Screening with Frames*

Franz Ostrizek[†] Denis Shishkin[‡]

December 3, 2018

Abstract

We analyze screening with frame-dependent valuations. A monopolist firm designs an extensive-form decision problem with frames at each stage. Consumers are sophisticated: they correctly anticipate their future choices, but lack commitment. The optimal extensive form has a simple three-stage structure: After an initial "sales pitch", consumers are given time to consider the offer and then presented with an extended tailored menu if they come back. Framing is used by the principal to create and exploit dynamic inconsistency to reduce information rents. If there are both sophisticated and naive consumers, the principal can perfectly screen by cognitive type and extract full surplus from naifs.

JEL Codes: D42, D84, L12, M31

Keywords: Screening, Framing, Extensive-Form Decision Problems, 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 and participants of the RSF Summer Institute 2018. Ostrizek gratefully acknowledges financial support from the Philip Bennett *79 Public Policy and Finance Fund.

[†]Department of Economics, Princeton University, ostrizek@princeton.edu

[‡]Department of Economics, Princeton University, shishkin@princeton.edu

1 Introduction

Ample evidence, casual empiricism and introspection suggest that framing effects are common in choice.[1] In particular, the way a product is presented and the setting of the sales interaction can have a strong impact on consumer valuations.[2]

Most of the literature focuses on framing in static decision situations.[3] 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, comtemplates his purchasing 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 monopolist designs not only the contracts, but also the structure of the sequential decision problem along with a frame at each stage. The interaction of framing and extensive forms provides a novel mechanism for surplus extraction, whether consumers are sophisticated or not. The monopolist induces changing tastes by varying framing throughout the interaction. This allows her to exploit dynamic inconsistency to reduce information rents.

We characterize the structure of an optimal extensive-form decision problem (EDP) with quasi-linear single crossing utility and any finite number of types and frames under regularity conditions. In particular, this EDP has the following key features:

1. *Short Interaction.* Every type interacts *at most three times* with the principal.

[1]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).

[2]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). Concordantly, many firms go to great expenses improving the presentation of their product in largely non-informative and payoff-irrelevant ways through packaging, in store presentation, the design of the sales pitch. 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).

[3]As a result, we might expect that a firm wants to present their product in the best light and quickly close the sale. This is, for example, how Loewenstein et al. (2003) explains the payoff from one-click shopping. This view is also reflected in legislation, mandating a right to return goods and cancel contracts, especially when the sale happened under pressure (e.g. door to door), eg. Directive 85/577/EEC in the EU.

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" (lower-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 faces a trade-off for each sophisticated consumer between maximizing the valuation by using only the highest frame and reducing information rents by using frames in a high-low-high pattern to induce dynamic inconsistency. There is no such trade-off for naive consumers.

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

Example 1. There are two equally likely types θ^1 (low) and θ^2 (high) and two frames, low and high. Preferences of type θ^i in frame $f \in \{l, h\}$ are represented by $u^i_f(p, q) = \theta^i_f q - p$, where the marginal utility θ^i_f 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 θ^1_h's participation constraint and θ^2_h's incentive compatibility constraint bind, which yields a profit of 20.[4]

The principal can do better. Consider the allocation that would arise if the 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 θ^1_l, $c^1 = (9, 3)$, and θ^2_h, $c^2 = (36, 6)$, obtaining a profit of $\Pi_{\text{extensive}} = 22.5 > 20$. We show that this is indeed possible in an EDP by varying the frames: $h \to l \to h$.

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

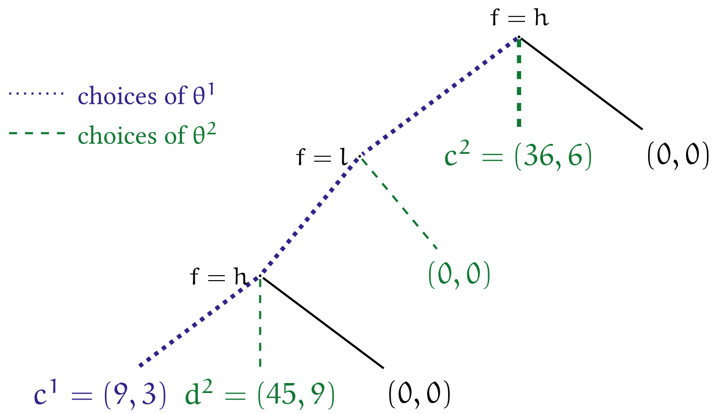
[4]In particular, the optimal contracts are $(p^1, q^1) = (8, 2)$ and $(p^2, q^2) = (8, 6)$. Note that with these functional forms, $q^i = \theta^i_f$ is efficient for frame f and the quality of type 2 is distorted downward compared to the efficient quality for *both* frames.

Figure 1: Example 1

(a) Payoff types.

	$f = l$	$f = h$
θ^1	3	4
θ^2	5	6

(b) The optimal extensive-form decision problem implementing c^1, c^2.



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. 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.

▲

For the main sections of this paper, we assume that consumers are sophisticated about the effect of framing: They correctly anticipate their choices, but cannot commit ex-ante to a course of action.[5] Consumers are exposed to sales pitches on a daily basis and, having had time and experience to learn, can correctly anticipate the flow of the interaction. Moreover, the optimal sales interaction has a simple, interpretable structure. Correctly anticipating behavior in this extensive form is relatively simple and requires little more than anticipating that you are more prone to choose expensive premium options when under pressure from the sales person. In addition, sophistication serves as a benchmark, by making it difficult

[5] Sophistication is a common modeling choice in the domain of time preference following the seminal work of Strotz (1955); Laibson (1997). Decision makers avoid choice situations they expect to lead to suboptimal decisions, e.g time inconsistent decision makers demand commitment (Ashraf et al., 2006).

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 radically changes the sales interaction and its outcomes.

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

Beyond the narrow 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. As the use of decoys to reduce information rents only relies on the presence of choice reversals and not on their source, our results hint at a general way to extract rents in these settings. Framing is a convenient framework to study, as it allows the principal to design the structure of taste changes. For the main sections, we focus on a literal interpretation of framing and the design of a sequential sales interaction and return to the more general perspective in the discussion.

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-quantity 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 and discuss consumer protection regulation. We conclude with discussions. Proofs are collected in the Appendix.

Related Literature

A growing literature in behavioral industrial organization 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. They restrict the principal to offer 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:[6] 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 this setting.

Our article is also related to behavioral contract theory more generally (for a recent survey, see Kőszegi, 2014), 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).[7] Off-path options that remain unchosen by every type ("decoys") are important in many of these contributions and central in our paper. Galperti (2015) shows that the commitment device for relatively consistent types contains unused options. These options make deviations less attractive and are thus analogous to the decoy contracts we introduce in the optimal extensive form for sophisticated agents. Heidhues and Kőszegi (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.

These papers consider situations when the preference reversals are *given* by the preferences of the agents (e.g. Gul and Pesendorfer (2001) or β-δ) and consequently consider the choice from a very restricted class of extensive forms as induced by the natural time structure of the problem, typically 2-stage. We study how a principal uses a behavioral choice pattern to optimally *induce* dynamic inconsistency to screen agents and designs the optimal EDP. Furthermore, we show that a only slightly more general class – 3-stage EDPs – are sufficient to achieve the optimum in our setting.

Glazer and Rubinstein (2012, 2014) consider models where the principal designs a pro-

[6]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 "basic" product that has to be offered and a less directly related model, 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.

[7]Eliaz and Spiegler (2006, 2008); Heidhues and Kőszegi (2017) 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.

cedure 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 (Kőszegi and Szeidl, 2013; Bordalo et al., 2013). We consider the case of framing through additional cues added to the choice situation, such as the sales pitch or the presentation format focusing consumer attention on quality. Thus, consumers in our model fit into the choice with frames framework of Salant and Rubinstein (2008), a model where frames are formally exogenous. Nevertheless, we expect similar constructions to be relevant in the case with context effects, as they mainly depend on the resulting failure of dynamic consistency.

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. de Clippel (2014) studies Nash implementation with general choice correspondences.

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 $q \geqslant 0$, 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 $|F| \geqslant 2$ and a finite type space Θ endowed with a full support prior μ. Each type is a function $\theta : F \to \mathbb{R}$ that maps frames into payoff types, denoted as $\theta_f \equiv \theta(f)$. For a given payoff type θ_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

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

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

$$u_{\theta_f}(c_\theta) \geqslant 0, \text{ and} \tag{P^f_θ}$$

$$u_{\theta_f}(c_\theta) \geqslant u_{\theta_f}(c_{\theta'}) \tag{$IC^f_{\theta\theta'}$}$$

as the participation constraint for θ and the incentive compatibility constraint (IC) from θ to θ' 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* $\theta, \theta' \in \Theta$*,* $f, f' \in F$ *:*

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

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, frames have 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 abut 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.

Formally, we define extensive-form decision problems (EDPs) by induction. For any S let $\mathcal{P}(S)$ denote the set of all finite subsets of S containing the outside option $\mathbf{0} := (0, 0)$. Call an extensive decision problem with k stages a k-EDP. A 1-EDP $e = (A, f)$ is a pair of a finite menu A and a frame f. The set of all 1-EDPs is $\mathcal{E}^1 := \mathcal{P}(\mathcal{C}) \times F$. Next, for each $k \geqslant 1$, a $k+1$-EDP $e = (E, f)$ is a pair of a finite set E of continuation problems and a frame f. The set of k-EDPs is $\mathcal{E}^k := \mathcal{P}(\cup_{l=0}^{k} \mathcal{E}^l) \times F$, where $\mathcal{E}^0 = \mathcal{C}$ to allow for terminal choices at all stages. Finally, the set $\mathcal{E}$ of all finite EDPs is given by

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

By construction, at each stage the outside option is available for consumers.

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 θ 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 e we assume that the consumers are *sophisticated.* Presented with a choice between several decision problems, the consumer correctly anticipates the corresponding frames and choices, and chooses the continuation problem according to her current frame. The current self has no commitment power other then the choice of a suitable continuation problem.

Formally, we define the sophisticated consumer's choice in an EDP by induction. Call $\sigma \in \mathcal{C}^\Theta$ an *outcome* to a 1-EDP $e = (A, f) \in \mathcal{E}^1$ if $\sigma(\theta)$ maximizes u_{θ_f} on A. Suppose the consumer is facing $e = (E, f) \in \mathcal{E}^k$. Choosing between EDPs in E, she anticipates her choice $\sigma^j(\theta) \in \mathcal{C}$ in each EDP $e^j \in E$, but evaluates the contracts $\{\sigma^j(\theta)\}_{e_j \in E}$ in the current frame f. Let Σ^e be the set of outcomes to an EDP $e \in \cup_{l=1}^{k-1} \mathcal{E}^l$. Then $\sigma \in \mathcal{C}^\Theta$ is an outcome to an EDP $e = (E, f) \in \mathcal{E}^k$ if for all $\theta \in \Theta$

$$\sigma(\theta) \in \underset{\{\sigma^j(\theta)\}_{e^j \in E}}{\operatorname{argmax}} \; u_{\theta_f}(\sigma^j(\theta))$$

with $\sigma^j \in \Sigma^{e^j}$.

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 $\forall \theta_f \in \mathbb{R}$, $\nu'_{\theta_f}(0) - \kappa'(0) > 0$, $\lim_{q\to\infty} \nu'_{\theta_f}(q) - \kappa'(q) < 0$.

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

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

Finally, the firm designs an EDP to maximize profits

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

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.[8]

We denote the efficient quantity for a payoff parameter θ_f by $\widehat{q}_{\theta_f}$ with

$$\nu'_{\theta_f}(\widehat{q}_{\theta_f}) = \kappa'(\widehat{q}_{\theta_f}).$$

The efficient quantity is unique, positive and strictly increasing in the payoff parameter by our assumptions on ν and κ. We denote the contract offering this quantity and extracting all surplus from the corresponding payoff type by $\widehat{c}_{\theta_f} := (\nu_{\theta_f}(\widehat{q}_{\theta_f}), \widehat{q}_{\theta_f})$.

[8]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 $\{\theta_h\}_{\theta\in\Theta}$.[9]

Observation 1. *Let* $\mathbf{c}^*$ *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* $(0 \cup \{c^*_\theta\}_{\theta\in\Theta}, h)$ *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:

$$h := \max F,$$
$$l := \max F \setminus \{l\}.$$

Definition 1. An EDP e is a *standard* EDP for a vector of contracts $\mathbf{c}$ if there exists a

[9]This is in contrast to (Salant and Siegel, 2018), where there is an ex-post participation constraint in an exogenously given frame, i.e. a modicum of extensive-form structure, or a default option, i.e. a restricted menu choice problem for the principal.

partition $\{\Theta_C, \Theta_R\}$ of Θ, and decoy contracts $\{d^{\theta}_{\theta'}\}_{\theta\in\Theta_C,\theta'>\theta}$, such that

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

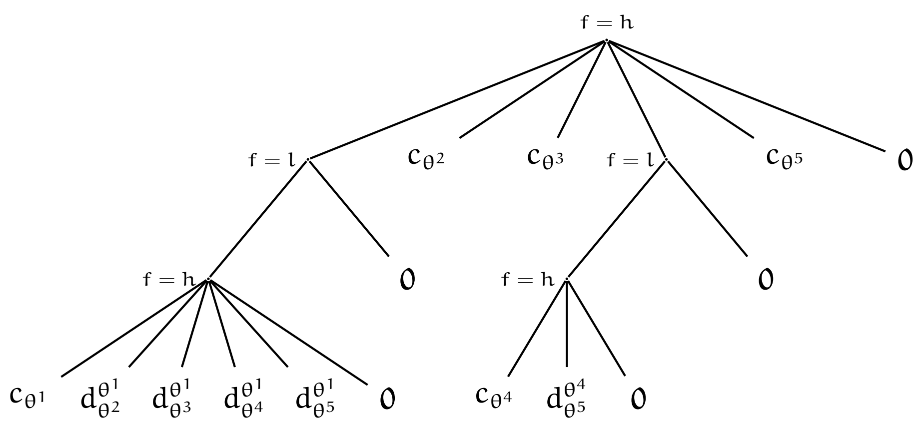
$$e_\theta = \left(\left\{\left(\{c_\theta, 0\} \cup \{d^{\theta}_{\theta'}\}_{\theta'>\theta}, h\right), 0\right\}, l\right), \forall\theta \in \Theta_C. \tag{2}$$

Put differently, for any standard EDP the set of types Θ is partitioned into two sets, as there are two ways to present the contract associated to a given type: Contracts c_θ 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_θ. Then, the three stages are (see Fig. 2):

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

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

Figure 2: A standard EDP for $(c_{\theta^1}, \ldots, c_{\theta^5})$ 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 $\mathbf{c}$.[10]

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

[10]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 $\mathbf{c}$.

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 sales person 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 sales person 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. Fianlly, only the highest two are used in the optimal EDP in order 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 e 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 θ there are two possibilities. If there exists a path from the root to c_θ with all decision nodes set in f_R, then θ is called revealed. Alternatively, if every path from the root to c_θ involves at leasts one $f \neq f_R$, then θ is called concealed. As usual, we will denote the sets of revealed and concealed types by Θ_R and Θ_C, respectively.

First, consider participation constraints. If the path from the root to c_θ passes through a node in frame f, then, since the outside option is always available, c_θ needs to satisfy the corresponding participation constraint P_θ^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 θ is revealed, c_θ 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 θ', c_θ must not be an attractive deviation:

$$u_{\theta'_{f_R}}(c_{\theta'}) \geqslant u_{\theta'_{f_R}}(c_\theta) \quad \forall\theta' \in \Theta. \tag{$IC_{\theta\theta'}^{f_R}$}$$

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

Note that we did not identify a single set of necessary conditions, but a family of such sets, indexed by the set of concealed types. Each implementable vector of contracts needs to satisfy at least one of these sets of constraints. The argument so far also allowed for arbitrary frames $f_R, \{f_\theta\}_{\theta\in\Theta}$. While the change of frame affects IC, the exact frame only matters for participation. Consequently, it is easy to show that we can assume that $f_R = h$ and $f_\theta = l$ for all concealed types as the two highest frames correspond to the least restrictive participation constraints.[11]

Proposition 1. *If* $\mathbf{c}$ *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* Θ.

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 total surplus. For concealed types, the participation constraint needs to be satisfied in the second highest frame. This reduces the total surplus from the interaction. The principal is compensated for this reduction through the removal of IC constraints into concealed types.

[11] Recall that $h, l \in F$ denote the highest and second highest frame. See the proof of Proposition 1 in Appendix A.2 for a full argument.

Sufficient Conditions for Implementation by Standard EDPs

Consider a vector of contracts $\mathbf{c}$. To construct a standard EDP that implements it, we first need to determine the set of concealed types and then construct decoys for the continuation problems of these types. Clearly, θ can be concealed, only if c_θ satisfies the participation constraint P^l_θ. Otherwise, θ would prefer to opt out when put into frame l en route to its contract, a contradiction with implementation.

If θ is concealed, the principal can design decoys in order to make *some* deviations to c_θ impossible. Whereas decoys cannot rule out all upward deviations, they can rule out all *downward* deviations into c_θ. 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 were attractive to imitate can be concealed.

Proposition 2. *If* $\mathbf{c}$ *satisfies the constraints* $\{P^h_\theta\}_{\theta\in\Theta_R}, \{P^l_\theta\}_{\theta\in\Theta_C}, \{IC^h_{\theta\theta'}\}_{\theta<\theta'}$, *and* $\{IC^h_{\theta\theta'}\}_{\theta\in\Theta,\theta'\in\Theta_R}$ *for some partition* $\{\Theta_C, \Theta_R\}$ *of* Θ, *then* $\mathbf{c}$ *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 this part of the proof of Lemma 1 in the text. The construction ensures that (ii) if θ is concealed, no type $\theta' > \theta$ can imitate θ. 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, θ chooses the intended contract (iii).

Lemma 1 (Decoy Construction)**.** *For any* $\theta \in \Theta$, c_θ *satisfies* P^l_θ *if and only if there exist decoys* $(d^\theta_{\theta'})_{\theta'>\theta}$, *such that the corresponding* e_θ *in* (2) *has an outcome* σ *that satisfies*

(i) $\sigma(\theta') \in \{0, c_\theta\}$ *for all* $\theta' \in \Theta$,

(ii) $\sigma(\theta') = 0$ *for all* $\theta' > \theta$, *and*

(iii) $\sigma(\theta) = c_\theta$.

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

$$u_{\theta^1_l}(0) = u_{\theta^1_l}(d^\theta_{\theta^1}). \tag{3}$$

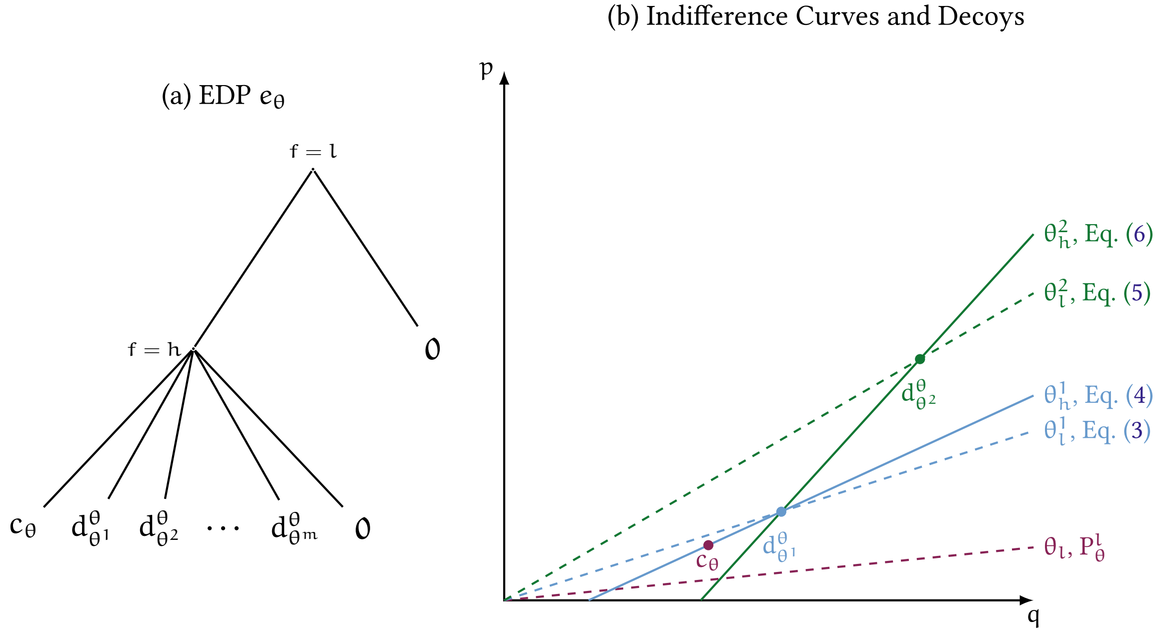
$$u_{\theta^1_h}(c_\theta) = u_{\theta^1_h}(d^\theta_{\theta^1}) \tag{4}$$

Then, decoy $d^\theta_{\theta^2}$ for the next type θ^2 solves

$$u_{\theta^2_l}(0) = u_{\theta^2_l}(d^\theta_{\theta^2}). \tag{5}$$

$$u_{\theta^2_h}(d^\theta_{\theta^1}) = u_{\theta^2_h}(d^\theta_{\theta^2}) \tag{6}$$

Figure 3: The construction of e_θ.



Proceeding by induction, we construct decoys for all $\theta' > \theta$. Now we define an outcome σ as follows. The single-crossing property ensures that each type $\theta' \geqslant \theta$ chooses their corresponding (decoy) contract out of the menu $\{c_\theta, d^\theta_{\theta_1}, \dots, d^\theta_{\theta_m}, 0\}$ in frame h. At the root, θ will choose its contract since it satisfies P^l_θ 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. □

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 conditional on the optimal set of concealed types.

An Equivalent Problem in Price-Quantity 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.

$$\Pi^R = \max_{\Theta_C \subseteq \Theta} \max_{(c_\theta)_{\theta\in\Theta}} \Pi(\mathbf{c}) \tag{RP}$$

$$\text{s.t.} \quad u_{\theta_h}(c_\theta) \geqslant 0, \quad \forall \theta \in \Theta_R := \Theta \setminus \Theta_C \tag{P^h_θ}$$

$$u_{\theta_l}(c_\theta) \geqslant 0, \quad \forall \theta \in \Theta_C \tag{P^l_θ}$$

$$u_{\theta_h}(c_\theta) \geqslant u_{\theta_h}(c_{\theta'}), \quad \forall \theta \in \Theta, \theta' \in \Theta_R \tag{$IC^h_{\theta\theta'}$}$$

It's important to reiterate the difference between the sufficient conditions for implementability in Proposition 2 and the necessary conditions in Proposition 1. In both cases, we identify concealed types for which we require a tighter participation constraint. extensive-form structure allows the principal to remove IC stemming from the imitation of these concealed types. The construction of decoys for standard EDPs in Lemma 1 only rules out downward deviations, we still require that upward IC into concealed types are satisfied. This is not the case for the necessary conditions: We cannot rule out that a more intricate EDP not fitting into the standard family can also eliminate upward deviations.

Theorem 1 shows that the capability of standard EDPs to discourage downward deviations is sufficient to attain the optimum. Consequently, the optimal vector of contracts satisfying the necessary conditions is implementable.

Theorem 2. *A pair* $(\Theta_C, \mathbf{c})$ *solves* (RP) *if and only if* $\mathbf{c}$ *solves* (GP). *Moreover, such a solution exits and* $\mathbf{c}$ *can be implemented by a standard EDP with a set of concealed types* Θ_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.[12] 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-quantity pairs, one for each potential set of concealed types and compare the attained values to find the optimum.[13] 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.

[12]It is not true that *every* vector of contracts satisfying the necessary conditions can be implemented, optimality of the contracts is a key step in the proof. However, the optimal contracts for any set of concealed types - even a suboptimal set - can be implemented.

[13]Indeed, the problem can be further simplified by noting that only local IC - those into the nearest revealed types - are binding. See Appendix A.2.

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_θ, q_θ) *for a type* θ *satisfies* $0 < \underline{q}_\theta \leqslant q_\theta \leqslant \widehat{q}_{\theta_h}$*, where* $v_{\theta_h}(\underline{q}_\theta) - \kappa(\underline{q}_\theta) := v_{\theta_l}(\widehat{q}_{\theta_l}) - \kappa(\widehat{q}_{\theta_l})$*. In particular, every type of consumer buys positive quantity.*

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 quantity and while extracting revenue from this type. Concealing a type consequently strictly dominates a crude shut-down.

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 a 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 quantity in the lower frame, $\widehat{q}_{\theta_l}$. 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.

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

$$(p_\theta, q_\theta) = \begin{cases} \widehat{c}_{\theta_l}, & \text{if } \Delta \leqslant v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l}), \\ (v_{\theta_l}(q^\star), q^\star), & \text{if } \Delta \in [v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l}), v_{\theta_h}(\widehat{q}_{\theta_h}) - v_{\theta_l}(\widehat{q}_{\theta_h})], \\ (v_{\theta_h}(\widehat{q}_{\theta_h}) - \Delta, \widehat{q}_{\theta_h}), & \text{if } \Delta \geqslant v_{\theta_h}(\widehat{q}_{\theta_h}) - v_{\theta_l}(\widehat{q}_{\theta_h}), \end{cases}$$

where $q^\star$ *solves* $v_{\theta_h}(q^\star) - v_{\theta_l}(q^\star) = \Delta$*,* $\Delta := \operatorname{argmax}_{\theta' \in \Theta_R} u_{\theta_h}(c_{\theta'})$ *denotes the rent delivered to type* $\theta \in \Theta_C$*, and* $\mathbf{c}$ *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 trade-off between rent extraction and concealing a type. Concealing a type reduces the information rent for other types, but it comes at a cost. From the perspective of the high frame, a concealed type always receives at least the minimum rent $\nu_{\theta_h}(\widehat{q}_{\theta_l}) - \nu_{\theta_l}(\widehat{q}_{\theta_l})$. This rent reduces the payoff of the principal in addition to the reduction in quality. These costs of concealing a type are decreasing in the information rent Δ. If it is sufficiently high (in the third regime of (4)), it is cost-less to conceal the type.

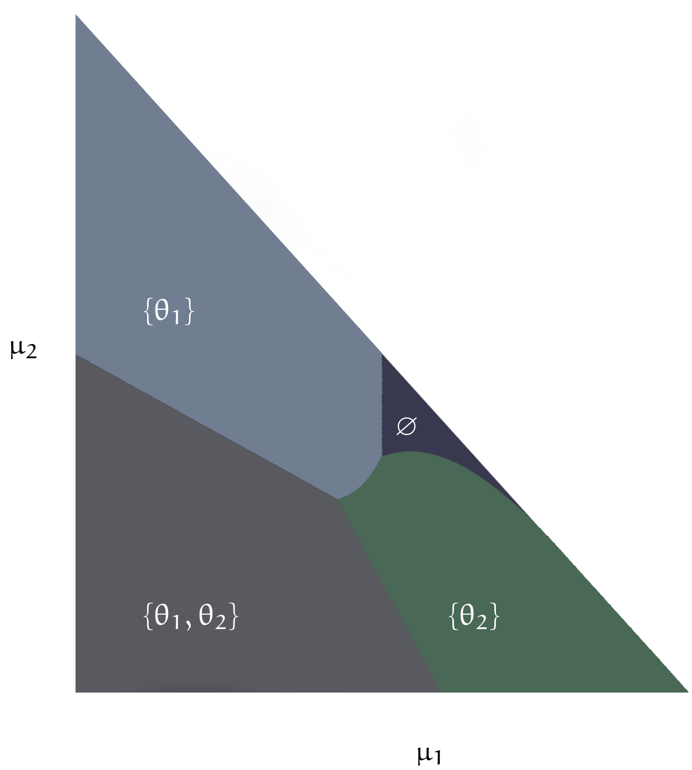
3.4 Optimal Concealed Types

So far, we characterized the structure of the optimal extensive form and – given the set of concealed types – the optimal contracts. Making general statements about the set of concealed types is more difficult.

One simple restriction is that concealing the highest type is never strictly optimal. Types are concealed in order to eliminate downward deviations into them. This is clearly not a concern for the highest type.

Observation 2. *Suppose* $(\Theta_C^*, \mathbf{c}^*)$ *solves* (RP) *and* $\overline{\theta} = \max\Theta \in \Theta_C^*$. *Then* $(\Theta_C^* \setminus \overline{\theta}, \mathbf{c}^*)$ *also solves* (RP).

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



Except for this restriction, we strongly conjecture that any set of types is optimally hidden for some set of probabilities. We illustrate this in an example with three types. In Fig. 4 we plot the regions of the probability simplex where particular sets of concealed types are optimal. Note that all four possibilities are realized for the same set of types, simply varying the probabilities.

Several conjectures one may formulate are rebutted by the example. First, the classic restriction to monotone virtual values that ensures monotonicity in the classic screening model doesn't rule out any configuration. Second, as concealing types is partially analogous to shut-down one may conjecture that it is always low types that are concealed. This is generally not correct. As we discussed in Proposition 4, concealing a type is more costly if it has a low information rent. Consequently, concealing only intermediate types may be optimal.

Loosely speaking, the concealed types are substitutes for the principal. Consider some type θ. By concealing a lower type, the principal reduces the rent θ obtains, increasing the costs of concealing θ. In addition, a lower rent implies that the gain from hiding θ is lower as well. The argument is similar for concealing a higher type. This pattern of substitutability is also illustrated in Fig. 4 as the regions $\Theta_C = \{\theta_1\}$ and $\Theta_C = \{\theta_2\}$ touch.

Still, we can give partial results about the set of concealed types and its' comparative statics.

Sufficiently Likely Types Are Revealed

Concealing a given type has two effects on profits. On the one hand, a tighter participation constraint decreases the profits extracted from this type. On the other hand, it decreases information rents of other types. As the next proposition shows, the former effect on total profits dominates the latter one when a type is sufficiently likely. Sufficiently likely types are always revealed.

Proposition 5. *For any type θ 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 θ.*

This proposition suggest 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 consumer types and available only on order. Note that the proposition

Monotonicity in θ_l

If a type is concealed and its low frame valuation goes up, the gain from concealing this type goes up. The following proposition shows that in this case the set of concealed types remains the same.

Proposition 6. *Let $\Theta_C = \{\theta^i\}_{i=1}^n$ and Θ_R be optimal sets of concealed and revealed types for Θ and $\tilde{\Theta}_C = \{\tilde{\theta}^i\}_{i=1}^n$ is such that $\forall i : \tilde{\theta}_l^i \geqslant \theta_l^i, \tilde{\theta}_h^i = \theta_h^i$. Then, for the set of types $\tilde{\Theta}_C \cup \Theta_R$ there exists a solution of the principal's problem RP with the set of concealed types $\tilde{\Theta}_C$.*

3.5 Discussion

Consumer Commitment and Direct Mechanisms We assume that consumers are sophisticated but lack commitment. This is crucial, as the power of the principal to relax IC constraints by concealing types relies on this failure of commitment. If consumers could commit to a path through the extensive form when choosing at the root, they would imitate lower types and could choose decoys. 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 the dynamic inconsistency to play out. 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, a direct approach is more convenient.

Principal Commitment The principal requires commitment. 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 decision problem. If consumers anticipate this, the extensive-form decision problem unravels.[14]

Assumptions 1 and 2: Relevant Frames and Comonotonicity Our assumptions can be relaxed at the cost of tractability. We assume that all frames can be ordered by their impact on the value of quality. This assumption can be relaxed as long as there is an unambiguously highest and second highest frame.[15] We also use the second highest frame only in order to eliminate incentive compatibility constraints from the problem. Consequently, we do not require that the lowest type is sensitive to framing at all, since there are now downward IC constraints from the lowest type. Hence, the second highest frame for the lowest type need not coincide with the second highest frame for all other types.[16]

We can also relax the assumption of comonotonic types. First, we only require comonotonicity in the two highest frames. Second, suppose two types are not comonotonic, i.e. $\theta^1_l < \theta^2_l < \theta^2_h < \theta^1_h$. This only creates a problem for the construction if type θ^1 is concealed.[17] If a type is sufficiently common that it is not profitable to conceal it, we can relax

[14]Characterizing the optimal decision problem without commitment is beyond the scope of this paper.

[15]This requires that there are two frames that satisfy comonotonicity with all other frames so they can be compared to all of them and that they induce higher valuations than all other frames. We do not require that the dominated frames are comparable among themselves.

[16]This generalizes if the optimal EDP doesn't conceal low types: We require framing effects and a well defined second highest frame only for types higher than the lowest concealed type.

[17]In this case, θ^1 receives a contract that satisfies the participation constraint in the low frame and hence gives a rent to θ^2. We cannot discourage this deviation. More precicely, it is impossible using the structure of

comononicity with this type.

Participation Constraint 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. An alternative specification 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.

Strength of Framing The gain for the principal from using different frames comes through the induced violation of dynamic consistency. The intensity of this violation is not relevant for the design of the contracts which are actually chosen. In particular, profits are not continuous in the limit as valuations in the low frame converge to valuations in the high frame. In light of this property, it is important to note that the different frames are to be interpreted as frames that can be induced *reliably* by the firm. The right way to think of the possibilities of framing vanishing is a vanishing probability that a given consumer will react to the frame. Profits are continuous in this limit. The strength of framing does impact the construction of the decoy contracts. As framing gets weaker, the quality of the decoys increases. The production cost of decoys is irrelevant, as the decoy contracts are never chosen. If consumers make mistakes and sometimes choose decoy contracts or quality is bounded, our construction requires sufficiently strong framing effects.

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 they prefer in their current frame.[18] They fail to take account of the fact that in this continuation problem, they may be

the decision problem above. We cannot rule out that there is a more complex EDP that can deal with some violtions of type comonotonicity.

[18]In other words, they suffer from complete projection bias (Loewenstein et al., 2003) in forecasting their future decisions. We can also extend our results to partial projection bias. Denote the parameter deteriming the intensity of projection bias by $\alpha \in [0, 1]$, with $\alpha = 0$ representing full sophistication. Assume for concreteness that a consumer with payoff type θ when predicting his own choices in a situation with taste θ' to be made according to $\widehat{\theta}(\theta, \theta', \alpha)$. This function is increasing in the first two arguments and satisfies $\widehat{\theta}(\theta, \theta, \alpha) = \theta$ for all α. Under full sophistication we have $\widehat{\theta}(\theta, \theta', 0) = \theta'$, under full naivete $\widehat{\theta}(\theta, \theta', 1) = \theta$. This structure ensures that predictions satisfy comonotonicity (Assumption 2). Whenever $\alpha < 1$, we can extend the sophisticated

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 $s_\theta : \mathcal{E} \to \mathcal{E}$ a naive strategy for θ if

$$\begin{aligned} s_\theta(E, f) &\in E \\ \mathcal{C}(s_\theta(E, f)) \cap \operatorname*{argmax}_{\mathcal{C}(e)} u_{\theta_f} &\neq \varnothing. \end{aligned}$$

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 one of the optima.

We call $\nu : \Theta \to \mathcal{C}$ a naive outcome of an EDP e if there exists a naive strategy profile s such that any type θ arrives at $\nu(\theta)$ by following s_θ, i.e. $\nu(\theta) = (s_\theta \underbrace{\circ \cdots \circ}_{k \text{ times}} s_\theta)(e)$ 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 this cognitive type. Let $\Theta = \Theta_S \sqcup \Theta_N$, where Θ_S and Θ_N denote the set of naive and sophisticated types, respectively. Note that we allow for the existence of $\theta^s \in \Theta_S$ and $\theta^n \in \Theta_N$ mapping frames into the same payoff types. These types differ only in their sophistication, but not in their tastes conditional on any frame. Define the optimal profits similarly to (GP) as

$$\begin{aligned} \Pi^*(\Theta, \mu) := \max_{e \in \mathcal{E}} & \ \Pi(\mathbf{p}, \mathbf{q}) \\ \text{s.t.} & \ (p_\theta, q_\theta) \in \Sigma^e(\theta), \forall \theta \in \Theta_S, \\ & \ (p_\theta, q_\theta) \in N^e(\theta), \forall \theta \in \Theta_N. \end{aligned} \tag{7}$$

Optimal Structure and Contracts

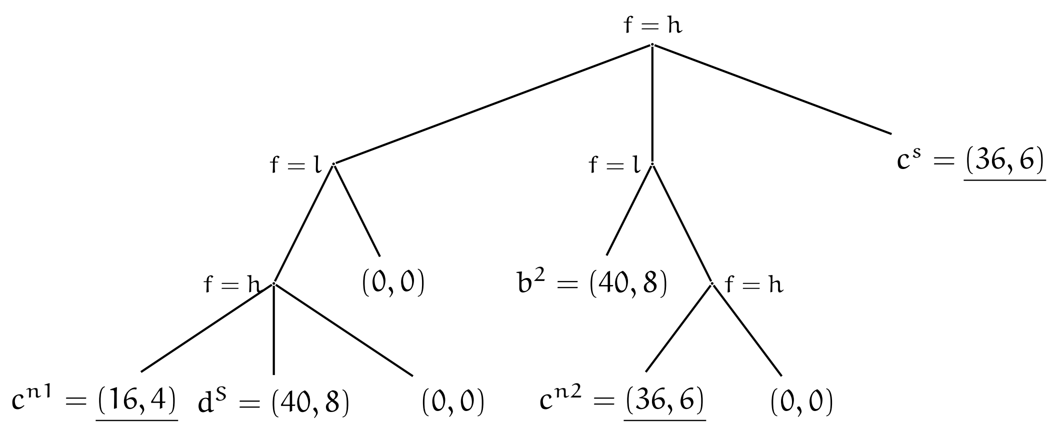
We return to Example 1 to illustrate how the principal can use decoys to screen when naive types are present.

construction by replacing θ_l^i by $\widehat{\theta}(\theta_h^i, \theta_l^i, \alpha)$ 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 increases the level of quality required in the decoys. If providing very high quality decoys is not entirely costless or quality is bounded, we expect to see the sophisticated construction for agents with weak projection bias and the naive construction for individuals with strong projection bias.

Example 2. Recall from the introduction that there are two frames, $f \in \{l, h\}$, and two payoff types, $\theta \in \{\theta^1, \theta^2\}$, with valuations given in Fig. 1a. The low payoff type is θ^1.

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 $\Theta = \Theta_S \sqcup \Theta_N = \{\theta^s\} \sqcup \{\theta^{n1}, \theta^{n2}\}$, where $\theta^{n1} = \theta^1, \theta^{n2} = \theta^s = \theta^2$. In this setting, the principal can give the efficient quantity for the high frame to naive consumers and fully extract their surplus. At the same time, 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. The optimal EDP is given in Fig. 5, where we omitted to outside option at the root for compactness. It implements $c^s = (36, 6)$, $c^{n1} = (16, 4)$, $c^{n2} = (36, 6)$.

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



First, consider the sophisticated type. As in Example 1, there is a contract that is more attractive than the implemented one, c^{n1}. It is concealed using the decoy d^s as in the previous example.

Let's turn to the naive consumers. The leftmost continuation problem is intended for θ^{n1}. Compared to the structure in Fig. 1b, there are no added features. Why can the principal extract all surplus in the high frame even though the path to c^1 passes through a node in the low frame? When at this node, c^{n1} indeed prefers the outside option over c^1. But he wrongly beliefs that he will choose the outside option after continuing.[19] Hence, he continues and – back in frame h – chooses c^1.

In order to implement the optimal vector of contracts for naive consumers, we also need to use decoys. This is illustrated by θ^{n2}. At the root, he strictly prefers c^1 to c^2. In order to lure him into the middle continuation problem, the principal introduces a decoy b^2.[20]

[19]Here, we are breaking the tie for the principal. By adding another decoy, the choice can be made strict at an arbitrarily small loss of profit.

[20]In this simple example, the two decoys, d^s and b^2, coincide. This is the case because they are designed to

This decoy works differently from the decoys used with sophisticated consumers. It serves as bait and is the most preferred contract out of the whole decision problem for θ^{n2}. As a consequence, he continues into the middle continuation problem. There, the switch happens: b^2 is unattractive from the perspective of the low frame and θ^{n2} continues, expecting to pick the outside option in the continuation problem. Like θ^{n1} he reconsiders at the terminal node and ends up with c^2.

In this particular example, the optimal EDP is independent of the probabilities of the three types. This is true in general for naive types, the optimal contracts for sophisticated types depend on the proportion of different payoff types among them. ▲

These constructions generalize. The optimal EDP achieves the same outcome as if the principal knows which consumers are naive and the types of those naive consumers. Naive types don't obtain information rents, they receive 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 Θ and σ and ν be its firm-preferred sophisticated and naive outcomes, respectively. Then there exists an EDP e_S that is optimal for the set of types Θ_S with conditional prior and its firm-preferred sophisticated outcome σ^s, such that*

$$\sigma(\theta) = \sigma^s(\theta), \quad \forall \theta \in \Theta_S$$
$$\nu(\theta) = \widehat{c}_{\theta_h}, \quad \forall \theta \in \Theta_N.$$

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.[21]

To achieve this result, 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 intended continuation problems. Agents wrongly believe that they will choose their respective decoy, which is the most attractive contract

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.

[21] If there are m different payoff types, there are at most $2m$ different continuation choices, one sophisticated and one naive per payoff type. This result generalizes to arbitrary finite support distributions over intensities of projection bias (see FN 18). The principal can separate them by cognitive type and implement the optimum as if their cognitive type is observable using at most $2m$ continuation problems (independently of the number of cognitive types).

in the whole EDP for them in their 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.[22] 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 a weak improvement*[23] *from the perspective of every frame over the contract under naivete.*

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.

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.[24] Interestingly, such legislation typically only applies to door-to-door sales and similar situations of high sales pressure to which the consumer 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

[22]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.

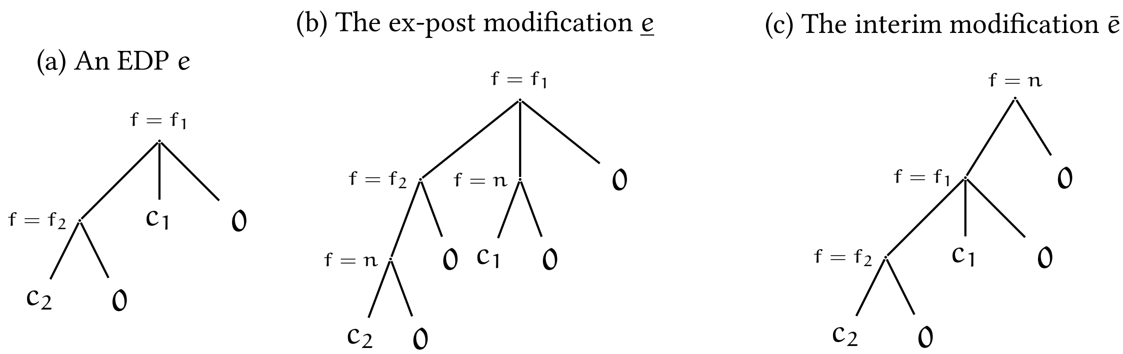
[23]This can be interpreted as a weak improvement in the sense of Bernheim and Rangel (2009) if the two contracts are not identical.

[24]E.g. directive 2011/83/EU: "the consumer should have the right of withdrawal because of the potential surprise element and/or psychological pressure".

can interpret this decision as an additional *ex-post* participation decision at every terminal decision stage.

Formally, denote the neutral frame by $n \in F$, $n < h$.[25] This is the frame the consumer is in when unaffected by direct sales pressure by the firm.[26] We call $\bar{e} := (\{e, 0\}, n)$ an interim modification[27] of e. Then e is an EDP with an interim participation constraint if it is an interim modification of some EDP. Next, 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 $e \in \mathcal{E}^0$. Having defined an ex-post 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 Intuitively, if consumers are sophisticated, both constraints are equivalent and imply that if a contract is chosen by type θ, then it must satisfy the additional participation constraint P^n_θ. 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*

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

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

[25]It is immediate that additional participation decisions in frame h don't change the result, as all implemented contracts satisfy them.

[26]One possible effect of marketing is influencing this neutral frame, but we won't consider this margin.

[27]Here, the notion of interim modification is defined on $\mathcal{E} \cup \mathcal{E}^0 \setminus \{0\}$. For simplicity, let $\bar{0} := 0$

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.[28]

If consumers are sophisticated, the two restrictions are equivalent. They 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 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 naive outcomes* $\bar{v}$ *and* $\underline{v}$*, respectively, are such that* $\forall \theta \in \Theta$

$$\bar{v}_\theta = \widehat{c}_{\theta_h},$$
$$\underline{v}_\theta = \widehat{c}_{\theta_n}.$$

The intuition underlying the design of regulation is faulty if consumers are naive. Naive consumers 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 and would make use of a right to return, if they had it. 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 the consumers decision in several ways, affecting consumer valuations as expressed by their choices. Such a setting naturally leads to extensive-form decision problems. We find that 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. This result is obtained despite full strategic sophistication of consumers. Indeed, the principal turns their

[28] 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.

sophistication against them. The optimal contracts can be implemented by an extensive-form decision problem with a simple and intuitive structure: 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.[29]

This simple extensive form allows the firm to eliminate information rents at the cost 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 a product in an exogenously given neutral frame, they only are protected against quality exaggeration, but not 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 sales person or the information material and assume that the properties of the choice set (such as an attribute being widely dispersed) are not relevant. This modeling choice is convenient, as the frame at every node is exogenously set by the principal and type independent.

Our construction exploits a violation of dynamic consistency. Those can also be caused by other factors, such as seasonal shifts in tastes or context effects. 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. Then, the car dealer can find the optimal EDP using our results and implement it as 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

[29]This simple structure also supports the assumption of sophistication. In the optimum, consumers only need to grasp relatively short and intuitively understandable extensive forms.

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. 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.

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. This is possible by introducing a high quality-high price option that remains unchosen by every type.[30] There is an important caveat, however. If it is possible to extend our construction into an analogous setting with endogenous frames, it may not be optimal. While with exogenous frames, the frames have to be chosen for every decision node independently of the agents 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. This can add to the implementation power of extensive-form decision problems.

[30] 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.

A Appendix

A.1 Preliminaries

Whenever types are enumerated by subscript i, we use notation $u_f^i := u_{\theta_f^i}$. For each θ_f^i, let $\succcurlyeq_f^i$ and (sometimes) $\succcurlyeq_{\theta_f^i}$ 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* $\bar{\theta} \geqslant \underline{\theta}$*, and contracts* $x, y \in \mathcal{C}$*, such that* $q^y \geqslant q^x$*, we have*

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

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

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

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* $\bar{p} > \underline{p} \geqslant 0$ *and payoff types* $\bar{\theta}_{f'} > \underline{\theta}_f$*, there exist* $q \geqslant 0$*, 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 ϕ is thrice differentiable and strictly increasing. Our assumption $\frac{\partial^3 v}{\partial q^2 \partial \theta} > 0$ implies that ϕ is strictly convex as

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

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

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

$$\bar{p} - \underline{p} = \phi(q) = \nu_{\bar{\theta}_{f'}}(q) - \nu_{\underline{\theta}_f}(q) \Leftrightarrow u_{\underline{\theta}_f}(\underline{p}, q) = u_{\bar{\theta}_f}(\bar{p}, q).$$

□

Lemma 4. *The efficient quantity for payoff type* θ_f *defined as*

$$\widehat{q}_{\theta_f} := \underset{q \geqslant 0}{\operatorname{argmax}}\, \nu_{\theta_f}(q) - \kappa(q)$$

exists, is unique and increasing in θ_f*.*

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

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

and note that it is continuous, twice differentiable and satisfies

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

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

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

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

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

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

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

Note that (12) implies

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

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

A.2 Proofs

Proof of Observation 1 on page 11: 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

$$\begin{aligned} p'_i &= p_i + \Delta_i \\ \Delta_1 &= \nu\left(\theta^1_h, q_1\right) - \nu\left(\theta^1_f, q_1\right) \\ \Delta_i &= \Delta_{i-1} + \nu(\theta^i_h, q_i) - \nu(\theta^i_f, q_i) - \left[\nu(\theta^i_h, q_{i-1}) - \nu(\theta^i_f, q_{i-1})\right] \end{aligned}$$

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

$$\begin{aligned} \nu(\theta^i_h, q_i) - p'_i &= \nu(\theta^i_f, q_i) + \nu(\theta^i_h, q_i) - \nu(\theta^i_f, q_i) - p_i - \Delta_i \\ &= \nu(\theta^i_f, q_{i-1}) + \nu(\theta^i_h, q_i) - \nu(\theta^i_f, q_i) - p_{i-1} - \Delta_i \\ &= \nu(\theta^i_f, q_{i-1}) - p_{i-1} - \Delta_{i-1} + \left[\nu(\theta^i_h, q_{i-1}) - \nu(\theta^i_f, q_{i-1})\right] \\ &= \nu(\theta^i_h, q_{i-1}) - p_{i-1} - \Delta_{i-1} \\ &= \nu(\theta^i_h, q_{i-1}) - p'_{i-1} \end{aligned}$$

Hence all other IC and P are satisfied by the usual arguments. Note that $\Delta_i \geqslant 0$ 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 $e = (E, h) \in \mathcal{E}^{k+1}$. σ is an outcome to e iff

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

Therefore, σ is an outcome to e if and only if

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

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

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 14: 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 θ' to θ 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. □

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 $\forall j \geqslant 1, \forall k \neq j$, we have*

1. $q_j^d \geqslant q_{j-1}^d$ *(decoy quantities are increasing),*
2. $d_j \sim_j^h d_{j-1}$,
3. $d_j \succcurlyeq_j^h x$ *and* $d_j \succcurlyeq_j^h d_k$ *(θ^j chooses d_j),*
4. $x \succcurlyeq_0^h d_j$ *(θ^0 chooses x),*
5. $0 \succcurlyeq_j^l d_j$ *(decoys are undesirable in l).*

Proof. Let $d_0 = x$ for brevity. We construct the decoys $d_i = (p_i, q_i)$ recursively.[31] For $i \in \{1, \dots, n\}$, pick

$$d_i \sim_{\theta_l^i} 0 \tag{15}$$

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

or equivalently

$$v(\theta_l^i, q_i) - p_i = 0$$
$$v(\theta_h^i, q_i) - p_i = v(\theta_h^i, q_{i-i}) - p_{i-i}$$

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: $q_i \geqslant q_{i-1}$.

Proof of Claim 1: By the two defining relations

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

[31] The present proof can be extended to the case without ordered types (but maintaining ordered frames).

Hence

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

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

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

△

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

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

Proof of Claim 2: We will show that for all j we have $d_i \succcurlyeq_{\theta_h^i} d_j$. 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

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

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

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

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

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

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

□

Proof of Lemma 1 on page 15: A continuation problem for type $\theta \in \Theta$ with contract c_θ satisfying all three properties is given by $e_\theta = (\{0, (\{0, c_\theta\} \cup \{d_{\theta'}\}_{\theta' > \theta}, h)\}, l)$, where the contracts $\{d_{\theta'}\}_{\theta' > \theta}$ are constructed in Lemma 5 as $(d_{\theta'})_{\theta' \geqslant \theta} := \mathcal{D}(c_\theta, (\theta')_{\theta' \geqslant \theta})$.

By construction, type θ chooses c_θ from the terminal problem and since c_θ 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 $\{d_{\theta'}, 0\}$ 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 $d_0 \succcurlyeq_{\theta_h} d_i$ and $q_i \geqslant q_0$, hence by single crossing $d_0 \succ_{\theta_h'} d_i$, which establishes that a lower type never chooses any of the decoys. □

Proof of Theorem 1 on page 12 and Theorem 2 on page 17: 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 Σ^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 θ 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 θ, and

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

the closest revealed type above a given type θ. 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.

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

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

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

$$\nu_{\theta_h}(q_\theta) - p_\theta \geqslant \nu_{\theta_h}(q_{\eta(\theta)}) - p_{\eta(\theta)} \quad \forall\theta \in \Theta \tag{18}$$

$$\nu_{\theta_h}(q_\theta) - p_\theta \geqslant \nu_{\theta_h}(q_{\chi(\theta)}) - p_{\chi(\theta)} \quad \forall\theta \in \Theta \tag{19}$$

We have the following

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

$$\theta, \theta' \in \Theta_R,\ \theta > \theta' \implies q_\theta \geqslant q_{\theta'} \tag{20}$$

and solves the relaxed problem.

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

$$\nu_{\theta_h}(q_\theta) - p_\theta \geqslant \nu_{\theta_h}(q_\eta) - p_\eta$$

$$\nu(\eta_h, q_\eta) - p_\eta \geqslant \nu(\eta_h, q_\theta) - p_\theta$$

and hence

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

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 Θ_R between the source of the $IC^h_{\theta\theta'}$ constraint θ and it's target θ'. 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

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

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). △

Lemma 7. *In the optimal contract of the relaxed problem the IC from any revealed type* θ *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 θ to $\eta(\theta)$. Suppose we increase the price in the contract of all revealed types greater than θ including θ by some $\epsilon > 0$. Note that this change doesn't affect any constraints between the affected types. Furthermore, θ isn't the lowest revealed type, hence the participation constraint of all revealed type is implied by the IC and not active since IC-$\theta \to \eta(\theta))$ 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. △

Lemma 8. *In the optimal contract of the relaxed problem,* $q_\theta \leqslant \widehat{q}_{\theta_h}$ *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 θ. Note that $q_\eta(\theta) \leqslant \widehat{q}_{\eta(\theta)_h} < \widehat{q}_{\theta_h}$ and denote the rent given to type θ as $\Delta := \nu(\theta^h, q_{\eta(\theta)}) - p_{\eta(\theta)}$. (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 θ by $(\widehat{q}_{\theta_h}, \nu_{\theta_h}(\widehat{q}_{\theta_h}) - \Delta)$. As θ receives the same utility in both contracts, no participation constraint is violated and all IC from α are still satisfied. The upward IC $\eta(\theta) \to \theta$ is still satisfied as it is implied by R-monotonicity (which is maintained) and the corresponding downward IC. Consider any higher type imitating θ. The amended contract gives the same utility to θ 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 θ in frame h is using quantity $\widehat{q}_{\theta_h}$. Hence, the initial set of contracts wasn't optimal. $\triangle$

Now we can show that there are no profitable feasible upward deviations in e^*. We proceed by induction. Order the types such that $\{\theta^1, \ldots, \theta^n\} = \Theta$, $\theta^i < \theta^{i+1}$. Clearly, the highest type has no feasible upward deviations. Suppose all upward deviations are either infeasible or unprofitable for types θ^i into types θ^j for $j > i > m$. We need to show that the required upward IC constraints out of type θ^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* $\Delta_{\theta^i} \leqslant \nu(\theta^i_h, \widehat{q}_{\theta^i_h}) - \nu(\theta^i_l, \widehat{q}_{\theta^i_h})$*:* Then the participation constraint of type θ^i is binding at the intermediate stage in frame l. But by single crossing

$$c_{\theta^i} \sim_{\theta^i_l} 0 \implies c_{\theta^i} \prec_{\theta^m_l} 0, \tag{21}$$

an imitation is infeasible.

2. *Deviations into a concealed type with rent* $\Delta_{\theta^i} > \nu(\theta^i_h, \widehat{q}_{\theta^i_h}) - \nu(\theta^i_l, \widehat{q}_{\theta^i_h})$*:* Note that in this case $q_{\theta^i} = \widehat{q}_{\theta^i_h}$ 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 $\Delta_{\theta^i} = \nu(\theta^i_h, q_\eta) - p_\eta$ for some $\eta < \theta^i$, $\eta \in \Theta_R$. But then the upward deviation isn't profitable unless the deviation into η is profitable, since $q_\eta \leqslant \widehat{q}_{\eta_h} \leqslant \widehat{q}_{\theta^i_h} = q_{\theta^i}$ and by single crossing

$$c_\eta \sim_{\theta^i_h} c_{\theta^i} \implies c_\eta \succ_{\theta^m_h} c_{\theta^i} \tag{22}$$

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 θ^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 θ^m has a strictly profitable deviation. Furthermore, since we impose downward incentive compatibility constraints, this lowest target type has to be greater than θ^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 $c_{\eta(\underline{\theta})} \sim_{\underline{\theta}_h} c_{\underline{\theta}}$ and by R-monotonicity $q_{\eta(\underline{\theta})} \leqslant q_{\underline{\theta}}$, but then by single crossing $c_{\eta(\underline{\theta})} \succcurlyeq_{\theta_h^m} c_{\underline{\theta}} \succ_{\theta_h^m} c_{\theta^m}$, 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 θ^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). □

Proof of Proposition 3 on page 18: 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 Θ_C and Θ_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 consuder. 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 θ and set the contract $(\widehat{q_{\theta_l}}, \widehat{p_{\theta_l}} - \Delta)$. Using the surplus function ζ_θ 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 Δ to type θ 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. □

Proof of Proposition 4 on page 18: 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 θ in (RP). The contract given to type θ solves

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

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

$$\Delta \leqslant [v_{\theta_h}(\widehat{q}_{\theta_l}) - v_{\theta_l}(\widehat{q}_{\theta_l})]\,. \tag{23}$$

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

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

In the intermediate case, both constraints are binding,

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

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

Proof of Proposition 5 on page 20: Take any type $\theta \in \Theta$. For each μ, consider (RP) with the constraint $\theta \in \Theta_C$ ($\theta \in \Theta_R$) and denote the corresponding optimal value by $\Pi^R_{C;\mu}$ ($\Pi^R_{R;\mu}$). Next, using the surplus function ζ_{θ_f} defined in (8), we can bound those values as

$$\Pi^R_{R;\mu} \geqslant \mu_\theta \zeta_{\theta_h}(\widehat{q}_{\theta_h}) \tag{25}$$

$$\Pi^R_{C;\mu} \leqslant \mu_\theta \zeta_{\theta_l}(\widehat{q}_{\theta_l}) + \sum_{\theta' \neq \theta} \mu_{\theta'} \zeta_{\theta'_h}(\widehat{q}_{\theta'_h}) \leqslant \mu_\theta \zeta_{\theta_l}(\widehat{q}_{\theta_l}) + (1-\mu_\theta)\zeta_{\tilde{\theta}_h}(\widehat{q}_{\tilde{\theta}_h}), \tag{26}$$

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

Note that Lemma 4 implies that

$$\zeta_{\theta_h}(\widehat{q}_{\theta_h}) > \zeta_{\theta_h}(\widehat{q}_{\theta_l}) > \zeta_{\theta_l}(\widehat{q}_{\theta_l}), \tag{27}$$

and define

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

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

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

Therefore, for any $\mu_\theta \in [\bar{\mu}_\theta, 1]$, it is optimal to reveal θ.[32]

□

Proof of Proposition 6 on page 20: Suppose that there exists j, such that $\tilde{\theta}^j > \theta^j$ and $\forall i \neq j, \tilde{\theta}^j = \theta^j$. Note that the case with multiple strict inequalities follows from it by induction.

First, we show that any $\tilde{\theta} \in \tilde{\Theta}_C$ can be optimally concealed. By contradiction, suppose there exists a subset $\tilde{\Theta}'_C \subseteq \tilde{\Theta}, \tilde{\Theta}'_C \not\ni \tilde{\theta}$ and a vector of contracts $\mathbf{c}'$, such that it yields strictly higher profits than concealing $\tilde{\theta}$. Then construct a subset $\Theta'_C := \tilde{\Theta}'_C$ of Θ and note that since $\theta \in \Theta \setminus \Theta'_C$, type θ is revealed. This implies that the vector $\mathbf{c}'$ is feasible under Θ'_C and yields strictly higher profits than the profits from concealing Θ_C, which is a contradiction.

Second, we show that any $\theta \in \tilde{\Theta} \setminus \tilde{\Theta}_C$ can be optimally revealed. By contradiction, suppose there exists a subset $\tilde{\Theta}'_C \subseteq \tilde{\Theta}, \tilde{\Theta}' \ni \tilde{\theta}$ and a vector of contracts $\mathbf{c}'$, such that it yileds strictly higher profits than revealing $\tilde{\theta}$. Then construct a subset $\Theta'_C := \{\theta_j\} \cup \tilde{\Theta}'_C \setminus \{\tilde{\theta}_j\}$ of Θ and note that $\mathbf{c}'$ is feasible under Θ'_C and yields strictly higher profits than the profits from concealing Θ_C, which is a contradiction.

□

Proof of Theorem 3 on page 25: 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,

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

To define E_i, let the most preferred alternative in e_{i-1} for type θ^i be $x_i := \operatorname{argmax}_{\mathcal{C}(e_{i-1})} u_{\theta^i_h}$. During the construction, we ensure that

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

[32] This bound is typically not tight, as we introduced slack in (26).

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

Let $E_i = \left\{ \left(\left\{ \left(\{N_i, \{d^{\theta'}_{N,i}\}_{\theta' > \theta^i : \theta' \in \Theta_S}, 0\}, h \right), \left(\{\widehat{c}_{\theta^i_h}, \{d^{\theta'}_i\}_{\theta' > \theta^i : \theta' \in \Theta_S}, 0\}, h \right), 0 \right\}, l \right) \right\}$.

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

$$N_i \sim_{\theta^i_h} x_i \tag{29}$$

$$q_{N_i} \geqslant q_{x_i} \tag{30}$$

$$N_i \preccurlyeq_{\theta^i_l} 0 \tag{31}$$

Second, define the decoys as

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

where $\Theta_S^{>\theta^i}$ is a vector of types in Θ_S that are strictly greater than θ^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 θ^i_h, hence continuing into E_i is part of a naive outcome for θ^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, $N_i \succcurlyeq_{\theta^I_h} d^{\theta'}_{N,i}$ and $q_{N_i} \leqslant q_{d^{\theta'}_{N,i}}$ and hence $N_i \succcurlyeq_{\theta^i_l} d^{\theta'}_{N,i}$. But $N_i \preccurlyeq_{\theta^i_l} 0$. As the decision problem containing $\widehat{c}_{\theta^i_h}$ also contains the outside option, continuing to this menu is part of a naive outcome. From the menu $\left(\{\widehat{c}_{\theta^i_h}, \{d^{\theta'}_i\}_{\theta' > \theta^i : \theta' \in \Theta_S}, 0\}, h \right)$, the DM chooses $\widehat{c}_{\theta^i_h}$ by the construction of the decoys. This establishes 3.

To see 2, note that all decoys have higher quality than N_i and $\widehat{c}_{\theta^i_h}$, respectively, and are less preferred according to θ^i_h. Hence, they are less preferred by lower naive types θ^j_h by single crossing. Furthermore, $\widehat{c}_{\theta^i_h}$ 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 $N_i \sim_{\theta^i_h} x_i$ and $q_{N_i} \geqslant q_{x_i}$ imply $N_i \preccurlyeq_{\theta^j_h} x_i$ for all $j < i$. By the induction hypothesis, $N_j \succcurlyeq_{\theta^j_h} \operatorname{argmax}_{\mathcal{C}(e_{i-1})} u_{\theta^j_h} \succcurlyeq_{\theta^j_h} x_i \succcurlyeq_{\theta^j_h} N_i$. 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 $v(\theta) \succcurlyeq_{\theta_n} 0$. The revenue maximal vector of contracts satisfying these contraints 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 n 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. □

Proof of Observation 3 on page 26: Let us denote the contract for type θ in the sophisticated problem as c^s_θ and note that the contract in the naive problem is $\widehat{c}_{\theta_h}$. Note that $c^s_\theta \succcurlyeq_{\theta_h} 0 \sim \widehat{c}_{\theta_h}$ and $q^s_\theta \leqslant \widehat{q}_{\theta_h}$. Hence by single crossing $c^s_\theta \succcurlyeq_{\theta_f} \widehat{c}_{\theta_h}$, strictly for $f \neq h$ if $c^s_\theta \neq \widehat{c}_{\theta_h}$. □

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.* 2

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, May 2006, *121* (2), 635–672. 4

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. 26

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," *Journal of Political Economy*, 2013, *121* (5), 803–843. 2, 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. 2

Busse, Meghan R., Devin G. Pope, Jaren C. Pope, and Jorge Silva-Risso, "The Psychological Effect of Weather on Car Purchases," *The Quarterly Journal of Economics*, February 2015, *130* (1), 371–414.

Conlin, Michael, Ted O'Donoghue, and Timothy J Vogelsang, "Projection Bias in Catalog Orders," *American Economic Review*, August 2007, *97* (4), 1217–1249.

de Clippel, Geoffroy, "Behavioral Implementation," *American Economic Review*, October 2014, *104* (10), 2975–3002. 7

DellaVigna, Stefano, John A. List, and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, January 2012, *127* (1), 1–56.

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. 2

Eliaz, Kfir and Ran Spiegler, "Contracting with Diversely Naive Agents," *The Review of Economic Studies*, January 2006, *73* (3), 689–714. 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

_ 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, 21

Glazer, Jacob and Ariel Rubinstein, "A Model of Persuasion with Boundedly Rational Agents," *Journal of Political Economy*, December 2012, *120* (6), 1057–1082. 6

_ and _ , "Complex Questionnaires," *Econometrica*, 2014, *82* (4), 1529–1541. 6

Gourville, John T., "Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation," *Journal of Consumer Research*, March 1998, *24* (4), 395–403. 2

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. 2

Heidhues, Paul and Botond Kőszegi, "Exploiting Naïvete about Self-Control in the Credit Market," *American Economic Review*, December 2010, *100* (5), 2279–2303. 6

— **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. 2

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, 30

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, May 1997, *112* (2), 443–478. 4

Lasaleta, Jannine D., Constantine Sedikides, and Kathleen D. Vohs, "Nostalgia Weakens the Desire for Money," *Journal of Consumer Research*, October 2014, *41* (3), 713–729.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," *The Quarterly Journal of Economics*, 2003, *118* (4), 1209–1248. 2, 22

Maskin, Eric and John Riley, "Monopoly with Incomplete Information," *The RAND Journal of Economics*, 1984, *15* (2), 171–196. 7

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. 2

— **, David J. Hargreaves, and Jennifer McKendrick**, "In-Store Music Affects Product Choice," *Nature*, November 1997, *390* (6656), 132–132. 2

— **, 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. 2

Piccione, Michele and Ran Spiegler, "Price Competition Under Limited Comparability," *The Quarterly Journal of Economics*, January 2012, *127* (1), 97–135. 5

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. 6, 11, 26, 28

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. 2

Spiegler, Ran, "Competitive Framing," *American Economic Journal: Microeconomics*, 2014, *6* (3), 35–58. 5

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. 2

Zhang, Wei, "Endogenous Preferences and Dynamic Contract Design," *The B.E. Journal of Theoretical Economics*, January 2012, *12* (1). 6