, including the “green” car, for others. Similarly, it rules out cases where the order of types by their payoff parameter depends on the frame. For example, the demand for health insurance coverage may be lower among smokers than nonsmokers if they are not reminded about the long run effects of their habit, but is higher for smokers than nonsmokers if the effects of smoking are made salient during the sale of insurance. Together with Assumption 1, it also rules out that certain frames are specific to certain types. We discuss how we can relax our assumptions in Section 3.5.

## 2.2 Extensive-Form Decision Problems

We model the sales interaction as an extensive-form decision problem (single-player game), with a frame attached to each decision node. For example, the following situation can be represented by a two-stage extensive-form decision problem. First, the consumer contemplates whether to visit the store and then purchases a product in the store. Perhaps, the consumer is initially affected by marketing materials (frame at the root) and then the consumer is affected by the sales pitch in the store (frame at the second stage).

We define extensive-form decision problems (EDPs) by induction. Call an extensive decision problem with  $k$  stages a  $k$ -EDP. For any set  $S$ , let  $\mathcal{P}(S)$  denote the set of all finite subsets of  $S$  containing the outside option  $\mathbf{0} := (0, 0)$ . The set of 1-EDPs is  $\mathcal{E}^1 := \mathcal{P}(\mathcal{C}) \times \mathcal{F}$ , that is, a 1-EDP  $e = (A, f)$  is a pair of a finite menu  $A$  and a frame  $f$ . For each  $k > 1$ , the set of  $k$ -EDPs is  $\mathcal{E}^k := \mathcal{P}(\cup_{l=0}^{k-1} \mathcal{E}^l) \times \mathcal{F}$ , so that  $e = (E, f) \in \mathcal{E}^k$  is a pair of a finite set of EDPs  $E$  and a frame  $f$ .<sup>9</sup> Finally, the set  $\mathcal{E}$  of all finite EDPs is given by

$$\mathcal{E} := \bigcup_{k=1}^{\infty} \mathcal{E}^k.$$

In other words, an EDP is a finite tree with a frame assigned to each decision node. Moreover, the outside option is available for consumers at each stage.

**Choice from Extensive-Form Decision Problems** The preferences of consumers are represented by a utility function defined on the set of contracts. Therefore, the choice of consumers with type  $\theta$  is well defined on the set  $\mathcal{E}^1$  of 1-EDPs which are simply menus with frames. To define consumer choice for any EDP we assume that the consumers are *sophisticated*. Presented with a choice between several decision problems, the consumer correctly anticipates future choices, and chooses the continuation problem according to her current frame. The current self has no commitment power other than the choice of a suitable continuation problem.

Formally, we define the sophisticated consumer's choice in an EDP by induction. Call  $\sigma : \Theta \rightarrow \mathcal{C}$  an *outcome* of a 1-EDP  $(A, f) \in \mathcal{E}^1$  if  $\sigma(\theta)$  maximizes  $u_{\theta, f}$  on  $A$ . Suppose the consumer is facing a  $k$ -EDP  $(E, f) \in \mathcal{E}^k$ . Choosing between continuation

---

<sup>9</sup>Set  $\mathcal{E}^0 = \mathcal{C}$  to allow for terminal choices at all stages. Also, note that the choice of frame at each stage is unrestricted. In particular, the frame at the root of an EDP does not place any constraints on subsequent frames. We discuss the role of this assumption in Section 3.5.

problems in  $E$ , she anticipates her choice  $\sigma^e(\theta) \in \mathcal{C}$  for each  $e \in E$ , but evaluates the contracts  $\{\sigma^e(\theta)\}_{e \in E}$  in the current frame  $f$ . Let  $\Sigma^e$  be the set of outcomes of an EDP  $e \in \cup_{l=0}^{k-1} \mathcal{E}^l$  with  $\Sigma^e = \{\theta \mapsto e\}, \forall e \in \mathcal{E}^0$ . Then  $\sigma$  is an outcome of  $(E, f)$  if there is a solution  $\sigma^e \in \Sigma^e$  for every  $e \in E$ , such that  $\forall \theta \in \Theta$

$$\sigma(\theta) \in \operatorname{argmax}_{\{\sigma^e(\theta): e \in E\}} u_{\theta_f}(\sigma^e(\theta)).$$

### 2.3 The Firm's Problem

The monopolist produces goods of quality  $q$  at convex cost  $\kappa(q)$ , that is twice-differentiable and satisfies boundary conditions to ensure interior efficient quantities:  $\kappa(0) = 0, \kappa' > 0, \kappa'' > 0$  and  $\forall \theta_f \in \mathbb{R}, v'_{\theta_f}(0) - \kappa'(0) > 0, \lim_{q \rightarrow \infty} v'_{\theta_f}(q) - \kappa'(q) < 0$ .

Given a vector of contracts  $\mathbf{c} = (c_\theta)_{\theta \in \Theta} = (p_\theta, q_\theta)_{\theta \in \Theta}$ , the profit of the firm is given by

$$\Pi(\mathbf{c}) := \sum_{\theta \in \Theta} \mu_\theta (p_\theta - \kappa(q_\theta)).$$

Finally, the firm designs an EDP to maximize profits

$$\Pi^* := \max_{e \in \mathcal{E}, \mathbf{c} \in \Sigma^e} \Pi(\mathbf{c}). \quad (\text{GP})$$

In analogy with the mechanism design literature, we say a vector of contracts  $\mathbf{c}$  is *implemented* by an EDP  $e$  if it has a solution  $\sigma = \mathbf{c}$ . We then call  $\mathbf{c}$  *implementable*. In these terms, the principal maximizes profits over the set of implementable contracts.<sup>10</sup>

We denote the efficient quality for a payoff parameter  $\theta_f$  by  $\hat{q}_{\theta_f}$  with

$$v'_{\theta_f}(\hat{q}_{\theta_f}) = \kappa'(\hat{q}_{\theta_f}).$$

The efficient quality is unique, positive and strictly increasing in the payoff parameter by our assumptions on  $v$  and  $\kappa$ . We denote the contract offering this quality and extracting all surplus from the corresponding payoff type by  $\hat{\mathbf{c}}_{\theta_f} := (v_{\theta_f}(\hat{q}_{\theta_f}), \hat{q}_{\theta_f})$ .

---

<sup>10</sup>Note that (GP) does not require  $\mathbf{c}$  to be the unique outcome of  $e$ . We only require partial implementation, as is customary in contract theory to ensure the compactness of the principal's problem. It can be shown that for any  $\epsilon > 0$  the firm can design an EDP of the same structure with a unique outcome that achieves  $\Pi^* - \epsilon$ .

### 3 Optimal Screening

Before we analyze the general problem (GP), we analyze two special cases. In both cases, the problem collapses to a simple static screening problem.

Consider a simpler problem, where the firm can only choose a 1-EDP, i.e. a menu and a frame. In this case, it is optimal to choose the highest frame  $h := \max F$ , maximizing consumer valuation. Alternatively, suppose there is only one frame:  $F = \{h\}$ . Consequently, any EDP must use the same frame at every stage. The extensive-form structure does not matter in this case: As consumers are perfectly rational and dynamically consistent, they pick the most preferred contract from the extensive form. Hence, an extensive form is equivalent to an unstructured menu offering the same set of contracts.

In both cases, the optimal menu corresponds to the solution of the classic monopolistic screening problem with the set of types  $\{\theta_h\}_{\theta \in \Theta}$ .<sup>11</sup>

**Observation 1.** *Let  $c^*$  be the vector of contracts obtained by maximizing profits subject to the participation constraint for the lowest type and all IC constraints, all in frame  $h$ . Then the 1-EDP  $(\mathbf{0} \cup \{c_\theta^*\}_{\theta \in \Theta}, h)$  solves (GP) if*

1. *the firm is constrained to 1-EDPs, or*
2. *there is only one frame:  $F = \{h\}$ .*

This shows that framing or extensive forms alone are not sufficient for our results. Only both features together allow the principal to use different frames at different stages of the decision and thereby generate violations of dynamic consistency that can be exploited.

#### 3.1 Optimal Structure of the Extensive Form

In this section, we show that the optimal EDP has a simple three-stage structure. Towards this result, let us define a class of EDPs which share these structural features. Let  $h$  and  $l$  denote the highest and second highest frame:

$$\begin{aligned} h &:= \max F, \\ l &:= \max F \setminus \{h\}. \end{aligned}$$

---

<sup>11</sup>This is in contrast to (Salant and Siegel, 2018), where there is an ex-post participation constraint in an exogenously given frame, or a default option, i.e. a restricted menu choice problem for the principal.

For any standard EDP the set of types  $\Theta$  is partitioned into two sets, as there are two ways to present the contract associated to a given type: Contracts  $c_\theta$  for revealed types ( $\theta \in \Theta_R$ ) are presented at the root, while contracts for concealed types ( $\theta \in \Theta_C$ ) are presented in separate continuation problems  $e_\theta$ . Then, the three stages are (see Fig. 2):

1. Root:  $f = h$ ; available choices: contracts  $c_\theta$  for  $\theta \in \Theta_R$  and EDPs  $e_\theta$  for  $\theta \in \Theta_C$ .
2. Continuation problem:  $f = l$ ; available choices: outside option and continue.
3. Terminal choice:  $f = h$ ; available choices:  $c_\theta$  and decoys  $d_{\theta'}^\theta$ , for  $\theta' > \theta$ .

Formally, we have the following definition.

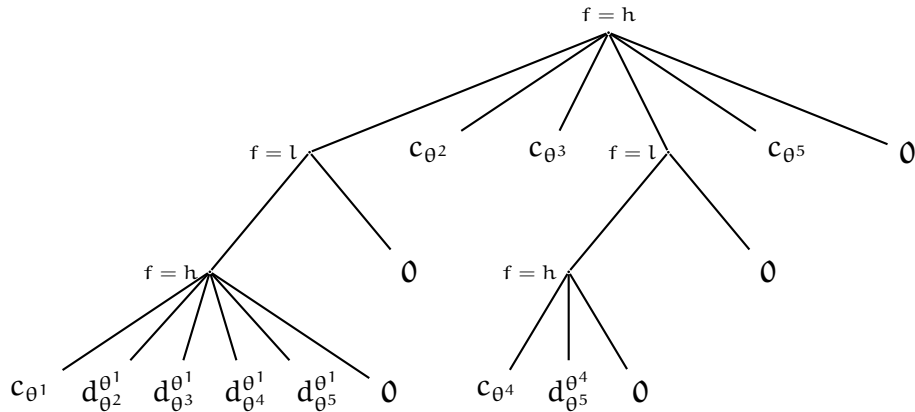
**Definition 1.** An EDP  $e$  is a *standard* EDP for a vector of contracts  $\mathbf{c}$  if there exists a partition  $\{\Theta_C, \Theta_R\}$  of  $\Theta$ , and decoy contracts  $\{d_{\theta'}^\theta\}_{\theta \in \Theta_C, \theta' > \theta}$ , such that

$$e = \left( \{e_\theta\}_{\theta \in \Theta_C} \cup \{c_\theta\}_{\theta \in \Theta_R} \cup \{0\}, h \right), \text{ where} \quad (1)$$

$$e_\theta = \left( \left\{ (\{c_\theta, 0\} \cup \{d_{\theta'}^\theta\}_{\theta' > \theta}, h), 0 \right\}, l \right), \forall \theta \in \Theta_C. \quad (2)$$

The extensive form in Example 1 is a standard EDP. Type  $\theta^1$  is concealed – his contract is available only after a continuation problem – while type  $\theta^2$  is revealed – his contract is available immediately at the root.

**Figure 2:** A standard EDP for  $(c_{\theta^1}, \dots, c_{\theta^5})$  with  $\Theta_R = \{\theta^2, \theta^3, \theta^5\}$  and  $\Theta_C = \{\theta^1, \theta^4\}$ .



Note that the notion of standard EDP is solely about the structure of the EDP. It puts no restrictions on the decoy contracts and is silent about choice. In particular,

a standard EDP for  $\mathbf{c}$  may not implement  $\mathbf{c}$ .<sup>12</sup>

Standard EDPs are sufficient to achieve the optimum.

**Theorem 1.** *If  $\mathbf{c}$  is an optimal vector of contracts in (GP), then it is implemented by a standard EDP.*

Several observations follow from this result. First, the optimum can be achieved in *three stages for an arbitrary number of agent types*, even though the principal has arbitrarily complicated and long extensive forms at her disposal. As the number of types increases, the structure and length of the decision problem stays the same, only the number of available contracts increases. Furthermore, the optimal EDP has a simple structure that we interpret as follows: At the beginning, the consumer is presented a range of contracts  $\{c_\theta\}_{\theta \in \Theta}$  while the salesperson focuses their attention on quality (high frame). Some of those contracts (those intended for revealed types) can be signed immediately, some others (those intended for concealed types) are only available after an additional procedure that gives the consumer some time to consider, while sales pressure is reduced (lower frame). This can be an explicit wait period, where the consumer is asked to think about the contract and recontact the seller. Alternatively, the change in frame could be achieved by a change in the salesperson or by acquiring a confirmation that this type of offer is even available for the consumer. If the consumer is still interested after this ordeal, she is presented with additional offers, the decoy contracts. On path, these offers remain unchosen, the consumer chooses the contract she initially intended to obtain.

Second, types are *separated at the root*. The principal does not use the extensive-form structure to discover the type of a consumer piecemeal, it is an implementation device to screen contracts against imitation.

Third, only the *two highest frames* are used. As we have seen in Observation 1, the principal desires to put everyone in the highest frame if there is no extensive-form structure. On the other hand, if every decision node uses the same frame, the extensive-form structure is irrelevant for agents choice. Consequently, the principal uses at least two frames in order to induce violations of dynamic consistency. As long as the principal induces such violations, the decoys can be constructed irrespective of the number of or cardinal differences between the frames used. Hence, two frames are sufficient for the principal to reap *all potential gains* from such violations. Finally, only the highest two are used in the optimal EDP in order

---

<sup>12</sup>Whenever we state that a vector of contracts  $\mathbf{c}$  is implemented by a standard EDP, however, it is understood that it is implemented by a standard EDP for  $\mathbf{c}$ .



to maximize valuations.

### 3.2 Necessary and Sufficient Conditions for Implementation

In order to provide foundations for Theorem 1, we proceed in two steps. First, we identify an upper bound on profits in any EDP by providing necessary conditions every implementable vector of contracts has to satisfy. Then, returning to standard EDPs, we provide sufficient conditions on a vector of contracts ensuring that it can be implemented in this class. In particular, we explicitly construct decoy contracts and show that the principal can thereby eliminate downward IC constraints into concealed types.

#### Necessary Conditions for Implementation by General EDPs

Consider an arbitrary EDP implementing a vector of contracts  $\mathbf{c} = (c_\theta)_{\theta \in \Theta}$ . Denote the frame at the root by  $f_R$ . Extending the notion of revealed and concealed types from standard EDPs, for each type  $\theta$  there are two possibilities: If there exists a path from the root to  $c_\theta$  with all decision nodes set in  $f_R$ , then  $\theta$  is called revealed. Alternatively, if every path from the root to  $c_\theta$  involves at least one  $f_\theta \neq f_R$ , then  $\theta$  is called concealed. As usual, we will denote the sets of revealed and concealed types by  $\Theta_R$  and  $\Theta_C$ , respectively.

First, consider participation constraints. If the path from the root to  $c_\theta$  passes through a node in frame  $f$ , then, since the outside option is always available,  $c_\theta$  needs to satisfy the corresponding participation constraint  $P_\theta^f$ . In particular, every contract has to satisfy the constraint at the root  $P_\theta^{f_R}$ .

We now turn to incentive compatibility constraints. If  $\theta$  is revealed,  $c_\theta$  can be reached by any type from the root, as consumers are dynamically consistent when the frame does not change along the path. Consequently, for any  $\theta'$ ,  $c_\theta$  must not be an attractive deviation:

$$u_{\theta'_{f_R}}(c_{\theta'}) \geq u_{\theta'_{f_R}}(c_\theta) \quad \forall \theta' \in \Theta. \quad (\text{IC}_{\theta\theta'}^{f_R})$$

If  $\theta$  is concealed, there is a change of frame along the path to  $c_\theta$ . This induces a violation of dynamic consistency, which may make deviations into  $c_\theta$  impossible. As we are looking for necessary conditions, we impose no incoming IC constraint in this case.

So far, we identified a family of conditions indexed by  $(f_R, \{f_\theta\}_{\theta \in \Theta}, \Theta_C)$ , such that a vector of contracts is implementable only if it satisfies at least one of them.

The following proposition shows that without loss of generality, we can set  $f_R = h$  and  $f_\theta = l$ . This is because the exact frame only matters for participation, while the change of frame affects IC. Consequently, the contract must satisfy the least restrictive participation constraints, i.e. in the highest and second highest frame.

**Proposition 1.** *If  $c$  is implemented by an EDP, then it satisfies constraints  $\{P_\theta^h\}_{\theta \in \Theta_R}$ ,  $\{P_\theta^l\}_{\theta \in \Theta_C}$ , and  $\{IC_{\theta\theta'}^h\}_{\theta \in \Theta, \theta' \in \Theta_R}$  for some partition  $\{\Theta_C, \Theta_R\}$  of  $\Theta$ .*

The necessary conditions illustrate the trade-off between using framing to increase consumer valuation and its use to reduce information rents. For revealed types, the participation constraint needs to be satisfied only in the highest frame, the frame resulting in the least restrictive constraint. This results in the greatest surplus. For concealed types, the participation constraint needs to be satisfied in the second highest frame. This reduces the surplus from the interaction. The principal is compensated for this reduction through the removal of IC constraints into concealed types.

### Sufficient Conditions for Implementation by Standard EDPs

To construct a standard EDP that implements a vector of contracts  $c$ , we proceed in two steps. First, we need to determine the set of concealed types. Then, we construct decoys for the continuation problems of these types. Clearly, type  $\theta$  can be concealed in a standard EDP only if  $c_\theta$  satisfies the participation constraint  $P_\theta^l$ , since otherwise he would prefer to opt out in the second stage.

If  $\theta$  is concealed, the principal can design decoys in order to make *some* deviations into  $c_\theta$  impossible. Whereas decoys cannot rule out all upward deviations, they can rule out *all downward* deviations into  $c_\theta$ . Consequently, a vector of contracts is implementable by a standard EDP even if it does not satisfy the downward IC constraints, as long as the types that are attractive to imitate can be concealed.

**Proposition 2.** *If  $c$  satisfies the constraints  $\{P_\theta^h\}_{\theta \in \Theta_R}$ ,  $\{P_\theta^l\}_{\theta \in \Theta_C}$ ,  $\{IC_{\theta\theta'}^h\}_{\theta < \theta'}$ , and  $\{IC_{\theta\theta'}^h\}_{\theta \in \Theta, \theta' \in \Theta_R}$  for some partition  $\{\Theta_C, \Theta_R\}$  of  $\Theta$ , then  $c$  is implemented by a standard EDP.*

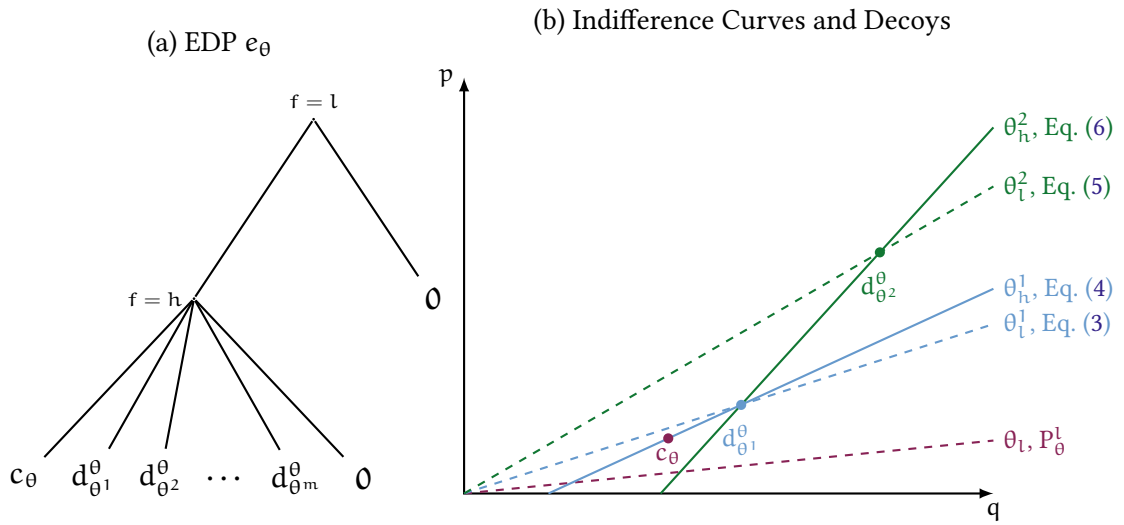
As in Example 1, the principal constructs decoys to render downward deviations into concealed types impossible in the extensive form. This construction is the central step in our results and we therefore present it in the text. The construction ensures that (ii) if  $\theta$  is concealed, no type  $\theta' > \theta$  can imitate  $\theta$ . The decoys don't interfere with the choices of any type at the root, as they will not be chosen from

the continuation problem (i). In particular,  $\theta$  chooses the intended contract (iii).

**Lemma 1** (Decoy Construction). *For any  $\theta \in \Theta$ ,  $c_\theta$  satisfies  $P_\theta^1$  if and only if there exist decoys  $(d_{\theta'}^\theta)_{\theta' > \theta}$ , such that the corresponding  $e_\theta$  in (2) has an outcome  $\sigma$  that satisfies*

- (i)  $\sigma(\theta') \in \{0, c_\theta\}$  for all  $\theta' \in \Theta$ ,
- (ii)  $\sigma(\theta') = 0$  for all  $\theta' > \theta$ , and
- (iii)  $\sigma(\theta) = c_\theta$ .

**Figure 3:** The construction of  $e_\theta$ .



*Construction.* The construction of the decoys and the continuation problem  $e_\theta$  is illustrated in Fig. 3. At the terminal stage, agents are presented with the choice between the contract  $c_\theta$ , the outside option and a set of decoys  $\{d_{\theta'}^\theta\}_{\theta' > \theta}$ , one for every type greater than  $\theta$ . Given a contract  $c_\theta$ , the decoy  $d_{\theta^1}^\theta$  for the next largest type  $\theta^1$  is implicitly defined by the system

$$u_{\theta^1}(\mathbf{0}) = u_{\theta^1}(d_{\theta^1}^\theta). \quad (3)$$

$$u_{\theta^1}(c_\theta) = u_{\theta^1}(d_{\theta^1}^\theta) \quad (4)$$

Then, decoy  $d_{\theta^2}^\theta$  for the next type  $\theta^2$  solves

$$u_{\theta^2}(\mathbf{0}) = u_{\theta^2}(d_{\theta^2}^\theta). \quad (5)$$

$$u_{\theta^2}(d_{\theta^1}^\theta) = u_{\theta^2}(d_{\theta^2}^\theta) \quad (6)$$

Proceeding by induction, we construct decoys for all  $\theta' > \theta$ . Now we define an outcome  $\sigma$  as follows. The single-crossing property ensures that each type  $\theta' \geq \theta$  chooses their corresponding (decoy) contract out of the menu  $\{c_\theta, d_{\theta_1}^\theta, \dots, d_{\theta_m}^\theta, 0\}$  in frame  $h$ . At the root,  $\theta$  will choose its contract since it satisfies  $P_\theta^l$  and any  $\theta' > \theta$  will choose the outside option. Finally, single crossing ensures that types  $\theta' < \theta$  prefer the outside option over the decoys as well. We formally verify the construction in Appendix A.2.  $\square$

### 3.3 Optimal Contracts

We show that the principal's problem (GP) over the space of extensive forms is equivalent to a two-step maximization problem based on the necessary conditions for implementation (Proposition 1). This relaxed problem characterizes the optimal vector of contracts.

#### An Equivalent Problem in Price-Quality Space

Let us summarize the necessary condition in an optimization problem. Recall that these conditions are indexed by the set of concealed types. This set is an additional choice variable for the principal in the relaxed problem.

$$\begin{aligned} \Pi^R &= \max_{\Theta_C \subseteq \Theta} \max_{(c_\theta)_{\theta \in \Theta}} \Pi(\mathbf{c}) && \text{(RP)} \\ \text{s.t. } & u_{\theta_h}(c_\theta) \geq 0, \quad \forall \theta \in \Theta_R := \Theta \setminus \Theta_C && \text{(P}_\theta^h\text{)} \\ & u_{\theta_l}(c_\theta) \geq 0, \quad \forall \theta \in \Theta_C && \text{(P}_\theta^l\text{)} \\ & u_{\theta_h}(c_\theta) \geq u_{\theta_h}(c_{\theta'}), \quad \forall \theta \in \Theta, \theta' \in \Theta_R && \text{(IC}_{\theta\theta'}^h\text{)} \end{aligned}$$

While it is not true that *every* vector of contracts satisfying the necessary conditions is implementable, the solution of (RP) is implementable, as it satisfies the sufficient conditions of Proposition 2.

**Theorem 2.** *A pair  $(\Theta_C, \mathbf{c})$  solves (RP) if and only if  $\mathbf{c}$  solves (GP). Moreover, such a solution exists and  $\mathbf{c}$  can be implemented by a standard EDP with a set of concealed types  $\Theta_C$ .*

In other words, the general problem (GP) attains the upper bound given by (RP)

$$\Pi^* = \Pi^R,$$

and the necessary conditions together with optimality is sufficient for implementation, even in the restricted class of standard EDPs.

Without the equivalent formulation, even verifying the existence of a solution to (GP) can be troublesome. Theorem 2 shows that instead of a complex optimization problem defined over extensive forms, the principal can solve well-behaved contracting problems over a menu of price-quality pairs, one for each potential set of concealed types and compare the attained values to find the optimum.<sup>13</sup> Once the principal found the (RP) optimal concealed types and vector of contracts, it is easy to construct a standard EDP implementing it using Lemma 1.

### No Shut-down

In the classic model of screening, it is sometimes optimal for the monopolist to exclude low types by selling the outside option to them. In our model, this is never the case, because concealing a type is always strictly better for the monopolist than excluding it.

**Proposition 3.** *The optimal contract  $(p_\theta, q_\theta)$  for a type  $\theta$  satisfies  $0 < \underline{q}_\theta \leq q_\theta \leq \hat{q}_{\theta_h}$ , where  $v_{\theta_h}(\underline{q}_\theta) - \kappa(\underline{q}_\theta) := v_{\theta_l}(\hat{q}_{\theta_l}) - \kappa(\hat{q}_{\theta_l})$ . In particular, every type of consumer buys positive quality.*

Indeed, concealing a type can be interpreted as a soft form of shut-down. In order to eliminate information rents, the principal reduces the revenue extracted from a type. The key difference is that it can be achieved at a strictly positive quality, while extracting revenue from this type.

### Optimal Contracts for Concealed Types

For concealed types, we can provide an additional lower bound on quality in the optimal contract. The contract for concealed types is subject to constraints in two frames: a participation constraint in the lower frame  $l$  and an IC constraint in the higher frame  $h$ . Since concealed types cannot be imitated, there is no reason to distort their quality downward below the efficient quality in the lower frame,  $\hat{q}_{\theta_l}$ . It can be optimal, however, to increase the quality above this level in order to deliver rent more cost-effectively in order to satisfy the IC constraint.

---

<sup>13</sup>Indeed, a stronger result holds. The (RP) optimal contracts for any, even suboptimal, set of concealed types is implementable. The problem can be further simplified by noting that only local IC - those into the nearest revealed types - are binding. See Appendix A.2.

**Proposition 4.** Consider a concealed type  $\theta \in \Theta_C$ . Then, the optimal quality is bounded between the efficient quality in frame  $l$  and  $h$ :  $\widehat{q}_{\theta_l} \leq q_\theta \leq \widehat{q}_{\theta_h}$ . In particular, the optimal contract is

$$(p_\theta, q_\theta) = \begin{cases} \widehat{c}_{\theta_l}, & \text{if } \Delta_\theta \leq v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l}), \\ (v_{\theta_l}(q^*), q^*), & \text{if } \Delta_\theta \in [v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l}), v_{\theta_h}(\widehat{q}_{\theta_h}) - v_{\theta_l}(\widehat{q}_{\theta_h})], \\ (v_{\theta_h}(\widehat{q}_{\theta_h}) - \Delta_\theta, \widehat{q}_{\theta_h}), & \text{if } \Delta_\theta \geq v_{\theta_h}(\widehat{q}_{\theta_h}) - v_{\theta_l}(\widehat{q}_{\theta_h}), \end{cases}$$

where  $q^*$  solves  $v_{\theta_h}(q^*) - v_{\theta_l}(q^*) = \Delta_\theta$ , and  $\Delta_\theta := \operatorname{argmax}_{\theta' \in \Theta_R} u_{\theta_h}(c_{\theta'})$  denotes the rent delivered to type  $\theta \in \Theta_C$ , and  $c$  is the optimal contract.

If the required rent is low, only the participation constraint in the low frame binds and the optimal contract is the efficient contract for the type in the low frame. As more rent needs to be delivered in the high frame, it becomes optimal to increase the quality of the product up to the efficient quality in the high frame.

The contract further illustrates the cost of concealing a type. From the perspective of the high frame, a concealed type always receives at least the minimum rent  $v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l})$ , reducing the payoff of the principal. The cost of concealing a type is decreasing in the information rent  $\Delta$ . If it is sufficiently high (in the third regime of (4)), it is costless to conceal the type.

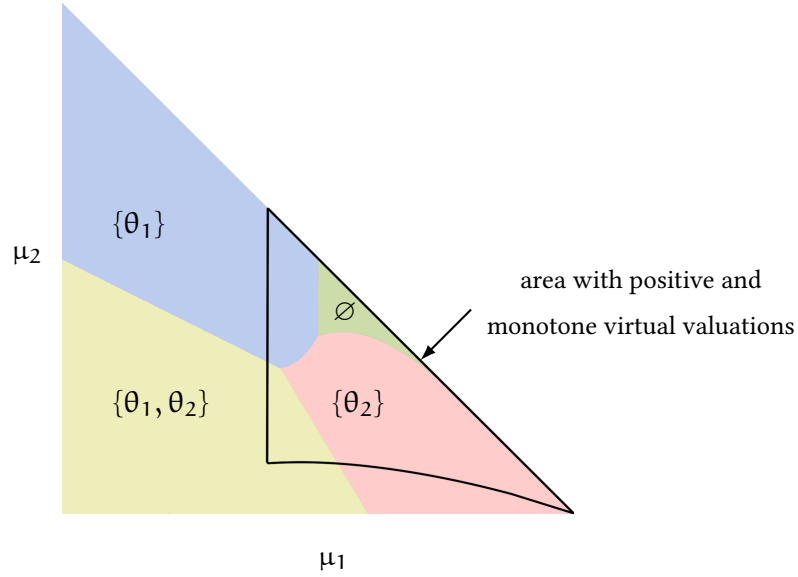
### 3.4 Optimal Concealed Types

One might conjecture that it is optimal for the principal to conceal low types and reveal high types. Even though this does not hold in general, this statement has a grain of truth: Revealing the highest type is always optimal. This is because types are concealed in order to eliminate downward deviations into them, which is not a concern for the highest type.

**Observation 2.** Suppose  $(\Theta_C^*, c^*)$  solves (RP) and the highest type  $\bar{\theta} = \max \Theta$  is concealed,  $\bar{\theta} \in \Theta_C^*$ . Then  $(\Theta_C^* \setminus \bar{\theta}, c^*)$  also solves (RP).

In general, there are no other restrictions on the optimal set of concealed types, as the following linear-quadratic three-type example illustrates. In Fig. 4 we plot the regions of the probability simplex where particular sets of concealed types are optimal. All four cases are realized for some probabilities. In addition, the restriction to monotone virtual values that ensures monotonicity in the classic screening model doesn't rule out any configuration.

**Figure 4:** Optimal  $\Theta_C$  for  $\theta^1 = (1, 3), \theta^2 = (4, 5), \theta^3 = (5, 6)$ .



Loosely speaking, the concealed types are substitutes for the principal. Consider two types  $\theta < \theta'$ . By concealing  $\theta$ , the principal reduces the rent  $\theta'$  obtains, increasing the costs of concealing  $\theta'$  (as it is more costly to conceal a type if it has a low information rent; Proposition 4). In addition, a lower rent implies that concealing  $\theta'$  has a lower gain as well, as information rents compound. Similarly, concealing  $\theta'$  reduces the benefit of concealing  $\theta$ . This pattern of substitutability is reflected in Fig. 4 as the regions  $\Theta_C = \{\theta_1\}$  and  $\Theta_C = \{\theta_2\}$  touch.

**Sufficiently Likely Types Are Revealed** It is not profitable to conceal very likely types, since the gain from the reduction of information rents for other types is outweighed by the loss of profits that can be extracted from them directly.

**Proposition 5.** *For any type  $\theta$  there exists a probability threshold  $\bar{\mu}_\theta \in (0, 1)$ , such that for any  $\mu_\theta \in [\bar{\mu}_\theta, 1]$ , an optimal set of revealed types contains  $\theta$ .*

This proposition suggests interpreting the contracts of revealed types as standard options that are relevant for common types of consumers and available immediately in the store, and the contracts for concealed types as specialty options relevant for rare types of consumers and available only on order.

**High  $\theta_l$  Favors Concealing** The difference between the valuations in frames  $h$  and  $l$  determines the cost of concealing. If we fix all types, but *increase* the  $l$ -frame



valuation of a concealed type, this cost is reduced and this type remains concealed.

**Proposition 6.** *Let  $\Theta_C$  be an optimal sets of concealed types for  $(\Theta, \mu)$  and let  $\theta \in \Theta_C$ . Define  $\tilde{\theta}$  such that  $\tilde{\theta}_l \geq \theta_l, \tilde{\theta}_h = \theta_h$ . Then, for the set of types  $(\Theta \setminus \{\theta\}) \cup \{\tilde{\theta}\}$  there exists a solution of the principal’s problem (RP) with the set of concealed types  $\tilde{\Theta}_C := (\Theta_C \setminus \{\theta\}) \cup \{\tilde{\theta}\}$ .*

Fixing the highest valuation the principal can achieve for each type, the cost of concealing is low if  $\theta_l$  is high. We can interpret this as a more precise control of the principal over consumer valuations. With sufficient control, she will conceal all types except for the highest.

**Proposition 7.** *Fix  $\theta_h, \forall \theta \in \Theta$ . There exists an  $\varepsilon > 0$  such that if  $\theta_h - \theta_l < \varepsilon, \forall \theta \in \Theta$ , then  $\Theta_C = \Theta \setminus \max \Theta$ .*

### 3.5 Discussion

**Commitment and Direct Mechanisms** Consumers are sophisticated but lack commitment. This is crucial, as the power of the principal to relax IC constraints by concealing types relies on the resulting dynamic inconsistency. In particular, this implies that our contracts cannot be implemented by a direct mechanism. Restricting to direct mechanisms effectively gives commitment as single-stage interaction does not allow for dynamic inconsistencies. As observed by Galperti (2015), with dynamically inconsistent agents, the revelation principle doesn’t apply directly. Instead, agents need to resubmit their complete private information at every stage. In our setting, an indirect mechanism is more convenient. Alternatively, one could construct an equivalent “quasi-direct” mechanism in which a reported type is mapped to a menu of menus instead of a contract.

The principal, by contrast, as the designer of the single-agent mechanism, has and requires commitment.<sup>14</sup>

**Weakening the Comonotonicity Assumption** Our assumptions can be relaxed at the cost of transparency. Suppose (i) *there exists a unique highest frame,*

---

<sup>14</sup>To see why, consider the terminal decision problem of a concealed type. Without commitment, the principal would increase quality on the contracts of concealed types and thereby violate the participation constraint in the low frame in the previous stage. If consumers anticipate this, the extensive-form decision problem unravels. Characterizing the outcome without commitment is beyond the scope of this paper.

i.e.

$$\exists h \in F, \{h\} = \bigcap_{\theta \in \Theta} \operatorname{argmax}_{f \in F} \theta_f$$

and (ii) *comonotonicity holds locally*, i.e. for each  $\theta$

$$\begin{aligned} \exists l(\theta) \in \operatorname{argmax}_{f \in F \setminus \{h\}} \theta_f, \text{ such that } \forall \theta', \\ \theta'_h < \theta_h \implies \theta'_{l(\theta)} \leq \theta_{l(\theta)}, \\ \theta_h < \theta'_h \implies \theta_{l(\theta)} < \theta'_{l(\theta)} < \theta'_h. \end{aligned}$$

Then our results generalize.<sup>15</sup> In particular, it is sufficient that there is an unambiguously highest and second highest frame. Note that we do not require the lowest type to be sensitive to framing, as framing is only used to place decoys.

Moreover, if  $\theta_{l(\theta)} = \theta_h, \forall \theta \in \Theta$ , then the principal achieves the first-best,  $\Pi^* = \Pi((\hat{c}_\theta)_{\theta \in \Theta})$ . This is because  $l(\theta)$  will be used to eliminate IC constraints, without tightening participation constraints as  $\theta_{l(\theta)} = \theta_h$ .

Consider the following example that violates Assumption 2, but satisfies the assumption above. A product has  $n$  flaws and there are  $n$  types of consumers, such that for type  $i$  flaw  $i$  is irrelevant. The sales person can either avoid discussing the flaws (high frame), or focus the attention on one of them. Formally, denote the types  $\theta^1, \dots, \theta^n$  and  $n + 1$  frames by  $h, l_1, \dots, l_n$  and suppose that

$$\begin{aligned} \forall i, \theta_h^i &= \theta_{l_i}^i, \\ \forall i, j, \theta_h^i &> \theta_{l_j}^i. \end{aligned}$$

The principal can implement the first-best using a standard EDP with  $\Theta_C = \Theta$  and type-specific low frame in the second stage.

**Participation Constraint At Every Stage** We assume that the agent can opt-out and choose the outside option at every stage of the decision problem. This is crucial for the trade-off between value exaggeration and rent extraction. A weaker restriction would be to require the outside option to be available only somewhere in the decision problem, for example at the root or in every terminal decision node. In this case, the principal can achieve full extraction at the efficient quantities in the high frame for some parameters.

<sup>15</sup>The principal only uses the frames  $h, \{l(\theta)\}_{\theta \in \Theta}$ . Furthermore, suppose the set of concealed types is  $\Theta_C$ . As long as local comonotonicity holds for all concealed types, our results generalize.

**Restriction on the Choice of Frames** We assume that the principal is unrestricted in the choice of frames and, in particular, that a change of framing is effective.<sup>16</sup> One might suppose that framing effects are instead partially “sticky”. That is, if the principal is choosing  $f'$  after  $f$ , then the consumer’s payoff type will be  $\alpha\theta_{f'} + (1 - \alpha)\theta_f$  for some  $\alpha \in (0.5, 1]$ . Then our results generalize.

**Random Mechanisms** We restrict the principal to use a deterministic extensive-form mechanism. One can show in examples that the principal can do strictly better by randomizing within the standard mechanism. Randomization allows the principal to smooth out the concealment of types. To see this, consider a situation with three types where it is optimal to conceal only the intermediate type and the  $P^1$ -constraint is binding in his contract. Then, the IC constraint from the highest to the intermediate type is slack at the root, as the intermediate type is concealed and the highest type obtains a strictly positive rent (from the IC to the lowest type). Consider a modification of the mechanism where the intermediate type is concealed with probability  $1 - \epsilon$  and revealed otherwise, obtaining the contract that is optimal ignoring the IC constraint of the highest type. The uncertainty resolves after the agent makes his decision at the root, but before the frame-change to  $l$ . In this mechanism, the highest type still strictly prefers not to imitate the intermediate type at the root if  $\epsilon$  is sufficiently small. Furthermore, ex-ante profit is strictly greater as the “revealed” contract for the intermediate type is more profitable than concealing him.

## 4 Extensions

### 4.1 Naivete

Naive consumers understand the structure of the extensive-form decision problem and the choices available to them, but they fail to anticipate the effect of framing. Faced with an EDP, they pick the continuation problem containing the contract

---

<sup>16</sup>This is in line with evidence showing that framing effects, such as gain-loss, are observed within-subject (Tversky and Kahneman, 1981) and even among philosophers who claim to be familiar with the notion of framing, to have a stable opinion about the answer to the manipulated question and were encouraged to consider a different framing from the one presented (Schwitzgebel and Cushman, 2015).

they prefer in their current frame.<sup>17</sup> They fail to take account of the fact that in this continuation problem, they may be in a different frame and end up choosing a different contract.

### Setup

Towards the definition of a naive outcome, let  $\mathcal{C}(e)$  denote the set of contracts in an EDP  $e$ . That is, letting  $\mathcal{C}(e) = e$  for  $e \in \mathcal{E}^0$ , define

$$\mathcal{C}(e) := \cup_{e' \in E} \mathcal{C}(e'), \text{ for } e = (E, f).$$

Now call  $s_\theta : \mathcal{E} \cup \mathcal{E}^0 \rightarrow \mathcal{E} \cup \mathcal{E}^0$  a naive strategy for  $\theta$  if

$$\begin{aligned} s_\theta|_{\mathcal{E}^0} &= \text{id} \\ s_\theta(E, f) &\in E \quad \forall (E, f) \in \mathcal{E} \\ \mathcal{C}(s_\theta(E, f)) &\cap \underset{\mathcal{C}(e)}{\text{argmax}} u_{\theta_f} \neq \emptyset. \end{aligned}$$

Put differently, when facing  $e = (E, f)$ , a consumer identifies the  $f$ -optima in the set of all contracts in  $e$ ,  $\mathcal{C}(e)$ , and chooses a continuation problem containing an optimum.

We call  $v : \Theta \rightarrow \mathcal{C}$  a naive outcome of an EDP  $e$  if there exists a naive strategy profile  $s$  such that any type  $\theta$  arrives at  $v(\theta)$  by following  $s_\theta$ , i.e.  $v(\theta) = (s_\theta \circ \dots \circ s_\theta)(e)$  for  $e \in \mathcal{E}^k$ . Let  $N^e$  be the set of all naive outcomes to an EDP  $e$ .

We consider the case when there are both naive and sophisticated consumers and the principal cannot observe their cognitive type. Let  $\Theta = \Theta_S \sqcup \Theta_N$  be the disjoint union of the set of sophisticated types  $\Theta_S$  and the set of naive types  $\Theta_N$ . That is, we allow for the existence of  $\theta^s \in \Theta_S$  and  $\theta^n \in \Theta_N$  which differ only in their sophistication, but not in their tastes conditional on any frame. Define the optimal profits similarly to (GP) as

$$\begin{aligned} \Pi^* &:= \max_{e \in \mathcal{E}} \Pi(c) & (7) \\ \text{s.t. } c_\theta &\in \Sigma^e(\theta), \forall \theta \in \Theta_S, \\ c_\theta &\in N^e(\theta), \forall \theta \in \Theta_N. \end{aligned}$$

---

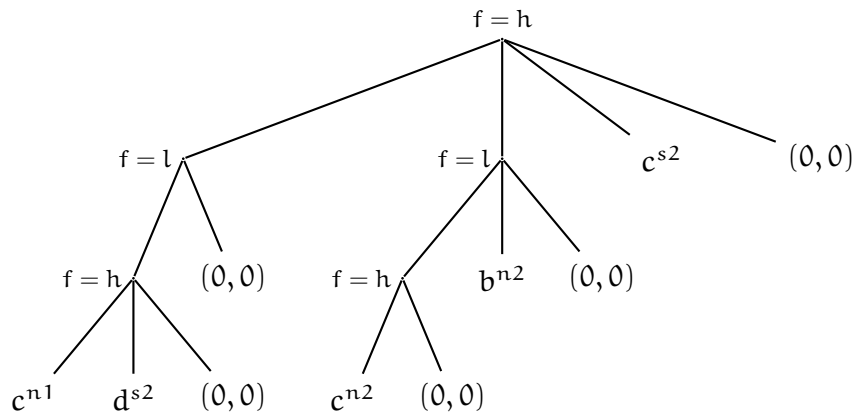
<sup>17</sup>A related idea is projection bias. (Loewenstein et al., 2003). The main difference is that our construction depends on the consumers' ability to forecast their future actions, not tastes. In this general sense, sophisticated consumers exhibit no projection bias, while naive consumers exhibit complete projection bias.

## Optimal Structure and Contracts

We illustrate in an example how the principal can use decoys to screen when naive types are present.

**Example 2.** Recall from Example 1 that there are two frames,  $\{l, h\}$ , and two payoff types,  $\{\theta^1, \theta^2\}$ . The key construction can be illustrated using three equally likely types, two naive and one sophisticated. There is a naive version of both payoff types, and a sophisticated high type, formally  $\Theta = \Theta_S \sqcup \Theta_N = \{\theta^{s2}\} \sqcup \{\theta^{n1}, \theta^{n2}\}$ . In this setting, the principal can sell the  $h$ -efficient quality to naive consumers and fully extract their surplus. This creates no information rents for the sophisticated type – screening by cognitive type is free. As a result, she can also implement the high-frame full-extraction contract for  $\theta^{s2}$ . The optimal EDP is given in Fig. 5. It implements  $c^{n1} = (16, 4)$ ,  $c^{n2} = (36, 6)$ ,  $c^{s2} = (36, 6)$ .

**Figure 5:** The optimal extensive-form decision problem in Example 2.



First, consider the sophisticated type. As in Example 1, the contract  $c^{n1}$  is more attractive than the implemented  $c^{s2}$ , but it is concealed using the decoy  $d^{s2}$ .

Let's turn to the naive types. The leftmost continuation problem is intended for  $\theta^{n1}$ . Even though  $\theta^{n1}$  is concealed, the principal extracts full surplus in the high frame. How is this possible? At the second stage in frame  $l$ , he indeed prefers the outside option over  $c^{n1}$ . But, he wrongly believes that he will choose the outside option after continuing. Hence, he continues and – back in frame  $h$  – chooses  $c^{n1}$ .

In order to implement the contract for  $\theta^{n2}$ , the principal needs to use a decoy. At the root, he strictly prefers  $c^{n1}$  to  $c^{n2}$ . In order to lure him into the middle continuation problem, the principal introduces a decoy  $b^{n2}$ . This decoy works

differently from the decoys used with sophisticated consumers.<sup>18</sup> It serves as bait and is the most preferred contract out of the whole decision problem for  $\theta^{n2}$ . As a consequence, he continues into the middle continuation problem. There, the switch happens:  $b^{n2}$  is unattractive from the perspective of the low frame and  $\theta^{n2}$  continues, expecting to pick the outside option in the continuation problem. Like  $\theta^{n1}$  he reconsiders at the terminal node and ends up with  $c^{n2}$ .  $\blacktriangle$

This construction generalizes. The optimal EDP achieves the same outcome as if the principal knows which consumers are naive and the types of the naive consumers. Naive types don't receive information rents, they obtain the full extraction contract in the high frame. Sophisticated consumers obtain the optimal contract according to Theorem 2.

**Theorem 3.** *Let  $e$  an optimal EDP with the set of types  $\Theta$  and  $\sigma$  and  $\nu$  be its firm-preferred sophisticated and naive outcomes, respectively. Then there exists an EDP  $e_S$  that is optimal for the set of types  $\Theta_S$  with conditional prior and its firm-preferred sophisticated outcome  $\sigma_S$ , such that*

$$\begin{aligned}\sigma(\theta) &= \sigma_S(\theta), \quad \forall \theta \in \Theta_S \\ \nu(\theta) &= \hat{c}_{\theta_n}, \quad \forall \theta \in \Theta_N.\end{aligned}$$

The optimal extensive-form decision problem retains the simple three-stage structure, we only add a continuation problem for each naive type to the extensive form described in Theorem 1. Consequently, the optimum can be achieved by a three-stage EDP with  $|\Theta|$  continuation choices at the root, similar to a standard EDP, but with additional second-stage decoys.

The principal also uses decoy contracts for naive consumers, but their role is reversed: In the construction for sophisticated consumers, we placed decoys in continuation problems to make sure that no other type wants to enter the continuation problem, as they correctly anticipate that they would choose the decoy. The construction for naive consumers is a mirror image: Instead of decoys to repel imitators, we introduce decoys in order to lure types into their corresponding continuation problems. Agents wrongly believe that they will choose their respective decoy, which is the most attractive contract in the whole EDP for them in their

---

<sup>18</sup>In this simple example, the two decoys,  $d^{s2} = (40, 8)$  and  $b^{n2} = (40, 8)$ , coincide. This is the case because they are designed to distract from the same option,  $c^2$ , and there are no other contracts in the decision problem. It doesn't hold true in general, even if naive and sophisticated consumers share the same payoff type.

current frame. Once types are separated at the root of the decision problem, the dynamic inconsistency introduced by changing frames allows the decision problem to reroute consumers from their decoy to the intended contract.

### **Welfare Gains from Sophistication**

Are consumers better off if they are sophisticated? Welfare statements in the presence of framing are generally fraught with difficulty. Still, we can rank the contracts obtained by sophisticated and naive agents from a consumer perspective without taking a stand on the welfare-relevant frame.<sup>19</sup> In the following sense sophistication partially protects consumers from exploitation through the use of framing.

**Observation 3.** *For all types, the contract under sophistication is weakly preferred to the contract under naivete from the perspective of every frame.*<sup>20</sup>

From an efficiency perspective, the two cases are not unambiguously ranked. For naive consumers, the principal implements the efficient quality from the perspective of the highest frame. Quality is lower for sophisticated consumers, an efficiency gain from the perspective of all frames except the highest one.

### **Discussion: Partial Naivete**

We can also extend our results to partial (magnitude) naivete. Denote the parameter determining the intensity of naivete by  $\alpha \in [0, 1]$ , with  $\alpha = 0$  representing full sophistication. Suppose a consumer with current payoff type  $\theta$  anticipates a future choice that will actually be made according to payoff type  $\theta'$ . Let  $\hat{\theta}(\theta, \theta', \alpha)$  denote what he currently perceives to be his future payoff type. Assume  $\hat{\theta}$  is increasing in the first two arguments and satisfies  $\hat{\theta}(\theta, \theta, \alpha) = \theta$  for all  $\alpha$ . Under full sophistication we have  $\hat{\theta}(\theta, \theta', 0) = \theta'$ , under full naivete  $\hat{\theta}(\theta, \theta', 1) = \theta$ . This structure ensures that predictions satisfy comonotonicity (Assumption 2). Whenever  $\alpha < 1$ , we can extend the sophisticated construction by replacing  $\theta_i^i$  by  $\hat{\theta}(\theta_h^i, \theta_l^i, \alpha)$  in (3) etc. Similarly, we can adjust the naive construction by modifying the decoy construction whenever  $\alpha > 0$ . In both cases, moving away from the baseline case

---

<sup>19</sup>The observation remains true if the choices in none of the frames are deemed welfare-relevant, as long as the welfare-relevant payoff parameters are weakly smaller than those induced by the highest frame.

<sup>20</sup>This can be interpreted as a weak improvement in the sense of Bernheim and Rangel (2009) if the two contracts are not identical.



increases the level of quality required in the decoys. If – contrary to what we assumed – providing very high quality decoys is not entirely costless or quality is bounded, we expect to see the sophisticated construction for agents with low  $\alpha$  and the naive construction for individuals with high  $\alpha$ .

## 4.2 Additional Participation Constraints and Cool-off Regulation

In many jurisdictions, regulation mandates a right to return a product for an extended period of time after the purchase. The express purpose of such regulation is to allow consumers to cool off and reconsider the purchase in a calm state of mind unaffected by manipulation by the seller.<sup>21</sup> Interestingly, such legislation typically only applies to door-to-door sales and similar situations of high sales pressure to which consumers did not decide to expose themselves. If consumers decide to enter a store or contact a seller, they are not protected by the law. This suggests that legislators consider the option to avoid the firm’s sales pressure entirely to be sufficient to protect consumers. Our framework allows us to evaluate this intuition.

Consider a situation when consumers decide whether or not to go to the store in the neutral frame. One can interpret this decision as an additional *interim* participation decision at the root. Alternatively, suppose that there is a regulation that allows consumers to return a product if they wish to do so ex-post in the neutral frame (as in [Salant and Siegel, 2018](#)). One can interpret this decision as an additional *ex-post* participation decision at every terminal decision stage.

Formally, denote the neutral frame by  $n \in F$ ,  $n < h$ .<sup>22</sup> This is the frame the consumer is in when unaffected by direct sales pressure by the firm.<sup>23</sup> We call  $\bar{e} := (\{e, 0\}, n)$  an *interim modification*<sup>24</sup> of  $e$ . Then  $e$  is an *EDP with an interim participation constraint* if it is an interim modification of some EDP. To define an *EDP with an ex-post participation constraint*, we define an *ex-post-modification*  $\underline{e}$  of an EDP  $e$  recursively. First, for any  $e \in \mathcal{E}^0$ , let  $\underline{e} := \bar{e}$ . Having defined an ex-post

---

<sup>21</sup>E.g. directive 2011/83/EU: "the consumer should have the right of withdrawal because of the potential surprise element and/or psychological pressure".

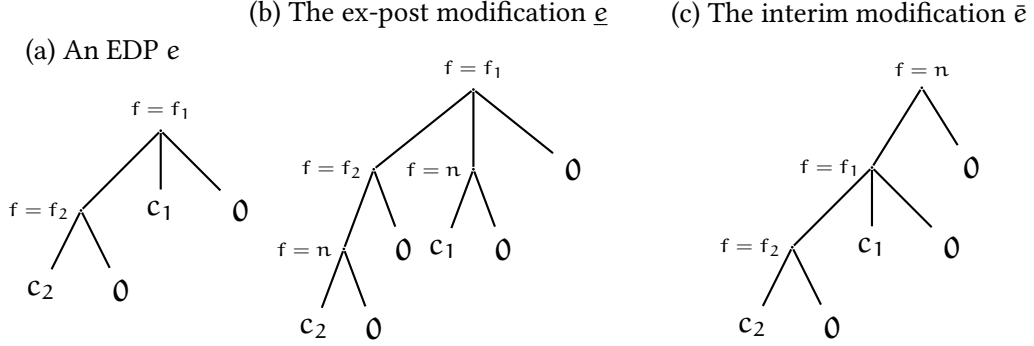
<sup>22</sup>It is immediate that additional participation decisions in frame  $h$  don’t change the result, as all implemented contracts satisfy them.

<sup>23</sup>One possible effect of marketing is influencing this neutral frame, but we won’t consider this margin.

<sup>24</sup>Here, the notion of interim modification is defined on  $\mathcal{E} \cup \mathcal{E}^0 \setminus \{0\}$ . For simplicity, let  $\bar{0} := 0$

modification on  $\mathcal{E}^j, \forall j = 0, \dots, k$ , for any  $e = (E, f) \in \mathcal{E}^{k+1}$ , define its ex-post modification as  $\underline{e} := (\{\underline{e}'\}_{e' \in E}, f)$ .

**Figure 6:** Interim and Ex-post Participation Constraints in Frame  $n$



**Sophisticated Consumers** If consumers are sophisticated, both constraints are equivalent and imply that if a contract is chosen by type  $\theta$ , then it must satisfy the additional participation constraint  $P_{\theta}^n$ . The following observation shows that the firm implements the efficient allocation associated with frame  $n$  and leaves no information rent to consumers.

**Observation 4.** *Suppose  $\Theta = \Theta_S$ . Let  $\bar{e}^*$  and  $\underline{e}^*$  be optimal EDPs with interim and ex-post participation constraints. Then their firm-preferred outcomes  $\bar{\sigma}$  and  $\underline{\sigma}$ , respectively, are such that*

$$\bar{\sigma}(\theta) = \underline{\sigma}(\theta) = \hat{c}_{\theta_n}.$$

This observation is immediate from Theorem 1. The principal can remove all incoming IC constraints at the cost of an additional participation constraint in a lower frame. As such a constraint is introduced in any case with interim or ex-post participation constraints in a neutral frame, the principal can conceal all types without additional cost.<sup>25</sup>

Both restrictions protect against overpurchases relative to the preferences in the neutral frame, but cannot protect against the extraction of all information rents by exploiting induced violations of dynamic consistency. If sophisticated

<sup>25</sup>Salant and Siegel (2018) show that the principal may not use framing when such a constraint is added to the problem of designing a framed menu. In particular, the principal cannot necessarily extract all rents without the use of an extensive form.

consumers can avoid the interaction with the firm, they indeed do not require additional protection by a right to return. They correctly anticipate their future actions and hence – given a choice – only interact with a seller, if the result will be acceptable to them from their current frame of reference.

**Naive Consumers** With naive consumers, we now need to distinguish between an interim choice to initiate the interaction and an ex-post right to return in the same neutral frame. While a right to return is still effective, naive consumers cannot protect themselves by avoiding the seller altogether.

**Observation 5.** *Suppose  $\Theta = \Theta_N$ . Let  $\bar{e}^*$  and  $\underline{e}^*$  be optimal EDPs with interim and ex-post participation constraints. Then their firm-preferred naive outcomes  $\bar{v}$  and  $\underline{v}$  satisfy  $\forall \theta \in \Theta$ ,*

$$\begin{aligned}\bar{v}_\theta &= \hat{c}_{\theta_n}, \\ \underline{v}_\theta &= \hat{c}_{\theta_n}.\end{aligned}$$

The intuition underlying the design of regulation does not apply for naive consumers. They are overly optimistic about the outcome of their interaction with the seller. As a result, the option to avoid the seller entirely is not sufficient to protect them from over-purchasing. In the optimal EDP, all consumers regret the purchase from the perspective of the neutral frame. A right to return even for in-store sales would offer them additional protection.

## 5 Conclusion and Discussion

We analyze the effect of framing in a model of screening. The principal can frame consumer decisions in several ways, affecting consumer valuations as expressed by their choices. Such a setting naturally leads to extensive-form decision problems. The firm uses framing not only to increase consumers valuations at the point of sale, but mainly to induce dynamic inconsistency and thereby reduce information rents, despite strategic sophistication of consumers. Our main result is that the optimal contracts can be implemented by an extensive-form decision problem with only three stages and two frames. At the initial interaction, only some contracts are immediately available, others are only available after the consumers' frame is

lowered – which we interpret as a cool-off period. Upon recall, the consumer is presented with an extended menu, but chooses the expected option.<sup>26</sup>

This simple extensive form allows the firm to eliminate information rents at the cost of lower surplus and thereby achieve a payoff that is strictly larger than full surplus extraction at all but the highest frame. Even if consumers are protected by a shop-entry decision or right to return the product in an exogenously given neutral frame, they are not protected against the full extraction of their information rents.

We also characterize the outcome with naive consumers. The structure of the optimal extensive form and the contracts of sophisticated agents are robust to the presence of naive types. Naive types can be screened without generating any additional information rents.

**Beyond Framing** Throughout the analysis, we assumed that choice depends on exogenous factors of the presentation of the product that are chosen by the principal (i.e. the frame), but satisfies the axioms of utility maximization given every frame. If framing affects choice through focusing the attention of consumers on certain attributes, for example, we consider the case where these attributes are emphasized by the salesperson or the information material and assume that the properties of the choice set (such as an attribute being widely dispersed) are not relevant. Our construction induces and exploits a violation of dynamic consistency.

Such violations can also be caused by other factors, such as seasonal shifts in tastes or context effects. Hence, our construction can in principle be extended to such a setting. Consider, for example, the sale of a convertible. The current weather affects the valuation consumers have for convertibles. It constitutes an exogenous frame that cannot be manipulated directly by the principal. Analogously to Assumption 2, assume that all consumer types have a higher valuation for convertibles if the weather is nice (comonotonic frames). Furthermore, if one consumer type has a higher valuation for quality convertibles than another when the sun is shining, this is still true when it rains – albeit the valuation of both types is reduced (comonotonic types). Sophisticated consumers expect these shifts but consider tastes different from their current ones as mistakes. In such a setting, one could ask which pattern of taste changes is required to achieve the optimum. Our results imply that a simple pattern of taste changes (high-low-high) is sufficient.<sup>27</sup>

---

<sup>26</sup>This simple structure also supports the assumption of sophistication. In the optimum, consumers only need to grasp relatively short and intuitively understandable extensive forms.

<sup>27</sup>In particular, the car dealer can find the optimal EDP using our results and implement it as

We expect that similar ideas can be applied to a setting with endogenous frames, e.g. the model of focusing (Kőszegi and Szeidl, 2013). There, a change of frame corresponds to introducing an option to the choice set that directs the focus more towards quality.<sup>28</sup> There is an important caveat, however. Even if it is possible to extend our construction into an analogous setting with endogenous frames, the resulting EDP may not be optimal. While in our setting, the frames have to be fixed for every decision node independently of the agent’s type, context effects depend on the choice set. The choice set is generated by backward induction and hence type dependent. In effect, the frame can be type dependent. Consequently, screening with choice-set dependent preferences is a considerably richer setting and left for future research.

## References

- Areni, Charles S. and David Kim**, “The Influence of Background Music on Shopping Behavior: Classical Versus Top-Forty Music in a Wine Store,” *ACR North American Advances*, 1993, NA-20. 1
- Bernheim, B. Douglas and Antonio Rangel**, “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” *The Quarterly Journal of Economics*, January 2009, 124 (1), 51–104. 28
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer**, “Salience and Consumer Choice,” *Journal of Political Economy*, 2013, 121 (5), 803–843. 1, 7
- Bushong, Benjamin, Lindsay M. King, Colin F. Camerer, and Antonio Rangel**, “Pavlovian Processes in Consumer Choice: The Physical Presence of a Good Increases Willingness-to-Pay,” *American Economic Review*, 2010, 100 (4), 1556–71. 1

---

follows. The cars intended for revealed types can be bought immediately when the sun is shining. Cars for concealed types need to be pre-ordered. The pre-order period is sufficiently long to contain a sustained period of rain and the order can be canceled at any time. When the car is ready, it can be picked up only when the sun is shining. The consumer is offered a range of (decoy) options at this point, which are immediately available.

<sup>28</sup>This is possible by introducing a high quality-high price option that remains unchosen by every type. Kőszegi and Szeidl (2013) argue that products that are extremely bad on all attributes are typically not taken into consideration. We don’t require such products, a high quality product that is too expensive for every consumer type is sufficient.

- DelVecchio, Devon, H. Shanker Krishnan, and Daniel C. Smith**, “Cents or Percent? The Effects of Promotion Framing on Price Expectations and Choice,” *Journal of Marketing*, 2007, 71 (3), 158–170. 1
- Eliaz, Kfir and Ran Spiegler**, “Contracting with Diversely Naive Agents,” *The Review of Economic Studies*, January 2006, 73 (3), 689–714. 4, 6
- **and** —, “Consumer Optimism and Price Discrimination,” *Theoretical Economics*, 2008, 3 (4), 459–497. 6
- Esteban, Susanna and Eiichi Miyagawa**, “Optimal Menu of Menus with Self-Control Preferences,” *mimeo*, 2006. 6, 7
- **and** —, “Temptation, Self-Control, and Competitive Nonlinear Pricing,” *Economics Letters*, March 2006, 90 (3), 348–355. 6
- , —, **and Matthew Shum**, “Nonlinear Pricing with Self-Control Preferences,” *Journal of Economic Theory*, July 2007, 135 (1), 306–338. 6
- Galperti, Simone**, “Commitment, Flexibility, and Optimal Screening of Time Inconsistency,” *Econometrica*, 2015, 83 (4), 1425–1465. 6, 7, 22
- Glazer, Jacob and Ariel Rubinstein**, “A Model of Persuasion with Boundedly Rational Agents,” *Journal of Political Economy*, December 2012, 120 (6), 1057–1082. 7
- **and** —, “Complex Questionnaires,” *Econometrica*, 2014, 82 (4), 1529–1541. 7
- Gourville, John T.**, “Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation,” *Journal of Consumer Research*, March 1998, 24 (4), 395–403. 1
- Gul, Faruk and Wolfgang Pesendorfer**, “Temptation and Self-Control,” *Econometrica*, 2001, 69 (6), 1403–1435. 6
- Hardisty, David J., Eric J. Johnson, and Elke U. Weber**, “A Dirty Word or a Dirty World? Attribute Framing, Political Affiliation, and Query Theory,” *Psychological Science*, 2010, 21 (1), 86–92. 1
- Heidhues, Paul and Botond Köszegi**, “Exploiting Naivete about Self-Control in the Credit Market,” *American Economic Review*, December 2010, 100 (5), 2279–2303. 6, 7

- **and** — , “Naivete-Based Discrimination,” *The Quarterly Journal of Economics*, May 2017, pp. 1019–1054. 6
- Herrero, Maria José and Sanjay Srivastava**, “Implementation via Backward Induction,” *Journal of Economic Theory*, 1992, 56 (1), 70–88. 7
- Johnson, Eric J., John Hershey, Jacqueline Meszaros, and Howard Kunreuther**, “Framing, Probability Distortions, and Insurance Decisions,” *Journal of Risk and Uncertainty*, 1993, 7, 35–51. 1
- Kőszegi, Botond**, “Behavioral Contract Theory,” *Journal of Economic Literature*, December 2014, 52 (4), 1075–1118. 6
- **and Adam Szeidl**, “A Model of Focusing in Economic Choice,” *The Quarterly Journal of Economics*, February 2013, 128 (1), 53–104. 7, 33
- Laibson, David**, “Golden Eggs and Hyperbolic Discounting,” *The Quarterly Journal of Economics*, May 1997, 112 (2), 443–478. 4
- Loewenstein, George, Ted O’Donoghue, and Matthew Rabin**, “Projection Bias in Predicting Future Utility,” *The Quarterly Journal of Economics*, 2003, 118 (4), 1209–1248. 25
- Maskin, Eric and John Riley**, “Monopoly with Incomplete Information,” *The RAND Journal of Economics*, 1984, 15 (2), 171–196. 7
- Moser, Christian and Pedro Olea de Souza e Silva**, “Optimal Paternalistic Savings Policies,” working paper, 2019. 6
- Mussa, Michael and Sherwin Rosen**, “Monopoly and Product Quality,” *Journal of Economic Theory*, August 1978, 18 (2), 301–317. 7
- North, Adrian C., Amber Shilcock, and David J. Hargreaves**, “The Effect of Musical Style on Restaurant Customers’ Spending,” *Environment and Behavior*, September 2003, 35 (5), 712–718. 1
- , **David J. Hargreaves, and Jennifer McKendrick**, “In-Store Music Affects Product Choice,” *Nature*, November 1997, 390 (6656), 132–132. 1
- , **Lorraine P. Sheridan, and Charles S. Areni**, “Music Congruity Effects on Product Memory, Perception, and Choice,” *Journal of Retailing*, March 2016, 92 (1), 83–95. 1



- Piccione, Michele and Ran Spiegler**, “Price Competition Under Limited Comparability,” *The Quarterly Journal of Economics*, January 2012, 127 (1), 97–135. 6
- Salant, Yuval and Ariel Rubinstein**, “(A, f): Choice with Frames,” *The Review of Economic Studies*, 2008, 75 (4), 1287–1296. 7
- **and Ron Siegel**, “Contracts with Framing,” *American Economic Journal: Microeconomics*, August 2018, 10 (3), 315–346. 4, 6, 12, 29, 30
- Schmitz, Hendrik and Nicolas R. Ziebarth**, “Does Price Framing Affect the Consumer Price Sensitivity of Health Plan Choice?,” *Journal of Human Resources*, 2017, 52 (1), 88–127. 1
- Schwitzgebel, Eric and Fiery Cushman**, “Philosophers’ Biased Judgments Persist despite Training, Expertise and Reflection,” *Cognition*, August 2015, 141, 127–137. 24
- Spiegler, Ran**, *Bounded rationality and industrial organization*, Oxford University Press, 2011. 5, 6
- , “Competitive Framing,” *American Economic Journal: Microeconomics*, 2014, 6 (3), 35–58. 6
- Strotz, R. H.**, “Myopia and Inconsistency in Dynamic Utility Maximization,” *The Review of Economic Studies*, 1955, 23 (3), 165. 4
- Tversky, Amos and Daniel Kahneman**, “The Framing of Decisions and the Psychology of Choice,” *Science*, 1981, 211 (4481), 453–458. 1, 24
- Yu, Pei Cheng**, “Optimal Retirement Policies with Present-Biased Agents,” working paper, 2018. 6
- Zhang, Wei**, “Endogenous Preferences and Dynamic Contract Design,” *The B.E. Journal of Theoretical Economics*, January 2012, 12 (1). 6

# A Online Appendix

## A.1 Preliminaries

Whenever types are enumerated by subscript  $i$ , we use notation  $u_f^i := u_{\theta_f^i}$ . For each  $\theta_f^i$ , let  $\succsim_f^i$  and (sometimes)  $\succ_{\theta_f^i}$  denote the corresponding preference relation. For any EDP  $e$ , let  $\mathcal{C}(e)$  denote the set of all contracts, available in  $e$ . Formally, for  $e = (A, f) \in \mathcal{E}^1$ ,  $\mathcal{C}(e) := A$  and, recursively, for  $e = (E, f) \in \mathcal{E}^k$ ,  $\mathcal{C}(e) := \cup_{e' \in E} \mathcal{C}(e')$ .

First, note that the consumer's preferences exhibit the single-crossing property, which is established in the following

**Lemma 2** (Single-crossing property). *For any two payoff types  $\bar{\theta}, \underline{\theta} \in \mathbb{R}$ , such that  $\bar{\theta} \geq \underline{\theta}$ , and contracts  $x, y \in \mathcal{C}$ , such that  $q^y \geq q^x$ , we have*

$$\begin{aligned} u_{\bar{\theta}}(y) \leq u_{\bar{\theta}}(x) &\implies u_{\underline{\theta}}(y) \leq u_{\underline{\theta}}(x), \\ u_{\underline{\theta}}(y) \geq u_{\underline{\theta}}(x) &\implies u_{\bar{\theta}}(y) \geq u_{\bar{\theta}}(x). \end{aligned}$$

*Proof.* Take any  $x, y \in \mathcal{C}$ , such that  $q^y \geq q^x$  and  $u_{\underline{\theta}}(y) - u_{\underline{\theta}}(x) \geq 0$ . Note that the increasing differences property ( $v_{\theta q} > 0$ ) imply that

$$\begin{aligned} u_{\bar{\theta}}(y) - u_{\bar{\theta}}(x) &= v(\bar{\theta}, y) - v(\bar{\theta}, x) + p^y - p^x \\ &= \int_x^y \frac{\partial v_{\bar{\theta}}}{\partial q}(q) dq + p^y - p^x \\ &= \int_x^y \left( \frac{\partial v_{\underline{\theta}}}{\partial q}(q) + \int_{\underline{\theta}}^{\bar{\theta}} v_{\theta q}(\theta, q) d\theta \right) dq + p^y - p^x \\ &\geq v_{\underline{\theta}}(y) - v_{\underline{\theta}}(x) + p^y - p^x \\ &= u_{\underline{\theta}}(y) - u_{\underline{\theta}}(x) \geq 0. \end{aligned}$$

The proof of the first implication is analogous. □

Second, we prove the following result that ensures existence of optimal EDP.

**Lemma 3.** *For any prices  $\bar{p} > \underline{p} \geq 0$  and payoff types  $\bar{\theta}_{f'} > \underline{\theta}_{f'}$ , there exist  $q \geq 0$ , such that*

$$u_{\underline{\theta}_{f'}}(\underline{p}, q) = u_{\bar{\theta}_{f'}}(\bar{p}, q).$$

*Or, equivalently, the function  $q \mapsto v_{\bar{\theta}_{f'}}(q) - v_{\underline{\theta}_{f'}}(q)$  is unbounded.*

*Proof.* Take any  $\bar{\theta}_{f'} > \underline{\theta}_{f'}$  and set  $\phi(q) := v_{\bar{\theta}_{f'}}(q) - v_{\underline{\theta}_{f'}}(q)$ . Note that  $\phi$  is thrice differentiable and strictly increasing. Our assumption  $\frac{\partial^3 v}{\partial q^2 \partial \theta} > 0$  implies that  $\phi$  is

strictly convex as

$$\phi''(q) = v''_{\hat{\theta}_f}(q) - v''_{\underline{\theta}_f}(q) = \int_{\underline{\theta}_f}^{\hat{\theta}_f} \frac{\partial^3 v_{\theta}}{\partial q^2 \partial \theta}(q) d\theta > 0.$$

Now, take any  $\hat{q} > 0$ , and note that  $\phi$  is weakly greater than its positively-sloped affine support function at  $\hat{q}$ , which is unbounded.

Finally, since  $\phi$  is unbounded, for any prices  $\bar{p} > \underline{p} \geq 0$ , there exists  $q \geq 0$ , such that

$$\bar{p} - \underline{p} = \phi(q) = v_{\hat{\theta}_f}(q) - v_{\underline{\theta}_f}(q) \Leftrightarrow u_{\theta_f}(\underline{p}, q) = u_{\theta_f}(\bar{p}, q).$$

□

**Lemma 4.** *The efficient quality for payoff type  $\theta_f$  defined as*

$$\hat{q}_{\theta_f} := \operatorname{argmax}_{q \geq 0} v_{\theta_f}(q) - \kappa(q)$$

*exists, is unique and increasing in  $\theta_f$ .*

*Proof.* Define the surplus function  $\zeta_{\theta_f} : \mathbb{R}_+ \rightarrow \mathbb{R}$  as

$$\zeta_{\theta_f}(q) := v_{\theta_f}(q) - \kappa(q) \tag{8}$$

and note that it is continuous, twice differentiable and satisfies

$$\zeta_{\theta_f}(0) = 0 \tag{9}$$

$$\zeta'_{\theta_f}(0) > 0 \tag{10}$$

$$\lim_{q \rightarrow \infty} \zeta'_{\theta_f}(q) < 0 \tag{11}$$

$$\zeta''_{\theta_f} > 0, \tag{12}$$

$$\frac{\partial \zeta'_{\theta_f}(\hat{q}_{\theta_f})}{\partial \theta_f} > 0. \tag{13}$$

Properties (10) and (11) together with the Mean Value Theorem ensure that there exists a unique global efficient quantify  $\hat{q}_{\theta_f}$  defined as

$$\zeta'_{\theta_f}(\hat{q}_{\theta_f}) = 0 \Leftrightarrow v'_{\theta_f}(\hat{q}_{\theta_f}) = \kappa'(\hat{q}_{\theta_f}).$$

Note that (12) implies

$$\operatorname{sgn}(\zeta'_{\theta_f}(q)) = \operatorname{sgn}(\hat{q}_{\theta_f} - q). \tag{14}$$

In addition, (13) implies that  $\hat{q}_{\theta_f}$  is increasing in  $\theta_f$ . □

## A.2 Proofs

*Proof of Observation 1 on page 12:* First, consider a 1-EDP: By the usual arguments, the revenue maximal menu satisfies monotonicity, participation at the bottom and local downward IC constraints, the latter two with equality. Conversely, these constraints jointly imply the full set of constraints. We index types in an increasing order, i.e.  $\Theta = \{\theta_1, \dots, \theta_n\}$  with  $\theta_i < \theta_{i+1}$ . Suppose towards a contradiction that the revenue under frame  $f < h$  is maximal with an optimal menu  $\{(p_i, q_i)\}_{i \in \{1, \dots, n\}}$ . Consider the menu  $\{(p'_i, q_i)\}_{i \in \{1, \dots, n\}}$  set in frame  $h$  with

$$\begin{aligned} p'_i &= p_i + \Delta_i \\ \Delta_1 &= v(\theta_h^1, q_1) - v(\theta_f^1, q_1) \\ \Delta_i &= \Delta_{i-1} + v(\theta_h^i, q_i) - v(\theta_f^i, q_i) - [v(\theta_h^i, q_{i-1}) - v(\theta_f^i, q_{i-1})] \end{aligned}$$

This menu still satisfies monotonicity, participation at the bottom and the local downward IC constraints are still binding, as

$$\begin{aligned} v(\theta_h^i, q_i) - p'_i &= v(\theta_f^i, q_i) + v(\theta_h^i, q_i) - v(\theta_f^i, q_i) - p_i - \Delta_i \\ &= v(\theta_f^i, q_{i-1}) + v(\theta_h^i, q_i) - v(\theta_f^i, q_i) - p_{i-1} - \Delta_i \\ &= v(\theta_f^i, q_{i-1}) - p_{i-1} - \Delta_{i-1} + [v(\theta_h^i, q_{i-1}) - v(\theta_f^i, q_{i-1})] \\ &= v(\theta_h^i, q_{i-1}) - p_{i-1} - \Delta_{i-1} \\ &= v(\theta_h^i, q_{i-1}) - p'_{i-1} \end{aligned}$$

Hence all other IC and P are satisfied by the usual arguments. Note that  $\Delta_i \geq 0$  by single crossing, hence expected revenue is strictly (weakly if all types pool at  $q = 0$ , but this is never optimal) higher under the modified contract in the higher frame.

The claim for  $F = \{h\}$  follows from the following:  $\forall \theta \in \Theta, \forall e \in \mathcal{E}^k, \Sigma^e(\theta) = \operatorname{argmax}_{\mathcal{C}(e)} u_{\theta_h}$  by induction on  $k$ . The base  $k = 1$  is by definition. Take any  $e = (E, h) \in \mathcal{E}^{k+1}$ .  $\sigma$  is an outcome to  $e$  iff

$$\sigma(\theta) \in \operatorname{argmax}_{\{\sigma^j(\theta)\}_{e^j \in E}} u_{\theta_h}(\sigma^j(\theta)) \text{ with } \sigma^j \in \Sigma^{e^j}(\theta) = \operatorname{argmax}_{\mathcal{C}(e^j)} u_{\theta_h}.$$

Therefore,  $\sigma$  is an outcome to  $e$  if and only if

$$\sigma(\theta) \in \operatorname{argmax}_{\bigcup_{e^j \in E} \operatorname{argmax}_{\mathcal{C}(e^j)} u_{\theta_h}} u_{\theta_h} = \operatorname{argmax}_{\mathcal{C}(e)} u_{\theta_h}$$

Hence  $e' := (\mathcal{C}(e), h)$  is outcome equivalent to  $e$  and the optimal EDP is equivalent to the optimal menu.  $\square$

The proof of Theorem 1 relies on many arguments that are required for the proofs of the following results as well. To avoid repetition, we prove Theorem 2 and Theorem 2 together.

*Proof of Proposition 1 on page 16:* The necessity of the constraints for a given set of frames  $f_R, \{f_\theta\}_{\theta \in \Theta^c}$  is derived in the main text. In particular, we saw that the incentive compatibility constraints are determined by the frames used on the path to the contract of the imitated type, not the imitating type and frames used on the path to  $c_{\theta'}$  cannot eliminate the IC constraints from  $\theta'$  to  $\theta$  for any  $\theta', \theta \in \Theta$

To prove the proposition, it remains to show that we can assume that  $f_R = h$  and  $f_\theta = l$  for all concealed types. Suppose  $f_R \neq h$ . But then we can set  $\Theta'_C = \Theta$ . As all contracts satisfy the participation constraint in  $f_R < h$ , they satisfy the participation constraint in  $l$ . As there are no incentive compatibility constraints with  $\Theta'_C = \Theta$ , all constraints associated to this set of hidden types are satisfied. Suppose instead that  $f_R = h$  but for some type  $\theta' \in \Theta_C$  we have  $f_{\theta'} \neq l$ . But then the participation constraint for  $f_{\theta'} = l$  is satisfied and hence the set of contracts is feasible in the relaxed problem.  $\square$

Before we proceed, we prove a more detailed decoy construction lemma.

**Lemma 5.** *For any ordered vector of types  $(\theta^j)_{j=0}^n$  and contract  $x$ , there exists a sequence of decoys  $\mathcal{D}(x, (\theta^j)_{j=0}^n) := (d_j)_{j=1}^n$ , such that  $\forall j \geq 1, \forall k \neq j$ , we have*

1.  $q_j^d \geq q_{j-1}^d$  (decoy quantities are increasing),
2.  $d_j \sim_j^h d_{j-1}$ ,
3.  $d_j \succ_j^h x$  and  $d_j \succ_j^h d_k$  ( $\theta^j$  chooses  $d_j$ ),
4.  $x \succ_0^h d_j$  ( $\theta^0$  chooses  $x$ ),
5.  $0 \succ_j^l d_j$  (decoys are undesirable in  $l$ ).

*Proof.* Let  $d_0 = x$  for brevity. We construct the decoys  $d_i = (p_i, q_i)$  recursively.<sup>29</sup> For  $i \in \{1, \dots, n\}$ , pick

$$d_i \sim_{\theta_i} 0 \tag{15}$$

$$d_i \sim_{\theta_i} d_{i-1} \tag{16}$$

or equivalently

$$v(\theta_i^i, q_i) - p_i = 0$$

$$v(\theta_h^i, q_i) - p_i = v(\theta_h^i, q_{i-1}) - p_{i-1}$$

<sup>29</sup>The present proof can be extended to the case without ordered types (but maintaining ordered frames).

Existence of such a  $d_i$  follows from Lemma 3. To verify this construction, we proceed through a series of claims.

*Claim 1.* Decoy quantities are increasing:  $q_i \geq q_{i-1}$ .

*Proof of Claim 1:* By the two defining relations

$$\begin{aligned} v(\theta_l^i, q_i) &= p_i \\ v(\theta_h^i, q_i) - p_i &= v(\theta_h^i, q_{i-1}) - p_{i-1} \end{aligned}$$

Hence

$$v(\theta_h^i, q_i) - v(\theta_l^i, q_i) = v(\theta_h^i, q_{i-1}) - v(\theta_{i-1}^l, q_{i-1}) > v(\theta_h^i, q_{i-1}) - v(\theta_l^i, q_{i-1})$$

and  $q_i > q_{i-1}$  is established as it is implied by single crossing from

$$\int_{\theta_l^i}^{\theta_h^i} v_\theta(\theta, q_i) d\theta > \int_{\theta_l^i}^{\theta_h^i} v_\theta(\theta, q_{i-1}) d\theta$$

△

This also shows that all  $(p, q)$  are positive, as  $q_i \geq q_0 = q_x$ .

*Claim 2.* The decoy intended for type  $\theta_i$  is chosen by this type:  $d_i \in \arg \max_{\theta_h^i} d_j$ .

*Proof of Claim 2:* We will show that for all  $j$  we have  $d_i \succ_{\theta_h^i} d_j$ . First, suppose  $j < i$ . Then for all  $k \in [j, i]$  we have

$$d_k \sim_{\theta_h^k} d_{k-1}$$

and  $q_k > q_{k-1}$ . This implies

$$d_k \succ_{\theta_h^k} d_{k-1}$$

and by since  $\theta_h^i \geq \theta_h^k$

$$d_k \succ_{\theta_h^i} d_{k-1}$$

The desired result follows by transitivity.

Second, suppose  $j > i$ . Again for all  $k \in [i, j]$ ,

$$d_k \sim_{\theta_h^k} d_{k-1}$$

and  $q_k > q_{k-1}$ . But then

$$d_k \preceq_{\theta_h^i} d_{k-1}$$

for every decoy since  $\theta_h^i \geq \theta_h^k$ . The desired result again follows by transitivity. △

□

*Proof of Lemma 1 on page 17:* A continuation problem for type  $\theta \in \Theta$  with contract  $c_\theta$  satisfying all three properties is given by  $e_\theta = (\{\mathbf{0}, (\{\mathbf{0}, c_\theta\} \cup \{d_{\theta'}\}_{\theta' > \theta}, h)\}, l)$ , where the contracts  $\{d_{\theta'}\}_{\theta' > \theta}$  are constructed in Lemma 5 as  $(d_{\theta'})_{\theta' \geq \theta} := \mathcal{D}(c_\theta, (\theta')_{\theta' \geq \theta})$ .

By construction, type  $\theta$  chooses  $c_\theta$  from the terminal problem and since  $c_\theta$  satisfies the participation constraint in the low frame,  $c_\theta \in \Sigma^{e_\theta}(\theta)$ . For higher types, the terminal decision problem resolves to the menu  $\{d_{\theta'}, \mathbf{0}\}$  and by construction the outside option is weakly preferred in the low frame. Having established (ii) and (iii), it remains to show (i). Consider a type  $\theta' < \theta$ . In the terminal decision problem, we have  $d_0 \succ_{\theta_h} d_i$  and  $q_i \geq q_0$ , hence by single crossing  $d_0 \succ_{\theta'_h} d_i$ , which establishes that a lower type never chooses any of the decoys.  $\square$

*Proof of Proposition 2 on page 16:* Suppose that  $c$  satisfies  $\{P_\theta^h\}_{\theta \in \Theta_R}, \{P_\theta^l\}_{\theta \in \Theta_C}, \{IC_{\theta\theta'}^h\}_{\theta < \theta'}, \{IC_{\theta\theta'}^h\}_{\theta \in \Theta, \theta' \in \Theta_R}$  for some partition  $\{\Theta_C, \Theta_R\}$  of  $\Theta$ . Then let  $e_\theta$  for each  $\theta \in \Theta_C$  be constructed as in Lemma 1 and consider a standard EDP

$$e^* = \left( \{e_\theta\}_{\theta \in \Theta_C} \cup \{c_\theta\}_{\theta \in \Theta_R} \cup \{\mathbf{0}\}, h \right).$$

Notice that from Lemma 1 it follows for each  $\theta \in \Theta_C$  that there exist an outcome  $\sigma^\theta$  of  $e_\theta$ , such that

$$\begin{aligned} \sigma^\theta(\theta') &\in \{\mathbf{0}, c_\theta\}, \forall \theta' \in \Theta, \\ \sigma^\theta(\theta') &= \mathbf{0}, \forall \theta' > \theta, \\ \sigma^\theta(\theta) &= c_\theta. \end{aligned}$$

Now let  $\sigma$  be such that  $\sigma(\theta) = c_\theta$ . To show that  $\sigma$  is an outcome of  $e^*$ , notice that constraints  $\{IC_{\theta\theta'}^h\}_{\theta \in \Theta, \theta' \in \Theta_R}$  and  $\{P_\theta^h\}_{\theta \in \Theta_R}$  imply that  $\forall \theta \in \Theta$ ,

$$\sigma(\theta) = c_\theta \in \operatorname{argmax}_{\{c_{\theta'}\}_{\theta' \in \Theta_R}} u_{\theta_h}.$$

Similarly, constraints  $\{IC_{\theta\theta'}^h\}_{\theta < \theta'}$  and  $\{P_\theta^h\}_{\theta \in \Theta_R}$  imply that  $\forall \theta \in \Theta$ ,

$$\sigma(\theta) = c_\theta \in \operatorname{argmax}_{\{c_{\theta'}\}_{\theta \leq \theta'}} u_{\theta_h}.$$

Therefore,  $\sigma$  satisfies

$$\sigma(\theta) \in \operatorname{argmax}_{\{\sigma^{\theta'}(\theta)\}_{e_\theta} \cup \{c_{\theta'}\}_{\theta' \in \Theta_R}} u_{\theta_h},$$

which means that it is an outcome of  $e^*$ .  $\square$

*Proof of Theorem 1 on page 14 and Theorem 2 on page 18:* Let  $(c_\theta)_{\theta \in \Theta}$ ,  $\Theta_C$  be a solution to the relaxed problem and  $e^*$  a standard EDP with decoys for all  $\theta \in \Theta_C$  constructed as in Lemma 5 and Lemma 1. We need to show that  $e^*$  implements  $(c_\theta)_{\theta \in \Theta}$ .

First, note that  $\Sigma^e$  is rectangular, i.e. if  $\sigma, \sigma' \in \Sigma^e$  with  $\sigma(\theta) \neq \sigma'(\theta)$  and  $\sigma(\theta') \neq \sigma'(\theta')$ , there exists a  $\sigma^* \in \Sigma^e$  with  $\sigma^* = \sigma$  except  $\sigma^*(\theta') = \sigma'(\theta')$ .

It follows from the IC constraints that there is no strictly profitable deviation into contracts of revealed types, i.e.  $\Sigma^{e^*}(\theta) \cap \{c_{\theta'}\}_{\theta' \in \Theta_R \setminus \theta} \neq \emptyset$  implies that  $c_\theta \in \Sigma^{e^*}(\theta)$ . From Lemma 1, it follows that type  $\theta$  cannot deviate downwards into concealed types and that no decoys are chosen, i.e.  $\Sigma^{e^*}(\theta) \subset \{c_{\theta'}\}_{\theta \setminus \{\theta' < \theta | \theta' \in \Theta_C\}}$ . It remains to show that there are no strictly profitable upwards deviations in  $e^*$  to complete the proof, establishing  $c_\theta \in \Sigma^{e^*}(\theta)$  for all  $(c_\theta)_{\theta \in \Theta}$ .

As the proof relies on properties of the solution to (RP), we start by simplifying the relaxed problem. Define

$$\eta(\theta) := \max\{\theta' \in \Theta_R | \theta' < \theta\}$$

the closest revealed type below a given type  $\theta$ , and

$$\chi(\theta) := \min\{\theta' \in \Theta_R | \theta' > \theta\}$$

the closest revealed type above a given type  $\theta$ . We now define the doubly relaxed problem, where we remove all but the downward IC constraints into the closest revealed type and the upwards IC constraints into the next largest revealed type.

$$\max_{\Theta_C \subseteq \Theta} \max_{\{(p_\theta, q_\theta)\}_{\theta \in \Theta}} \sum_{\theta \in \Theta} \mu_\theta (p_\theta - \kappa(q_\theta)) \quad (17)$$

$$\text{s.t. } v_{\theta_h}(q_\theta) - p_\theta \geq 0 \quad \forall \theta \in \Theta_R$$

$$v_{\theta_l}(q_\theta) - p_\theta \geq 0 \quad \forall \theta \in \Theta_C$$

$$v_{\theta_h}(q_\theta) - p_\theta \geq v_{\theta_h}(q_{\eta(\theta)}) - p_{\eta(\theta)} \quad \forall \theta \in \Theta \quad (18)$$

$$v_{\theta_h}(q_\theta) - p_\theta \geq v_{\theta_h}(q_{\chi(\theta)}) - p_{\chi(\theta)} \quad \forall \theta \in \Theta \quad (19)$$

We have the following

**Lemma 6.** *The solution to the doubly relaxed problem satisfies R-monotonicity*

$$\theta, \theta' \in \Theta_R, \theta > \theta' \implies q_\theta \geq q_{\theta'} \quad (20)$$

and solves the relaxed problem.



*Proof.* Consider  $\theta \in \Theta_R$ ,  $\eta := \eta(\theta) < \theta$ . Then  $\theta = \chi(\eta)$  and we have

$$\begin{aligned} v_{\theta_h}(q_\theta) - p_\theta &\geq v_{\theta_h}(q_\eta) - p_\eta \\ v(\eta_h, q_\eta) - p_\eta &\geq v(\eta_h, q_\theta) - p_\theta \end{aligned}$$

and hence

$$\begin{aligned} v_{\theta_h}(q_\theta) - v(\eta_h, q_\theta) &\geq v_{\theta_h}(q_\eta) - v(\eta_h, q_\eta) \\ \int_{\eta_h}^{\theta_h} v_\theta(t, q_\theta) dt &\geq \int_{\eta_h}^{\theta_h} v_\theta(t, q_\eta) dt \end{aligned}$$

which implies  $q_\theta > q_\eta$ , establishing R-monotonicity by transitivity.

Then, we need to show that all IC are implied by the local IC. Let us proceed by induction on the number of types in  $\Theta_R$  between the source of the  $IC_{\theta\theta'}^h$ , constraint  $\theta$  and it's target  $\theta'$ . If there are no revealed types between, then  $\theta' = \eta(\theta)$  (resp.  $\chi(\theta)$ ) and we are done. Suppose that all constraints with up to  $n$  intermediate revealed types are implied and let  $\theta > \theta'$ ,  $\theta' \in \Theta_R$  with  $n+1$  intermediate revealed types. The argument for  $\theta' > \theta$  is identical. Then

$$\begin{aligned} v_{\theta_h}(q_\theta) - p_\theta &\geq v_{\theta_h}(q_{\eta(\theta)}) - p_{\eta(\theta)} \\ &= (v_{\theta_h}(q_{\eta(\theta)}) - v(\eta(\theta)_h, q_{\eta(\theta)})) + v(\eta(\theta)_h, q_{\eta(\theta)}) - p_{\eta(\theta)} \\ &\geq (v_{\theta_h}(q_{\eta(\theta)}) - v(\eta(\theta)_h, q_{\eta(\theta)})) + v(\eta(\theta)_h, q_{\theta'}) - p_{\theta'} \\ &\geq (v_{\theta_h}(q_{\theta'}) - v(\eta(\theta)_h, q_{\theta'})) + v(\eta(\theta)_h, q_{\theta'}) - p_{\theta'} \\ &= v_{\theta_h}(q_{\theta'}) - p_{\theta'} \end{aligned}$$

where we used the local IC, the induction hypothesis and monotonicity. Hence all constraints of (RP) are implied and hence satisfied at the solution to (DRP).  $\triangle$

**Lemma 7.** *In the optimal contract of the relaxed problem the IC from any revealed type  $\theta$  to the closest lower revealed type  $\eta(\theta)$  is active.*

*Proof.* As the relaxed problem and the doubly relaxed problem are equivalent, it is sufficient to show that local downward IC between revealed types are active in the doubly relaxed problem. Suppose towards a contradiction that one of them is not active, say from type  $\theta$  to  $\eta(\theta)$ . Suppose we increase the price in the contract of all revealed types greater than  $\theta$  including  $\theta$  by some  $\epsilon > 0$ . Note that this change doesn't affect any constraints between the affected types. Furthermore,  $\theta$  isn't the lowest revealed type, hence the participation constraint of all revealed type is

implied by the IC and not active since  $IC-\theta \rightarrow \eta(\theta)$  isn't active. As we can pick epsilon sufficiently small, this IC is still slack and we strictly increased revenue, contradiction the optimality of the initial contract.  $\Delta$

**Lemma 8.** *In the optimal contract of the relaxed problem,  $q_\theta \leq \hat{q}_{\theta_h}$  for all  $\theta \in \Theta$ .*

*Proof.* As the relaxed problem and the doubly relaxed problem are equivalent, we can work on the doubly relaxed problem. The result follows from Proposition 4 for concealed types. Suppose towards a contradiction that this property is violated for some subset of revealed types. Pick the smallest revealed type for which this is the case and denote it as  $\theta$ . Note that  $q_\eta(\theta) \leq \hat{q}_{\eta(\theta)_h} < \hat{q}_{\theta_h}$  and denote the rent given to type  $\theta$  as  $\Delta := v(\theta^h, q_{\eta(\theta)}) - p_{\eta(\theta)}$ . (This is the correct expression, because the local downward IC is active by the above Lemma.) Consider the set of contracts where we replaced the initial contract for type  $\theta$  by  $(\hat{q}_{\theta_h}, v_{\theta_h}(\hat{q}_{\theta_h}) - \Delta)$ . As  $\theta$  receives the same utility in both contracts, no participation constraint is violated and all IC from  $\alpha$  are still satisfied. The upward IC  $\eta(\theta) \rightarrow \theta$  is still satisfied as it is implied by R-monotonicity (which is maintained) and the corresponding downward IC. Consider any higher type imitating  $\theta$ . The amended contract gives the same utility to  $\theta$  at a lower quality, hence it gives a strictly lower deviation payoff to higher types. In particular, all IC are satisfied. The revised contract is also more profitable for the principal as the most profitable way to transfer rent to type  $\theta$  in frame h is using quality  $\hat{q}_{\theta_h}$ . Hence, the initial set of contracts wasn't optimal.  $\Delta$

Now we can show that there are no profitable feasible upward deviations in  $e^*$ . We proceed by induction. Order the types such that  $\{\theta^1, \dots, \theta^n\} = \Theta$ ,  $\theta^i < \theta^{i+1}$ . Clearly, the highest type has no feasible upward deviations. Suppose all upward deviations are either infeasible or unprofitable for types  $\theta^i$  into types  $\theta^j$  for  $j > i > m$ . We need to show that the required upward IC constraints out of type  $\theta^m$  are satisfied. We will proceed case by case, in addition showing that the upward IC from concealed to revealed types are always slack:

1. *Deviations into a concealed type with rent  $\Delta_{\theta^i} \leq v(\theta_h^i, \hat{q}_{\theta_h^i}) - v(\theta_l^i, \hat{q}_{\theta_l^i})$ :* Then the participation constraint of type  $\theta^i$  is binding at the intermediate stage in frame l. But by single crossing

$$c_{\theta^i} \sim_{\theta^i} \mathbf{0} \implies c_{\theta^i} \prec_{\theta^m} \mathbf{0}, \quad (21)$$

an imitation is infeasible.

2. *Deviations into a concealed type with rent*  $\Delta_{\theta^i} > v(\theta_h^i, \widehat{q}_{\theta_h^i}) - v(\theta_l^i, \widehat{q}_{\theta_l^i})$ : Note that in this case  $q_{\theta^i} = \widehat{q}_{\theta_h^i}$  and this rent has to be the result of a possible deviation that is discouraged by a constraint of the problem and hence by the induction hypothesis this is a downward deviation into a revealed type. Hence  $\Delta_{\theta^i} = v(\theta_h^i, q_\eta) - p_\eta$  for some  $\eta < \theta^i$ ,  $\eta \in \Theta_R$ . But then the upward deviation isn't profitable unless the deviation into  $\eta$  is profitable, since  $q_\eta \leq \widehat{q}_{\eta_h} \leq \widehat{q}_{\theta_h^i} = q_{\theta^i}$  and by single crossing

$$c_\eta \sim_{\theta_h^i} c_{\theta^i} \implies c_\eta \succ_{\theta_h^m} c_{\theta^i} \quad (22)$$

so all we have to show is that deviations into revealed types are not profitable. If  $\eta < \theta^m$ , this is achieved already by the maintained IC constraints, if  $\theta^m \in \Theta_R$  it is by the upward IC. The case we need to consider are deviations from concealed types upwards into revealed types.

3. *Deviations from a concealed into a revealed type*: Consider a concealed type  $\theta^m$  with a profitable upwards deviation into a revealed type. As the set of types is finite, there has to exist a lowest revealed type into which  $\theta^m$  has a strictly profitable deviation. Furthermore, since we impose downward incentive compatibility constraints, this lowest target type has to be greater than  $\theta^m$ . We will show that such a lower bound cannot exist, hence there can be no profitable upward deviation.

Suppose such a lower bound exists,  $\underline{\theta} = \min\{\theta \in \Theta_R | \theta_h^m q_\theta - p_\theta > \theta_h^m q_{\theta^m} - p_{\theta^m}\}$ . But then, consider type  $\eta(\underline{\theta})$ . A deviation into this type is also strictly profitable since  $c_{\eta(\underline{\theta})} \sim_{\theta_h} c_{\underline{\theta}}$  and by R-monotonicity  $q_{\eta(\underline{\theta})} \leq q_{\underline{\theta}}$ , but then by single crossing  $c_{\eta(\underline{\theta})} \succ_{\theta_h^m} c_{\underline{\theta}} \succ_{\theta_h^m} c_{\theta^m}$ , contradicting the minimality of  $\underline{\theta}$ .

Hence there can be no strictly profitable upward deviation.

And we established that there can be no upward deviation by type  $\theta^m$ . By induction, no type prefers any attainable contract offered to higher types in  $e^*$  and hence we found an EDP that attains the upper bound to the solution of (GP) and therefore (RP)=(GP).  $\square$

*Proof of Proposition 3 on page 19*: Let  $\mathbf{c} = (c_\theta)_\theta = ((p_\theta, q_\theta))_\theta$  be an optimal vector of contracts implemented by some EDP. By Theorem 1, we can construct a standard EDP  $e$  with that implements it. Let  $\Theta_C$  and  $\Theta_R$  be the sets of revealed and concealed types in  $e$ . If  $\theta \in \Theta_C$ , the statement follows from Proposition 4. We proved that  $q_\theta < \widehat{q}_{\theta_h}$  as Lemma 8. Therefore there is only one case left to consider.

Assume that  $\theta \in \Theta_R$  and towards a contradiction that  $q_\theta < \underline{q}_\theta$ , where  $\underline{q}_\theta$  satisfies  $\zeta_{\theta_h}(\underline{q}_\theta) = \zeta_{\theta_l}(\widehat{q}_{\theta_l})$ . Denote the rent in this contract by  $\Delta := v_{\theta_h}(q_\theta) - p_\theta$ .

We will construct a vector of contracts with strictly higher revenue. Starting from the old EDP, we now conceal type  $\theta$  and set the contract  $(\widehat{q}_{\theta_l}, \widehat{p}_{\theta_l} - \Delta)$ . Using the surplus function  $\zeta_\theta$  defined in (8), note that since  $q_\theta < \bar{q}_\theta$

$$\zeta_{\theta_h}(q_\theta) < \zeta_{\theta_l}(\widehat{q}_{\theta_l})$$

and consequently

$$\begin{aligned} \zeta_{\theta_h}(q_\theta) - \Delta &< \zeta_{\theta_l}(\widehat{q}_{\theta_l}) - \Delta \\ p_\theta - \kappa(q_\theta) &< \widehat{p}_{\theta_l} - \Delta - \kappa(\widehat{q}_{\theta_l}) \end{aligned}$$

and the principal receives weakly higher profit in the modified contract.

Clearly, this contract satisfies the participation constraint in frame l and delivers rent greater than  $\Delta$  to type  $\theta$  in the high frame, hence there is no deviation by this type. There is no downward deviation into this contract since the type is concealed. Furthermore, we don't have to worry about upward deviations. The optimal concealed contract – which is strictly better for profit – is never subject to them and we will establish that even a sub-optimal concealed contract delivers an improvement in profits. Hence the original vector was not optimal, a contradiction.  $\square$

*Proof of Proposition 4 on page 20:* Note that there are no IC constraints into a type  $\theta \in \Theta_C$ . Hence we can separate the principals problem and solve for the optimal contract of  $\theta$  in (RP). The contract given to type  $\theta$  solves

$$\begin{aligned} \max_{(p,q)} \quad & p - \kappa(q) \\ \text{s.t.} \quad & v_{\theta_l}(q) - p \geq 0 \\ & v_{\theta_h}(q) - p \geq \Delta \end{aligned}$$

Dropping the second constraint, the optimal contract is  $\widehat{c}_{\theta_l}$ , which delivers rent  $v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l})$ , hence the second constraint is satisfied if

$$\Delta \leq [v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l})]. \quad (23)$$

Similarly, note that the optimal contract dropping the first constraint is  $(v_{\theta_h}(\widehat{q}_{\theta_h}) - \Delta, \widehat{q}_{\theta_h})$ , which gives utility  $v_{\theta_l}(\widehat{q}_{\theta_h}) - v_{\theta_h}(\widehat{q}_{\theta_h}) + \Delta$  in the low frame. Hence the first constraint is satisfied if

$$\Delta \geq v_{\theta_h}(\widehat{q}_{\theta_h}) - v_{\theta_l}(\widehat{q}_{\theta_h}). \quad (24)$$

In the intermediate case, both constraints are binding,

$$\begin{aligned} v_{\theta_l}(q^*) &= p \\ v_{\theta_h}(q^*) - v_{\theta_l}(q^*) &= \Delta \end{aligned}$$

and the optimal contract is  $(v_{\theta_l}(q^*), q^*)$ . Note that  $q^* \in (\hat{q}_{\theta_l}, \hat{q}_{\theta_h})$  by single crossing.  $\square$

*Proof of Proposition 5 on page 21:* Take any type  $\theta \in \Theta$ . For each  $\mu$ , consider (RP) with the constraint  $\theta \in \Theta_C$  ( $\theta \in \Theta_R$ ) and denote the corresponding optimal value by  $\Pi_{C;\mu}^R$  ( $\Pi_{R;\mu}^R$ ). Next, using the surplus function  $\zeta_{\theta_f}$  defined in (8), we can bound those values as

$$\Pi_{R;\mu}^R \geq \mu_\theta \zeta_{\theta_h}(\hat{q}_{\theta_h}) \quad (25)$$

$$\Pi_{C;\mu}^R \leq \mu_\theta \zeta_{\theta_l}(\hat{q}_{\theta_l}) + \sum_{\theta' \neq \theta} \mu_{\theta'} \zeta_{\theta'_h}(\hat{q}_{\theta'_h}) \leq \mu_\theta \zeta_{\theta_l}(\hat{q}_{\theta_l}) + (1 - \mu_\theta) \zeta_{\tilde{\theta}_h}(\hat{q}_{\tilde{\theta}_h}), \quad (26)$$

where  $\tilde{\theta} := \max\{\Theta \setminus \theta\}$ .

Note that Lemma 4 implies that

$$\zeta_{\theta_h}(\hat{q}_{\theta_h}) > \zeta_{\theta_h}(\hat{q}_{\theta_l}) > \zeta_{\theta_l}(\hat{q}_{\theta_l}), \quad (27)$$

and define

$$\bar{\mu}_\theta := \frac{\zeta_{\tilde{\theta}_h}(\hat{q}_{\tilde{\theta}_h})}{\zeta_{\theta_h}(\hat{q}_{\theta_h}) - \zeta_{\theta_l}(\hat{q}_{\theta_l}) + \zeta_{\tilde{\theta}_h}(\hat{q}_{\tilde{\theta}_h})} \in (0, 1).$$

Finally, combining (25), (26) and (27) yields

$$\begin{aligned} \Pi_{R;\mu}^R - \Pi_{C;\mu}^R &\geq \mu_\theta \zeta_{\theta_h}(\hat{q}_{\theta_h}) - \mu_\theta \zeta_{\theta_l}(\hat{q}_{\theta_l}) - (1 - \mu_\theta) \zeta_{\tilde{\theta}_h}(\hat{q}_{\tilde{\theta}_h}) \\ &= \mu_\theta \left[ \zeta_{\theta_h}(\hat{q}_{\theta_h}) - \zeta_{\theta_l}(\hat{q}_{\theta_l}) + \zeta_{\tilde{\theta}_h}(\hat{q}_{\tilde{\theta}_h}) \right] - \zeta_{\tilde{\theta}_h}(\hat{q}_{\tilde{\theta}_h}) \\ &\geq 0. \end{aligned}$$

Therefore, for any  $\mu_\theta \in [\bar{\mu}_\theta, 1]$ , it is optimal to reveal  $\theta$ .<sup>30</sup>

$\square$

*Proof of Proposition 6 on page 22:* It is easy to see that the optimal contract and set of concealed types before the change of valuations is still feasible after the change. Hence  $\Pi_\Theta^* \leq \Pi_{\tilde{\Theta}}^*$ . If  $\tilde{\theta}$  is concealed in the optimum, we are done. Suppose that

<sup>30</sup>This bound is typically not tight, as we introduced slack in (26).

instead it is not concealed. Then, since  $\tilde{\theta}_l$  is the only difference to the initial problem and doesn't affect the constraints unless  $\tilde{\theta}$  is concealed, the optimal contract under  $\tilde{\Theta}$  is feasible under  $\Theta$  and  $\Pi_{\tilde{\Theta}}^* \leq \Pi_{\Theta}^*$ . Hence, the original vector of contracts is still optimal, establishing the claim.  $\square$

*Proof of Proposition 7 on page 22:* Consider the profit from concealing all types except the highest,  $\Pi_C(\epsilon) := \sum_{\theta < \bar{\theta}} \mu_{\theta} v_{\theta_l}(\hat{q}_{\theta_l}) + \mu_{\bar{\theta}} v_{\bar{\theta}_h}(\hat{q}_{\bar{\theta}_h})$ , where we consider the  $\theta_l$  as a function of  $\epsilon$ . It is easy to see that  $\hat{q}_{\theta_l} \rightarrow \hat{q}_{\bar{\theta}_h}$  as  $\theta_l \rightarrow \bar{\theta}_h$ . By continuity of  $v$ ,  $\Pi_C(\epsilon) \rightarrow \Pi((\hat{c}_{\theta_h})_{\theta \in \Theta})$ .

Suppose that in the optimum  $\theta < \bar{\theta}$  is revealed. Then  $q_{\theta} > \underline{q}_{\theta} > 0$  by Proposition 3 and  $p_{\theta} \leq v_{\theta_h}(q_{\theta})$ . But then, by the incentive compatibility constraint of  $\bar{\theta}$ ,  $\Pi < \Pi((\hat{c}_{\theta_h})_{\theta \in \Theta}) - \mu_{\bar{\theta}}(v_{\bar{\theta}_h}(q_{\theta}) - v_{\bar{\theta}_h}(q_{\theta})) < \Pi((\hat{c}_{\theta_h})_{\theta \in \Theta})$ . Hence, there exists an  $\epsilon_{\theta} > 0$  such that  $\Pi_C(\epsilon) > \Pi$  for  $\epsilon < \epsilon_{\theta}$ , so it cannot have been optimal to reveal  $\theta$  for sufficiently small  $\epsilon$ . The result follows by taking the maximum over  $\{\epsilon_{\theta} : \theta \in \Theta \setminus \bar{\theta}\}$ .  $\square$

*Proof of Theorem 3 on page 27:* Let  $e_0$  denote an EDP constructed for sophisticated types in Theorem 1. Order naive types  $\Theta_N = \{\theta^1, \dots, \theta^m\}$  with  $\theta^i < \theta^{i+1}$ . We will construct an optimal EDP for the mixed case inductively.

Starting from  $e_0 = (E_0, h)$ , we add one continuation problem at the root for every naive type,

$$e_{n+1} = \left( \bigcup_{i=0}^{n+1} E_i, h \right). \quad (28)$$

To define  $E_i$ , let the most preferred alternative in  $e_{i-1}$  for type  $\theta^i$  be  $x_i := \arg\max_{e \in (e_{i-1})} u_{\theta^i}$ . During the construction, we ensure that

1. no sophisticated type prefers to continue to  $E_i$ ,
2. no naive type  $\theta^j$  with  $j < i$  prefers to continue to  $E_i$ , and
3. type  $\theta^i$  indeed proceeds to  $E_i$  and chooses  $\hat{c}_{\theta^i}$  eventually.

If we ensure this during our construction, all sophisticated types choose as in  $e_0$  and all naive types choose their efficient contract  $\hat{c}_{\theta^i}$  and we establish the theorem.

Let  $E_i = \left\{ \left( \left( \{N_i, \{d_{N,i}^{\theta'}\}_{\theta' > \theta^i: \theta' \in \Theta_S}, \mathbf{0}\}, h \right), \left( \hat{c}_{\theta^i}, \{d_i^{\theta'}\}_{\theta' > \theta^i: \theta' \in \Theta_S}, \mathbf{0}\}, h \right), \mathbf{0} \right\}, l \right\}$ .

We now have to specify  $N_i$  and the decoys and verify 1-3 above. First, use the

mapping from Lemma 5 to set  $N_i := \mathcal{D}(x, (\theta_{i-1}, \theta^i))$  so that

$$N_i \sim_{\theta_h^i} x_i \quad (29)$$

$$q_{N_i} \geq q_{x_i} \quad (30)$$

$$N_i \preceq_{\theta_h^i} \mathbf{0} \quad (31)$$

Second, define the decoys as

$$\begin{aligned} (d_{N,i}^{\theta'})_{\theta' > \theta^i: \theta' \in \Theta_S} &:= \mathcal{D}(N_i, (\theta^i, \Theta_S^{>\theta^i})), \\ (d_i^{\theta'})_{\theta' > \theta^i: \theta' \in \Theta_S} &:= \mathcal{D}(\widehat{c}_{\theta_h^i}, (\theta^i, \Theta_S^{>\theta^i})), \end{aligned}$$

where  $\Theta_S^{>\theta^i}$  is a vector of types in  $\Theta_S$  that are strictly greater than  $\theta^i$ . By construction, every sophisticated type  $\theta > \theta^i$  prefers the outside option to the contract chosen from the continuation problems. Hence they have no incentive to enter. Furthermore, all contracts are, by construction, worse in frame  $l$  than the outside option for all types  $\theta < \theta^i$ , hence lower sophisticated types have no incentive to enter. Hence, we established 1.

By construction,  $E_i$  contains a most preferred option for  $\theta_h^i$ , hence continuing into  $E_i$  is part of a naive outcome for  $\theta^i$ . At the subsequent decision node, the decision problem containing  $N_i$  is as attractive as the outside option: By the construction of the decoys,  $N_i \succ_{\theta_h^i} d_{N,i}^{\theta'}$  and  $q_{N_i} \leq q_{d_{N,i}^{\theta'}}$  and hence  $N_i \succ_{\theta_h^i} d_{N,i}^{\theta'}$ . But  $N_i \preceq_{\theta_h^i} \mathbf{0}$ . As the decision problem containing  $\widehat{c}_{\theta_h^i}$  also contains the outside option, continuing to this menu is part of a naive outcome. From the menu  $(\{\widehat{c}_{\theta_h^i}, \{d_i^{\theta'}\}_{\theta' > \theta^i: \theta' \in \Theta_S}, \mathbf{0}\}, h)$ , the DM chooses  $\widehat{c}_{\theta_h^i}$  by the construction of the decoys. This establishes 3.

To see 2, note that all decoys have higher quality than  $N_i$  and  $\widehat{c}_{\theta_h^i}$ , respectively, and are less preferred according to  $\theta_h^i$ . Hence, they are less preferred by lower naive types  $\theta_h^j$  by single crossing. Furthermore,  $\widehat{c}_{\theta_h^i}$  is not attractive to lower naive types, as it is worse than the outside option. It remains to check whether  $N_i$  is attractive. But note that  $N_i \sim_{\theta_h^i} x_i$  and  $q_{N_i} \geq q_{x_i}$  imply  $N_i \preceq_{\theta_h^j} x_i$  for all  $j < i$ . By the induction hypothesis,  $N_j \succ_{\theta_h^j} \operatorname{argmax}_{c \in (e_{i-1})} u_{\theta_h^j} \succ_{\theta_h^j} x_i \succ_{\theta_h^j} N_i$ . Consequently,  $N_i$  is not attractive to lower naive types, and there is a naive outcome where types  $\theta^j < \theta^i$  choose  $E_j$ .

Clearly, the contract implemented for naive types is optimal given the participation constraint in the high frame any implemented contract needs to satisfy. Furthermore, suppose there is an EDP implementing contracts for sophisticated types that are not implemented by an optimal EDP in Theorem 1. Then the contracts

don't solve (RP), so we can find a strictly better set of contracts and use the above construction. Hence every optimal EDP in (7) satisfies Theorem 3. From that, the decomposition theorem is immediate.

If there is an ex-post participation constraint, any naive outcome needs to satisfy  $v(\theta) \succ_{\theta_n} 0$ . The revenue maximal vector of contracts satisfying these constraints is  $(\widehat{c}_{\theta_n})_{\theta \in \Theta_N}$ . It is immediate from the proof of Theorem 3 that this set of contracts can be implemented using an analogous construction.

The outside option trivially satisfies the participation constraint in every frame, so continuing at the root is always part of a valid naive strategy in the interim modification of any extensive-form decision problem. Hence,  $N^e \subseteq N^{\bar{e}}$ , which establishes the observation.  $\square$

*Proof of Observation 3 on page 28:* Let us denote the contract for type  $\theta$  in the sophisticated problem as  $c_\theta^s$  and note that the contract in the naive problem is  $\widehat{c}_{\theta_h}$ . Note that  $c_\theta^s \succ_{\theta_h} 0 \sim \widehat{c}_{\theta_h}$  and  $q_\theta^s \leq \widehat{q}_{\theta_h}$ . Hence by single crossing  $c_\theta^s \succ_{\theta_f} \widehat{c}_{\theta_h}$ , strictly for  $f \neq h$  if  $c_\theta^s \neq \widehat{c}_{\theta_h}$ .  $\square$

*Proof of Observation 4 on page 30:* To implement the vector of contracts  $(\widehat{c}_{\theta_n})_{\theta \in \Theta}$ , the principal can simply conceal *all* types using neutral frame  $n$ .

Notice that  $(\widehat{c}_{\theta_n})_{\theta \in \Theta}$  satisfies all the constraints of (RP) for  $\Theta_R = \emptyset, \Theta_C = \Theta$ . Therefore, by Theorem 2, there exists a standard EDP  $e^*$  that implements it. Notice that since the contract  $\widehat{c}_{\theta_n}$  for type  $\theta$  satisfy  $P_\theta^n$ , the interim and ex-post modifications  $\bar{e}^*$  and  $\underline{e}^*$  also implement  $(\widehat{c}_{\theta_n})_{\theta \in \Theta}$ .  $\square$

*Proof of Observation 5 on page 31:* First, consider the ex-post modification. To implement  $(\widehat{c}_{\theta_n})_{\theta \in \Theta}$ , the principal can use the same construction as in Theorem 3, but with  $n$ . Notice that because any contract that is implemented with ex-post participation constraints must satisfy those constraints, the principal cannot do better.

Second, consider the interim modification. Suppose that the optimal EDP without the modification is  $e^*$  and notice that  $e^*$  implements  $(\widehat{c}_{\theta_h})_{\theta \in \Theta}$ . Now consider its interim modification  $\bar{e}^*$ . Since naive consumers think they would get a better option than 0, they would proceed to  $e^*$ . Therefore, there exists a naive solution  $\bar{v}$  to  $\bar{e}^*$ , such that  $\bar{v}_\theta = \widehat{c}_{\theta_h}$  for all  $\theta$ .  $\square$