LISTING SPECS: THE EFFECT OF FRAMING ATTRIBUTES ON CHOICE[1]

Consistent evidence across important domains shows that people's decisions can depend on the order or emphasis with which the attributes of the available options are presented to them. We introduce the first model of such framing effects, which we characterize in terms of observable behavior. We apply the model to study how the strategic use of attribute framing affects the outcomes of negotiations and competition in markets. We extend the model to stochastic-choice frameworks, which are often used in practice.

KEYWORDS: Attribute, Framing, Order, Multi-Attribute Choice, Primacy, Recency, Emphasis.

JEL CLASSIFICATION: D01, D11, D90

"In order to construct rich economic models one often needs a model of choice with frames." (Salant and Rubinstein (2008))

1. Introduction

Can the *order* in which information is presented to people affect their behavior? Consistent evidence shows that such framing effects exist: Decisions (sometimes) depend on the order in which the attributes of available alternatives are presented, often in important domains. People's willingness to pay for medical treatment depends on the presentation position of its price (Kjær et al. (2006)). Choices of health plans depend on how attributes like copay, deductibles, and premium are presented (Ericson and Starc (2016)). Doctors' diagnoses depend on the order in which pieces of information are encountered (e.g., Bergus et al. (1995), Cunnington et al. (1997), Chapman and Elstein (2000)). Some police investigations and jury decisions depend on the presentation order of alibi and eyewitness evidence (Dahl et al. (2009)). Consumers' evaluation of some products depends on the presentation order of their attributes (Kumar and Gaeth (1991), Levav et al. (2010) and references therein). Blake et al. (2021) show that the "purchase funnel"—the order of steps to buy a product—can affect consumers' decisions. Other papers in marketing and psychology report related evidence.[1]

[1] See, e.g., Cornelissen and Werner (2014) and Auspurg and Jäckle (2017) for recent reviews as well as Chrzan (1994), Day and Prades (2010) and Day et al. (2012)). Even voters' support for political candidates may depend on the presentation order of their "attributes." For instance, in the 2016 U.S. Presidential race Hillary Clinton presented herself moving her maiden name (Rodham) to her middle name so as to *emphasize* her independence from her husband (Shafer (2017)).

Current economic theory cannot capture these framing phenomena. This calls for a new framework that allows the ordering of attributes to affect choice behavior, can be broadly applied, and can be falsified. To this end, we introduce the first explicit, decision-theoretic, model of framing of choice items' attributes and its effects on decisions. We take the physical attributes of an item as the given information to be framed; different presentation orders correspond to different frames. We interpret the presentation position as the *observable* emphasis given to the attribute. As such, our model can be used in other settings where the emphasis is given by some graphical means, such as font size or color.

The paper has three main parts. After introducing the model, we first present an application with "large" domains of items: We study how negotiators may frame their offers to strike a deal when each attribute of the negotiation can vary freely in a rich set. Second, we consider an application with "small" domains of items: We analyze how competing firms may strategically frame their products to influence their competitive landscape when the constituent attributes of the products (excluding the price) are fixed. Third, we offer an axiomatization of our model that clarifies various general properties and shows that the model can be identified and falsified. We do not assume or describe psychological processes that may generate our framing effects—although understanding such mechanisms seems an important avenue for future research.[2] Instead, in line with modern decision theory, our goal is to develop

[2] See, e.g., Schrift et al. (2018) for experimental evidence on possible channels.

a model which is consistent with choice data that exhibits those framing effects and that can be easily used in economic applications.[3]

To fix ideas, consider an example. Health plans are often presented in tables where each row is an attribute (copay, deductibles, premium, etc.) and each column is a plan. Let N be the number of attributes (hence, rows) and let $f(i)$ be the attribute in row i. The assignment f of attributes to rows is our frame. A plan is then a vector $x_f = (x_{f(i)})_{i=1}^N$, where $x_{f(i)}$ is the level of attribute $f(i)$. Mainstream choice theory assumes f is irrelevant. We allow f to affect which plan customers choose (hence, each plan's market share). For instance, this may change if the premium is moved from the first to the last row.

To study these framing effects, in Section 2, we introduce a baseline model called the *attribute-framing model*. Given a set of items all framed according to f, the decision-maker chooses the x_f that maximizes

$$\sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}). \tag{1}$$

Each $u_{f(i)}$ is a utility function that captures the decision-maker's (stable) underlying tastes for each attribute. The weight function α depends on the attribute's

[3] For recent prominent examples of this approach in economics, see Rubinstein and Salant (2006), Salant and Rubinstein (2008), Masatlioglu et al. (2012), Manzini and Mariotti (2014), Masatlioglu and Ok (2014), Ok et al. (2015), and Cattaneo et al. (2020).

presentation position and is the heart of our model. Depending on its shape, we can capture several empirical regularities in how attribute orders influence choice by changing the marginal rate of substitution between attributes. We characterize which α give rise to *recency* (*primacy*) effects, whereby attributes presented later influence more (less) the evaluation of items than do earlier attributes.[4] One interpretation is that α reveals whether the decision-maker perceives the attributes presented later or earlier as being emphasized. Primacy effects are consistent with the old adage "first impressions matter" and with the "leader-driven" effect: An item that starts ahead in terms of the first attribute is more likely to be chosen (Carlson et al. (2006)). We also characterize what it means to be more susceptible to these effects. Other forms of α are possible, depending on which presentation position carries *relatively* more weight for the decision-maker.

Section 3 presents two of the many possible applications of our model. The first application of our model analyzes framing in negotiations. Framing is often regarded as an important negotiation technique, which can help break an impasse and reach an agreement. Despite this, modeling framing in negotiations has been challenging. We study negotiations that involve multiple attributes whose ideal levels differ between parties. We show how the proposing party chooses and frames an offer to strike the best deal based on the attributes' importance, degree of conflict between parties, and status-quo level for the receiving party. Emphasizing attributes that are important

[4] For evidence on the primacy and recency effects, see Kardes and Herr (1990), Haugtvedt and Wegener (1994), Payne et al. (2000), Bond et al. (2007), Ge et al. (2011).

or have a bad status quo allows the proposer to highlight the benefits of reaching a deal. But if such attributes involve greater conflict, emphasizing them also highlights the concessions the receiver has to make. Our model allows us to uncover and study this trade-off. A basic insight is that de-emphasizing conflict is the main force driving the proposer's framing strategy. In fact, sometimes the proposer strategically de-emphasizes some attribute, despite its importance, to weaken the impact of strong disagreement with the receiver on that attribute. Moreover, we show that frames can be a tool to break an impasse; but when they become more powerful at influencing preferences, they can make it *harder* for the proposer to find a good compromise.

Our second application analyzes how a firm can frame products to influence the competition it faces. We show that by ordering attributes—and thus giving each different emphasis—a firm can create *fictitious* product differentiation that results in higher profits both for itself and for its competitors. Sometimes incumbents can also use framing to deter entry into a market, but this involves trade-offs. In a nutshell, the incumbent has to make its product "look good, but not too much," which results in lower profits than in uncontested markets. We find that the incumbent is more likely to use framing to deter entry when its product is more similar to that of a potential entrant or the consumers' tastes are less heterogeneous. We relate our findings to the industrial-organization literature on obfuscation strategies.

The axiomatic characterization of our model rests on the assumption that we can observe which items a decision-maker chooses as well as how their attributes are framed.[5] We rely exclusively on choices from menus whose items are all framed in the same way. We consider the rich domain of lotteries over items to identify the weight function α. Our main axiom delivers this identification by considering appropriately calibrated lotteries and swaps of attribute positions.

Section 4 generalizes our theory in several ways. We consider a non-separable model where how much the decision-maker weighs attributes presented later depends on how good earlier attributes are. For example, she may overlook later attributes if the first ones already give her high utility. We axiomatize this model in the domain of stochastic choice, as this offers more structure for this task and allows for simpler identification of the model's parameters. This extension also allows us to showcase how to introduce framing effects in the domain of stochastic choice, which is often used in practice for structural analysis and hence broadens the applicability of our theory. We focus on a Luce framework, but the theory can be extended to richer frameworks such as the perturbed-utility model of Fudenberg et al. (2015) and the rational-inattention model of Matějka and McKay (2015).

The flexibility and tractability of our model allow one to formalize and study other questions regarding attribute framing. For reasons of space, we leave these for future research, but briefly discuss them in an Online Appendix. These include how to study choice from menus whose items are presented with different frames, how to understand

[5] For other axiomatic characterizations of choice behavior that depends on observable framing see, e.g., Masatlioglu and Ok (2005); Salant and Rubinstein (2008); Ahn and Ergin (2010).

self-serving motivated framing (which can be related to the endowment effect), and how to conduct welfare analysis in the presence of attribute-framing effects.

Related Literature. The importance of framing for decision-making has been recognized at least since the seminal work of Kahneman and Tversky (1979) and Tversky and Kahneman (1981). The literature has then evolved in two directions. Some papers developed general frameworks to think about framing (Salant and Rubinstein (2008), Bernheim and Rangel (2009), Salant (2011)). Others have focused on modeling specific ways in which frames can influence choice, especially for applications. Our paper belongs to this second strand.

This literature considers several forms of framing. Following Kahneman and Tversky (1979), many papers have studied presenting choices as gains or losses. Another form is 'mental accounting' in relation to saving and investment decisions (Thaler (1985), Thaler (1990)). Several papers have modeled salience, where the weights the decision-maker gives to attributes can depend on how each *stands out* from the others in a menu (e.g., Kőszegi and Szeidl (2012), and Bordalo et al. (2013)). To the best of our knowledge, our paper is the first to study theoretically framing as the presentation order of items' attributes. As such, it offers a complementary theory of salience: While in those papers salience depends on how much an attribute varies across items and in relation to other attributes, in our paper it depends on the presentation emphasis given to an attribute. This can be a useful addition: For instance, in their study of health-insurance decisions, Ericson and Starc (2016) argue that their evidence "suggests that theories of salience that only rely on the attributes of choice (rather than how they are presented) miss important elements of salience." In some settings, attribute-order effects can drive list-order effects (see Section 4 and the Online Appendix). Such connection relates our analysis to Rubinstein and Salant (2006) where items' position on a list affects choice.

The cognitive-science literature has studied how people seem to form their preferences at the moment of elicitation (Lichtenstein and Slovic (2006)). One interpretation of our model is that the decision-maker has well-defined tastes for each attribute. However, when it comes to combining them to evaluate an item, the attributes' framing influences her evaluation. In this way, elicitation methods can influence her choices right when she makes them. Although this may seem to undermine the discovery of decision-makers' true tastes, we show how observing choices across frames can overcome this issue.

Finally, the contract-theory literature has examined strategic framing in buyer-seller relationships, where framing is *assumed* to influence the buyer's willingness to pay (see, e.g., Salant and Siegel (2018), Ostrizek and Shishkin (2023)). Our model can provide a foundation for this influence. In a related paper, Piccione and Spiegler (2012) study how firms can influence market competition and profits by limiting consumers' ability to compare prices (see also Spiegler (2014)). Our work adds to this literature in several ways. It considers markets with heterogeneous consumers. It shows how framing can complement and manipulate standard vertical differentiation and can be used to dampen competition by hindering product comparisons as well as deterring entry. It offers specific predictions about which frames firms will adopt.

2. The Model

The choice objects are called *items*. Each is described by the attributes in a set A. For example, cars are described by make, model, year, color, style, size, power train, etc.. We assume that $|A| = N$ is finite and $N \geq 2$. Each attribute $a \in A$ can take multiple levels, denoted by the set L_a. An item consists of a list of the level of all its attributes and is denoted by $x = (x_{a'}, \ldots, x_{a''}) \in X = \times_{a \in A} L_a$.

We want to allow the order in which attributes are presented to affect choice. To this end, we introduce the notion of *attribute-frame*. Let F be the set of all bijections from $\{1, \ldots, N\}$ to A. For every frame $f \in F$, $f(i)$ is the attribute presented in the ith position of the item description. We later discuss other interpretations of f, for instance in terms of observable emphasis. We denote an item x under frame f as

$$x_f = (x_{f(i)})_{i=1}^{N},$$

where $x_{f(i)} \in L_{f(i)}$ for all i. The set of all items under frame f is

$$X_f = \times_{i \in N} L_{f(i)}.$$

A set collecting multiple items is called a menu. If all items in a menu are described according to f, we call it an f-*menu* and denote it by $M_f \subseteq X_f$. We will focus on f-menus, which are widespread in practice. In online stores, items are often organized in tables—whether they are health plans, investment products like ETFs, or electronic devices. Nevertheless, our framework allows for the analysis of general menus as explained in Online Appendix F.

EXAMPLE 1. *Suppose items are health plans described by copay, deductibles, and premium:* $A = \{c, d, p\}$*. Each attribute can be high or low:* $L_a = \{h, l\}$*. A frame is the order in which a plan description presents its attributes. This order may be* (d, c, p) *for frame* f *and* (p, c, d) *for* f'*. Thus, the same plan with a high premium, a high copay, and low deductibles may be presented as* $x_f = (l_d, h_c, h_p)$ *or* $x_{f'} = (h_p, h_c, l_d)$*. Health plans are often presented in a table where, say, the rows are the attributes and the columns are the plans. Viewed as menus, such tables always present all items using the same frame.*

We can now introduce our baseline model of frame-dependent choice, whose axiomatization appears in Section 4. Let $c(M_f)$ be the set of choices from menu M_f, which we assume to be nonempty for every M_f.

DEFINITION 1. *An attribute-framing (AF) choice model is defined by a pair* (α, u)*, where* $u = (u_a)_{a \in A}$*, each* $u_a : L_a \to \mathbb{R}$ *is an attribute utility function, and* α :

$\{1, \ldots, N\} \to \mathbb{R}_{++}$ *is a weight function that together satisfy, for all* $f \in F$ *and* M_f*,*

$$c(M_f) = \underset{x_f \in M_f}{\arg\max} \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}).$$

The interpretation is that the decision-maker derives utility from each attribute, which he has to aggregate somehow. In the model, he does so linearly in a way that depends on the presentation order of the attributes through the weights α. Thus, attributes presented early can receive higher or lower weight than later attributes, so marginal rates of substitution between attributes can depend on their presentation position. For simplicity, hereafter we will normalize α so that $\sum_{i=1}^{N} \alpha(i) = 1$.[6] The additive structure of our AF model is intuitive and tractable. It is also widely used in studies of multi-attribute decision making.[7] We will relax it in Section 4.

In the AF model, α controls the effects of the attributes' presentation order. We will focus on the main effects consistently found in the empirical literature: primacy and recency effects.[8] We define them here and give a behavioral characterization in Section 4.

[6] This model is related to the so-called "expectancy value model" of framing in psychology (e.g., Ajzen and Fishbein (1980); Nelson et al. (1997)).

[7] See, e.g., Lancaster (1966), McFadden (1973), Gorman (1980), Smith and Brynjolfsson (2001), Allen and Rehbeck (2023).

[8] See Kardes and Herr (1990), Haugtvedt and Wegener (1994), Payne et al. (2000), Bond et al. (2007), Ge et al. (2011).

Definition 2. *Given the AF model (α, u), the decision-maker exhibits primacy (recency) effects if α is strictly decreasing (increasing).*

We may want to compare decision-makers in terms of how susceptible they are to attribute framing. For this comparison to be meaningful, their tastes over attributes have to be the same.

Definition 3. *Let (α^1, u^1) and (α^2, u^2) be AF models of decision-makers 1 and 2. Suppose for all $a \in A$, $u^1_a = \gamma u^2_a + \zeta_a$ for some $\gamma > 0$ and $\zeta_a \in \mathbb{R}$. Decision-maker 1 is more susceptible to recency (primacy) effects than decision-maker 2 is if*

$$\frac{\alpha^1(i+1)}{\alpha^1(i)} \geq (\leq) \frac{\alpha^2(i+1)}{\alpha^2(i)}, \quad \forall i = 1, \ldots, N-1.$$

In words, decision-maker 1 is more susceptible to recency (primacy) effects than decision-maker 2 if α^1 increases (decreases) faster than does α^2.

Model Discussion. It is worth clarifying a premise of our framework. When we study f-menus, it may look as if we are assuming the decision-maker rigidly follows the order in which attributes are presented. In fact, what we assume is that the exogenously given f influences the decision-makers's preference in a *consistent* way. In this sense, we are not interested in why this phenomenon happens and do not make assumptions about the channels through which it arises. Our goal is to develop a framework to capture this phenomenon in line with the evidence, as in classical decision theory. In this sense, our model relates to the framing idea of "shrouded attributes" despite differing from it (see, e.g., Ellison (2005); Gabaix and Laibson (2006)). While shrouding an attribute is usually intended as excluding it entirely from the decision-maker's consideration, the type of framing this paper is interested in assumes that frames cannot change the set of attributes considered by the decision-maker. Instead, our frames can change how asymmetrically the decision-maker treats the existing attributes. At the same time, one can view placing an attribute in a position to which the decision-maker assigns a very low weight α as akin to *partially* shrouding that attribute, thus making our model a possible generalization of shrouding.

Finally, while our primary interpretation of frames is the order in which attributes are physically presented, other interpretations are possible. One is to view each position $i = 1, \ldots, N$ as a degree of emphasis that the presentation of an item gives to its attributes. For instance, emphasizing may involve highlighting an attribute with color and font size or by placing it in a prominent position on an ad page. The key assumptions are that such attribute frames should (1) create an objective observable order and (2) work in terms of relative effects. That is, for instance, increasing all fonts proportionately does not change anything because the relative emphasis stays the same.[9] With these assumptions in mind, one can apply our model in a variety of settings where the emphasis given to attributes is part of the observable data.

[9] To see that salience is a relative concept, see Milosavljevic et al. (2012) and references therein.

3. Framing at Work: Applications

Our model can be applied in a variety of settings. For space reasons, we focus on two applications, but develop them in some detail. The first showcases how to incorporate framing into the analysis of negotiation and its implications. The second application analyzes how firms can frame products to influence the competition they face, in a classic industrial-organization setting. We also focus on and highlight the key novel aspects and omit aspects that, though realistic and important, would distract from the main point. The goal here is to illustrate the tractability of our approach, its ability to capture real phenomena, and its potential to offer novel insights that can be embedded in more general settings.

3.1. Break the Impasse: Framing in Negotiations

Framing can play an important role in negotiations.[10] This is based on the fact that the way a party describes his offer affects how others view it. Framing often occurs in negotiations, whether parties are aware of it or not. The party controlling the framing process can define a negotiation to its advantage. Positioning a proposal advantageously at the outset of a negotiation is viewed as essential for achieving favorable outcomes. Sometimes re-framing problems helps break an impasse. One framing technique often used is focusing attention on some aspects of a problem and

[10] For related evidence and general discussions about framing and negotiation, see, e.g., Pinkley and Northcraft (1994), Schweitzer and DeChnrch (2001), Levav et al. (2010), Donohue et al. (2011), Druckman and Wagner (2021), and references therein.

leaving others in the background, thereby shaping what other parties pay attention to. Negotiators usually emphasize what they believe are important and advantageous aspects for them. They may also take others' viewpoints into account so as to offer win-win solutions.

Despite its importance, modeling framing in negotiations and how it is used has been challenging. We believe our model provides a step forward in tractability and offers some insights into how the proposing party may select which aspects are important and advantageous in framing a negotiation. This section aims to provide an illustration and pave the way for more studies of framing in bargaining problems. We start with a simple specification of the problem to highlight the mechanics of how to deal with disagreement about attribute levels. We then move to the more general case that allows for richer preferences over attributes.

3.1.1. A Warm-up Two agents, called the proposer P (she) and the receiver R (he), negotiate over a problem that involves several attributes. Let $L_a = \mathbb{R}$ for all $a \in A$. A specification of these attributes under f defines an item x_f in our model. Let agent j's payoff from x_f be

$$-\sum_{i=1}^{N} \alpha(i)(x_{f(i)} - \bar{x}^j_{f(i)})^2,$$

where $\bar{x}^j_{f(i)} \in \mathbb{R}$ is the bliss point of attribute i for agent j. That is, agents differ only with respect to their bliss points. To avoid trivial cases, we assume that $\bar{x}^P_a \neq \bar{x}^R_a$ for all $a \in A$. We assume that α is strictly decreasing.

The negotiation proceeds as follows. The proposer chooses a frame $f \in F$ and an item x_f to maximize her payoff. The receiver accepts x_f if and only if his payoff is at least as large as the reservation utility $\overline{u}^R \in \mathbb{R}_-$. We assume that the receiver would not accept the proposer's bliss item $\bar{x}^P$ under any frame:

$$\overline{u}^R > \max_{f \in F} \Big\{ -\sum_{i=1}^{N} \alpha(i)(\bar{x}^P_{f(i)} - \bar{x}^R_{f(i)})^2 \Big\}. \tag{2}$$

For now, let the proposer's reservation utility be $\overline{u}^P = -\infty$.

Before solving the model, a few remarks are in order. First, we assume that the agents disagree only in their bliss points, but have otherwise the same preferences. This allows us to focus on how the proposer exploits differences across attributes to frame the problem by emphasizing some attributes over others, which is the core of our paper. This aspect would be obfuscated by differences in the agents' preferences. Second, a literal interpretation of the model is that the proposer chooses the frame and all the attribute levels of an item, offers it to the receiver, and *then* he evaluates the whole item to accept or reject the offer. However, we can also think about the proposer and the receiver as negotiating one attribute at a time in the order specified by the frame, and then again the receiver evaluates the whole proposal at the end. The key premise is that the order of the negotiation affects the receiver's final evaluation as specified by our AF model. Third, the assumption that $L_a = \mathbb{R}$ means that the proposer and receiver are negotiating over a rich set of alternatives, similar to settings where parties bargain on how to share a pie. Finally, one interpretation of why the proposer's payoff depends on framing is that she is a third party who negotiates on behalf of a client and hence internalizes how her client will perceive the outcome based on its presentation.

We start from the proposer's offer given any frame. By standard steps (provided in Appendix A), the optimal level of each attribute is

$$x_{f(i)}(\lambda) = \frac{\lambda \bar{x}^R_{f(i)} + \bar{x}^P_{f(i)}}{\lambda + 1}, \tag{3}$$

$$\lambda = \frac{1}{\sqrt{-\overline{u}^R}} \sqrt{\sum_{i=1}^{N} \alpha(i)(\bar{x}^P_{f(i)} - \bar{x}^R_{f(i)})^2} - 1,$$

where λ is the Lagrange multiplier of the receiver's participation constraint. Note that condition (2) ensures that λ is strictly positive.

Given this, it is easy to see how the proposer will choose to frame her offer. The utilities she gets from the optimal offer is

$$u^P(x_f(\lambda)) = \overline{u}^R \lambda^2.$$

Since $\overline{u}^R$ is negative, the proposer wants to minimize λ. She can do this by choosing a frame that orders the attributes from the one on which the two agents agree the most—i.e., their bliss points are closest—to the one on which they agree the least. That is, the proposer wants to start with and thus emphasize the attributes that generate less disagreement, while leaving the most contentious attributes for the end. This insight is certainly intuitive, but the point here is that our model can generate it while existing models cannot. This insight holds in more general settings with richer functional forms. Yet, the next section shows that other forces also determine optimal frames once we allow attributes to differ in relative importance and frames to affect reservation utilities.

3.1.2. General Case: Important versus Sensitive Attributes The setting is the same as before, except for two aspects. First, agent j's payoff from x_f is now

$$\sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}; \bar{x}^j_{f(i)}), \qquad \text{where} \qquad u_a(x_a; \bar{x}^j_a) = \beta_a - \gamma_a (x_a - \bar{x}^j_a)^2,$$

$\bar{x}_a^j, \beta_a \in \mathbb{R}$, and $\gamma_a > 0$ for all $a \in A$. Again, we interpret $\bar{x}_a^j$ as agent j's bliss point for a and we assume that $\bar{x}_a^P \neq \bar{x}_a^R$ for all $a \in A$. We can interpret $u_a(\cdot; \bar{x}_a^j)$ as a second-order Taylor approximation of a single-peaked function around the bliss point. Importantly, β_a and γ_a can differ across attributes. Note that we can replace each u_a with

$$\hat{u}_a(x_a; \bar{x}_a^j) = \frac{\beta_a}{\sum_{a' \in A} \beta_{a'}} \left[1 - \frac{\gamma_a}{\beta_a}(x_a - \bar{x}_a^j)^2\right] = \hat{\beta}_a[1 - \hat{\gamma}_a(x_a - \bar{x}_a^j)^2],$$

without changing the agents' preferences. Thus, β_a is directly related to the relative *importance* of attribute a for the agents and γ_a to their *sensitivity* to deviations from the bliss point. Finally, α is strictly decreasing as before.

The second generalization is that we allow the receiver's reservation utility to depend on the frame. To this end, let d_i^R be the default level of attribute i, which the receiver will get if he and the proposer do not reach a deal. Then, given f, the receiver's reservation utility is

$$\overline{u}^R(f) = \sum_{i=1}^{N} \alpha(i) \left[\beta_{f(i)} - \gamma_{f(i)}(d_{f(i)}^R - \overline{x}_{f(i)}^R)^2\right].$$

This specification subsumes the special case in which $\bar{u}^R$ does not depend on f if $\beta_a - \gamma_a(d_a^R - \overline{x}_a^R)^2$ takes the same value for all a. We assume the receiver would never accept the proposer's bliss item $\bar{x}^P$, but there are proposals other than the bliss item $\bar{x}^R$ that he would accept. To be specific, we assume that for all $f \in F$

$$\sum_{i=1}^{N} \alpha(i)\beta_{f(i)} > \overline{u}^R(f) > \sum_{i=1}^{N} \alpha(i)[\beta_{f(i)} - \gamma_{f(i)}(\bar{x}^P_{f(i)} - \bar{x}^R_{f(i)})^2]. \tag{4}$$

For now, we again ignore the proposer's reservation utility (i.e., $\overline{u}^P = -\infty$). For the same reasons as before, the agents disagree only in their bliss points, but have otherwise the same preferences (in terms of β_a or γ_a).

The negotiation proceeds as before. We start from the proposer's offer given any frame. As in equation (3), she optimally offers a compromise between bliss points for every attribute. The easier it is to convince the receiver to accept—as captured by a lower λ—the more this compromise caters to the proposer's bliss point. Indeed, $x_f(\lambda)$ also has to satisfy

$$\sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}(\lambda); \bar{x}^R_{f(i)}) = \bar{u}^R(f),$$

which can be written as

$$\sum_{i=1}^{N} \alpha(i)[\beta^R_{f(i)} - \gamma_{f(i)}(x_{f(i)}(\lambda) - \bar{x}^R_{f(i)})^2] = 0, \tag{5}$$

where $\beta_a^R = \gamma_a(d_a^R - \overline{x}_a^R)^2$ for all a. Note that the left-hand side of (5) increases in λ. This is where the choice of f matters, as it can help satisfy (5) and thus lower λ. Moreover, the proposer's problem is equivalent to a situation where the receiver's reservation utility is $\bar{u}^R = 0$, but the relative importance of attribute a for the receiver is β_a^R instead of β_a.

To examine optimal framing, we proceed as follows (see Appendix A for details). Using (3) and (5), we can substitute $x_f(\lambda)$ and λ in the proposer's payoff function and derive a reduced objective that depends only on f:

$$U^P(f) = B(f) - \left[\sqrt{\Gamma(f)} - \sqrt{B^R(f)}\right]^2, \tag{6}$$

where

$$B(f) = \sum_{i=1}^{N} \alpha(i)\beta_{f(i)}, \qquad B^R(f) = \sum_{i=1}^{N} \alpha(i)\beta_{f(i)}^R,$$

$$\Gamma(f) = \sum_{i=1}^{N} \alpha(i)\gamma_{f(i)}(\bar{x}_{f(i)}^P - \bar{x}_{f(i)}^R)^2.$$

Crucially, $U^P(f)$ increases as $\Gamma(f)$ decreases and as $B(f)$ or $B^R(f)$ increases. Thus, the proposer faces a trade-off between emphasizing important attributes (high β_a), attributes involving little disagreement (low $\gamma_a|\bar{x}_a^R - \bar{x}_a^P|$), and attributes for which the receiver has a lot to gain relative to the status quo (high β_a^R). We may then conclude the following.

Proposition 1 (Optimal Framing in Negotiations). *If attributes a and a' satisfy $\beta_a \geq \beta_{a'}$, $\beta_a^R \geq \beta_{a'}^R$, and $\gamma_a|\bar{x}_a^R - \bar{x}_a^P| \leq \gamma_{a'}|\bar{x}_{a'}^R - \bar{x}_{a'}^P|$ with at least one strict inequality, then every optimal frame f presents a before a' (i.e., $f^{-1}(a) < f^{-1}(a')$).*

The takeaway here is that the proposer should present earlier attributes that are important, involve little disagreement, and offer large potential gains for the receiver; she should present later attributes that are unimportant, involve significant disagreement, and offer small potential gains for the receiver. Things are more subtle for important but highly conflictual attributes, for instance, which should be presented in middle positions. The point, however, is that the proposer may strategically de-emphasize some attribute of the negotiation, despite its being very important for her or promising for the receiver, so as to weaken the impact of their disagreement on that attribute. For this to be the case, the gain through $\Gamma(f)$ has to dominate the loss through $B(f)$ and $B^R(f)$.

In our model, framing can also emerge as a tool to break an impasse. Consider again the interpretation of agent P as acting on behalf of a client. Suppose the client's reservation utility is now equal to $\overline{u}^P > -\infty$. By assumption (4), for any frame f the proposer can always find a deal that the receiver accepts. Yet, this deal may be unacceptable for the proposer's client, unless framed in the right way. Consider the generic case in which framing matters:

$$U^P(\underline{f}^P) = \min_{f \in F} U^P(f) < \max_{f \in F} U^P(f) = U^P(\bar{f}^P).$$

Corollary 1 (Breaking the Impasse). *If the reservation utility of the proposer satisfies* $U^P(\underline{f}^P) < \overline{u}^P < U^P(\bar{f}^P)$*, then using frame* $\underline{f}^P$ *leads to an impasse, while using* $\bar{f}^P$ *leads to an agreement.*

Thus, our model captures the common intuition that successful negotiators are those who also have the skill of finding the right way to frame things. Moreover, by being explicit about how framing works, the model offers insights into strategies to break an impasse. Corollary 1 follows from the above characterization of $U^P(f)$ in (6) and the observation that, for every f, the corresponding optimal offer leads the receiver's participation constraint to bind, which implies that only $\overline{u}^P$ determines whether there is an impasse.

Finally, one may wonder whether more susceptibility to framing effects always helps the proposer in negotiations. Consider two proposer-receiver pairs that differ only in α, denoted by α^1 for the first pair and α^2 for the second. Suppose α^1 exhibits more susceptibility to primacy effects than α^2 (Definition 3).

Corollary 2. *Suppose more important attributes also involve less disagreement and larger potential gains for the receiver (i.e., $\beta_a > \beta_{a'}$ if and only if $\gamma_a|\bar{x}_a^R - \bar{x}_a^P| < \gamma_{a'}|\bar{x}_{a'}^R - \bar{x}_{a'}^P|$ and $\beta_a^R > \beta_{a'}^R$ for all $a, a' \in A$). Then, more susceptibility to framing always leads to a higher $U^P(\bar{f}^P)$.*

With this inverse relationship between importance and disagreement across attributes, more susceptibility to framing also always helps the proposer break an impasse between her client and the receiver. But otherwise, more susceptibility to framing can hinder reaching a deal. We illustrate this possibility using the next simple example with two attributes. The key is that the more important attribute also involves more disagreement. The proposer then faces a trade-off when deciding which attribute to present first, and this trade-off can make it harder to find a proposal both her client and the receiver can agree on when they are more susceptible to framing.

Example 2. *There are two attributes, $\hat{a}$ and a', which satisfy $\beta_{\hat{a}} = 0$, $\beta_{a'} = 1$, $\gamma_{\hat{a}}(\bar{x}_{\hat{a}}^R - \bar{x}_{\hat{a}}^P)^2 = 1$, and $\gamma_{a'}(\bar{x}_{a'}^R - \bar{x}_{a'}^P)^2 = 1 + z$ where $z > 0$. For simplicity, suppose reservation utilities are frame independent and that $\overline{u}^P > -\infty$ and $\overline{u}^R = 0$ (this is just a normalization). Slightly abusing notation, let $\alpha(1) = \alpha \in \left(\frac{1}{2}, 1\right)$ and $\alpha(2) = 1 - \alpha$. Thus, more susceptibility to primacy effects here means higher α. Let $\hat{f}$ present first $\hat{a}$ and f' present first a'. Using (6), we obtain*

$$U^P(\hat{f}) \quad = \quad 1 - \alpha - \left[\sqrt{1 + (1-\alpha)z} - \sqrt{1-\alpha}\right]^2$$

$$U^P(f') \quad = \quad \alpha - \left[\sqrt{1+\alpha z} - \sqrt{\alpha}\right]^2.$$

For the proposer's client to agree to a proposal, we need $\max\{U^P(\hat{f}), U^P(f')\} \geq \overline{u}^P$*.*

We will show that there exist z *and* $\overline{u}^P$ *such that a mutually acceptable deal may be reached when susceptibility to framing is weak (low* α*), but not when it is strong (high* α*). This follows from showing that both* $U^P(\hat{f})$ *and* $U^P(f')$ *are strictly decreasing in* α *for sufficiently large* α*. Intuitively, a higher* α *penalizes both putting attribute* a' *first because the already large disagreement between parties is magnified, and putting* a' *last because its importance is diluted. By simple steps,* $\frac{\partial U^P(f)}{\partial \alpha} < 0$ *and* $\frac{\partial U^P(f')}{\partial \alpha} < 0$ *if and only if*

$$\frac{\frac{1}{\alpha} + z}{\sqrt{\frac{1}{\alpha} + z} - 1} < z < \frac{\frac{1}{1-\alpha} + z}{\sqrt{\frac{1}{1-\alpha} + z} - 1}.$$

Note that the left term is decreasing in α*, the right term is increasing in* α*, and*

$$\lim_{\alpha \to 1} \frac{\frac{1}{1-\alpha} + z}{\sqrt{\frac{1}{1-\alpha} + z} - 1} = +\infty.$$

Evaluated at $\alpha = 1$*, the first inequality holds if*

$$\frac{1}{z} + 2 < \sqrt{1+z},$$

hence for sufficiently large but finite z*. Given this* z*, there exists* $\bar{\alpha}$ *such that both derivatives are strictly negative for* $\alpha > \bar{\alpha}$*.*

3.2. Look Good, But Not Too Much: Strategic Framing and Market Competition

In this section, we show that by ordering their attributes—and thus giving each different emphasis—firms can create *fictitious* product differentiation that results in higher profits. Sometimes incumbents can also use framing to deter entry, but doing so involves some trade-offs. Our model allows us to provide insights into when and how firms achieve these outcomes.

We start from a canonical model of vertical differentiation (Tirole (1988), Ch 7.5.1). Each of two firms, the incumbent and the entrant, manufactures a product. Their equal marginal cost is normalized to 0. Entry costs $K > 0$. Each consumer demands one product. The payoff of a product with intrinsic value $v > 0$ and price t is

$$\theta v - t.$$

The taste parameter θ is uniformly distributed across consumers between $1/h$ and $1+1/h$, where we assume $h > 1$. Note that the higher h is, the more heterogeneous the consumers are in relative terms. The payoff of buying nothing is 0.

We modify this model as follows. The products have three attributes: price (p), reliability (r), and build quality (b). The payoff of a product under frame f is

$$\theta[\alpha_f(r)x_r + \alpha_f(b)x_b] - \alpha_f(p)x_p,$$

where $\alpha_f(a) = \alpha(f^{-1}(a))$ for $a \in \{b, p, r\}$. Thus, the products' intrinsic value depends on the level of r and b. The presentation order affects each attribute's weight in the payoff.[11] We continue to assume that only θ differs across consumers, while α, u_b, and u_r are the same. These assumptions allow us to focus on the interaction between framing and vertical differentiation. We will analyze primacy effects: $\alpha(1) > \alpha(2) > \alpha(3)$. The analysis for other forms of α is analogous. To simplify notation, denote the incumbent's and the entrant's product by $I = (I_b, I_p, I_r)$ and $E = (E_b, E_p, E_r)$.[12]

To focus on the role of framing and avoid uninteresting cases, we make the following assumptions. First, each product's reliability and build quality are exogenous. Second, $I_r > E_r > 0$ and $E_b > I_b > 0$. Third, the differences in reliability

[11] Note that we can write this payoff in terms of our AF model in Definition 1 as $\sum_{i=1}^{3} \alpha(i)u_{f(i)}(x_{f(i)}; \theta)$, where $u_r(x_r; \theta) = \theta x_r$, $u_b(x_b; \theta) = \theta x_b$, and $u_p(x_p; \theta) = -x_p$.

[12] We use this lighter notation rather than $x^I = (x_b^I, x_p^I, x_r^I)$ and $x^E = (x_b^E, x_p^E, x_r^E)$.

and build quality between products offset each other:

$$I_r - E_r = E_b - I_b \equiv \delta > 0.$$

This property implies that the products are overall equivalent for consumers not affected by framing (i.e., if α is constant). Thus, in this benchmark case frames play no role and entry would lead to standard Bertrand competition.

The timing is as follows: First, the incumbent chooses f, which also applies to the entrant's product. Second, the entrant decides whether to enter. If it does, the firms compete in prices à la Bertrand; otherwise, the incumbent sets its monopoly price.

Before continuing, it is worth discussing two assumptions. First, we can interpret the exogeneity of attributes in two ways. One is that the engineers of each firm have been able to develop its product to a certain degree by leveraging its comparative advantage and know-how and it is now time for the marketing team to choose how to sell that product. Another interpretation is that we are considering a localized market (e.g., some specific country), while the incumbent and entrant are active on a global scale and have already designed their products' attributes globally.[13] Second, we assume that the incumbent controls how to frame both products because we are

[13]We already know from Tirole (1988), Ch 7.5.1, what incentives vertical differentiation creates for firms' choices of their intrinsic quality. Nonetheless, endogenizing the attributes in the present setting and studying its relation to framing remains an important avenue for future research.

interested in how it can use framing to influence its competitive landscape. Also, since the incumbent is established in the market, it alone may have the resources to run ads that fix f. Section 3.2.1 extends the results by allowing the entrant to choose f for its product.

In this setting, framing allows the incumbent to differentiate its product by emphasizing its superior reliability and de-emphasizing its inferior build quality. This is a realistic and expected strategy, of course. The point is that this strategy is limited by the consumers' susceptibility to framing, and our model allows us to describe and analyze this. Let the difference in intrinsic value between the incumbent's and the entrant's product under f be

$$\delta_f = [\alpha_f(r) - \alpha_f(b)]\delta,$$

which is positive if and only if f presents attribute r before b (i.e., $f^{-1}(r) < f^{-1}(b)$). As in Tirole (1988), we assume that after entry both firms have a positive market share in equilibrium.[14] We begin by characterizing the continuation equilibrium after entry. All proofs for this section appear in Appendix B.

[14] A sufficient condition is that

$$\delta\,(h-1) \leq 3 \min_{f \in F} \frac{\alpha_f(r)E_r + \alpha_f(b)E_b}{|\alpha_f(r) - \alpha_f(b)|},$$

which holds if the products' intrinsic difference or the consumers' heterogeneity is sufficiently small (i.e., δ or h are low).

Lemma 1 (Framing-Driven Differentiation Equilibrium). *Fix f. After entry the equilibrium prices and profits (denoted by π^o for oligopoly) of products x and y satisfy the following properties:*

$$x_p = \frac{|\delta_f|}{3\alpha_f(p)}(2+h^{-1}) \qquad \pi^o(x_f) = \frac{x_p}{3}(2+h^{-1})$$

$$y_p = \frac{|\delta_f|}{3\alpha_f(p)}(1-h^{-1}) \qquad \pi^o(y_f) = \frac{y_p}{3}(1-h^{-1}).$$

If $\delta_f > 0$, then $x = I$ and $y = E$. If $\delta_f < 0$, then $x = E$ and $y = I$.

Lemma 1 offers several insights. First, the differentiation created by framing allows the incumbent to make higher profits, of course by presenting its product as superior to the competitor's product. In particular, the incumbent captures the top of the market (i.e., the consumers with high θ). Thus, by controlling the product frame, firms can not only boost their appeal with all consumers, but also capture the most profitable ones.[15]

A second insight is that de-emphasizing prices can raise profits only if products are differentiated. Suppose products are homogeneous (i.e., $\delta = 0$). Then, even if f presents p at the end, the equilibrium profits are zero—despite consumer

[15] Note that differentiation can depend on framing only if there are at least two attributes other than the price, which is often the case for most products.

heterogeneity in θ. Even if framing nudges them to weigh prices less, Bertrand competition neutralizes this by erasing any profit.

Several papers find consistent evidence about how changing the emphasis on prices affects product choice. In Lynch and Ariely (2000), consumers buy higher quality wine when prices are displayed not alongside product descriptions, but only later at checkout. Also, price elasticities are higher for undifferentiated wines (akin to small δ) independently of price presentation, and when it is harder to notice product differentiation (akin to small $|\delta_f|$). In Blake et al. (2021), postponing purchase fees for concert tickets until checkout induces consumers to buy higher quality tickets and increases revenues from such tickets. In Smith and Brynjolfsson (2001), perceived differences between otherwise homogeneous goods help explain markups and price dispersion in online markets.

Through the lens of primacy effects, postponing prices may be interpreted as akin to obfuscation strategies that weaken price sensitivity by creating search frictions. In Ellison and Ellison (2009), firms endogenously create such frictions to soften price competition and raise markups. Ellison and Ellison (2009) argue that "obfuscation could [...] involve [...] altering [the consumers'] utility functions in a way that raises equilibrium profits," which is what happens in our model. They also find that obfuscation raises the price elasticity for low-quality products, but lowers it for high-quality products. In our model, if $\delta_f > 0$, the price elasticities of the entrant's and incumbent's demand are (see Appendix B)

$$\frac{E_p}{I_p - E_p - \frac{\delta_f}{h\alpha_f(p)}} \qquad \text{and} \qquad \frac{I_p}{(1+h)\frac{\delta_f}{h\alpha_f(p)} - I_p + E_p}.$$

Note that lowering $\alpha_f(p)$ raises the first, but lowers the second. Ellison and Ellison (2009) note that it is hard to know what the elasticities would be absent obfuscation. Our model could provide such counterfactuals given estimates of α.

A third insight of Lemma 1 is that the incumbent creates a positive *framing externality* on the entrant. By making its product "look better," the incumbent weakens the competition from the entrant and charges higher prices. This leaves the bottom consumers exclusively for the entrant, which can then earn a profit. Thus, an incumbent faces a trade-off in emphasizing strengths and de-emphasizing weaknesses of its product. Doing so maximizes its value for all consumers—hence, the monopoly profits. However, it also emphasizes differences from potential competitors, thereby rendering entry more attractive for them. The best framing strategy may then differ between contested (low K) and uncontested (high K) markets.

To characterize the incumbent's optimal framing strategy, we need to know how it ranks frames as a monopolist. As we will see, it suffices to focus on three frames:

i	f^m	f^*	f_*
1	r	r	p
2	b	p	r
3	p	b	b

Letting π^m denote the monopoly profits, we get (see Lemma B.B.1 in Appendix B)

$$\pi^m(I_{f^m}) > \pi^m(I_{f^*}) > \pi^m(I_{f_*}).$$

We will focus on settings where the incumbent always prefers to remain a monopolist: $\pi^m(I_{f^m}) > \max_{f\in F} \pi^o(I_f)$, which holds if the products' intrinsic difference δ or the consumers' heterogeneity h is sufficiently small.[16] A monopolist simply uses framing to emphasize what its product delivers and de-emphasize what one has to pay for it.

By contrast, under the threat of competition framing becomes a tool to emphasize strengths and de-emphasize weaknesses *relative* to the entrant. This tool can be used to make the market less attractive to entry. Thus, the optimal strategy is more nuanced and depends on how consumers respond to framing. Define

$$\overline{\alpha}(2) = \frac{[\alpha(1)]^2 + [\alpha(3)]^2}{\alpha(1) + \alpha(3)} \qquad \text{and} \qquad \underline{\alpha}(2) = \alpha(1) - \alpha(3),$$

[16]Indeed, $\pi^m(I_{f^m}) > \max_{f\in F} \pi^o(I_f)$ is equivalent to

$$\frac{\alpha(1)I_r + \alpha(2)I_b}{\alpha(3)} > \delta\frac{4}{9}\Big(\frac{2h+1}{h+1}\Big)^2 \max_{f\in F} \frac{|\alpha_f(r) - \alpha_f(b)|}{\alpha_f(p)}. \tag{7}$$

which satisfy $\overline{\alpha}(2) > \underline{\alpha}(2)$. We first characterize the cases where framing can never help deter entry: Either entry is not a threat and the incumbent uses f^m, or entry is inevitable and the incumbent uses f^m or f^*.

Proposition 2 (Optimal Framing without Entry Deterrence).
If $K > \pi^o(E_{f^m})$, *the incumbent chooses frame* f^m *and remains a monopolist. If* $K \leq \min\{\pi^o(E_{f^m}), \pi^o(E_{f_*})\}$, *the incumbent cannot deter entry; it chooses* f^m *if* $\alpha(2) < \underline{\alpha}(2)$ *and* f^* *if* $\alpha(2) > \underline{\alpha}(2)$.

Note that $\pi^o(E_{f_*}) < \pi^o(E_{f^*})$, while $\pi^o(E_{f_*}) > \pi^o(E_{f^m})$ if and only if $\alpha(2) > \overline{\alpha}(2)$.

When deterring entry is impossible, the incumbent presents its strengths first, but may present its price *before* its weaknesses. If consumers underweight the second attribute only a little ($\alpha(2) > \underline{\alpha}(2)$), the incumbent is forced to present its weakness after its price to optimally differentiate its product from the entrant's—which constrains its ability to weaken price elasticity in its market segment. If instead consumers underweight a lot the second attribute ($\alpha(2) < \underline{\alpha}(2)$), the incumbent can effectively de-emphasize both its weakness and price, thus presenting the price last.

Next, we describe when the incumbent uses framing to deter entry. This always involves frames that are suboptimal from the monopolist's viewpoint. The incumbent shows its strengths before its weaknesses, but may again emphasize its price by presenting it earlier—even first. In so doing, the incumbent forgoes some of its appeal to all consumers in exchange for saving its monopolistic position, by rendering the market less attractive for the entrant. For instance, by emphasizing its price, the incumbent intentionally makes the consumers more price sensitive, which renders it harder for the entrant to compete. More generally, framing is used to *weaken* the power of differentiation, should entry occur.

PROPOSITION 3 (Optimal Framing with Entry Deterrence).

I: *If* $\overline{\alpha}(2) > \alpha(2) > \underline{\alpha}(2)$*, then* $\pi^o(E_{f^*}) > \pi^o(E_{f^m}) > \pi^o(E_{f_*})$*. In this case, if* $\pi^o(E_{f^m}) \geq K > \pi^o(E_{f_*})$*, the incumbent chooses* f_* *and remains a monopolist when the products' intrinsic difference* δ *or the consumers' heterogeneity* h *is sufficiently small.*[17] *Otherwise, it chooses* f^* *and the entrant enters.*

II: *If* $\alpha(2) < \underline{\alpha}(2)$*, then* $\pi^o(E_{f^m}) > \pi^o(E_{f^*}) > \pi^o(E_{f_*})$*. In this case, we have that*

- *if* $\pi^o(E_{f^m}) \geq K > \pi^o(E_{f^*})$*, the incumbent chooses* f^* *and remains a monopolist when* δ *or* h *is sufficiently small;*
- *if* $\pi^o(E_{f^*}) \geq K > \pi^o(E_{f_*})$*, the incumbent chooses* f_* *and remains a monopolist when* δ *or* h *is sufficiently small;*
- *otherwise, the incumbent chooses* f^m *and the entrant enters.*

Note that the incumbent presents its price first only when this successfully deters entry: This strategy forces the entrant to suffer strong price competition if it enters, but does not punish the incumbent excessively if it retains its monopoly.

[17] The proof expresses this and the following conditions on δ and h as precise inequalities that they have to satisfy.

To better understand the effects of competition and the incumbent's framing responses, we compare the total industry profits under the monopoly regime and the potential entry regime.[18] If the incumbent successfully deters entry but has to use a frame different from f^m, its profits and, hence, the industry profits must fall relative to no competition. Things are more subtle when entry occurs. If the products' intrinsic difference δ or susceptibility to framing are weak, competition can clearly lead to lower total profits. Recall that in this industry total profits would be zero without framing effects as in standard Bertrand settings. In some cases, however, the incumbent's framing can allow the industry to achieve higher profits than under monopoly.

Corollary 3 (Profits Comparison). *Suppose the incumbent chooses frame f^m and entry occurs. Then, the difference in industry profits*

$$\pi^o(I_{f^m}) + \pi^o(E_{f^m}) - \pi^m(I_{f^m})$$

is positive if susceptibility to framing is sufficiently strong (i.e., $\alpha(2)$ is sufficiently smaller than $\alpha(1)$) and the firms' comparative advantages are sufficiently large (i.e., E_r and I_b are sufficiently small).

[18] We thank an anonymous referee for suggesting this comparison.

The reason is that, under these conditions, frame f^m creates a strong—albeit fictitious—differentiation between the two firms, giving them enough market power in their individual segments to reach higher total profits.

We conclude with how the incumbent's framing strategy depends on the primitives of the market, in particular the consumers' susceptibility to framing.

PROPOSITION 4 (Comparative Statics).

– *The incumbent is more likely to remain a monopolist and to use framing to deter entry when intrinsic product difference δ or consumer heterogeneity h is smaller.*
– *A weaker susceptibility to primacy effects implies that $\pi^o(E_{f^m})$ and $\pi^o(E_{f^*})$ are lower and that the incumbent is more likely to use f_* to deter entry. Otherwise, it has ambiguous effects on $\pi^o(E_{f_*})$ and the incumbent's use of f^* to deter entry.*

If the differentiation allowed by framing is smaller due to lower δ, the post-entry market is more competitive and less profitable. Thus, entry has to cost less to be a threat. The incumbent also has more to lose and so is more willing to deter entry, even if this requires forgoing some monopoly profit. A lower h has similar effects, as it curbs the benefits of splitting the market between the top and bottom consumers.

Optimal framing depends in more intricate ways on the consumers' susceptibility to primacy effects. Weakening it curbs the frames' ability to create fictitious differentiation—lowering post-entry profits—but also to deter entry. Either way, weaker primacy effects can render entry less likely, as frames are less effective at stifling competition after entry *and* doing so benefits entrants less.

Our results offer some novel insights into advertisement. These complement the view that its function is to provide information about available products to consumers who have fixed tastes. Here, we keep that information fixed and change *how* it is framed, which is an important part of advertisement. The discussed benefits of controlling frames suggest another reason why firms seek to be presented in prominent positions to consumers (like in web searches or e-commerce stores). The logic of our results is also related to the so-called *pioneering advantage*: Carpenter and Nakamoto (1989) find a gap between the market shares of pioneers and later entrants that cannot be explained by switching costs and seems to arise from the process whereby consumers form their preferences.

As a final note, the main insights of this section would carry over to settings with more than two attributes (in addition to the price). Having more attributes would give the incumbent more ways to render the market less attractive for entry, that is, to manipulate the fictitious differentiation δ_f. This means that, in such settings, the incumbent may still prefer *not* to present its price last if there are weaknesses that can be de-emphasized more to deter entry, without losing as much potential profit. At the same time, the incumbent may not have to present its price first so as to deter entry.

3.2.1. Extensions We consider two extensions of the baseline model: In the first, the entrant also chooses a frame; in the second, some consumers are unaffected by framing.[19]

Both Firms Choose Frames. Let f^I and f^E be the incumbent's and entrant's frames. Suppose the consumers compare products using either frame before buying one. That is, we assume each consumer uses f^I with probability μ and f^E with probability $1-\mu$, where $0.5 \leq \mu \leq 1$. Equivalently, we can view μ and $1-\mu$ as the share of consumers who use f^I and f^E. One interpretation is that each consumer uses whichever frames he or she encounters first (as under hypothesis H_3 in Online Appendix F), and this is more likely to be the incumbent's frame.[20] The timing is as follows:

1. The incumbent chooses f^I.
2. The entrant decides whether to enter at cost $K > 0$.
3. If the entrant enters, it chooses f^E.
4. If the entrant enters, the firms compete in prices knowing f^I and f^E; otherwise, the incumbent sets its monopoly price.
5. Consumers make their comparisons (if any) and purchase decisions.

---

[19] We thank anonymous referees for suggesting these extensions.

[20] For similar ways of modeling which frame dominates in competitive settings, see Piccione and Spiegler (2012) and Spiegler (2014).

The ability of the entrant to influence which frame consumers use complicates the analysis. However, we will show that the predictions and insights of the previous section are robust for μ close to 1. We interpret high levels of μ as capturing a defining feature of incumbents, namely a long presence in the market that gives them strong clout on consumers' perception of products.

The choice of a frame for the entrant is actually simple. This choice is irrelevant if the consumers use f^I. If instead they use f^E, the entrant clearly wants to make its product look as good as possible. That is, it faces the same situation as the incumbent when entry cannot be deterred. Therefore, the optimal frame of the entrant is either f^{Em} with attribute order (b, r, p) or f^{E*} with attribute order (b, p, r), by the same logic that makes f^m or f^* optimal for the incumbent given entry.

The complications arise at the pricing stage conditional on entry. When $\mu = 1$, the incumbent can guarantee for itself the top of the market by choosing a frame that emphasizes its superior attributes relative to the entrant's product. As a result, the incumbent can safely charge a high price and force the entrant to charge a low price. With $\mu < 1$, however, the incumbent can end up in a situation where consumers perceive its product as dominated by the entrant's if they adopt f^E, and will buy only from the entrant if its price is lower than the incumbent's price. This creates an incentive for the incumbent to lower its price and for the entrant to raise its price, which could reverse who charges the higher price. This incentive remains weak for sufficiently high μ, which allows us to conclude the following.

Proposition 5. *Fix the entrant's optimal frame f^E. Generically with respect to the other parameters of the model, there exists $\underline{\mu} < 1$ such that if $\mu > \underline{\mu}$ the following holds:*

- *the incumbent's equilibrium choice of its frame $f^I \in \{f^m, f^*, f_*\}$ is the same as when $\mu = 1$;*
- *the incumbent deters entry for intermediate entry cost K if and only if it deters entry at $\mu = 1$;*
- *conditional on entry, the incumbent charges a high price and the entrant charges a low price (i.e., $I_p > E_p$).*

The intuition is that, given $f^I \in \{f^m, f^*, f_*\}$, $f^E \in \{f^{Em}, f^{E*}\}$, and entry, the equilibrium prices are unique in a neighborhood of $\mu = 1$. Hence, by standard results they change continuously in μ, which renders the continuation payoff of the incumbent from each $f^I \in \{f^m, f^*, f_*\}$ also continuous in μ. Finally, one frame in $\{f^m, f^*, f_*\}$ strictly dominates the others for the incumbent when $\mu = 1$ generically.

Frame-insensitive Consumers. One may wonder how our results change if different consumers respond to framing in different ways. In particular, suppose that there is a share of consumers who are insensitive to framing (i.e., their function α is constant). Since such consumers view the two firms' products as delivering the same total value (given $I_r - E_r = E_b - I_b$), they will always buy the cheaper of the two, joining the bottom segment of the market. Thus, frame-insensitive consumers boost the demand for the entrant if it undercuts the incumbent's price. But if they are relatively few in the market, this boost is not enough to change our qualitative results. As long as they represent a sufficiently small share of the overall population of consumers, the incumbent and the entrant continue to have the same incentives to use framing to create fictitious differentiation so as to dampen competition. The incumbent can still use its ability to set the dominant frame to prevent entry as discussed above.

4. Axiomatizations and Extensions

4.1. Behavioral Characterization of AF Models

To characterize our model, we enrich the choice domain by allowing for (simple) *lotteries* over items. This provides enough structure for us to identify the weights α in an intuitive and direct calibration exercise. This identification is essential to then characterize the behavioral properties of the model, such as primacy and recency effects. The idea is that each item involves some risk: Its attributes are presented in a specific order, but (the consequences of) their levels can be uncertain at the time of choice. For instance, when choosing between a new sedan or a used SUV, a buyer may not know which will better serve his needs.[21]

We will rely only on lotteries whose support involves items all framed in the same way. Such lotteries belong to $\Delta(X_f)$ and are denoted by p_f, q_f, and r_f. To simplify notation, we denote lotteries that yield x_f with probability p and y_f with probability

[21] Another possibility is to use a consumption space $\mathbb{R}^N$ as the choice domain. While this space shares some features with the lottery domain, it unnecessarily complicates the axiomatization. The main reason is that the characterization of α will have to work through the attribute utilities u and, hence, will depend on assumptions about u.

$1 - p$ by

$$(x_f, y_f; p)$$

An f-menu M_f is a subset of $\Delta(X_f)$. We assume that $|A| = N \geq 3$; one can allow for $N = 2$ at the cost of stronger separability axioms.

As the primitive data, we assume to observe the decision-maker's choices from all f-menus. This choice set is denoted by $c(M_f)$ and has the usual interpretation. Note that the frames of each item in a menu are part of the dataset.

Our basic assumption is that we can describe the decision-maker's choices from f-menus as the maximization of some utility function that can depend at most on f. We go one step further and assume that she is an expected-utility maximizer. We present these properties directly as an assumption because they follow from standard axioms.

Assumption 1 (f-EU Representation). *For every $f \in F$, there exists a function $w_f : X_f \to \mathbb{R}$ such that for every M_f*

$$c(M_f) = \arg\max_{q_f \in M_f} \; v_f(q_f), \qquad \textit{where} \qquad v_f(q_f) = \sum_{x_f \in supp\, q_f} w_f(x_f) q_f(x_f).$$

To characterize our AF model, we need to find properties of c that correspond to each w_f taking the form

$$w_f(x_f) = \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)})$$

for some $\alpha : \{1, \dots, N\} \to \mathbb{R}_{++}$ and $u_a : L_a \to \mathbb{R}$ for all $a \in A$. We organize these properties in four axioms.

Axiom 1 is a simple non-triviality condition: For no attribute, the decision-maker is indifferent between all its levels. To formalize this, let $x_{f(-i)}$ be the description of item x_f excluding position i.[22]

Axiom 1 (Non-triviality). *For every $a \in A$, there exists $x_a, y_a \in L_a$ such that, if $f(1) = a$ and $x_{f(-1)} = y_{f(-1)}$, then $c(x_f, y_f) = \{x_f\}$.*

Axiom 2 is inspired by standard separability axioms as in Debreu (1960): How the decision maker trades off the levels of any two attributes does not depend on the levels of other attributes. We relax this in Section 4.2.

Axiom 2 (Separability). *Fix $f \in F$ and any $j, k \in \{1, \dots, N\}$. For all x_f, x'_f, y_f, y'_f that satisfy $x_{f(i)} = y_{f(i)}$ and $x'_{f(i)} = y'_{f(i)}$ for $i = j, k$ and $x_{f(i)} = x'_{f(i)}$ and $y_{f(i)} = y'_{f(i)}$*

[22] We write $c(p_f, \dots, q_f)$ for $c(\{p_f, \dots, q_f\})$ to simplify notation.

for all $i \neq j, k$, *we have*

$$c(x_f, x'_f) = \{x_f\} \Leftrightarrow c(y_f, y'_f) = \{y_f\}.$$

Axiom 3 captures the property that the decision-maker's tastes for each attribute do not depend on the position in which the attributes are presented.

Axiom 3 (Taste Framing Independence). *For every* $i, j = 1, 2, \ldots, N$, $a \in A$, *and* $f, f' \in F$ *such that* $f(i) = a$ *and* $f'(j) = a$, *the following holds: If* $p_{f(i)}, q_{f(i)} \in \Delta(L_{f(i)})$, $\hat{p}_{f'(j)} = p_{f(i)}$, $\hat{q}_{f'(j)} = q_{f(i)}$, $x_{f(-i)} = y_{f(-i)}$, *and* $\hat{x}_{f'(-j)} = \hat{y}_{f'(-j)}$, *then*

$$c((p_{f(i)}, x_{f(-i)}), (q_{f(i)}, y_{f(-i)})) = \{(p_{f(i)}, x_{f(-i)})\} \Leftrightarrow c((\hat{p}_{f'(j)}, \hat{x}_{f'(-j)}), (\hat{q}_{f'(j)}, \hat{y}_{f'(-j)})) = \{(\hat{p}_{f'(j)}, \hat{x}_{f'(-j)})\}.$$

Axiom 4 exploits the cardinality of expected utility to identify how the decision-maker weights attributes based on their presentation. To this end, we need to find an observable way to elicit how changing the presentation position of attributes affects how the decision-maker trades them off. The idea is to consider three items—x, y, and z—that differ in only two attributes. For x, both attributes are "good;" for y, both attributes are "bad;" for z, one attribute is "good" and the other is "bad." Given this, consider a lottery between x and y that the decision-maker deems indifferent to z. This lottery reveals how she trades off the "good" and "bad" attributes in z.[23] Formally, for every $a \in A$ we say that x_a is strictly preferred to y_a—written $x_a \succ y_a$—if $c(x_f, y_f) = \{x_f\}$ whenever $f(1) = a$, $x_{f(1)} = x_a$, $y_{f(1)} = y_a$, and $x_{f(-1)} = y_{f(-1)}$.

DEFINITION 4 (Calibration Lottery). *Fix any* $i, j \in \{1, \ldots, N-1\}$. *Let* x_f, y_f, *and* z_f *be such that* $x_{f(i)} = z_{f(i)} \succ y_{f(i)}$, $x_{f(j)} \succ z_{f(j)} = y_{f(j)}$, *and* $x_{f(k)} = z_{f(k)} = y_{f(k)}$ *for* $k \neq i, j$. *Then, define* p_{xyz_f} *as the indifference probability that satisfies*

$$\{(x_f, y_f; p_{xyz_f}), z_f\} = c((x_f, y_f; p_{xyz_f}), z_f).$$

Intuitively, these calibration lotteries may depend on the presentation positions of the two different attributes if order frames affect the decision-maker. However, if these effects take the weighting structure of our AF model, this dependence has to occur in a specific way, formalized by our last and key axiom. Note that, when assessing a calibration lottery, the decision-maker compares the *relative likelihood* of getting the "good" attributes of x and the "bad" attributes of y vis-à-vis the trade-off between the "good" and "bad" attributes in different positions of z.

[23] Note that such a lottery necessarily exists because of the expected-utility structure of the model. This is also true generally for continuous models.

Axiom 4 (Position Dependence). *Let $f, \hat{f} \in F$ satisfy $f(1) = \hat{f}(1)$ and $f(j) = \hat{f}(j')$. Let p_{xyz_f} and $p_{xyz_{\hat{f}}}$ be the probabilities in the calibration lottery just defined with $i = 1$. If $x_{f(j)} = x_{\hat{f}(j')}$, $y_{f(j)} = y_{\hat{f}(j')}$, and $z_{f(j)} = z_{\hat{f}(j')}$, then the ratio*

$$\frac{p_{xyz_f}}{1 - p_{xyz_f}} \Big/ \frac{p_{xyz_{\hat{f}}}}{1 - p_{xyz_{\hat{f}}}}$$

can depend only on i and j.

Axiom 4 surely imposes significant structure on c. However, note that a standard model without framing effects is characterized by the more stringent condition $p_{xyz_f} = p_{xyz_{\hat{f}}}$.

Theorem 1 (AF Representation). *Under Assumption 1, Axioms 1–4 hold if and only if c has an AF representation: There exists $\alpha : \{1, \ldots, N\} \to \mathbb{R}_{++}$ and non-constant $u_a : L_a \to \mathbb{R}$ for every $a \in A$ that satisfy, for all $f \in F$ and M_f,*

$$c(M_f) = \underset{x_f \in M_f}{\arg\max} \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}).$$

Proposition 6 (Uniqueness). *Suppose (α, u) and $(\hat{\alpha}, \hat{u})$ describe two AF models. Then, (α, u) and $(\hat{\alpha}, \hat{u})$ represent the same choice function $c(\cdot)$ if and only if there* *exists scalars* $\chi > 0$, $\chi_a > 0$, *and* $\zeta_a \in \mathbb{R}$ *such that* $\hat{\alpha} = \chi\alpha$ *and* $\hat{u}_a = \chi_a u_a + \zeta_a$ *for all* $a \in A$.

The proofs of this section appear in Appendix C, which also shows that the lotteries p_{xyz_f} offer a method to directly identify α without requiring any assumptions about the attribute-specific utility functions. For some applications, it may be useful to know how to identify the model parameters without having to rely on lotteries—for instance, because the available data involves items in a standard consumption space. Appendix D outlines this alternative identification. Another possibility is to rely on random-choice data as in the analysis of Section 4.2, which also does not use calibration lotteries.

4.1.1. Behavioral Characterization of Primacy and Recency Effects We now characterize the attribute-framing effects and the comparison between decision-makers in Definitions 2 and 3. To this end, we define primacy and recency effects in terms of observable data using our calibration lotteries.

Definition 5 (Revealed Primacy/Recency Effects). *For* $i = 1, \ldots, N-1$, *define* $p^i_{xyz_f}$ *and* $p^i_{xyz_{f'}}$ *as the probabilities of the calibration lotteries such that* $j = i+1$ *and* f' *swaps only the attributes in position* i *and* $i+1$. *Then,* c *exhibits primacy (recency) effects if*

$$p^i_{xyz_f} > (<)\ p^i_{xyz_{f'}} \quad \textit{for all } i = 1, \ldots, N-1.$$

Intuitively, z_f dominates y_f in an earlier attribute, while $z_{f'}$ dominates $y_{f'}$ in a later attribute. Thus, $p^i_{xyz_f}$ should be higher (lower) than $p^i_{xyz_{f'}}$ if the decision-maker is affected by primacy (recency) effects.

PROPOSITION 7. *Let (α, u) be an AF representation of c. Then, c exhibits a primacy (recency) effect if and only if α is strictly decreasing (increasing).*

We now turn to comparing individuals' susceptibility to attribute framing. For this to be meaningful, we have to compare individuals who exhibit the same tastes for attributes.

DEFINITION 6 (Revealed Same Tastes). *Decision-maker 1 and 2 exhibit the same tastes for attributes if c^1 and c^2 have the following property. For every $a \in A$, $f \in F$ that satisfies $f(1) = a$, and $p_{f(1)}, q_{f(1)} \in \Delta(L_{f(1)})$,*

$$c^1((p_{f(1)}, x_{f(-1)}), (q_{f(1)}, y_{f(-1)})) = c^2((p_{f(1)}, x_{f(-1)}), (q_{f(1)}, y_{f(-1)})).$$

This explains Definition 3 because if c^1 and c^2 have this property, u^1_a and u^2_a represent the same vN-M preference over $\Delta(L_a)$, hence $u^1_a = \gamma_a u^2_a + \zeta_a$ for $\gamma_a > 0$ and $\zeta_a \in \mathbb{R}$.

Even if two decision-makers have the same tastes for attributes, they may still trade off attributes differently. This complicates isolating the effects due exclusively to framing. To this end, let $p^{1i}_{xyz_f}$ and $p^{1i}_{xyz_{f'}}$ be the probabilities of the calibration lotteries in Definition 5 for decision-maker 1; define $p^{2i}_{xyz_f}$ and $p^{2i}_{xyz_{f'}}$ similarly for decision-maker 2. We know that for both primacy and recency effects there exist scalars $\lambda^{1i}_{xyz_f}$ and $\lambda^{2i}_{xyz_f}$ that satisfy

$$p^{1i}_{xyz_{f'}} = (1 + \lambda^{1i}_{xyz_f}) \cdot p^{1i}_{xyz_f} \quad \text{and} \quad p^{2i}_{xyz_{f'}} = (1 + \lambda^{2i}_{xyz_f}) \cdot p^{2i}_{xyz_f},$$

where the scalars are positive (negative) for primacy (recency) effects. These scalars λ directly capture how much one has to adjust the calibration lotteries when postponing an attribute that the decision-maker likes better so as to maintain indifference. Then, intuitively, a decision-maker who is more susceptible to framing effects should require stronger adjustments—appropriately normalized.

Definition 7 (Revealed Comparative Primacy/Recency). *Suppose decision-makers 1 and 2 exhibit the same tastes for attributes. Then, decision-maker* 1 *is more susceptible to primacy (recency) effect than decision-maker 2 is if, for all* $i = 1, \ldots, N-1$,

$$\frac{\lambda^{1i}_{xyz_f}}{1 - p^{1i}_{xyz_f}} \geq (\leq) \frac{\lambda^{2i}_{xyz_f}}{1 - p^{2i}_{xyz_f}}.$$

The next result maps these behavioral comparisons into properties of our AF representation, thereby providing the foundations for Definition 3.

PROPOSITION 8. *Suppose decision-makers 1 and 2 can be represented by AF models* (α^1, u^1) *and* (α^2, u^2) *and exhibit the same tastes for attributes. Decision-maker* 1 *is more susceptible to primacy (recency) effect than decision-maker 2 is if and only if*

$$\frac{\alpha^1(i)}{\alpha^1(i+1)} \geq (\leq) \frac{\alpha^2(i)}{\alpha^2(i+1)}, \quad i = 1, \ldots, N-1.$$

Given the structure of any AF model, this tight characterization follows immediately from observing that

$$\frac{\alpha(i)}{\alpha(i+1)} = \sqrt{\frac{\lambda^i_{xyz_f}}{1 - p^i_{xyz_f}} + 1}.$$

4.2. Framing without Separability

In this section, we relax the additive structure of AF models. To this end, it helps to transition to a random-choice framework due to its additional structure. This allows us to develop a model that can be more easily applied to empirical analysis, which often examines framing in terms of how it affects the probability of choosing an item. In the literature on random choice, several papers include observable attributes (Lancaster (1966), McFadden (1973), Gorman (1980), Allen and Rehbeck (2023)).

We also include their framing as part of the dataset.[24] For every finite $M_f \subset \Delta(X_f)$, we assume to observe the probability that the decision-maker chooses each $q_f \in M_f$, denoted by

$$\varphi(q_f, M_f).$$

This has the usual interpretation of the random-choice literature.[25]

We again start by assuming that we can describe choices from f-menus using a standard general model. We use a canonical Luce representation and later consider more general models of random choice. We again introduce this representation directly as an assumption, because it follows from well-known axioms.

Assumption 2 (f-EU Luce Representation). *For every $f \in F$, there exists a function $w_f : X_f \to \mathbb{R}$ such that for every finite $M_f \subset \Delta(X_f)$*

$$\varphi(q_f, M_f) = \frac{e^{v_f(q_f)}}{\sum_{q'_f \in M_f} e^{v_f(q'_f)}}, \qquad \textit{where} \qquad v_f(q_f) = \sum_{x_f \in supp\, q_f} w_f(x_f) q_f(x_f). \quad (8)$$

[24] Gul et al. (2014) propose a related, but different, approach where the decision-maker *subjectively* frames multi-attribute items. Their elegant analysis identifies how she treats items as more or less substitutes based on the subjective similarity of attributes. This approach is silent about the role of objective and exogenous frames. It seems possible that exogenous and subjective frames interact, opening an interesting connection between our and their work. For a study of how ordering of *alternatives* might affect choice in the Luce model, see Tserenjigmid (2021).

[25] See, e.g., Luce (1959), Block and Marschak (1960), Marschak (1974), Gul and Pesendorfer (2006), Manzini and Mariotti (2014), and Apesteguia and Ballester (2018).

The basic premise of this paper is that people often encounter attributes of items in an exogenous order and this may affect their choices. One way to keep this premise while relaxing additivity is to allow the weight a decision maker assigns to an attribute to depend on its presentation position as well as the attributes that come before it. That is, this decision maker—let's call her Ann—may aggregate the utilities across attributes as follows:

$$w_f(x_f) = \sum_{i=1}^{N} u_{f(i)}(x_{f(i)}) Q(i, x_{f(i-1)}, \ldots, x_{f(1)}).$$

On practical grounds, it is valuable to impose more structure on the dependence of Q on earlier attributes. We therefore introduce and characterize the form

$$Q(i, x_{f(i-1)}, \ldots, x_{f(1)}) = \alpha(i) \exp\left\{ \sum_{k=1}^{i-1} \varphi_k(u_{f(k)}(x_{f(k)})) \right\}, \tag{9}$$

where $\alpha : \{1, \ldots, N\} \to \mathbb{R}_{++}$, $\varphi_i : \mathcal{U} \to \mathbb{R}$ for all $i = 1, \ldots, N$ with $\mathcal{U} = \cup_{a \in A} u_a(L_a)$ (i.e., the union of the ranges of all attributes' utility functions), and by convention $\sum_{k=1}^{0} \varphi_k(u_{f(k)}(x_{f(k)})) \equiv 0$. We refer to this model by the triplet (α, u, φ), where

$u = (u_a)_{a \in A}$ and $\varphi = (\varphi_i)_{i=1}^N$. If each φ_i is constant, we obtain a random-choice version of our AF model.[26]

The idea behind expression (9) is that the *utility* from attributes presented earlier affects the weight assigned to later attributes. The first impression left by early attributes matters also because it affects the responsiveness to later impressions. For example, suppose each φ_k is decreasing. Then, the more Ann likes early attributes, the less she weighs later attributes. Put differently, she may underweight later attributes not just because they come later, but also because earlier attributes are already pretty good. Other possible interpretations are that Ann pays less attention to later attributes if earlier ones are good enough; if instead early attributes are not good, she may look for reasons to like an item by carefully inspecting later attributes (decreasing φ_k), or she may lose interest (increasing φ_k). In this way, the model allows for smooth forms of satisficing across attributes.

Our characterization involves four axioms. First, suppose items x and y differ only in attribute a and Ann prefers x_a to y_a. Then, we would expect that she also prefers x to y—in probabilistic terms, she is more likely to choose x than y—no matter what the frame is.

Axiom 5 (Attribute Monotonicity). *For every $i = 1, \ldots, N-1$, $a \in A$, and $f, f' \in F$ such that $f(N) = a$ and $f'(i) = a$, the following holds: If $x_{f(N)} = \hat{x}_{f'(i)} = x_a$,*

[26]The model defined by (9) is related to Epstein (1983), to which we owe significant inspiration for our characterization.

$y_{f(N)} = \hat{y}_{f'(i)} = y_a$, $x_{f(-N)} = y_{f(-N)}$, and $\hat{x}_{f'(-i)} = \hat{y}_{f'(-i)}$, then

$$\varphi(x_f|\{x_f, y_f\}) \geq \frac{1}{2} \quad \Rightarrow \quad \varphi(\hat{x}_{f'}|\{\hat{x}_{f'}, \hat{y}_{f'}\}) \geq \frac{1}{2}.$$

This intuitive property rules out some predictions that are possible under expression (9) without further restrictions, but are highly unrealistic. If attribute a appears in position $k < N$ and φ_k is decreasing, the better x_a reduces more the weight Ann assigns to later attributes than does y_a. If this effect is strong enough, Ann's overall value of x may be smaller than that of y, leading her to choose y more often. Such violations of simple dominance seem implausible.

Axiom 6 considers the comparison of items whose attributes are identical *up to* some position i. It states that the levels of such attributes do not affect how Ann trades off the attributes after position i. Given any x_f, let $x_f^i = (x_{f(1)}, \ldots, x_{f(i)})$. Note that for $p_f \in \Delta(\times_{k=i+1}^N L_{f(k)})$, the object (x_f^i, p_f) defines a lottery in $\Delta(X_f)$.

Axiom 6 (Common-Root Independence). *Fix any $f \in F$ and $i = 1, \ldots, N-1$. For all (x_f^i, p_f), (y_f^i, p_f), (x_f^i, q_f), and (y_f^i, q_f) in $\Delta(X_f)$, we have*

$$\varphi((x_f^i, p_f)|\{(x_f^i, p_f), (x_f^i, q_f)\}) = \varphi((y_f^i, p_f)|\{(y_f^i, p_f), (y_f^i, q_f)\}).$$

Axiom 7 considers the comparison of items whose attributes are identical *after* some position. It requires that how these identical attributes are ordered does not affect Ann's choice.

Axiom 7 (Tail Frame Invariance). *Fix $i \geq 2$ and any $f, f' \in F$ that satisfy $f(k) = f'(k)$ for $k \leq i-1$. Let x_f and y_f satisfy $x_{f(k)} = y_{f(k)}$ for $k \geq i$. Let $\hat{x}_{f'}$ and $\hat{y}_{f'}$ satisfy $x_{f(k)} = \hat{x}_{f'(k)}$ and $y_{f(k)} = \hat{y}_{f'(k)}$ for $k \leq i-1$ and $x_{f'(k)} = \hat{x}_{f'(k)}$ and $y_{f'(k)} = \hat{y}_{f'(k)}$ for $k \geq i$. Then, the following holds*

$$\varphi(x_f|\{x_f, y_f\}) = \varphi(\hat{x}_{f'}|\{\hat{x}_{f'}, \hat{y}_{f'}\}).$$

Finally, Axiom 8 exploits comparisons between frames to identify their effects. It allows the effect of postponing an attribute in the presentation order to depend on the level of the attributes that precede it.

Axiom 8 (Lasting Impressions). *For all $f, f' \in F$ that satisfy $f(i) = f'(1)$ for $i \neq 1$, the following holds: If $x_{f(i)} \neq \hat{x}_{f(i)}$, $x_{f(i)} = y_{f'(1)}$, $\hat{x}_{f(i)} = \hat{y}_{f'(1)}$, $x_{f(-i)} = \hat{x}_{f(-i)}$, $y_{f'(-1)} = \hat{y}_{f'(-1)}$, and $\varphi(y_{f'}, \{y_{f'}, \hat{y}_{f'}\}) \neq \varphi(\hat{y}_{f'}, \{y_{f'}, \hat{y}_{f'}\})$, then*

$$\frac{\varphi(x_f, \{x_f, \hat{x}_f\})}{\varphi(\hat{x}_f, \{x_f, \hat{x}_f\})} \Big/ \frac{\varphi(y_{f'}, \{y_{f'}, \hat{y}_{f'}\})}{\varphi(\hat{y}_{f'}, \{y_{f'}, \hat{y}_{f'}\})}$$

can depend at most on i and $(x_{f(1)}, \ldots, x_{f(i-1)})$.

This is where we relax additive separability. In fact, if we required the ratio in Axiom 8 to depend at most on i, we would obtain that each φ_i is constant and hence a random-choice version of our AF model. In this case, the choice probabilities in Axiom 8 can also be used to directly identify the decision-maker's α: Normalizing $\alpha(1) = 1$, we have

$$\frac{\ln(\varphi(x_f, \{x_f, \hat{x}_f\})) - \ln(\varphi(\hat{x}_f, \{x_f, \hat{x}_f\}))}{\ln(\varphi(y_{f'}, \{y_{f'}, \hat{y}_{f'}\})) - \ln(\varphi(\hat{y}_{f'}, \{y_{f'}, \hat{y}_{f'}\}))} = \frac{\alpha(i)[u_{f(i)}(x_{f(i)}) - u_{f(i)}(\hat{x}_{f(i)})]}{u_{f'(1)}(y_{f'(1)}) - u_{f'(1)}(\hat{y}_{f'(1)})} = \alpha(i).$$

Before stating our result, we introduce the restrictions on (α, u, φ) implied by our axioms (in particular Axiom 5). For all $f \in F$, $x_f \in X_f$, and $i = 1, \ldots, N-1$, let

$$R^i_{\alpha,u,\varphi}(x_f) = \sum_{j=i+1}^{N} u_{f(j)}(x_{f(j)})\alpha(j) \exp\left\{\sum_{k=i+1}^{j-1} \varphi_k(u_{f(k)}(x_{f(k)}))\right\},$$

which is the residual value of x_f after position i. For all $a \in A$ and $x_a \in L_a$, let

$$\sigma^i_{u,\varphi}(x_a, y_a) = -\frac{e^{\varphi_i(u_a(x_a))} - e^{\varphi_i(u_a(y_a))}}{u_a(x_a) - u_a(y_a)},$$

which measures the relative strength of the framing effect and direct utility effect on an item's value of changing x_a with y_a in position i. The next condition ensures that this direct effect always dominates, taking into account the residual value of an item. This condition refines the general formulation in expression (9) in a way that precisely guarantees that the model will not generate the implausible violations of dominance ruled out by Axiom 5.

DEFINITION 8 (Regularity). *The model (α, u, φ) is regular if for all $i = 1, \ldots, N$, $f \in F$, and $x_f \in X_f$*

$$\alpha(i) \geq \sup_{y_{f(i)} \in L_{f(i)}} \sigma^i_{u,\varphi}(x_{f(i)}, y_{f(i)}) R^i_{\alpha,u,\varphi}(x_f).$$

This condition looks complex due to its generality, but is intuitive. It exactly characterizes our model as the next theorem shows. Note that it holds automatically if all φ_i are increasing and all u_a are positive, or if all φ_i are decreasing and all u_a are negative. In applications, it is easy to select regular (α, u, φ). If all u_a are bounded and all φ_i are differentiable, we can ensure regularity by assuming an appropriate bound on each derivative φ'_i.

Theorem 2. *Axioms 5–8 hold if and only if there exist regular* (α, u, φ) *such that for every* $f \in F$ *the function* w_f *in expression (8) satisfies*

$$w_f(x_f) = \sum_{i=1}^{N} u_{f(i)}(x_{f(i)})\alpha(i) \exp\left\{\sum_{k=1}^{i-1} \varphi_k(u_{f(k)}(x_{f(k)}))\right\}.$$

It is possible to extend our theory to richer choice frameworks and provide more general ways of modeling stochastic choice influenced by attribute framing effects. Online Appendix E.1 shows how to do so for the perturbed-utility model of Fudenberg et al. (2015) and the rational-inattention model of Matějka and McKay (2015). In particular, the latter model overcomes some of the well-known limitations of the Luce model. More interestingly for us, it allows for interactions between the framing effects of the order of attributes and the framing effects of the order of items on a list.

5. Final Remarks

We introduced a model of framing effects that explicitly takes into account how alternatives are presented to people. The order or emphasis given to the attributes of available items can influence which is chosen. This is at odds with mainstream choice theory, for which the presentation of the attributes should be irrelevant, but is in line with rich empirical evidence suggesting that such effects should be taken into account when studying choice behavior.

The model provides a first theoretical structure to understand such attribute-framing effects. It provides testable predictions and the possibility to compare framing effects across individuals. It can be easily generalized to allow for richer framing effects. In particular, it may open a bridge between attribute-order effects and list-order effects. Moreover, the model has several interesting implications, which we illustrated in applications to competition among firms and negotiation.

Furthermore, our model offers a stepping stone to formulating and addressing other questions about the effects of framing the attributes of choice alternatives. We briefly discuss some of these questions in the Online Appendix. The first is how framing affects choice when the available alternatives are framed in different ways. Our model allows us to formulate testable hypotheses of such effects on behavior and turn them into usable choice models. The second question relates to a large body of evidence showing that people often engage in motivated reasoning, rationalization, self-deception, self-justification, and reduction of cognitive dissonance by strategically presenting to themselves situations and decisions in the most favorable *perspective* (Bénabou and Tirole (2016)). We argue that our framework can provide a way to capture self-serving perspective manipulation in a disciplined manner. We connect this point with the ideas of *decision utility* and *experienced utility* (Kahneman et al. (1997); Kahneman et al. (1999)) and with the well-known phenomenon of the endowment effect (Thaler (1980)). Finally, we discuss how one may conduct welfare analysis in the presence of attribute-framing effects.

Appendix

Appendix A: Framing in Negotiations – Proofs

A.1. Proof of Quadratic-Loss Case

Given f, the problem P solves in the second period is the following:

$$\begin{aligned} \max_{x_f} \quad & -\sum_{i=1}^{N} \alpha(i)(x_{f(i)} - \bar{x}^P_{f(i)})^2 \\ \text{s.t.} \quad & -\sum_{i=1}^{N} \alpha(i)(x_{f(i)} - \bar{x}^R_{f(i)})^2 \geq \overline{u}^R. \end{aligned} \tag{A.1}$$

That leads to the following Lagrangian, with standard positivity conditions for the variables of interest and complementary slackness conditions:

$$\max_{x_f, \lambda} \quad -\sum_{i=1}^{N} \alpha(i)(x_{f(i)} - \bar{x}^P_{f(i)})^2 + \lambda \left(-\sum_{i=1}^{N} \alpha(i)(x_{f(i)} - \bar{x}^R_{f(i)})^2 - \overline{u}^R \right). \tag{A.2}$$

Thus, we get the following necessary and sufficient FOC for $i = 1, \ldots, N$:

$$\begin{aligned} x_{f(i)} : \quad & 2\alpha(i) \left[-\lambda(x_{f(i)} - \bar{x}^R_{f(i)}) - (x_{f(i)} - \bar{x}^P_{f(i)}) \right] = 0 \\ \lambda : \quad & -\sum_{i=1}^{N} \alpha(i)(x_{f(i)} - \bar{x}^R_{f(i)})^2 - \overline{u}^R = 0 \end{aligned} \tag{A.3}$$

and the result follows.

A.2. Proof of Proposition 1

Given f, agent P solves

$$\max_{x_f} \quad \sum_{i=1}^{N} \alpha(i)[\beta_{f(i)} - \gamma_{f(i)}(x_{f(i)} - \bar{x}^P_{f(i)})^2]$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \alpha(i)[\beta^R_{f(i)} - \gamma_{f(i)}(x_{f(i)} - \bar{x}^R_{f(i)})^2] \geq 0.$$

Following the standard Lagrangian method, we get the following necessary and sufficient FOC for $i = 1, \ldots, N$:

$$x_{f(i)}: \quad -2\alpha(i)\left[\gamma_{f(i)}(x_{f(i)} - \bar{x}^P_{f(i)}) + \gamma_{f(i)}\lambda(x_{f(i)} - \bar{x}^R_{f(i)})\right] = 0$$

$$\lambda: \quad \sum_{i=1}^{N} \alpha(i)[\beta^R_{f(i)} - \gamma_{f(i)}(x_{f(i)} - \bar{x}^R_{f(i)})^2] = 0$$

So we get the same optimal proposal as before:

$$x_{f(i)} = \frac{1}{1+\lambda}\bar{x}^P_{f(i)} + \frac{\lambda}{1+\lambda}\bar{x}^R_{f(i)}.$$

Substituting the optimal x_f into agent P's objective function and agent R's participation constraint, we obtain

$$\sum_{i=1}^{N} \alpha(i)\Big[\beta_{f(i)} - \left(\frac{\lambda}{1+\lambda}\right)^2 \gamma_{f(i)}(\bar{x}^R_{f(i)} - \bar{x}^P_{f(i)})^2\Big]$$

$$= \sum_{i=1}^{N} \alpha(i)\beta_{f(i)} - \left(\frac{\lambda}{1+\lambda}\right)^2 \sum_{i=1}^{N} \alpha(i)\gamma_{f(i)}(\bar{x}^R_{f(i)} - \bar{x}^P_{f(i)})^2.$$

and

$$\left(\frac{1}{1+\lambda}\right)^2 \sum_{i=1}^{N} \alpha(i)\gamma_{f(i)}(\bar{x}^P_{f(i)} - \bar{x}^R_{f(i)})^2 = \sum_{i=1}^{N} \alpha(i)\beta^R_{f(i)}$$

$$\Rightarrow \lambda = \frac{\sqrt{\sum_{i=1}^{N} \alpha(i)\gamma_{f(i)}(\bar{x}^P_{f(i)} - \bar{x}^R_{f(i)})^2} - \sqrt{\sum_{i=1}^{N} \alpha(i)\beta^R_{f(i)}}}{\sqrt{\sum_{i=1}^{N} \alpha(i)\beta^R_{f(i)}}}$$

$$= \sqrt{\frac{\Gamma(f)}{B^R(f)}} - 1.$$

Due to the inequalities in (4), we have $\Gamma(f) > B^R(f)$ for all f and hence $\lambda > 1$. We can then use these last conditions to replace λ in agent P's objective and obtain

$$U^P(f) = \sum_{i=1}^{N} \alpha(i)\beta_{f(i)} - \left(\frac{\lambda}{1+\lambda}\right)^2 \sum_{i=1}^{N} \alpha(i)\gamma_{f(i)}(\bar{x}^R_{f(i)} - \bar{x}^P_{f(i)})^2$$

$$= \quad B(f) - \lambda^2 B^R(f)$$

$$= \quad B(f) - \left[\sqrt{\Gamma(f)} - \sqrt{B^R(f)}\right]^2 .$$

Since the quantity in squared brackets is always positive, $U^P(f)$ is increasing in $B^P(f)$ and $B^R(f)$ and decreasing in $\Gamma(f)$. We can then conclude the following:

- If a and a' satisfy $\beta_a > \beta_{a'}$, then a should be presented before a'.
- If a and a' satisfy $\gamma_a|\overline{x}_a^P - \overline{x}_a^R| \leq \gamma_{a'}|\overline{x}_{a'}^P - \overline{x}_{a'}^R|$ and $\gamma_a|d_a^R - \overline{x}_a^R| \geq \gamma_{a'}|d_{a'}^R - \overline{x}_{a'}^R|$, and one of the inequalities is strict, then a should be presented before a'.

A.3. Proof of Corollary 2

We can view α^1 and α^2 as probability distributions over the positions $\{1, \ldots, N\}$. Then, the condition of Definition 3 can be read as α^2 MLR dominates α^1, which in turn implies that α^2 FOSD α^1. Given the optimal framing strategy f^* in Proposition 1, we have that $\beta_{f^*(i)}$ and $\beta^R_{f^*(i)}$ are decreasing functions of i and $\gamma_{f^*(i)}(\bar{x}^R_{f^*(i)} - \bar{x}^P_{f^*(i)})$ is an increasing function of i under both α^1 and α^2. Standard results imply that $\Gamma^2(f^*) > \Gamma^1(f^*)$, $B^2(f^*) < B^1(f^*)$, and $B^{R2}(f^*) < B^{R1}(f^*)$. Therefore, using the expression of U^P, we get that the proposer is better off when payoffs are defined by α^1 than by α^2.

Appendix B: Strategic Framing – Proofs

B.1. Proof of Lemma 1

Fix f and suppose $\delta_f > 0$—the other case follows similarly. Let θ_f^* identify the type of consumer indifferent between I_f and E_f:

$$\theta_f^*[\alpha_f(r)I_r + \alpha_f(b)I_b] - \alpha_f(p)I_p = \theta_f^*[\alpha_f(r)E_r + \alpha_f(b)E_b] - \alpha_f(p)E_p$$

and therefore

$$\theta_f^* = \frac{I_p - E_p}{\delta_f}\alpha_f(p).$$

The demand for the incumbent's and the entrant's product is then respectively

$$1 + \frac{1}{h} - \frac{I_p - E_p}{\delta_f}\alpha_f(p) \quad \text{and} \quad \frac{I_p - E_p}{\delta_f}\alpha_f(p) - \frac{1}{h}.$$

The firms' profit-maximization problems are

$$\max_{I_p} \Big(1 + \frac{1}{h} - \frac{I_p - E_p}{\delta_f}\alpha_f(p)\Big) I_p \qquad \text{and} \qquad \max_{E_p} \Big(\frac{I_p - E_p}{\delta_f}\alpha_f(p) - \frac{1}{h}\Big) E_p.$$

They result in the following best response functions:

$$I_p = \frac{1}{2}\Big[E_p + \frac{(1+h)\delta_f}{h\alpha_f(p)}\Big] \qquad \text{and} \qquad E_p = \frac{1}{2}\Big[I_p - \frac{\delta_f}{h\alpha_f(p)}\Big].$$

Solving this system of equations leads to the claimed equilibrium prices, which we can substitute in the profit functions to derive $\pi^o(I)$ and $\pi^o(E)$.

B.2. Lemma B.B.1

LEMMA B.B.1. *Under monopoly (i.e., $K = +\infty$), the incumbent's optimal frame is f^m and $\pi^m(I_{f^m}) > \pi^m(I_{f^*}) > \pi^m(I_{f_*})$.*

Proof. Suppose $K = +\infty$ and fix f. Given I_p, the type of consumers that is indifferent between buying I_f or nothing is

$$\theta_f^m = \frac{\alpha_f(p)}{\alpha_f(r)I_r + \alpha_f(b)I_b} I_p.$$

Thus, the monopolist maximizes

$$I_p\Big(1 + \frac{1}{h} - \frac{\alpha_f(p)}{\alpha_f(r)I_r + \alpha_f(b)I_b} I_p\Big),$$

which leads to the optimal monopolistic price and profit

$$I_p = \Big[\frac{\alpha_f(r)}{\alpha_f(p)} I_r + \frac{\alpha_f(b)}{\alpha_f(p)} I_b\Big] \frac{1+h}{2h} \qquad \text{and} \qquad \pi^m(I_f) = \Big[\frac{\alpha_f(r)}{\alpha_f(p)} I_r + \frac{\alpha_f(b)}{\alpha_f(p)} I_b\Big] \frac{(1+h)^2}{4h^2}.$$

It is easy to see that f^m maximizes $\pi^m(I_f)$ and $\pi^m(I_{f^m}) > \pi^m(I_{f^*}) > \pi^m(I_{f_*})$. ∎

B.3. Proof of Propositions 2, 3, and 4

Recall that $\pi^o(E_f)$ is proportional to $\pi^o(I_f)$, so the incumbent and the entrant rank frames in the same way.

If $K > \pi^o(E_{f^m})$, then the incumbent can choose the monopoly-optimal frame f^m and deter entry. Under condition (7), this is the best strategy for the incumbent.

If instead $K \leq \min\{\pi^o(E_{f^m}), \pi^o(E_{f_*})\}$, then for every choice of f we have $\pi^o(E_f) \geq K$. Thus, the incumbent cannot prevent entry. In this case, by Lemma 1 the incumbent will always choose f such that $\delta_f > 0$: The optimal f always presents attribute r before attribute b. Given this, to maximize $\pi^o(I_f)$, $f \in \{f^m, f^*, f_*\}$ has to maximize

$$\frac{\delta_f}{\alpha_f(p)} = \Big[\frac{\alpha_f(r)}{\alpha_f(p)} - \frac{\alpha_f(b)}{\alpha_f(p)}\Big]\delta.$$

This implies that $\pi^o(I_{f^*}) > \pi^o(I_{f_*})$, so the optimal frame is either f^m or f^*. However, $\pi^o(I_{f^m}) > \pi^o(I_{f^*})$ if and only if

$$\frac{\alpha(1)}{\alpha(3)} - \frac{\alpha(2)}{\alpha(3)} > \frac{\alpha(1)}{\alpha(2)} - \frac{\alpha(3)}{\alpha(2)} \qquad \Leftrightarrow \qquad \alpha(2) < \alpha(1) - \alpha(3) = \underline{\alpha}(2).$$

This completes the proof of Proposition 2.

To prove Proposition 3, note that if $\alpha(2) > \underline{\alpha}(2)$, then $\pi^o(I_{f^m}) > \pi^o(I_{f_*})$ if and only if

$$\frac{\alpha(1)}{\alpha(2)} - \frac{\alpha(2)}{\alpha(3)} > \frac{\alpha(2)}{\alpha(1)} - \frac{\alpha(3)}{\alpha(1)} \qquad \Leftrightarrow \qquad \alpha(2) < \frac{[\alpha(1)]^2 + [\alpha(3)]^2}{\alpha(1) + \alpha(3)} = \overline{\alpha}(2).$$

Thus, if $\overline{\alpha}(2) > \alpha(2) > \underline{\alpha}(2)$, then $\pi^o(E_{f^*}) > \pi^o(E_{f^m}) > \pi^o(E_{f_*})$ and the optimal frame under oligopoly is f^*. Given this, the only case where the incumbent can use framing to deter entry is if $K \in (\pi^o(E_{f_*}), \pi^o(E_{f^m})]$, which requires to switch to frame f_*—every other frame in F yields $\pi^o(E_f) \geq \pi^o(E_{f^m})$ and hence cannot deter entry. The incumbent prefers deterring entry with f_* to allow entry by choosing f^* if and only if $\pi^m(I_{f_*}) \geq \pi^o(I_{f^*})$, which is equivalent to

$$\frac{\alpha(2)I_r + \alpha(3)I_b}{\alpha(1)} \geq \left[\frac{\alpha(1)}{\alpha(2)} - \frac{\alpha(3)}{\alpha(2)}\right]\delta\frac{4}{9}\left(\frac{2h+1}{h+1}\right)^2. \tag{B.1}$$

This inequality holds if and only if either δ or h are sufficiently small.

If instead $\alpha(2) < \underline{\alpha}(2)$, then $\pi^o(E_{f^m}) > \pi^o(E_{f^*}) > \pi^o(E_{f_*})$ and the optimal frame under oligopoly is f^m. Thus, there are two cases in which the incumbent can use framing to deter entry. The first is if $K \in (\pi^o(E_{f^*}), \pi^o(E_{f^m})]$, which implies that best entry-deterrent frame is f^*. Every other frame in F either yields $\pi^o(E_f) \geq \pi^o(E_{f^m})$—hence cannot deter entry—or it yields $\pi^m(I_f) < \pi^m(I_{f^*})$. Given this, the incumbent prefers deterring entry with f^* to allowing entry by choosing f^m if and only if $\pi^m(I_{f^*}) \geq \pi^o(I_{f^m})$, which is equivalent to

$$\frac{\alpha(1)I_r + \alpha(3)I_b}{\alpha(2)} \geq \frac{\alpha(2)}{\alpha(3)}\Big[\frac{\alpha(1)}{\alpha(2)} - 1\Big]\delta\frac{4}{9}\left(\frac{2h+1}{h+1}\right)^2. \tag{B.2}$$

The second case is if $K \in (\pi^o(E_{f_*}), \pi^o(E_{f^*})]$, which implies that best entry-deterrent frame is f_*. Again, every other frame in F either yields $\pi^o(E_f) \geq \pi^o(E_{f^*})$—hence cannot deter entry—or it yields $\pi^m(I_f) < \pi^m(I_{f_*})$. Given this, the incumbent prefers deterring entry with f_* to allowing entry by choosing f^m if and only if $\pi^m(I_{f_*}) \geq \pi^o(I_{f^m})$, which is equivalent to

$$\frac{\alpha(2)I_r + \alpha(3)I_b}{\alpha(1)} \geq \frac{\alpha(2)}{\alpha(3)}\Big[\frac{\alpha(1)}{\alpha(2)} - 1\Big]\delta\frac{4}{9}\left(\frac{2h+1}{h+1}\right)^2. \tag{B.3}$$

Consider now Proposition 4. It is easy to see that if either δ is lower or h is, then $\pi^o(E_{f^*})$, $\pi^o(E_{f^m})$, and $\pi^o(E_{f_*})$ are all lower. Thus, the range of entry costs where the incumbent faces no entry threat (i.e., $K > \pi^o(E_{f^m})$) expands, and the range of costs where the incumbent can never prevent entry (i.e., $K \leq \min\{\pi^o(E_{f^m}), \pi^o(E_{f_*})\}$) shrinks. In addition, the right-hand side of conditions (B.1), (B.2), and (B.3) are all smaller if either δ or h is lower. Thus, for $K \in (\pi^o(E_{f_*}), \max\{\pi^o(E_{f^m}), \pi^o(E_{f^*})\}]$, the incumbent is more likely to use framing to deter entry.

Finally, consider α and α' such that $\alpha(1)/\alpha(2) \geq \alpha'(1)/\alpha'(2)$ and $\alpha(2)/\alpha(3) \geq \alpha'(2)/\alpha'(3)$. It is easy to see that this implies that $\pi^o(E_{f^m})$ and $\pi^o(E_{f^*})$ are both lower under α than under α'. Moreover, the left-hand side of conditions (B.1) and (B.3) is higher under α' than under α, while the right-hand side of conditions (B.1) and (B.3) is lower under α' than under α. Thus, whenever K falls in the region where deterring entry requires to use f_*, if the incumbent finds this optimal under α, it also finds it optimal under α'. Regarding $\pi^o(E_{f_*})$ and condition B.2, their ranking under α and α' is ambiguous.

B.4. Proof of Corollary 3

The incumbent's monopoly profits are

$$\pi^m(I_{f^m}) = \Big[\frac{\alpha(1)}{\alpha(3)} I_r + \frac{\alpha(2)}{\alpha(3)} I_b\Big] \frac{(1+h)^2}{4h^2}.$$

Since $\delta_{f^m} > 0$, the industry profits after entry are

$$\pi^o(I_{f^m}) + \pi^o(E_{f^m}) = \frac{\delta_{f^m}}{9\alpha_{f^m}(p)} \Big[(2+h^{-1})^2 + (1+h^{-1})^2\Big].$$

Hence, the difference in total profits can be written as

$$\frac{\delta_f}{9\alpha_f(p)} \Big[(2+h^{-1})^2 + (1+h^{-1})^2\Big] - \frac{\alpha(1)I_r + \alpha(2)I_b}{4\alpha(3)}(1+h^{-1})^2$$

$$= \frac{(\alpha(1)-\alpha(2))(I_r - E_r)}{9\alpha(3)}(3+2h^{-1})$$

$$- \frac{(\alpha(1)+8\alpha(2))I_r + 9\alpha(2)I_b + 8(\alpha(1)-\alpha(2))E_r}{36\alpha(3)}(1+h^{-1})^2,$$

which is decreasing in E_r and I_b. Therefore, consider the limit at $E_r = I_b = 0$, where the last expression becomes

$$\frac{(\alpha(1)-\alpha(2))I_r}{9\alpha(3)}(3+2h^{-1}) - \frac{(\alpha(1)+8\alpha(2))I_r}{36\alpha(3)}(1+h^{-1})^2,$$

which is proportional to

$$4(\alpha(1)-\alpha(2))(3+2h^{-1}) - (\alpha(1)+8\alpha(2))(1+h^{-1})^2.$$

If $\alpha(2)$ (and $\alpha(3)$) is sufficiently small, this expression will be positive because $\alpha(1)(1+h^{-1})^2$ is smaller than $4\alpha(1)(3+2h^{-1})$.

B.5. Proof of Proposition 5

Given any $f_I \in \{f^m, f^*, f_*\}$ and $f_E \in \{f^{Em}, f^{E*}\}$, the firms compete in prices. We will refer to the resulting equilibrium as *continuation equilibrium under* f_I *and* f_E. We will refer to the expected profits in this equilibrium as *continuation profits under* f_I *and* f_E. At $\mu = 1$, we know by Lemma 1 that this continuation equilibrium exists, is unique; moreover, the overall equilibrium is generically unique and results in the incumbent taking the top of the market while the entrant takes the bottom of the market. The proof will rely on showing that the incumbent's and the entrant's continuation profits (and prices) under f_I and f_E are continuous in μ. This will imply that there exists $\underline{\mu} < 1$ such that for $\mu > \underline{\mu}$, the incumbent's optimal choice of frame will be the same as at $\mu = 1$. The entrant's decision to enter based on K will generically also be the same as at $\mu = 1$. These hold as long as the incumbent and the entrant have a unique optimal choice (of f_I, f_E and whether to enter or not) at $\mu = 1$, which generically holds.

When we consider $\mu < 1$, the firms' demand functions will become complicated for arbitrary prices. However, locally (around the original equilibrium prices at $\mu = 1$)[27] the demand functions are given by

$$D^I = \mu \left(1 + h^{-1} - \frac{I_p - E_p}{\delta_{f_I}} \alpha_{f_I}(p)\right),$$

$$D^E = \mu \left(\frac{I_p - E_p}{\delta_{f_I}} \alpha_{f_I}(p) - h^{-1}\right) + (1-\mu)\left(1 + h^{-1} - \frac{\alpha_{f_E}(p)}{\alpha_{f_E}(r)E_r + \alpha_{f_E}(b)E_b}\right).$$

There are two reasons for this. First, the indifferent consumer under f_I is strictly between h^{-1} and $1 + h^{-1}$ at the original equilibrium. This implies that small changes in either price will not change the demand expression under $f = f_I$, as both firms will still have positive market shares. Second, the expression for the indifferent consumer under f_E is negative at the original equilibrium (due to $\delta_{f_E} < 0$), which implies that the entrant will get the whole market under f_E. Again, this will hold under small changes in either price.

[27] Formally, there is a neighborhood in the space of (E_p, I_p) around the original equilibrium prices where the demand functions are given by the expressions that follow.

Fix any $f_I \in \{f^m, f^*, f_*\}$ and $f_E \in \{f^{Em}, f^{E*}\}$. It is straightforward to derive the firms' best-response price functions from the demands above:

$$E_p = \frac{1}{2} \cdot \frac{(1-\mu) + (1-2\mu)h^{-1} + \mu \frac{\alpha_{f_I}(p)}{\delta_{f_I}} I_p}{\mu \frac{\alpha_{f_I}(p)}{\delta_{f_I}} + (1-\mu)\frac{\alpha_{f_E}(p)}{\alpha_{f_E}(r)E_r + \alpha_{f_E}(b)E_b}},$$

$$I_p = \frac{1}{2}\left[E_p + \frac{\delta_{f_I}}{\alpha_{f_I}(p)}(1+h^{-1})\right].$$

These best-response functions are continuous in μ. It is straightforward to derive the equilibrium prices under these best responses and show that they are continuous in μ. We can also confirm that the equilibrium prices converge to the prices in Lemma 1 as $\mu \to 1$. Denote these equilibrium prices by $I_p^*(f_I, f_E, \mu)$ and $E_p^*(f_I, f_E, \mu)$.

We now have to argue that $I_p = I_p^*(f_I, f_E, \mu)$ and $E_p = E_p^*(f_I, f_E, \mu)$ is an equilibrium once we consider arbitrary I_p and E_p. For that, we need to prove that for sufficiently high $\mu < 1$ neither firm has an incentive to deviate to a drastically different price, where the local demand functions from above no longer hold. For this, we will rely on the fact that the continuation equilibrium at $\mu = 1$ is strict in prices.

Suppose that the entrant picks $E_p = E_p^*(f_I, f_E, \mu)$ when $\mu < 1$ and consider the incumbent. Consider an arbitrary price $\hat{I}_p$ that is significantly lower than $I_p^*(f_I, f_E, 1)$, so that the local demand functions from above do not apply to it. Since the equilibrium is strict at $\mu = 1$, it follows that the incumbent gets a strictly higher profit under $I_p = I_p^*(f_I, f_E, 1)$ than under $I_p = \hat{I}_p$ when $\mu = 1$. Let that profit difference be equal to Δ. Since the demand functions for any E_p and I_p are continuous in μ (as a weighted sum of two fixed demand values), it follows that the profit functions are also continuous in μ for all E_p and I_p. Additionally, recall that $E_p^*(f_I, f_E, \mu)$ is continuous in μ, and the firms' profits are continuous in both prices. Hence, there is an $\varepsilon > 0$ and $\psi > 0$ such that for $\mu > 1 - \varepsilon$, $E_p^*(f_I, f_E, \mu)$ changes by less than $\psi > 0$, which in turn makes the profit under $I_p = I_p^*(f_I, f_E, 1)$ fall by less than $\frac{1}{2}\Delta$ and the profit under $I_p = \hat{I}_p$ rise by less than $\frac{1}{2}\Delta$. As a result, the incumbent will still prefer the original equilibrium price $I_p^*(f_I, f_E, 1)$ to $\hat{I}_p$ when $\mu > 1 - \varepsilon$. Thus, $\hat{I}_p$ cannot be a best response for $\mu > 1 - \varepsilon$.

This allows us to rule out all non-local prices $\hat{I}_p$ as best responses when $\mu > 1 - \varepsilon$, for an appropriately chosen $\varepsilon > 0$. Hence, the incumbent's best response must be local and must be equal to $I_p^*(f_I, f_E, \mu)$. A similar argument shows the same is true for the entrant.

Therefore, there is a $\underline{\mu} < 1$ such that for $\mu > \underline{\mu}$, the continuation equilibrium prices are described by $I_p = I_p^*(f_I, f_E, \mu)$ and $E_p = E_p^*(f_I, f_E, \mu)$. Since these functions are continuous in μ, and the firms' profit functions are continuous in both μ and prices, it follows that the continuation equilibrium profits are continuous in μ at $\mu = 1$.

Since the continuation equilibrium profits are continuous in μ for each $f_I \in \{f^m, f^*, f_*\}$, incentives that are strict at $\mu = 1$ remain strict in its neighborhood. If the continuation profit under a certain frame f is strictly better than under other frames at $\mu = 1$, perturbing μ will change the profits continuously, and thus keep f more profitable than the other frames. Thus, the incumbent's equilibrium choice of f_I will remain the same as when $\mu = 1$.

In the case of entry deterrence, this logic still holds. If a particular frame f strictly deters the entrant from entering—i.e., the continuation profit is strictly lower than K—then the continuity of equilibrium profit in μ ensures that f will still strictly deter entry in the neighborhood of $\mu = 1$. If the incumbent strictly prefers entry-deterring frame f to another frame $\hat{f}$ at $\mu = 1$, this will still be the case in its neighborhood. The continuation profit under f is locally constant in μ due to the entry being deterred, and the continuation profit under $\hat{f}$ is either also locally constant (if $\hat{f}$ deters entry) or is continuous in μ. Thus, the continuation profit under f will still be strictly bigger than under $\hat{f}$ for $\mu > \underline{\mu}$.

To show the last part of the proposition, note that the equilibrium prices $I_p^*(f_I, f_E, \mu)$ and $E_p^*(f_I, f_E, \mu)$ are continuous in μ, and since $I_p^*(f_I^*, f_E^*, 1) > E_p^*(f_I^*, f_E^*, 1)$, it must be the case that for $\mu > \underline{\mu}$ we have $I_p^*(f_I^*, f_E^*, \mu) > E_p^*(f_I^*, f_E^*, \mu)$.

Appendix C: Characterization of the AF Model

C.1. Proof of Theorem 1 and Proposition 6

We will prove sufficiency of Axioms 1–4; necessity is easy to verify and is thus omitted. Given Assumption 1, the condition $c(x_f, x'_f) = \{x_f\} \Leftrightarrow c(y_f, y'_f) = \{y_f\}$ implies that

$$v_f(x_f) > (=)\ v_f(x'_f) \quad \Leftrightarrow \quad v_f(y_f) > (=)\ v_f(y'_f).$$

Given the restrictions on x_f, x'_f, y_f, and y'_f in Axiom 2, this means that how v_f ranks the attributes in positions j and k is independent of the other attributes' levels. By Axiom 1 each position of f can matter for choice. By standard arguments (Debreu (1960)), we can write v_f in an additive form in positions and the attribute assigned by f to each position:

$$v_f(x_f) = \sum_{i=1}^{N} w_{i,f(i)}^f(x_{f(i)}), \tag{C.1}$$

where for every $f \in F$ there exist non-constant $w_{i,f(i)}^f : L_{f(i)} \to \mathbb{R}$ for every $i \in \{1, \ldots, N\}$. Additive forms are unique up to positive affine transformations, which in this case can depend on f: If we have two such representations v_f and $\hat{v}_f$ of choice under f, we must have $v_f = \beta_f \hat{v}_f + \xi_f$, where $\beta_f > 0$ and $\xi_f \in \mathbb{R}$.

Given this, Axiom 3 implies that for every $a \in A$, if $f(i) = f'(j) = a$, then $w^f_{i,f(i)}$ and $w^{f'}_{j,f'(j)}$ represent the same vN-M utility function over L_a. Therefore, for every $a \in A$, fix any $f_a \in F$ such that $f_a(1) = a$. Let $u_a = w^{f_a}_{1,a}$. For any other $f \in F$ and $i = 1, \dots, N$ such that $f(i) = a$, we have that $w^f_{i,a} = \gamma^f_i u_a + \zeta^f_i$, where $\gamma^f_i > 0$ and $\zeta^f_i \in \mathbb{R}$.[28] Letting $\gamma^{f_a}_1 = 1$ and $\zeta^{f_a}_1 = 0$, we can write

$$v_f(x_f) = \sum_{i=1}^{N} w^f_{i,f(i)}(x_{f(i)}) = \sum_{i=1}^{N} \gamma^f_i u_{f(i)}(x_{f(i)}) + \sum_{i=1}^{N} \zeta^f_i.$$

By affine uniqueness of v_f, for all $f \in F$ we can let $\gamma^f_1 = 1$ and $\zeta^f_i = 0$ for all $i = 1, \dots, N$. Therefore, we obtain the representation

$$v_f(x_f) = u_{f(1)}(x_{f(1)}) + \sum_{i=2}^{N} \gamma^f_i u_{f(i)}(x_{f(i)}).$$

Next, we want to show that γ^f_i depends only on i. By definition of x_f, y_f, and z_f in Axiom 4, we have

$$p_{xyz_f}[u_{f(1)}(x_{f(1)}) + \gamma^f_i u_{f(i)}(x_{f(i)})] + (1 - p_{xyz_f})[u_{f(1)}(y_{f(1)}) + \gamma^f_i u_{f(i)}(y_{f(i)})]$$

[28]Note that γ^f_i cannot also depend on $f(i)$ in addition to i and f because there is no $\gamma^f_{i,b}$ for $b \neq f(i)$. If any, the superscript already allows for the dependence on $f(i)$.

$$=u_{f(1)}(y_{f(1)}) + \gamma_i^f u_{f(i)}(x_{f(i)}).$$

This implies that

$$\frac{p_{xyz_f}}{1 - p_{xyz_f}} = \gamma_i^f \frac{u_{f(i)}(x_{f(i)}) - u_{f(i)}(y_{f(i)})}{u_{f(1)}(x_{f(1)}) - u_{f(1)}(y_{f(1)})}.$$

Similarly, for any f',

$$\frac{p_{xyz_{f'}}}{1 - p_{xyz_{f'}}} = \gamma_j^{f'} \frac{u_{f'(j)}(x_{f'(j)}) - u_{f'(j)}(y_{f'(j)})}{u_{f'(1)}(x_{f'(1)}) - u_{f'(1)}(y_{f'(1)})} = \gamma_j^{f'} \frac{u_{f(i)}(x_{f(i)}) - u_{f(i)}(y_{f(i)})}{u_{f(1)}(x_{f(1)}) - u_{f(1)}(y_{f(1)})}.$$

Therefore, for any f',

$$\frac{p_{xyz_f}}{1 - p_{xyz_f}} = \frac{\gamma_i^f}{\gamma_j^{f'}} \cdot \frac{p_{xyz_{f'}}}{1 - p_{xyz_{f'}}}.$$

By Axiom 4, we have

$$\frac{\gamma_i^f}{\gamma_j^{f'}} = g(i,j)$$

and therefore $\gamma_i^f = \gamma_i > 0$ and $\gamma_j^{f'} = \gamma_j > 0$ for all f, f'.

We conclude that for every $f \in F$ and $x_f \in X_f$

$$v_f(x_f) = u_{f(1)}(x_{f(1)}) + \sum_{i=2}^{N} \gamma_i u_{f(i)}(x_{f(i)}).$$

Representation Uniqueness. The uniqueness of α up to a scalar multiplication follows from the fact, shown above, that choices identify only $\alpha(i)/\alpha(j)$ through the calibration lotteries p_{xyz_f} and $p_{xyz_{f'}}$. The uniqueness up to affine transformations of each u_a follows from the standard uniqueness of vN-M utility functions.

C.2. Proof of Proposition 7

Consider primacy effect—the argument is the same for recency effect. Recall that $x_a \succ_a y_a$ if and only if $u_a(x_a) > u_a(y_a)$. Fix any $i = 1, \ldots, N-1$. Using the AF representation, we have that

$$\{(x_f, y_f; p^i_{xyz_f}), z_f\} = c((x_f, y_f; p^i_{xyz_f}), z_f)$$

is equivalent to the following equality between the expected utilities derived from z_f and $(x_f, y_f; p^i_{xyz_f})$:

$$\sum_{j=1}^{N} \alpha(j) u_{f(j)}(z_{f(j)}) = p^i_{xyz_f} \left\{ \sum_{j=1}^{N} \alpha(j) u_{f(j)}(x_{f(j)}) \right\} + (1 - p^i_{xyz_f}) \left\{ \sum_{j=1}^{N} \alpha(j) u_{f(j)}(y_{f(j)}) \right\}.$$

Since $x_{f(j)} = y_{f(j)} = z_{f(j)}$ for all $j \neq i, i+1$, this condition becomes

$$\alpha(i)u_{f(i)}(z_{f(i)}) + \alpha(i+1)u_{f(i+1)}(z_{f(i+1)}) =$$

$$p^i_{xyz_f} \left\{\alpha(i)u_{f(i)}(x_{f(i)}) + \alpha(i+1)u_{f(i+1)}(x_{f(i+1)})\right\}$$

$$+(1 - p^i_{xyz_f}) \left\{\alpha(i)u_{f(i)}(y_{f(i)}) + \alpha(i+1)u_{f(i+1)}(y_{f(i+1)})\right\}.$$

Using $x_{f(i)} = z_{f(i)}$ and $y_{f(i+1)} = z_{f(i+1)}$, we obtain

$$\frac{p^i_{xyz_f}}{1 - p^i_{xyz_f}} = \frac{\alpha(i)[u_{f(i)}(x_{f(i)}) - u_{f(i)}(y_{f(i)})]}{\alpha(i+1)[u_{f(i+1)}(x_{f(i+1)}) - u_{f(i+1)}(y_{f(i+1)})]}.$$

By similar calculations, using $x_{f'(i+1)} = z_{f'(i+1)}$ and $y_{f'(i)} = z_{f'(i)}$, we have

$$\frac{p^i_{xyz_{f'}}}{1 - p^i_{xyz_{f'}}} = \frac{\alpha(i+1)[u_{f'(i+1)}(x_{f'(i+1)}) - u_{f'(i+1)}(y_{f'(i+1)})]}{\alpha(i)[u_{f'(i)}(x_{f'(i)}) - u_{f'(i)}(y_{f'(i)})]}.$$

Since $x_{f'}$, $y_{f'}$, and $z_{f'}$ are obtained from x_f, y_f, and z_f by swapping the attributes in positions i and $i+1$, we have

$$\frac{p^i_{xyz_{f'}}}{1 - p^i_{xyz_{f'}}} = \frac{\alpha(i+1)[u_{f(i)}(x_{f(i)}) - u_{f(i)}(y_{f(i)})]}{\alpha(i)[u_{f(i+1)}(x_{f(i+1)}) - u_{f(i+1)}(y_{f(i+1)})]}.$$

It follows that

$$\frac{p^i_{xyz_f}}{1-p^i_{xyz_f}} = \left[\frac{\alpha(i)}{\alpha(i+1)}\right]^2 \cdot \frac{p^i_{xyz_{f'}}}{1-p^i_{xyz_{f'}}}.$$

This implies that $\alpha(i) > \alpha(i+1)$ if and only if $p^i_{xyz_f} > p^i_{xyz_{f'}}$ as desired.

Appendix D: Identification in Consumption Spaces

Taking our AF model as given, we now outline an approach to identifying its parameters (α, u) in a standard consumption space. We make the following assumptions: One of the items' attributes is the price, denoted by $p \in \mathbb{R}$; for each attribute $a \in A$ other than p the set of its levels is $L_a = \mathbb{R}_+$; the attribute-specific utility of the price satisfies $u(p) = -p$; and the utility of the outside-option of not consuming any item is constant and normalized to zero. Also, hereafter we normalize α so that $\alpha(1) = 1$.

The identification of the other attribute-specific utility functions follows standard steps. To do this, for each attribute a focus on items that present this attribute in the first position and hold all other attributes constant. Varying the levels of x_a, we can identify u_a as we identify any utility function in classic choice analysis.

We now proceed to identify α. First, suppose there is only one attribute other than the price. Fix any value c of this attribute. Define p and $\hat{p}$ as satisfying the indifference

$$u(c) - \alpha(2)p = 0 = -\hat{p} + \alpha(2)u(c).$$

That is, we can interpret p and $\hat{p}$ as the willingness to pay (WTP) for the item when presenting the price second and first. We can then solve for $\alpha(2)$ using the data $\hat{p} - p$ and $u(c)$.

Now suppose there are multiple attributes other than the price, where N is the total number of attributes and, hence, presentation positions. Fix the level of two attributes, c_a and c_b, so that $\Delta u = u_a(c_a) - u_b(c_b) \neq 0$. First, place these attributes in positions 1 and 2 and place the price in the last position N. Define p and $\hat{p}$ again as WTPs so that

$$u_a(c_a) + \alpha(2)u_b(c_b) + \cdots - \alpha(N)p = 0 = u_b(c_b) + \alpha(2)u_a(c_a) + \cdots - \alpha(N)\hat{p}.$$

Rearranging, we obtain the equation

$$\alpha(2) = 1 + \alpha(N)\frac{p - \hat{p}}{\Delta u}.$$

Now, place the price in position 1 and place attributes a and b in positions 2 and N. Define p' and $\hat{p}'$ by

$$-p' + \alpha(2)u_a(c_a) + \cdots + \alpha(N)u_b(c_b) = 0 = -\hat{p}' + \alpha(2)u_b(c_b) + \cdots + \alpha(N)u_a(c_a).$$

Rearranging, we obtain the equation

$$\alpha(2) = \alpha(N) + \frac{\hat{p}' - p'}{\Delta u}.$$

These two equations allow us to identify $\alpha(2)$ and $\alpha(N)$ using the data Δu, $p - \hat{p}$, and $p' - \hat{p}'$.

Now, suppose there is any position i such that $2 < i < N$. Place the price in position N and attributes a and b in positions i and $i-1$. Define p and $\hat{p}$ as

$$\alpha(i-1)u_a(c_a) + \alpha(i)u_b(c_b) + \cdots - \alpha(N)p = 0 = \alpha(i-1)u_b(c_b) + \alpha(i)u_a(c_a) + \cdots - \alpha(N)\hat{p},$$

which leads to

$$\alpha(i) = \alpha(i-1) + \alpha(N)\frac{p - \hat{p}}{\Delta u}.$$

Since we know $\alpha(2)$ and $\alpha(N)$, we can use this equation for $i = 3$ to find $\alpha(3)$. Iterating this step, we can find $\alpha(i)$ for all the remaining positions i.

References

David S Ahn and Haluk Ergin. Framing contingencies. Econometrica, 78(2):655–695, 2010.

Icek Ajzen and Martin Fishbein. Understanding attitudes and predicting social behavior. Englewood cliffs NJ: Prentice–Hall, 1980.

Roy Allen and John Rehbeck. Revealed stochastic choice with attributes. Economic Theory, 75(1):91–112, 2023.

Jose Apesteguia and Miguel A Ballester. A measure of rationality and welfare. Journal of Political Economy, 123(6):1278–1310, 2015.

Jose Apesteguia and Miguel A Ballester. Monotone stochastic choice models: The case of risk and time preferences. Journal of Political Economy, 126(1):74–106, 2018.

Katrin Auspurg and Annette Jäckle. First equals most important? order effects in vignette-based measurement. Sociological Methods & Research, 46(3):490–539, 2017.

Ryan K Beasley and Mark R Joslyn. Cognitive dissonance and post-decision attitude change in six presidential elections. Political psychology, 22(3):521–540, 2001.

James K Beggan. On the social nature of nonsocial perception: The mere ownership effect. Journal of Personality and Social Psychology, 62(2):229, 1992.

Roland Bénabou and Jean Tirole. Mindful economics: The production, consumption, and value of beliefs. Journal of Economic Perspectives, 30(3):141–64, 2016.

George R Bergus, Gretchen B Chapman, C Gjerde, and Arthur S Elstein. Clinical reasoning about new symptoms despite preexisting disease: sources of error and order effects. Family medicine, 27(5):314–320, 1995.

B Douglas Bernheim and Antonio Rangel. Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. Quarterly Journal of Economics, 124(1):51–104, 2009.

Tom Blake, Sarah Moshary, Kane Sweeney, and Steve Tadelis. Price salience and product choice. Marketing Science, 40(4):619–636, 2021.

Henry David Block and Jacob Marschak. Random orderings and stochastic theories of response. Contributions To Probability And Statistics, 1960. Stanford University Press.

Samuel D Bond, Kurt A Carlson, Margaret G Meloy, J Edward Russo, and Robin J Tanner. Information distortion in the evaluation of a single option. Organizational Behavior and Human Decision Processes, 102(2):240–254, 2007.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience and consumer choice. Journal of Political Economy, 121(5):803–843, 2013.

Kurt A Carlson, Margaret G Meloy, and J Edward Russo. Leader-driven primacy: Using attribute order to affect consumer choice. Journal of Consumer Research, 32(4):513–518, 2006.

Gregory S Carpenter and Kent Nakamoto. Consumer preference formation and pioneering advantage. Journal of Marketing Research, pages 285–298, 1989.

Matias D Cattaneo, Xinwei Ma, Yusufcan Masatlioglu, and Elchin Suleymanov. A random attention model. Journal of Political Economy, 128(7):2796–2836, 2020.

Gretchen B Chapman and Arthur S Elstein. Cognitive processes and biases in medical decision making. Decision making in health care: Theory, psychology, and applications, pages 183–210, 2000.

Keith Chrzan. Three kinds of order effects in choice-based conjoint analysis. Marketing Letters, 5(2):165–172, 1994.

Joep P Cornelissen and Mirjam D Werner. Putting framing in perspective: A review of framing and frame analysis across the management and organizational literature. The Academy of Management Annals, 8(1):181–235, 2014.

John PW Cunnington, Jeffrey M Turnbull, Glenn Regher, Michael Marriott, and Geoffrey R Norman. The effect of presentation order in clinical decision making. Academic Medicine, 1997.

Leora C Dahl, CA Elizabeth Brimacombe, and D Stephen Lindsay. Investigating investigators: How presentation order influences participant-investigators' interpretations of eyewitness identification and alibi evidence. Law and Human Behavior, 33(5):368, 2009.

Brett Day and Jose-Luis Pinto Prades. Ordering anomalies in choice experiments. Journal of Environmental Economics and Management, 59(3):271–285, 2010.

Brett Day, Ian J Bateman, Richard T Carson, Diane Dupont, Jordan J Louviere, Sanae Morimoto, Riccardo Scarpa, and Paul Wang. Ordering effects and choice set awareness in repeat-response stated preference studies. Journal of Environmental Economics and Management, 63(1):73–91, 2012.

Gerard Debreu. Topological methods in cardinal utility theory. Mathematical Methods in the Social Sciences, 1959:16–26, 1960.

William Donohue, Randall G Rogan, and Sanda Kaufman. Framing matters: Perspectives on negotiation research and practice in communication. Peter Lang Publishing, 2011.

Daniel Druckman and Lynn Wagner. The role of issues in negotiation: Framing, linking, and ordering. Negotiation Journal, 37(2):249–278, 2021.

Glenn Ellison. A model of add-on pricing. Quarterly Journal of Economics, 120(2):585–637, 2005.

Glenn Ellison and Sara Fisher Ellison. Search, obfuscation, and price elasticities on the internet. Econometrica, 77(2):427–452, 2009.

Larry G Epstein. Stationary cardinal utility and optimal growth under uncertainty. Journal of Economic Theory, 31(1):133–152, 1983.

Keith M Marzilli Ericson and Amanda Starc. How product standardization affects choice: Evidence from the massachusetts health insurance exchange. Journal of Health Economics, 50:71–85, 2016.

Drew Fudenberg, Ryota Iijima, and Tomasz Strzalecki. Stochastic choice and revealed perturbed utility. Econometrica, 83(6):2371–2409, 2015.

Xavier Gabaix and David Laibson. Shrouded attributes, consumer myopia, and information suppression in competitive markets. Quarterly Journal of Economics, 121(2):505–540, 2006.

Xin Ge, Gerald Häubl, and Terry Elrod. What to say when: influencing consumer choice by delaying the presentation of favorable information. Journal of Consumer Research, 38 (6):1004–1021, 2011.

William M Gorman. A possible procedure for analysing quality differentials in the egg market. Review of Economic Studies, 47(5):843–856, 1980.

Faruk Gul and Wolfgang Pesendorfer. Random expected utility. Econometrica, 74(1):121–146, 2006.

Faruk Gul, Paulo Natenzon, and Wolfgang Pesendorfer. Random choice as behavioral optimization. Econometrica, 82(5):1873–1912, 2014.

Curtis P Haugtvedt and Duane T Wegener. Message order effects in persuasion: An attitude strength perspective. Journal of Consumer Research, 21(1):205–218, 1994.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. Econometrica, pages 263–291, 1979.

Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. Anomalies: The endowment effect, loss aversion, and status quo bias. Journal of Economic Perspectives, 5(1):193–206, 1991.

Daniel Kahneman, Peter P Wakker, and Rakesh Sarin. Back to bentham? explorations of experienced utility. Quarterly Journal of Economics, 112(2):375–406, 1997.

Daniel Kahneman et al. Objective happiness. Well-being: The foundations of hedonic psychology, 3(25):1–23, 1999.

Frank R Kardes and Paul M Herr. Order effects in consumer judgment, choice, and memory: The role of initial processing goals. ACR North American Advances, 1990.

Trine Kjær, Mickael Bech, Dorte Gyrd-Hansen, and Kristian Hart-Hansen. Ordering effect and price sensitivity in discrete choice experiments: need we worry? Health economics, 15(11):1217–1228, 2006.

Botond Kőszegi and Matthew Rabin. A model of reference-dependent preferences. Quarterly Journal of Economics, 121(4):1133–1165, 2006.

Botond Kőszegi and Adam Szeidl. A model of focusing in economic choice. Quarterly journal of economics, pages 53–104, 2012.

Jon A Krosnick and Duane F Alwin. An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opinion Quarterly, 51(2):201–219, 1987.

V Kumar and Gary J Gaeth. Attribute order and product familiarity effects in decision tasks using conjoint analysis. International Journal of Research in Marketing, 8(2):113–124, 1991.

Kelvin J Lancaster. A new approach to consumer theory. Journal of Political Economy, 74 (2):132–157, 1966.

Jonathan Levav, Mark Heitmann, Andreas Herrmann, and Sheena S Iyengar. Order in product customization decisions: Evidence from field experiments. Journal of Political Economy, 118(2):274–299, 2010.

Irwin P Levin and Gary J Gaeth. How consumers are affected by the framing of attribute information before and after consuming the product. Journal of Consumer Research, 15 (3):374–378, 1988.

Sarah Lichtenstein and Paul Slovic. The construction of preference. Cambridge University Press, 2006.

Joan Lindsey-Mullikin. Beyond reference price: understanding consumers' encounters with unexpected prices. Journal of Product & Brand Management, 2003.

John A List. Does market experience eliminate market anomalies? Quarterly Journal of Economics, 118(1):41–71, 2003.

R Duncan Luce. Individual choice behavior: A theoretical analysis. John Wiley, 1959.

John G Jr Lynch and Dan Ariely. Wine online: Search costs affect competition on price, quality, and distribution. Marketing Science, 19(1):83–103, 2000.

Paola Manzini and Marco Mariotti. Stochastic choice and consideration sets. Econometrica, 82(3):1153–1176, 2014.

Jacob Marschak. Binary-choice constraints and random utility indicators (1960). In Economic Information, Decision, and Prediction, pages 218–239. Springer, 1974.

Yusufcan Masatlioglu and Efe A Ok. Rational choice with status quo bias. Journal of Economic Theory, 121(1):1–29, 2005.

Yusufcan Masatlioglu and Efe A Ok. A canonical model of choice with initial endowments. Review of Economic Studies, 81(2):851–883, 2014.

Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Y Ozbay. Revealed attention. American Economic Review, 102(5):2183–2205, 2012.

Filip Matějka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. American Economic Review, 105(1):272–98, 2015.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. 1973.

Milica Milosavljevic, Vidhya Navalpakkam, Christof Koch, and Antonio Rangel. Relative visual saliency differences induce sizable bias in consumer choice. Journal of Consumer Psychology, 22(1):67–74, 2012.

Thomas E Nelson, Zoe M Oxley, and Rosalee A Clawson. Toward a psychology of framing effects. Political Behavior, 19(3):221–246, 1997.

Efe A Ok, Pietro Ortoleva, and Gil Riella. Revealed (p) reference theory. American Economic Review, 105(1):299–321, 2015.

Franz Ostrizek and Danis Shishkin. Screening with frames: Implementation in extensive form. Review of Economic Studies, forthcoming, 2023.

John W Payne, David A Schkade, William H Desvousges, and Chris Aultman. Valuation of multiple environmental programs. Journal of Risk and Uncertainty, 21(1):95–115, 2000.

Michele Piccione and Ran Spiegler. Price competition under limited comparability. Quarterly Journal of Economics, 127(1):97–135, 2012.

Robin L Pinkley and Gregory B Northcraft. Cognitive interpretations of conflict: Implications for dispute processes and outcomes. Academy of Management Journal, 37 (1):193–205, 1994.

Ariel Rubinstein and Yuval Salant. A model of choice from lists. Theoretical Economics, 1 (1):3–17, 2006.

Ariel Rubinstein and Yuval Salant. Eliciting welfare preferences from behavioural data sets. Review of Economic Studies, 79(1):375–387, 2011.

Yuval Salant. Procedural analysis of choice rules with applications to bounded rationality. American Economic Review, 101(2):724–48, 2011.

Yuval Salant and Ariel Rubinstein. (a, f): choice with frames. Review of Economic Studies, 75(4):1287–1296, 2008.

Yuval Salant and Ron Siegel. Contracts with framing. American Economic Journal: Microeconomics, 10(3):315–46, 2018.

Rom Y Schrift, Jeffrey R Parker, Gal Zauberman, and Shalena Srna. Multistage decision processes: The impact of attribute order on how consumers mentally represent their choice. Journal of Consumer Research, 44(6):1307–1324, 2018.

Maurice E Schweitzer and Leslie A DeChnrch. Linking frames in negotiations: Gains, losses and conflict frame adoption. International Journal of Conflict Management, 2001.

Emily Fitzgibbons Shafer. Hillary rodham versus hillary clinton: Consequences of surname choice in marriage. Gender Issues, 34(4):316–332, 2017.

Shaul Shalvi, Francesca Gino, Rachel Barkan, and Shahar Ayal. Self-serving justifications: Doing wrong and feeling moral. Current Directions in Psychological Science, 24(2):125–130, 2015.

Michael D Smith and Erik Brynjolfsson. Consumer decision-making at an internet shopbot: Brand still matters. Journal of Industrial Economics, 49(4):541–558, 2001.

Ran Spiegler. Competitive framing. American Economic Journal: Microeconomics, 6(3): 35–58, 2014.

Michal A Strahilevitz and George Loewenstein. The effect of ownership history on the valuation of objects. Journal of Consumer Research, 25(3):276–289, 1998.

Richard Thaler. Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1):39–60, 1980.

Richard Thaler. Mental accounting and consumer choice. Marketing Science, 4(3):199–214, 1985.

Richard H Thaler. Anomalies: Saving, fungibility, and mental accounts. The Journal of Economic Perspectives, 4(1):193–205, 1990.

Jean Tirole. The Theory of Industrial Organization. MIT press, 1988.

Gerelt Tserenjigmid. The order dependent luce model, 2021.

Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. Science, 211:30, 1981.

Ying Zhang and Ayelet Fishbach. The role of anticipated emotions in the endowment effect. Journal of Consumer Psychology, 15(4):316–324, 2005.

Supplemental Material

(For Online Publication)

Appendix E: Characterization of the Non-separable AF Model – Proof of Theorem 2

The proof proceeds in five steps. We seek to obtain a regular representation of the form

$$w_f(x_f) = \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}) \prod_{k=1}^{i-1} B_k(u_{f(k)}(x_{f(k)})), \tag{E.1}$$

where $\prod_{k=1}^{0} B_k(u_{f(k)}(x_{f(k)})) \equiv 1$ and $B_i : \mathcal{U} \to \mathbb{R}_{++}$ for every $i = 1, \ldots, N$ and $\mathcal{U} = \cup_{a \in A} u_a(L_a)$. The representation in Theorem 2 follows from the change of variables

$$\varphi_k(u_{f(k)}(x_{f(k)})) = \ln\{B_k(u_{f(k)}(x_{f(k)}))\}.$$

For every (q, M), let

$$\ell(q, M) = \ln(\varphi(q, M)).$$

Step 1. Axiom 6 implies that for all x^i_f, p_f, y^i_f, p_f, x^i_f, q_f, and y^i_f, q_f in $\Delta(X_f)$, we have that

$$\frac{\varphi((x^i_f, p_f)|\{(x^i_f, p_f), (x^i_f, q_f)\})}{\varphi((x^i_f, q_f)|\{(x^i_f, p_f), (x^i_f, q_f)\})} = \frac{\varphi((y^i_f, p_f)|\{(y^i_f, p_f), (y^i_f, q_f)\})}{\varphi((y^i_f, q_f)|\{(y^i_f, p_f), (y^i_f, q_f)\})}.$$

Therefore,

$$\ell((x^i_f, p_f)|\{(x^i_f, p_f), (x^i_f, q_f)\}) - \ell((x^i_f, q_f)|\{(x^i_f, p_f), (x^i_f, q_f)\}) = v_f(x^i_f, p_f) - v_f(x^i_f, q_f)$$

is equal to

$$\ell((y^i_f, p_f)|\{(y^i_f, p_f), (y^i_f, q_f)\}) - \ell((y^i_f, q_f)|\{(y^i_f, p_f), (y^i_f, q_f)\}) = v_f(y^i_f, p_f) - v_f(y^i_f, q_f).$$

It follows that

$$v_f(x^i_f, p_f) \geq v_f(x^i_f, q_f) \Leftrightarrow v_f(y^i_f, p_f) \geq v_f(y^i_f, q_f).$$

This implies that $v_f(x^i_f, \cdot)$ and $v_f(y^i_f, \cdot)$ represent the same preference over $\Delta(\times_{k=i+1}^N L_{f(k)})$ for all $x^i_f, y^i_f \in \times_{k=1}^i L_{f(k)}$ and all $i = 1, \ldots, N-1$.

To unpack the consequences of this property, consider first $i = 1$ and fix any level $\overline{x}_{f(1)} \in L_{f(1)}$. By the uniqueness properties of vN-M utility representations, there exists $u^f_{f(1)}(x_{f(1)}; \overline{x}_{f(1)}) \in \mathbb{R}$ and $B^f_{f(1)}(x_{f(1)}; \overline{x}_{f(1)}) > 0$ such that, for all $x_{f(1)} \in L_{f(1)}$ and all

$p_f \in \Delta(\times_{k=2}^{N} L_{f(k)})$, we have

$$v_f(x_f^1, p_f) = u_{f(1)}^f(x_{f(1)}; \overline{x}_{f(1)}) + B_{f(1)}^f(x_{f(1)}; \overline{x}_{f(1)}) v_f(\overline{x}_f^1, p_f).$$

Therefore, clearly, $u_{f(1)}^f(\overline{x}_{f(1)}; \overline{x}_{f(1)}) = 0$ and $B_{f(1)}^f(\overline{x}_{f(1)}; \overline{x}_{f(1)}) = 1$.

Now consider $i = 2$ and focus on the elements (x_f^2, p_f) with the property that $x_{f(1)} = \overline{x}_{f(1)}$. Using Axiom 6, we conclude that $v_f(\overline{x}_{f(1)}, x_{f(2)}, \cdot)$ and $v_f(\overline{x}_{f(1)}, y_{f(2)}, \cdot)$ represent the same EU preference over $\Delta(\times_{k=3}^{N} L_{f(k)})$. Fix any level $\overline{x}_{f(2)} \in L_{f(2)}$. By the same uniqueness argument as before, there exists $u_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}, \overline{x}_{f(2)}) \in \mathbb{R}$ and $B_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}, \overline{x}_{f(2)}) > 0$ such that, for all $x_{f(2)} \in L_{f(2)}$ and all $p_f \in \Delta(\times_{k=3}^{N} L_{f(k)})$, we have

$$v_f(\overline{x}_{f(1)}, x_{f(2)}, p_f) = u_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}, \overline{x}_{f(2)}) + B_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}, \overline{x}_{f(2)}) v_f(\overline{x}_{f(1)}, \overline{x}_{f(2)}, p_f).$$

If we now replace in the expression for $v_f(x_f^1, p_f)$, we have

$$\begin{aligned} v_f(x_f^2, p_f) =& u_{f(1)}^f(x_{f(1)}; \overline{x}_{f(1)}) + B_{f(1)}^f(x_{f(1)}; \overline{x}_{f(1)}) \Big\{ u_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}, \overline{x}_{f(2)}) \\ & + B_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}, \overline{x}_{f(2)}) v_f(\overline{x}_{f(1)}, \overline{x}_{f(2)}, p_f) \Big\} \\ =& u_{f(1)}^f(x_{f(1)}; \overline{x}_f^1) + B_{f(1)}^f(x_{f(1)}; \overline{x}_f^1) u_{f(2)}^f(x_{f(2)}; \overline{x}_{f(1)}^2) \end{aligned}$$

$$+ B^f_{f(1)}(x_{f(1)};\overline{x}^1_f)B^f_{f(2)}(x_{f(2)};\overline{x}^2_f)v_f(\overline{x}^2_f, p_f).$$

Iteratively repeating this argument, we obtain that for all $x_f \in X_f$

$$v_f(x_f) = u^f_{f(1)}(x_{f(1)};\overline{x}^1_f) + \sum_{k=2}^{N} u^f_{f(k)}(x_{f(k)};\overline{x}^k_f)\prod_{j=1}^{k-1} B^f_{f(k)}(x_{f(k)};\overline{x}^k_f),$$

which becomes after suppressing the dependence on the arbitrary $\overline{x}_f$,

$$v_f(x_f) = u^f_{f(1)}(x_{f(1)}) + \sum_{k=2}^{N} u^f_{f(k)}(x_{f(k)})\prod_{j=1}^{k-1} B^f_{f(k)}(x_{f(k)}). \tag{E.2}$$

Step 2. Now consider Axiom 5. Note that $\varphi(x_f|\{x_f, y_f\}) \geq \frac{1}{2}$ is equivalent to

$$\ell(x_f|\{x_f, y_f\}) - \ell(y_f|\{x_f, y_f\}) = v_f(x_f) - v_f(y_f) \geq 0.$$

Using the representation in (E.2), we have that

$$v_f(x_f) \geq v_f(y_f) \Leftrightarrow u^f_a(x_a)\prod_{j=1}^{i-1} B^f_{f(j)}(x_{f(j)}) \geq u^f_a(y_a)\prod_{j=1}^{i-1} B^f_{f(j)}(y_{f(j)}) \Leftrightarrow u^f_a(x_a) \geq u^f_a(y_a),$$

because $\prod_{j=1}^{i-1} B^f_{f(j)}(x_{f(j)}) = \prod_{j=1}^{i-1} B^f_{f(j)}(y_{f(j)}) > 0$.

By similar reasoning, the axiom says that $v_{f'}(\hat{x}_{f'}) \geq v_{f'}(\hat{y}_{f'})$ for every specification of $\hat{x}_{f'(k)} = \hat{y}_{f'(k)}$. Also, recall that there exists $\overline{x}_{f'} \in X_{f'}$ such that $u^{f'}_{f'(k)}(\overline{x}_{f'(k)}) = 0$ for all k. Therefore, letting $\hat{x}_{f'(k)} = \hat{y}_{f'(k)} = \overline{x}_{f'(k)}$ for all $k \neq i$, we have

$$v_{f'}(\hat{x}_{f'}) \geq v_{f'}(\hat{y}_{f'}) \quad \Leftrightarrow \quad u^{f'}_a(x_a) \prod_{j=1}^{i-1} B^{f'}_{f'(j)}(\overline{x}_{f'(j)}) \geq u^{f'}_a(y_a) \prod_{j=1}^{i-1} B^{f'}_{f'(j)}(\overline{x}_{f'(j)})$$

$$\Leftrightarrow \quad u^{f'}_a(x_a) \geq u^{f'}_a(y_a),$$

because $\prod_{j=1}^{i-1} B^{f'}_{f'(j)}(\overline{x}_{f'(j)}) > 0$. We conclude that u^f_a and $u^{f'}_a$ represent the same ranking over L_a. In particular, this means that $u^f_a(x_a) = u^f_a(y_a)$ if and only if $u^{f'}_a(x_a) = u^{f'}_a(y_a)$.

Towards our goal of showing that each $B^f_{f(k)}$ depends on $x_{f(k)}$ only via $u^f_{f(k)}(x_{f(k)})$, consider first the simple case where $u^f_{f(k)}(x_{f(k)}) \neq u^f_{f(k)}(y_{f(k)})$ for all $x_{f(k)}, y_{f(k)} \in L_{f(k)}$ (i.e., $u^f_{f(k)}$ is injective). Then, we can just re-define

$$\hat{B}^f_{f(k)}(u^f_{f(k)}(x_{f(k)})) = B^f_{f(k)}((u^f_{f(k)})^{-1}(u^f_{f(k)}(x_{f(k)}))).$$

Thus, the desired property of the weights holds trivially.

Now consider the less immediate case where $u^f_{f(k)}(x_{f(k)}) = u^f_{f(k)}(y_{f(k)})$ for some $x_{f(k)}, y_{f(k)} \in L_{f(k)}$, where $f(k) = a \in A$ and $k < N$. Then, by the previous argument, for every frame f' with $f'(N) = a$, we must have $u^{f'}_a(x_{f(k)}) = u^{f'}_a(y_{f(k)})$ and so

$$v_{f'}(\hat{x}_{f'(1)}, \ldots, \hat{x}_{f'(N-1)}, x_{f(k)}) = v_{f'}(\hat{x}_{f'(1)}, \ldots, \hat{x}_{f'(N-1)}, y_{f(k)}).$$

By the axiom, we must also have

$$v_f(\overline{x}_{f(1)},\ldots,\overline{x}_{f(k-1)},x_{f(k)},z_{f(k+1)},\ldots,z_{f(N)}) = v_f(\overline{x}_{f(1)},\ldots,\overline{x}_{f(k-1)},y_{f(k)},z_{f(k+1)},\ldots,z_{f(N)}),$$

for every $(z_{f(k+1)},\ldots,z_{f(N)}) \in \times_{j=k+1}^{N} L_{f(j)}$. Therefore, after simplifying the term $\prod_{j=1}^{k-1} B_{f(j)}^{f}(\overline{x}_{f(j)}) > 0$ and using $u_{f(j)}^{f}(\overline{x}_{f(j)}) = 0$ for $j < k$, we have

$$\begin{aligned} 0 =& u_{f(k)}^{f}(x_{f(k)}) - u_{f(k)}^{f}(y_{f(k)}) \\ &+ \left[B_{f(k)}^{f}(x_{f(k)}) - B_{f(k)}^{f}(y_{f(k)})\right] \sum_{j=k+1}^{N} u_{f(j)}^{f}(z_{f(j)}) \prod_{i=k+1}^{j-1} B_{f(i)}^{f}(z_{f(i)}) \\ =& \left[B_{f(k)}^{f}(x_{f(k)}) - B_{f(k)}^{f}(y_{f(k)})\right] \sum_{j=k+1}^{N} u_{f(j)}^{f}(z_{f(j)}) \prod_{i=k+1}^{j-1} B_{f(i)}^{f}(z_{f(i)}), \end{aligned}$$

where $\prod_{i=k+1}^{k} B_{f(i)}^{f}(z_{f(i)}) \equiv 1$. Since $u_{f(j)}^{f}(z_{f(j)}) \neq 0$ for some $z_{f(j)} \in L_{f(j)}$ and some $j \geq k+1$, we must have

$$u_{f(k)}^{f}(x_{f(k)}) = u_{f(k)}^{f}(y_{f(k)}) \quad \Rightarrow \quad B_{f(k)}^{f}(x_{f(k)}) = B_{f(k)}^{f}(y_{f(k)}).$$

This implies that $B_{f(k)}^{f}$ cannot depend on $x_{f(k)}$ other than through $u_{f(k)}^{f}(x_{f(k)})$, as desired.

To recap, we now have the following representation: For all $f \in F$ and $x_f \in X_f$,

$$v_f(x_f) = u^f_{f(1)}(x_{f(1)}) + \sum_{j=2}^{N} u^f_{f(j)}(x_{f(j)}) \prod_{k=1}^{j-1} B^f_{f(k)}(u^f_{f(k)}(x_{f(k)})).$$

We concluded earlier that u^f_a for $f(N) = a$ and $u^{f'}_a$ for any other $f' \in F$—where $f'(N)$ may be different from a—represent the same ranking over L_a. By Axiom 6, u^f_a is also a vN-M utility function over L_a. Therefore, there exists $\gamma^{f'}_a > 0$ and $\zeta^{f'}_a \in \mathbb{R}$ such that, for every f' different from a fixed f^* with $f^*(N) = a$, we must have

$$u^{f'}_a = \gamma^{f'}_a u_a + \zeta^{f'}_a,$$

where we define $u_a = u^{f^*}_a$. Note that this implies that without loss of generality each $B^f_{f(k)}$ function depends on $x_{f(k)}$ only through $u_{f(k)}$: We can simply define

$$\hat{B}^f_{f(k)}(u_{f(k)}) = B^f_{f(k)}(\gamma^f_{f(k)} u_{f(k)} + \zeta^f_{f(k)}),$$

so we will omit $\gamma^f_{f(k)}$ and $\zeta^f_{f(k)}$ hereafter.

Therefore, (simplifying notation) we have

$$v_f(x_f) = \gamma^f_{f(1)} u_{f(1)}(x_{f(1)}) + \zeta^f_{f(1)} + \sum_{j=2}^{N} [\gamma^f_{f(j)} u_{f(j)}(x_{f(j)}) + \zeta^f_{f(j)}] \prod_{k=1}^{j-1} B^f_{f(k)}(u_{f(k)}(x_{f(k)}))$$

$$= \gamma^f_{f(1)} u_{f(1)}(x_{f(1)}) + \zeta^f_{f(1)} + \sum_{j=2}^{N} \gamma^f_{f(j)} u_{f(j)}(x_{f(j)}) \prod_{k=1}^{j-1} B^f_{f(k)}(u_{f(k)}(x_{f(k)}))$$

$$+ \sum_{j=2}^{N} \zeta^f_{f(j)} \prod_{k=1}^{j-1} B^f_{f(k)}(u_{f(k)}(x_{f(k)})).$$

Step 3. We now would like to show that each $B^f_{f(k)}$ depends only on the position k for all $f \in F$. To this end, we use Axiom 8, which implies the following. First, note that

$$\ell(x_f, \{x_f, \hat{x}_f\}) - \ell(\hat{x}_f, \{x_f, \hat{x}_f\}) = [u^f_{f(i)}(x_{f(i)}) - u^f_{f(i)}(\hat{x}_{f(i)})] \prod_{k=1}^{i-1} B^f_{f(k)}(u_{f(k)}(x_{f(k)}))$$

and

$$\ell(y_{f'}, \{y_{f'}, \hat{y}_{f'}\}) - \ell(\hat{y}_{f'}, \{y_{f'}, \hat{y}_{f'}\}) = u^{f'}_{f'(1)}(x_{f(i)}) - u^{f'}_{f'(1)}(\hat{x}_{f(i)}).$$

Using this, we have

$$\ell(x_f, \{x_f, \hat{x}_f\}) - \ell(\hat{x}_f, \{x_f, \hat{x}_f\}) = [u_{f(i)}(x_{f(i)}) - u_{f(i)}(\hat{x}_{f(i)})] \prod_{k=1}^{i-1} \gamma^f_{f(i)} B^f_{f(k)}(u_{f(k)}(x_{f(k)}))$$

and

$$\ell(y_{f'}, \{y_{f'}, \hat{y}_{f'}\}) - \ell(\hat{y}_{f'}, \{y_{f'}, \hat{y}_{f'}\}) = [u_{f'(1)}(x_{f(i)}) - u_{f'(1)}(\hat{x}_{f(i)})] \gamma^{f'}_{f(1)}.$$

The axiom requires that

$$\frac{\prod_{k=1}^{i-1}\gamma^{f}_{f(i)}B^{f}_{f(k)}(u_{f(k)}(x_{f(k)}))}{\gamma^{f'}_{f(1)}} = r(i, x_{f(1)}, \ldots, x_{f(i-1)}).$$

This implies that $\gamma^{f'}_{f(1)} = \gamma_1$ for all $f, f' \in F$ and some $\gamma_1 > 0$, $\gamma^{f}_{f(i)} = \gamma_i$ for all $f \in F$ and some $\gamma_i > 0$, and $B^{f}_{f(k)}(u_{f(k)}(x_{f(k)})) = B_k(u_{f(k)}(x_{f(k)}))$ for all $f \in F$ and some real number $B_k(u_{f(k)}(x_{f(k)})) > 0$. Thus, we can define $\mathcal{U} = \cup_{a\in A} u_a(L_a)$ and the function $B_k : \mathcal{U} \to \mathbb{R}_{++}$ as taking the values just defined.

These steps refine the representation of v_f to the following:

$$v_f(x_f) = \gamma_1 u_{f(1)}(x_{f(1)}) + \zeta^{f}_{f(1)} + \sum_{j-2}^{N} \gamma_j u_{f(j)}(x_{f(j)}) \prod_{k=1}^{j-1} B_k(u_{f(k)}(x_{f(k)}))$$

$$+ \sum_{j=2}^{N} \zeta^{f}_{f(j)} \prod_{k=1}^{j-1} B_k(u_{f(k)}(x_{f(k)})).$$

Step 4. By the uniqueness of v_f as a Luce value up to adding constants, we can set $\zeta^{f}_{f(1)} = 0$ for every f without loss. We would like to also show that $\zeta^{f}_{f(j)} = 0$ for every f and $j > 1$. To this end, we exploit Axiom 7 to further refine the representation as follows. For $i = 2$, its conclusion is equivalent to the equality between

$$\ell(x_f|\{x_f, y_f\}) - \ell(y_f|\{x_f, y_f\}) = v_f(x_f) - v_f(y_f)$$

$$= \gamma_1[u_{f(1)}(x_{f(1)}) - u_{f(1)}(y_{f(1)})]$$

$$+\zeta^f_{f(2)} \left[B_1(u_{f(1)}(x_{f(1)})) - B_1(u_{f(1)}(y_{f(1)}))\right]$$

and

$$\begin{aligned} \ell(x'_{f'}|\{x'_{f'}, y'_{f'}\}) - \ell(y'_{f'}|\{x'_{f'}, y'_{f'}\}) &= v_{f'}(x'_{f'}) - v_{f'}(y'_{f'}) \\ &= \gamma_1[u_{f(1)}(x_{f(1)}) - u_{f(1)}(y_{f(1)})] \\ &\quad +\zeta^{f'}_{f'(2)} \left[B_1(u_{f(1)}(x_{f(1)})) - B_1(u_{f(1)}(y_{f(1)}))\right]. \end{aligned}$$

This implies that

$$\left[\zeta^f_{f(2)} - \zeta^{f'}_{f'(2)}\right] \left[B_1(u_{f(1)}(x_{f(1)})) - B_1(u_{f(1)}(y_{f(1)}))\right] = 0.$$

Since B_1 is not constant, we must have $\zeta^f_{f(2)} = \zeta^{f'}_{f'(2)} = \zeta_{f(1)}$ for all f and f' that satisfy $f(1) = f'(1)$. Now suppose that, for all $k = 2, \ldots, j$, we have $\zeta^f_{f(k)} = \zeta^{f'}_{f'(k)} = \zeta_{f(1),\ldots,f(k-1)}$ for all f and f' that satisfy $f(m) = f'(m)$ for $m \leq k-1$. Let $i = j+1$ in Axiom 7. Its conclusion is equivalent to the equality between

$$\begin{aligned}
& \ell(x_f|\{x_f, y_f\}) - \ell(y_f|\{x_f, y_f\}) \\
= \quad & v_f(x_f) - v_f(y_f) \\
= \quad & \sum_{k=1}^{j} \gamma_k u_{f(k)}(x_{f(k)}) \prod_{m=1}^{k-1} B_m(u_{f(m)}(x_{f(m)})) \\
& - \sum_{k=1}^{j} \gamma_k u_{f(k)}(y_{f(k)}) \prod_{m=1}^{k-1} B_m(u_{f(m)}(y_{f(m)})) \\
& + \sum_{k=2}^{j} \zeta_{f(1),\dots,f(k-1)} \left\{ \prod_{m=1}^{k-1} B_m(u_{f(m)}(x_{f(m)})) - \prod_{m=1}^{k-1} B_m(u_{f(m)}(y_{f(m)})) \right\} \\
& + \zeta^f_{f(j+1)} \left\{ \prod_{m=1}^{j} B_m(u_{f(m)}(x_{f(m)})) - \prod_{m=1}^{j} B_m(u_{f(m)}(y_{f(m)})) \right\}
\end{aligned}$$

and

$$\begin{aligned}
& \ell(x'_{f'}|\{x'_{f'}, y'_{f'}\}) - \ell(y'_{f'}|\{x'_{f'}, y'_{f'}\}) \\
= \quad & v_{f'}(x'_{f'}) - v_{f'}(y'_{f'}) \\
= \quad & \sum_{k=1}^{j} \gamma_k u_{f(k)}(x_{f(k)}) \prod_{m=1}^{k-1} B_m(u_{f(m)}(x_{f(m)}))
\end{aligned}$$

$$-\sum_{k=1}^{j}\gamma_k u_{f(k)}(y_{f(k)})\prod_{m=1}^{k-1}B_m(u_{f(m)}(y_{f(m)}))$$

$$+\sum_{k=2}^{j}\zeta_{f(1),\ldots,f(k-1)}\left\{\prod_{m=1}^{k-1}B_m(u_{f(m)}(x_{f(m)}))-\prod_{m=1}^{k-1}B_m(u_{f(m)}(y_{f(m)}))\right\}$$

$$+\zeta^{f'}_{f'(j+1)}\left\{\prod_{m=1}^{j}B_m(u_{f(m)}(x_{f(m)}))-\prod_{m=1}^{j}B_m(u_{f(m)}(y_{f(m)}))\right\}.$$

This implies that

$$\left[\zeta^{f}_{f(j+1)}-\zeta^{f'}_{f'(j+1)}\right]\left\{\prod_{m=1}^{j}B_m(u_{f(m)}(x_{f(m)}))-\prod_{m=1}^{j}B_m(u_{f(m)}(y_{f(m)}))\right\}=0.$$

Since the quantity in brackets is again not constant, we must have $\zeta^{f}_{f(j+1)}=\zeta^{f'}_{f'(j+1)}=\zeta_{f(1),\ldots,f(j)}$ for all f and f' that satisfy $f(m)=f'(m)$ for $m\leq j$. By induction, we can extend this property to every $j=2,\ldots,N$.

Now consider any f and any x_f that satisfies $x_{f(i)}=\overline{x}_{f(i)}$ for all $i\geq 2$, so that $\gamma_i u_{f(i)}(x_{f(i)})+\zeta_{f(1),\ldots,,f(i-1)}=0$. In this case, we have that $v_f(x_{f(1)},\overline{x}_{f(-1)})$ is a vN-M utility function over $L_{f(1)}$ and it takes the form

$$v_f(x_{f(1)},\overline{x}_{f(-1)})=\gamma_1 u_{f(1)}(x_{f(1)})+\zeta_{f(1)}B_1(u_{f(1)}(x_{f(1)})).$$

Since $u_{f(1)}$ is also a vN-M utility function over $L_{f(1)}$, we must have

$$v_f(x_{f(1)}, \overline{x}_{f(-1)}) = \hat{\gamma}_{f(1)} u_{f(1)}(x_{f(1)}) + \hat{\zeta}_{f(1)}.$$

This implies that

$$[\hat{\gamma}_{f(1)} - \gamma_1] u_{f(1)}(x_{f(1)}) + \hat{\zeta}_{f(1)} = \zeta_{f(1)} B_1(u_{f(1)}(x_{f(1)})).$$

There are several cases to consider, which all yield $\zeta_{f(1)} = 0$. First, if $\hat{\gamma}_{f(1)} = \gamma_1$, then the equality can hold if and only if $\hat{\zeta}_{f(1)} = \zeta_{f(1)} = 0$ because B_1 is not constant. Given this, suppose that $\hat{\gamma}_{f(1)} > \gamma_1$ without loss of generality. If B_1 is not an affine function of $u_{f(1)}$, the equality can only hold if and only if $\zeta_{f(1)} = 0$. Finally, suppose that B_1 is indeed affine in $u_{f(1)}$, that is, there exist $\overline{\gamma}_1 > 0$ and $\overline{\zeta}_1 \in \mathbb{R}$ such that

$$B_1(u_{f(1)}) = \overline{\gamma}_1 u_{f(1)} + \overline{\zeta}_1.$$

In this case, we have

$$\begin{aligned} v_f(x_{f(1)}, \overline{x}_{f(-1)}) &= \gamma_1 u_{f(1)}(x_{f(1)}) + \zeta_{f(1)}[\overline{\gamma}_1 u_{f(1)}(x_{f(1)}) + \overline{\zeta}_1] \\ &= [\gamma_1 + \zeta_{f(1)} \overline{\gamma}_1] u_{f(1)}(x_{f(1)}) + \zeta_{f(1)} \overline{\zeta}_1. \end{aligned}$$

By the uniqueness properties of any Luce value function, we can let $\overline{\zeta}_1 = 0$ without loss of generality. Finally, by Axiom 8, $\gamma_1 + \zeta_{f(1)}\overline{\gamma}_1$ cannot depend on $f(1)$ but only on the position $i = 1$. Thus, without loss of generality, we can let $\zeta_{f(1)} = 0$ and adjust γ_1 accordingly.

Now, suppose we established that $\zeta_{f(1),\ldots,f(j-1)} = 0$ for all $j = 1, \ldots, i-1$. Consider any f and any x_f that satisfies $x_{f(j)} = \overline{x}_{f(j)}$ for all $j \neq i$, so that $\gamma_j u_{f(j)}(x_{f(j)}) + \zeta_{f(1),\ldots,,f(j-1)} = 0$. Again, we have that $v_f(x_{f(i)}, \overline{x}_{f(-i)})$ is a vN-M utility function over $L_{f(i)}$ and it takes the form

$$v_f(x_{f(i)}, \overline{x}_{f(-i)}) = [\gamma_i u_{f(i)}(x_{f(i)}) + \zeta_{f(1),\ldots,f(i-1)} B_i(u_{f(i)}(x_{f(i)}))]\overline{B}_i,$$

where $\overline{B}_i = \prod_{k=1}^{i-1} B_k(u_{f(k)}(\overline{x}_{f(k)})) > 0$. Since $u_{f(i)}$ is also a vN-M utility function over $L_{f(i)}$, we must have

$$v_f(x_{f(i)}, \overline{x}_{f(-i)}) = \hat{\gamma}_{f(i)} u_{f(i)}(x_{f(i)}) + \hat{\zeta}_{f(i)}.$$

This implies that

$$[\hat{\gamma}_{f(i)} - \gamma_i \overline{B}_i] u_{f(i)}(x_{f(i)}) + \hat{\zeta}_{f(i)} = \zeta_{f(1),\ldots,f(i-1)} B_i(u_{f(i)}(x_{f(i)})).$$

Repeating the previous reasoning, we again conclude that $\zeta_{f(1),\ldots,f(i-1)} = 0$ without loss of generality.

Step 5. Combining all these steps, we obtain the representation

$$w_f(x_f) = \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}) \prod_{k=1}^{i-1} B_k(u_{f(k)}(x_{f(k)})),$$

where the function α is defined by $\alpha(i) = \gamma_i$ for all $i = 1, \ldots, N$.

We conclude by showing that (α, u, φ) must satisfy regularity. Using the representation, for any $x_{f'}$ and $y_{f'}$ such that $x_{f'(-N)} = y_{f'(-N)}$ we have that $\varphi(x_{f'}|\{x_{f'}, y_{f'}\}) \geq \frac{1}{2}$ if and only if $u_{f'(N)}(x_{f'(N)}) \geq u_{f'(N)}(y_{f'(N)})$. By Axiom 5, we must have $\varphi(x_f|\{x_f, y_f\}) \geq \frac{1}{2}$ for any f that presents attribute $f'(N)$ in any position $i < N$. This holds if $u_{f(i)}(x_{f(i)}) = u_{f(i)}(y_{f(i)})$, so hereafter assume that $u_{f(i)}(x_{f(i)}) > u_{f(i)}(y_{f(i)})$. In this case, $\varphi(x_f|\{x_f, y_f\}) \geq \frac{1}{2}$ if and only if

$$\sum_{j=i}^{N} u_{f(j)}(x_{f(j)}) \alpha(j) \prod_{k=1}^{j-1} B_k(u_{f(k)}(x_{f(k)})) \geq \sum_{j=i}^{N} u_{f(j)}(y_{f(j)}) \alpha(j) \prod_{k=1}^{j-1} B_k(u_{f(k)}(y_{f(k)})).$$

Since $\prod_{k=1}^{i-1} B_k(u_{f(k)}(x_{f(k)})) = \prod_{k=1}^{i-1} B_k(u_{f(k)}(y_{f(k)})) > 0$, the last condition is equivalent to

$$u_{f(i)}(x_{f(i)}) \alpha(i) + \sum_{j=i+1}^{N} u_{f(j)}(x_{f(j)}) \alpha(j) \prod_{k=i}^{j-1} B_k(u_{f(k)}(x_{f(k)}))$$

$$\geq u_{f(i)}(y_{f(i)}) \alpha(i) + \sum_{j=i+1}^{N} u_{f(j)}(y_{f(j)}) \alpha(j) \prod_{k=i}^{j-1} B_k(u_{f(k)}(y_{f(k)})).$$

Using now $\prod_{k=i+1}^{i-1} B_k(u_{f(k)}(x_{f(k)})) = \prod_{k=i+1}^{i-1} B_k(u_{f(k)}(y_{f(k)}))$ and $u_{f(i)}(x_{f(i)}) > u_{f(i)}(y_{f(i)})$, we obtain

$$\alpha(i)$$

$$\geq -\frac{B_i(u_{f(i)}(x_{f(i)})) - B_i(u_{f(i)}(y_{f(i)}))}{u_{f(i)}(x_{f(i)}) - u_{f(i)}(y_{f(i)})} \sum_{j=i+1}^{N} u_{f(j)}(x_{f(j)})\alpha(j) \prod_{k=i+1}^{j-1} B_k(u_{f(k)}(x_{f(k)})).$$

Note that this same condition is required if we started with $u_{f(i)}(x_{f(i)}) < u_{f(i)}(y_{f(i)})$. Since this has to hold for all $f \in F$, $x_f \in X_f$, and $y_{f(i)}$, it is equivalent to the condition in Definition 8.

E.1. Framing in Perturbed-Utility Models

Fudenberg et al. (2015) describe the choice probabilities as resulting from a maximization problem: Ann maximizes the expected utility of her choices net of some cost that is convex in probabilities. Formally, for every M_f

$$(\varphi(x_f|M_f), \ldots, \varphi(y_f|M_f)) \in \arg \max_{\varphi \in \Delta(M_f)} \Big\{ \sum_{z_f \in M_f} v(z_f)\varphi(z_f) - \chi(\varphi(z_f)) \Big\},$$

where χ is a perturbation function that may reward Ann for randomizing.

We connect the two models building on Fudenberg et al.'s (2015) elegant analysis. Given any continuous and strictly increasing function $h : (0,1) \to \mathbb{R}_+$, define the marginal cost as

$$\chi'(\varphi) = \ln(h(\varphi)).$$

Fudenberg et al. (2015) show that the utility of any two items x_f and y_f satisfies

$$v_f(x_f) - v_f(y_f) = \chi'(\varphi(x_f|\{x_f, y_f\})) - \chi'(\varphi(y_f|\{x_f, y_f\})).$$

Our previous characterizations involved specifying properties of payoff differences of the form $v_f(x_f) - v_f(y_f)$ through our axioms. Thus, to specify similar properties in the perturbed-utility framework, we only need to reformulate our axioms in terms of the "rescaled" probabilities $h(\varphi(x_f|\{x_f, y_f\}))$.

E.2. Framing in Rational-Inattention Models

We can also connect our theory of attribute framing with the random-choice model based on rational inattention proposed by Matějka and McKay (2015). Recall that we can think of M_f as a table, where attributes correspond to the rows $i = 1, \ldots, N$ and items to the columns $j = 1, \ldots, |M_f|$. Then, choosing an item leads to the consequence of getting the attributes $(x^j_{f(i)})_{i=1}^N$ in the corresponding column j.

The rational-inattention model is based on the idea that the decision-maker is uncertain about the consequences of his choices and spends costly attention to learn about them. In our case, suppose Bob is uncertain about the entries of the table (i.e., M_f) and so about the levels of the attributes obtained by selecting a specific item (i.e., column). Let G be his prior about the entries of menus. As in Matějka and McKay (2015), suppose Bob allocates attention to the items in a menu incurring a cost in the form of entropy reduction. Following Matějka and McKay (2015), the solution to Bob's optimal attention-allocation problem leads to the choice probabilities

$$\varphi(x_f^j|M_f) = \frac{\overline{\varphi}^j e^{v_f(x_f^j)}}{\sum_{j'} \overline{\varphi}^{j'} e^{v_f(x_f^{j'})}}, \qquad \text{where } \overline{\varphi}^i = \mathbb{E}_G[\varphi(x_f^i|M_f)] \tag{E.3}$$

for every M_f. Thus, for every realization of the entries in table M_f, the probability that Bob chooses the item in column j is similar to our Luce model (8), except for the additional weights $\overline{\varphi}^j$. Each $\overline{\varphi}^j$ equals the ex-ante probability of choosing the item in column j averaging over all realizations of M_f.

We can connect our theory with this model as follows. Note that expression (E.3) implies that, for fixed column j,

$$v_f(x_f^j) - v_f(y_f^j) = \ln(\varphi(x_f^j|M_f)) - \ln(\varphi(y_f^j|M_f)).$$

This again suggests a simple way to adapt our axioms to specialize the function v_f in the present context as we did in the previous characterizations.

The flexibility of Matejka and Makay's (2015) framework allows for several extensions of our theory of attribute-framing effects. First, it allows to overcome some of the well-known limitations of the Luce model (like unrealistic responses to item copies). More interestingly for us, it allows for interactions between attribute-order and list-order effects. To illustrate the point, suppose that as Bob goes through the columns from left to right, he gets tired and pays less attention to later columns. In this case, no matter what, he is overall less likely to choose items in later columns (i.e., $\overline{\varphi}^j > \overline{\varphi}^{j+1}$ for all j). This can be formalized by assuming a prior G specifying that later columns are very likely to have sufficiently bad attribute realizations. Importantly, the average weights $\overline{\varphi}^j$ have to be consistent with the actual choice frequencies, which are affected by the attribute frames. Therefore, those frames ultimately influence the list-order effects. We leave studying such interactions for future research.

Appendix F: General Menus and Frame-Driven List Effects

Sometimes decision-makers face items whose attributes are presented with different orders—for example, on the shelves of grocery stores or when sellers independently choose how to best frame their products. This leads to general menus where different frames are present simultaneously. How do frames affect choice in such cases?

The question is non-trivial, but our model can help organize the discussion and provide some answers. The idea is to let the data speak. We can identify our model using f-menus only (see Section 4). We can then use it to predict choices under different hypotheses on how the decision-maker responds to general menus—we present a few shortly. We can test these hypotheses against the data and select which we judge to be the best fit. Formally, let H be an hypothesis and $\mathbf{M}$ a collection of menus. The decision-maker's choices from these menus form the actual dataset $c(\mathbf{M})$. Using our model with (α, u) calibrated to this decision-maker, we can calculate his utility from each item and his choices under H, which form the predicted dataset $\hat{c}(\mathbf{M}; H)$. We can then compare $c(\mathbf{M})$ and $\hat{c}(\mathbf{M}; H)$ in any standard way. For example, H can be falsified if $\hat{c}(\mathbf{M}; H) \neq c(\mathbf{M})$, or, more realistically, if $d(\hat{c}(\mathbf{M};H), c(\mathbf{M})) > \tau$ for some distance function d and tolerance threshold $\tau > 0$. Among multiple hypotheses, we may select the one that minimizes $d(\hat{c}(\mathbf{M};H), c(\mathbf{M}))$.[29]

To illustrate this approach, we consider three hypotheses.

Own-frame hypothesis (H_1): Suppose that, when facing a general menu, Bob chooses as if he evaluates each item following its own order of attributes. That is, for every M, the predicted choice is

$$\hat{c}(M;H_1) = \arg\max_{x_f \in M} \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}).$$

Note that for f-menus this boils down to the model in Definition 1.

Single-reframe hypothesis (H_2): A second possibility is that, when facing a general menu, Bob chooses as if he reframes all its items using the same f. Formally, let φ be a function that maps every M to some $f \in F$. Let $M_{\varphi(M)}$ be the $\varphi(M)$-menu that contains all items in the original M presented with frame $\varphi(M)$. In this case, Bob's choice from M should coincide with that predicted by the AF model from $M_{\varphi(M)}$:

$$\hat{c}(M;H_2) = \arg\max_{x_{\varphi(M)} \in M_{\varphi(M)}} \sum_{i=1}^{N} \alpha(i) u_{\varphi(M)(i)}(x_{\varphi(M)(i)}).$$

This hypothesis is very flexible. Bob may adopt different frames for different menus. Alternatively, he may adopt the same frame for all general menus. That is, H_2 covers

[29] Standard methods can be used to define d. For example, one can use the *swap index* in Apesteguia and Ballester (2015). Given H, the model generates a preference relation over the items in each menu. One can measure the distance between the actual and predicted choices by the number of swaps in the preference relation needed to make the actual choice preferred to the predicted choice. One can then aggregate this measure across menus and choose the hypothesis that minimizes it.

the possibility that general menus cause framing effects to disappear. This may happen if the effort to organize and make sense of the various items causes the different emphasis put on attributes to wane.

Anchor-frame hypothesis (H_3): A special case of H_2 is that Bob chooses as if he uses the frame of one item in the menu as an anchor (see Krosnick and Alwin (1987) for consistent evidence). For example, the f of the first listed item, the last listed item, or the most frequent f could cue Bob to use f to compare all items in the menu. This anchoring may introduce links between attribute-order framing effects and list framing effects, which we can formalize and test with our model (see Section E.2 for further discussion). In fact, list framing effects are sometimes viewed as the outcome of particular ways of processing the attributes of the listed items (see, e.g., Rubinstein and Salant (2006) and references therein). Knowing what determines framing anchors can be valuable for sellers. For instance, if the anchor is the item listed first, sellers have an incentive to try to put their product in that position and frame it in the most favorable way. This mechanism may contribute to explaining why firms pay a premium to be listed first, say, by search engines.

The role of items as cues for how to reframe other items in general menus renders choice and preferences menu dependent. This can lead to failures of standard axioms, such as the independence of irrelevant alternatives (IIA). This is intuitive. Consider the menus $\{z_{f''}, x_f, y_{f'}\}$ and $\{x_f, y_{f'}\}$. Suppose for general menus Bob uses the item listed first to reframe the others. Then, it is possible that for him x dominates y and z under frame f'', but y dominates x under f. Thus, through the lens of our model we can understand violations of IIA as resulting from attribute-framing effects. Of course, IIA may fail for many other reasons. The example also illustrates that, due to the same mechanism, larger menus may increase the likelihood that some item is chosen by cuing a frame that favors it. This violates other regularity axioms that characterize standard choice models.

Appendix G: Self-Serving Rationalization via Framing

A large body of evidence shows that people often engage in motivated reasoning, rationalization, self-deception, self-justification, and reduction of cognitive dissonance by strategically presenting to themselves situations and decisions in the most favorable *perspective* (Bénabou and Tirole (2016)). One way is to emphasize some of their aspects over others. Such habits can be conscious or automatic, affective (to feel better) or functional (to achieve goals), and depend on emotions. For instance, rationalization can serve to avoid disappointment, guilt, or regret. Cognitive dissonance may result in a strategy called minimization, namely, reducing the importance of elements of dissonance (Lindsey-Mullikin (2003), Beasley and Joslyn (2001)). Self-serving justification aims to make questionable behaviors appear less unethical. It can occur ex ante—to paint violations as excusable in the eye of one's moral self—or ex post—to lessen the experienced threat for one's moral self (Shalvi et al. (2015)).

A key question is how to capture self-serving perspective manipulations in a disciplined manner. We argue that our framework can provide a way. Our premise is that, when making decisions, some individuals are susceptible to frames set by someone else, like salespeople or experimenters. In a similar logic, the choosing self of such individuals may also be influenced by frames set by their rationalizing self. This dual-self view is consistent with leading models of motivated reasoning (Bénabou and Tirole (2016)). Imagine we can describe the choosing self with our AF model. The rationalizing self can set f to manipulate the perspective under which the choosing self makes decisions, emphasizing certain aspects with their presentation order. Introspection suggests that when facing a decision—especially new and complex ones—we first try to organize its aspects, thereby forming a specific presentation order. This order may depend on our motivations and affect our choice.

We distinguish between two scenarios: *ex-ante* and *ex-post* self-serving framing. In the first, f is set *before* the choosing self makes a decision; in the second, f is set *after* a decision. The rationalizing self may want to maximize or minimize the evaluation of an item depending on her motivation in the situation of the moment (Bénabou and Tirole (2016)). For every item $x \in X$, define $\overline{f}_x$ and $\underline{f}_x$ as

$$\overline{f}_x \in \arg\max_{f \in F} \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}) \qquad \text{and} \qquad \underline{f}_x \in \arg\min_{f \in F} \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}).$$

Ex ante, Ann may adopt $\overline{f}_x$ to motivate herself to do x—say, exercise—or to justify doing x—say, violating some norm; she may adopt $\underline{f}_x$ if for example she is going to bargain on the price of x. Ex post, Bob may adopt $\overline{f}_x$ if he bought x and wants to justify the expenditure to himself, or $\underline{f}_x$ if he could not get x and wants to lessen his feeling of regret or disappointment.

Importantly, our model imposes trade-offs and constraints on the rationalizing self. An individual cannot deceive herself without limits. Emphasizing some aspects requires de-emphasizing others: It is not possible to simply increase or decrease the weight on all attributes. Also, de-emphasizing has bounded effects: It is not possible to 'forget' bad aspects since $\alpha > 0$. Finally, our model assigns a precise meaning to frames, namely, the order in which the rationalizing self describes the aspects of an item.

The possibility that a decision-maker may frame items in a self-serving manner is consistent with the possibility that he is influenced by frames set by others. Bob's choosing self may buy x under the influence of some f in the store, which can differ from the frame $\overline{f}_x$ her rationalizing self sets once at home (recall Proposition 3). This relates to and offers a formalization of the distinction between *decision utility* and *experienced utility* (Kahneman et al. (1997); Kahneman et al. (1999)). The first is the utility that drives decisions in the heat of the moment—for instance, in the store under the f crafted by a skillful salesperson. The second is the hedonic utility experienced in the cold state of the rationalizing moment—for instance, at home after calmly thinking about the bought item. For the above reasons, the experienced utility may be determined by $\overline{f}_x$. Our model provides a tool to calculate both decision and experienced utility knowing (α, u).

G.1. A Framing Perspective on the Endowment Effect

To illustrate the logic of self-serving framing, we apply it in the context of a well-known phenomenon: the endowment effect (Thaler (1980)). This phenomenon relates the willingness to pay (WTP) for acquiring an object and the willingness to accept (WTA) for giving up possession of the same object. Standard choice theory predicts that $WTA = WTP$. Yet, evidence shows that subjects often exhibit $WTA > WTP$ (Kahneman et al. (1991)). Here we sketch one angle to think about this phenomenon, which may complement the leading explanation based on expectation-based reference dependence (Kőszegi and Rabin (2006)).

Imagine the following situation. The choice items of interest have multiple attributes. Ann can be described by an AF model (α, u), where α is not constant. Her value of having no item is zero. We offer Ann the possibility of acquiring x under frame f, which determines her decision utility for x_f. Assuming quasi-linearity in money, we can define Ann's willingness to pay for x as

$$WTP(x) = \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}).$$

If when asked to give up x Ann evaluates it under f, we observe $WTA(x) = WTP(x)$. But she may frame x differently at this stage, which leads to $WTP(x) \neq WTA(x)$.

Even so, why should reframing of acquired items systematically lead to $WTA \geq WTP$? We already mentioned reasons for self-serving framing suggested by cognitive science and psychology that may lead Ann to try to avoid negative feelings ex post. In addition, according to Beggan (1992) the desire to see oneself favorably may induce people to overvalue objects associated with the self, namely, owned objects. Thus, Ann may tend to use $\overline{f}_x$ when considering giving up x. This implies

$$WTA(x) = \sum_{i=1}^{N} \alpha(i) u_{\overline{f}_x(i)}(x_{\overline{f}_x(i)}) \geq WTP(x)$$

for every initial f, with strict inequality for some f. The avoidance of negative feelings seems to be a potential cause of the endowment effect (Zhang and Fishbach (2005)). Other evidence shows that the longer an individual owns the item, the bigger is the WTA-WTP gap (Strahilevitz and Loewenstein (1998)). Presumably, the longer ownership allows Ann to figure out the best frame for x.

It is worth noting that even if sellers can select f to maximize profits, a gap between WTP and WTA may still arise. Section 3.2 showed that in competitive settings it may not be optimal for sellers to select $f = \overline{f}_x$.

One final insight of our model is that experience should eliminate the endowment effect. If Ann remembers how she reframed x after experiencing it a few times, the WTP–WTA gap should disappear because she cannot be manipulated again by changing the presentation of x. Consistent with this, evidence shows that market experience seems to eliminate the endowment effect (List (2003)) and that the effects of attribute-order framing disappear for subjects who had experience with the choice items (Levin and Gaeth (1988), Kumar and

Gaeth (1991)). This does not reduce the importance of studying framing effects, as many and consequential choices in life happen infrequently and with little to no feedback.

Appendix H: Welfare Analysis

Our framework—by being explicit about what frames are and how they work—provides several ways to approach the thorny question of running welfare analysis in the presence of framing effects (Bernheim and Rangel (2009); Rubinstein and Salant (2011)). Each way has some merits and flaws. Since these are often well-known, we limit ourselves to discuss what these are.

H.1. Choice-Based Welfare

One way is to apply the choice-based approach proposed by Bernheim and Rangel (2009). They define a generalized choice situation as a constraint set paired with an ancillary condition. In our setting, the constraint set corresponds to a set of choice items $D \subset X$; the ancillary condition is the collection $\mathbf{f}$ of all frames f with which the items are presented. Following Bernheim and Rangel (2009), we say that it is possible to strictly improve upon $x \in \hat{D}$ if there exists $x' \in \hat{D}$ such that, for all $(D, \mathbf{f})$ which satisfy $x, x' \in D$, the decision-maker never chooses x. We can then say that x is a weak individual welfare optimum when a strict improvement is not possible.[30] Bernheim and Rangel (2009) show that this welfare criterion is the most discerning criterion that never overrules choice.

Bernheim and Rangel's criterion has specific implications for our model (see also their Theorem 3). Suppose that for every $\mathbf{f}$ Ann compares $x, x' \in D$ using a common frame $f \in \mathbf{f}$ (like in the case of f-menus or hypothesis H_2 above). Then, she never chooses x when x' is

[30] We refer the reader to Bernheim and Rangel (2009) for the definitions of weak improvement and strict welfare optimum.

available if and only if for all $f \in F$

$$\sum_{i=1}^{N} \alpha(i) u_{f(i)}(x'_{f(i)}) > \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}).$$

Therefore, x is (weakly) welfare optimal if for all x' there exists *some* f such that

$$\sum_{i=1}^{N} \alpha(i) u_{f(i)}(x_{f(i)}) \geq \sum_{i=1}^{N} \alpha(i) u_{f(i)}(x'_{f(i)}).$$

In other words, x is welfare optimal if for every x' Ann prefers x_f to x'_f for some appropriately chosen f based on x'. Note that this f need not be the one that maximizes the utility from x (i.e., $\overline{f}_x$).

H.2. Frame-Free Welfare

The last alternative we discuss uses the properties of our model to entirely remove frames from welfare measures. It is based on the premise that frames should not matter for decision-makers' choices and hence, a fortiori, for a planner's welfare analysis.

The idea is to exploit our model's identification of the tastes for each attribute. We can define the frame-free welfare generated by an item as the sum of the utilities of its attributes. That is, given a decision-maker described by (α, u), this measure is

$$U^o(x) = \sum_{a \in A} u_a(x_a), \quad x \in X.$$

Of course, this way of removing framing effects involves some degree of paternalism. For another interpretation of this approach, note that U^o is equivalent, in terms of ranking, to taking the average across all frames of the total utility of an item. Indeed, since each attribute can be presented in each position, we have

$$\frac{\sum_{f\in F}\sum_{i=1}^{N}\alpha(i)u_{f(i)}(x_{f(i)})}{|F|} = \left(\frac{\sum_{i=1}^{N}\alpha(i)}{|A|}\right)\sum_{a\in A}u_a(x_a).$$

H.3. Experienced-Utility Welfare

Another typical approach of behavioral economics to welfare analysis involves the distinction between decision utility and experienced (or true) utility, where the latter should be used to measure well-being (Kahneman et al. (1997); Kahneman et al. (1999); Bernheim and Rangel (2009)). As noted in Section G, our model provides a way and a rationale for defining experienced utility for decision-makers affected by attribute-framing. Suppose that, for the reasons discussed above, Ann re-frames each owned item x according to $\overline{f}_x$ after acquiring it. Then, her experienced utility is

$$\overline{U}(x) = \sum_{i=1}^{N}\alpha(i)u_{\overline{f}_x}(x).$$

This welfare measure is based on the idea that the important source of well-being for Ann is the utility she experiences once owning x, not the utility she used to choose x.

It is worth noting a couple of properties of the experienced utility $\overline{U}$. First, it defines a ranking over items that is frame-independent. The original dependence is removed by considering the best frame $\overline{f}_x$ for each x. Second, although the underlying AF model is additively separable across attributes for every f, the induced $\overline{U}$ need not be separable as the whole x determines the best frame $\overline{f}_x$. Thus, there can be interdependences between attributes that are exclusively driven by self-serving framing considerations. This structural difference between decision utility and experienced utility may suggest a way to test this approach to welfare analysis.