Transportation Research Part B 93 (2016) 318–337

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Modeling the decoy effect with context-RUM Models: Diagrammatic analysis and empirical evidence from route choice SP and mode choice RP case studies

C. Angelo Guevara[a,*], Mitsuyoshi Fukushi[b]

[a] *Departamento de Ingeniería Civil, Universidad de Chile, Casilla 228-3, Santiago, Chile*
[b] *Facultad de Ingeniería y Ciencias Aplicadas, Universidad de los Andes, Monseñor Álvaro del Portillo 12.455, Las Condes, Chile*

ARTICLE INFO

Article history:
Available online 9 August 2016

Keywords:
Regularity
Independence of irrelevant alternatives (IIA)
Random regret minimization (RRM)
Cross-validation

ABSTRACT

Evidence outside transportation has suggested that the introduction of a decoy to the choice-set could increase the share of other alternatives. This evidence breaks the regularity assumption, which is at the root of the classical Random Utility Maximization (RUM) model with utilities that ignore the choice context. This article assesses the suitability of various context-RUM choice models that could overcome this limitation. For this we use a diagrammatic analysis, as well as Stated Preference (SP) and Revealed Preference (RP) transportation choice evidence. We begin confirming that the reported decoy outcomes cannot be replicated with the classical RUM models and that such a goal could be achieved instead using a set of five context-RUM models. We then show, for the first time, that the Asymmetrically Dominated (AD) and Compromise (CP) decoy effects were present in an SP route choice setting. We also show that, for a subset of individuals, the relative strength of the different decoy types was coherent with a Data Generation Process (DGP) defined by the Random Regret Minimization (RRM) or by the Regret by Aspects (RBA) parsimonious models. Then, we use cross-validation analysis where we found that RRM and RBA were superior to a classical Logit for all decoy types. Nevertheless, the ad-hoc Emergent Value (EV) model was consistently superior to all models suggesting that, although the parsimonious models may in theory replicate all decoy types, they seem to still make an incomplete representation of the DGP behind the overall decoy effect. We finally consider an RP mode choice experiment with which we detect, for the first time, an AD decoy effect in this choice setting. We also use this experiment to illustrate how to handle the decoy phenomena in a real context with various alternatives and variables. The article concludes summarizing the main contributions of this research and suggesting future lines of investigation for it.

1. Introduction

Proper modeling of choice behavior is crucial for the analysis of transportation systems because systematic experimentation is not feasible in this area. Classical choice models, such as the Logit or the Nested Logit (see e.g. Ben-Akiva and Lerman, 1985), are based on the assumption that individuals are rational, what means that their choices are based in what is known as the Random Utility Maximization (RUM) principle. Nevertheless, various experiments performed in the areas

* Corresponding author.
E-mail address: crguevar@ing.uchile.cl (C.A. Guevara).

http://dx.doi.org/10.1016/j.trb.2016.07.012
0191-2615/

of psychology and behavioral economics seem to have defied the rationality principle (see. e.g. Ariely, 2008). The essential need for suitable transportation models and the apparent refutation of the key assumption of currently used tools, calls for a review of their building blocks and, potentially, for their adaptation to overcome their possible limitations. Consequently, the main purpose of this article is to analyze practical context-RUM models that could achieve such a goal.

One of the so called “irrational behaviors” that have been described in the literature is known as the decoy effect (Huber et al., 1982). It consists in that, under some circumstances, the introduction of a new alternative in a choice-set may increase the probability of choosing one of the former options. The decoy effect is a direct violation of the regularity assumption (Luce, 1977), which is at the root of rational choice models.

Consider, for example, the case of a Logit model representing the choice of individual n among the alternatives j within the choice-set C_n. The rational individual chooses the alternative j with the largest utility $U_{jn} = V_{jn} + \varepsilon_{jn}$, which systematic part $V_{jn} = \sum_k \beta_k x_{kjn}$ depends linearly on attributes x_{kjn} with coefficients β_k, and which random part ε_{jn} is assumed to be *iid* Extreme Value I (0, μ). Under this setting, if the individual faces the choice among two alternatives, O (Objective) and C (Competitor), with systematic utilities V_{On} and V_{Cn} respectively, the probability that n would choose O will be

$$P_n(O|C_n = \{O, C\}) = \frac{e^{\mu V_{On}}}{e^{\mu V_{On}} + e^{\mu V_{Cn}}}.$$

Consider now that a third alternative D (Decoy), with systematic utility V_{Dn}, is included in the choice-set. In such a case, the probability that n would choose O would become

$$P_n(O|C_n = \{O, C, D\}) = \frac{e^{\mu V_{On}}}{e^{\mu V_{On}} + e^{\mu V_{Cn}} + e^{\mu V_{Dn}}},$$

which cannot be larger than $P_n(O|C_n = \{O, C\})$ for any value of V_{Dn}. This implies that it would be impossible for the Logit model to replicate a decoy outcome[1].

The decoy effect has been detected in various areas, including stated purchases (Huber et al., 1982), real purchases (Doyle et al., 1999), choices made by birds and bees (Shafir et al., 2002), and even in choices made by amoebas (Latty and Beekman, 2011). However, to the best of our knowledge, this effect has not been detected so far in transportation. In spite of recent controversy about its robustness (Yang and Lynn, 2014; Frederick et al, 2014; Huber et al., 2014; Tsetsos et al, 2015; Trueblood et al, 2015), the great body of evidence suggesting the prevalence of the decoy effect in many areas motivates the quest undertaken in this article.

The article is structured as follows. After this introduction, in Section 2 we characterize the different types of decoy outcomes that have been described in the literature and the conceptual cognitive behaviors that have been proposed as possible sources for them. Then, in Section 3 we adapt those conceptual cognitive behaviors into practical context-RUM models and assess their feasibility for being the underlying Data Generation Process (DGP) that results in observing the different types of decoy effects. Section 4 conveys the development, collection and analysis of a Stated Preferece (SP) convenience survey of route choice specially crafted to account for different types of decoy effects. This case study is then used to detect the presence of the decoy effect in this context and to characterize the relative strength of its different types. Then, Section 5 assesses the estimation and cross-validation performance of the different context-RUM models proposed in Section 3. Section 6 studies the detection of the decoy effect in Revealed Preferece (RP) mode choice data and, finally, Section 7 summarizes the main contributions of the article and proposes future lines of research in this area.

2. Decoys types described in the literature and conceptual cognitive models proposed to account for them

For expositional purposes, consider the hypothetical problem of choosing mode between Train and Car in a given origin-destination pair and time period. Suppose that the alternatives are only described by their travel time and travel cost and that it takes 55 min to reach the destination by Train and 60 min by Car. In contrast, the fare of the Train is 10 $US, and the cost of the Car is 5 $US. These alternatives are illustrated in Fig. 1.

For increasing train's patronage, the authority changes the fare scheme as follows. Now, the (55, 10) Train option can only be bought by the holders of a free “loyalty card” (LC) and the price for other customers in the same train becomes 12 $US. This new fare scheme is illustrated in Fig. 2.

Since the LC is free, it will be used in practice by all interested customers. Thus, almost nobody would choose the (55, 12) alternative. If individuals are fully rational, the inclusion of this new unchosen alternative should have no effect on the choices, and the shares attained in the fare scenarios depicted in Figs. 1 and 2 should remain the same. Nevertheless, according to the decoy effect, the introduction of the unattractive alternative (55, 12) would result in an increase of the share of the original Train option (55,10), compared to the Car. In this example, (55, 12) is the Decoy (D) alternative that favors the (55,10), which is the Objective (O), in detriment of the Car, which is known as the Competitor (C).

The decoy types are defined by the location of D, with respect to O. The decoy effect and the behavioral mechanisms behind it were first studied by Huber et al. (1982), who termed it as the **Asymmetrically Dominated** (AD) decoy effect. In this case, the decoy (D) is worse in every attribute, if compared to the Objective (O) alternative, and superior in some,

[1] This results from Logit's IIA property, but an equivalent finding also holds for non-IIA models such as the Nested Logit. Note that it is being assumed that the scale μ does not change with the inclusion of the decoy. This issue is further discussed in Section 3.

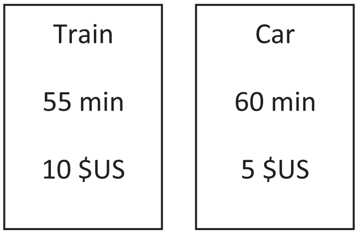

Fig. 1. Train versus car example with no decoy alternative.

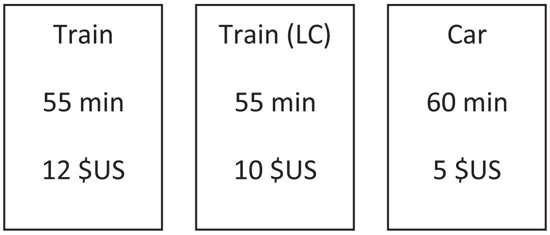

Fig. 2. Train versus car example including decoy alternative.

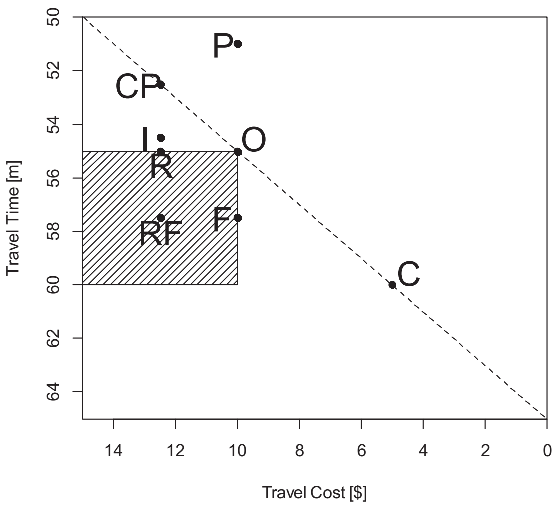

Fig. 3. Decoy types that have been described in the literature.

but inferior on other attributes, if compared with the Competitor (C) alternative. The intuition for the AD decoy is that the inclusion of a dominated alternative creates a simpler comparison between O (dominant) and the decoy D that makes O look better, not only compared with D, but also overall (Ariely 2008).

The potential locations of the AD decoys that have been reported in the literature are shown in the shaded area of Fig. 3. In it, each alternative is defined by a point in the attribute space, which in the Car/Train example corresponds to a combination of travel time and travel cost. Because people prefer to travel less time and for less money, the best alternatives will be those located to the right and up in Fig. 3. Besides, it is assumed that O and C are located over the trade-off line, which is depicted with a dashed line in Fig. 3 and, in this example, corresponds to a value of time of 1$US per minute. Under a formal micro-economic framework, instead of the trade-off line, one would have to assume that both O and C lie instead over the same indifference curve, which will then not necessarily be a straight line, but instead a convex function over the attributes' space. We will maintain the trade-of line concept throughout the article for illustration purposes and to be consistent with previous decoy literature.

Following Weddell and Pettibone (1996) AD decoys can be further classified in three types. When the AD decoy is (almost) equal to the objective in its best attribute, the decoy is said to be of the **Range** (R) type. This is the type of decoy used for the Car/Train example depicted in Figs. 1 and 2. When the Decoy is (almost) equal to the objective in its worst attribute, it is said to be of the **Frequency** (F) type. Finally, when the AD decoy differs in all attributes from the objective, it

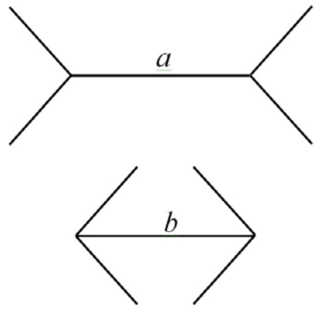

Fig. 4. Müller-Lyer's (1889) illusion. An example of the importance of context.

is said to be a **Range-Frequency** (RF) AD decoy. The locations of the three types of AD decoys are depicted with the letters R, F and RF, respectively, in Fig. 3.

When the decoy is located outside the AD area it is said to be a **Non-Dominated** (ND) decoy. When it is outside but adjacent to the AD area, it is known as an **Inferior** (I) decoy. This decoy is depicted with the letter I in Fig. 3. In this case the decoy is not worse in every attribute compared to O, but it is clearly below it in the trade-off line, and also close to O. This type o decoy has been described by Huber and Putto (1983) and Pettibone and Wedell(2000).

Two additional types of ND decoys have been described in the literature. The first is known as **Compromise** (CP), which is depicted over the trade-off line in Fig. 3. This decoy is in a location that places the objective (O) in the center. The CP decoy has been studied, among others, by Simonson (1989) and Pettibone and Wedell (2000). The intuition behind the CP decoy is that, because people tend to avoid extreme options, deploying a CP decoy will make the O alternative look less hazardous and, consequently, more appealing.

A final type of ND decoy that has been described in the literature is known as the **Phantom** (P) decoy. This decoy is above the trade-off line and close to the Objective (O). However when the individual wants to get it, s/he is told that the alternative is no longer available. The intuition for this decoy is that, after finding out that the desired alternative is no longer available, the closer option, in this case O, becomes more appealing to the choice-maker. This type of decoy has been described by Pratkanis and Farquhar (1992) and Pettibone and Weddel (2000) and is depicted in Fig. 3 with the letter P.

The R, F, RF and CP decoys are going to be studied later in the article. We will first analyze the suitability of various context-RUM choice models for reproducing them, and then we will measure their relative impact, and assess the estimation and forecasting performance of various context-RUM models that may account for them.

Capturing the decoy effect with a discrete choice model calls for a formulation that accounts for changes in the choice context. This feature is not embedded in the classical RUM models, but it seems to be an inherent characteristic of the way in which humans assess their environment.

To illustrate the importance of the choice-context consider, for example, the task of choosing the largest horizontal line in Fig. 4. At first sight, it may seem that line *a* is larger than line *b*, but that can be shown to be wrong by measuring them. Lines *a* and *b* have the same length, but our brain deceive us when the context is changed just by changing the angle of the diagonal lines that start from each end. This experiment is known as Müller-Lyer's (1889) illusion.

The utility maximization principle can be modified to account for the choice-context following various cognitive assumptions. Four of these modifications have been suggested as possible explanations for the decoy effect. In this section we will describe these proposed changes and in Section 3 we will first transform them into practical context-RUM choice models and then qualitatively assess their capacity for replicating the R, F, RF and CP decoy types. Later in Section 5 we will assess their performance in estimation and forecasting using real SP data.

The first cognitive assumption that results in a modification of the utility maximization principle to account for the choice context is known as **weight change (WC)**. Huber et al (1982) and Ariely and Wallsten (1995) suggested that the decoy may change choice-makers' perception of the relative importance of the attributes of the alternatives. In particular, Ariely and Wallsten (1995) developed an experiment in which interviewees were confronted with decoys in various choice contexts and were asked to declare the importance of the attributes of the alternatives for each choice. The authors found evidence suggesting that the dimension with the largest weight was the one in which D was dominated by O. This result was although later criticized by Pettibone and Wedell (2000) because the respondents were asked to judge the value of each attribute, instead of being asked to perform an actual choice. We provide in Section 5 further evidence against the WC hypothesis as being the DGP behind the decoy outcomes.

A second proposed modification of the utility maximization principle to account for context changes is known **emergent values (EV)**, which was suggested by Simonson (1989), Wedell (1991) and Pettibone and Wedell (2000). In this case, the cognitive assumption is that the inclusion of a new option impacts the utility of the other alternatives in the form of a term that is added to their utility. Simonson (1989) suggests that the emergent value could represent the easiness to justify the choice or the reduction in choosing anxiety that may result from the inclusion of the decoy. Wedell (1991) shows evidence suggesting that this additional term could be related to dominance relations. Among this class of models one could also classify Rooderkerk's et al. (2011) Unifying Model of Contextual Effects (UMCE). We provide in Section 5 further evidence in favor of the EV hypothesis as being the DGP behind the decoy outcomes.

A third modification of the utility maximization principle, so that it could account for the decoy effect, is known as **value shift** (Wedell, 1991; Pechtl, 2009). In this case, the cognitive assumption is that the decoy influences the way in which the individual perceives the attributes of the alternatives, analogously to what happens with the Müller-Lyer's (1889) illusion depicted in Fig. 4. Wedell (1991) tried to link Parduccy's (1974) Range-Frequency theory to the weight-change cognitive assumption for the decoy effect, but without finding conclusive evidence.

A final cognitive explanation for finding the decoy effect in practice comes from what is known as **prospect theory** (Kahneman and Tversky, 1979). This theory postulates that human decisions are based in two principles: reference-dependence and loss-aversion. The former postulates that utility is evaluated compared to a reference point and the latter states that people weight more heavily the losses than the wins. Simonson and Tversky (1992) justify the decoy using the loss aversion principle. Highhouse (1996) and Herne (1998) suggest a link between prospect theory in general and the decoy effect, while the first author adjusts its principles to explain phantom decoys in particular. In Section 3 we show that three prospect theory practical choice models could be behind all decoy types, but empirical evidence analyzed later in Section 5 suggests that their performance differ substantially in forecasting.

Among all the decoy types only CP, to the best of our knowledge, has implicitly or indirectly received prior attention in a transportation context. The work of Chorus and Bierlaire (2013) is the main antecedent to this respect. The authors used a route choice SP experiment specially designed to account for compromise alternatives, but were "not concerned with empirically exploring the potential presence of … a 'compromise effect' in [their] … data". However, their approach effectively permits "testing the empirical performance of models that allow for capturing the (potential) popularity of compromise alternatives", in a way that is similar to what we do in Section 5. Under such a context, the authors explored the performance of models that can be classified among the emergent value and the prospect theory classes. Another antecedent is the work of Leon and Hensher (2012) who explored the impact of embedding multiple heuristics to model the choice of route in an SP experiment that was however, not only, not explicitly designed to detect the decoy effect, but neither designed to represent compromise alternatives. The link of this work with the decoy phenomena is that some of the heuristics studied by the authors could be classified among the prospect theory class.

3. Diagrammatic analysis of the feasibility of context-RUM choice-models for replicating the decoy effect

In this section we analyze the suitability of various practical choice models as potential DGP than may result in observations of what has been described as the decoy effect. The study is done diagrammatically, rather than analytically, to ensure its viability and comparability for all of context-RUM models under analysis.

The diagrammatic analysis consisted in setting up a simulation in which each respective model (M) under study defines the DGP. Under this setting, we first simulated a case in which only two alternatives, the Objective (O) and the Competitor (C) were available. The alternatives were defined by two negatively valuated attributes, e.g. travel time and travel cost (*tt, tc*), and the model parameters were chosen such that alternatives O and C get a 50% choice probability for the given binary model M.

$$P_n^M(O|C_n = \{O, C\}) \equiv \frac{1}{2}$$

Afterwards, a Decoy alternative D_r, defined by a given combination r of travel time and cost (tt_r, tc_r), was added to the choice-set. Then, a new choice-probability for O was calculated, assuming that the model M still defines the DGP. The choice probabilities with and without the decoy were used to calculate the indicator

$$\theta_r^M = \frac{P_n^M(O|C_n = \{O, C, D_r\})}{P_n^M(O|C_n = \{O, C\})}$$

for each possible value of D_r mapped over (tt_r, tc_r). The simulated values of θ_r^M were finally plotted in a diagram similar to Fig. 3.

Using the resulting diagram, the model M was judged a feasible DGP for producing decoy outcomes if, for some combination of the model parameters, $\theta_r^M > 1$ in a region that is similar to the ones that been reported in previous literature (Fig. 3). In turn, if $\theta_r^M \leq 1$ for all possible values of r, or $\theta_r^M > 1$ but for a region not coherent with Fig. 3, the method M was judged an infeasible DGP for decoy outcomes.

We begin considering the Logit model. As it was discussed in Section 2, if the scale of the Logit model is not modified by the inclusion of the decoy, $\theta_r^{Logit} \leq 1$ for all possible locations r. Hence, a diagrammatic analysis for the Logit model would show that the area for which $\theta_r^{Logit} > 1$, is the empty set. This occurs because the Logit model complies not only with regularity assumption but also with the stronger property of Proportionality (see. e.g. Luce, 1977), also known as Independence of Irrelevant Alternatives (IIA) (see, e.g. Ben-Akiva and Lerman, 1985).

Relaxing the fix scale assumption for the Logit, one peculiar case when θ_r^{Logit} may be larger than 1 could occur when, e.g., $V_{On} << V_{Cn} = V_{Dn}$, and the introduction of the decoy somehow produces a tremendous increase in the variance of ε_{jn}making the scale μ to go close to zero. In such bizarre case $\theta_r^{Logit} > 1$ because $P_n(O|C_n = \{O, C\}) < 1/3 \approx P_n(O|C_n = \{O, C, D\})$. Note however that such a case may materialize in the whole attributes' space, an area that would have no link at all to the

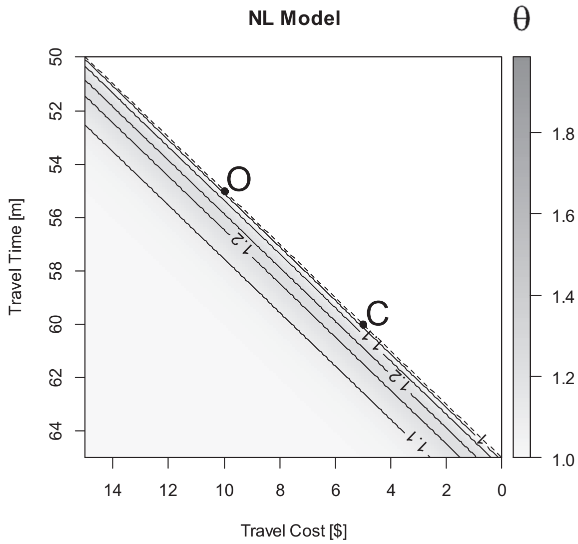

Fig. 5. Diagrammatic analysis of the decoy region for the NL that does not comply with RUM.

regions in which the decoy has been detected in practice, as shown in Fig. 3. Discarding such unrealistic case, we can sustain that the Logit cannot be claimed as the DGP behind the observation of the decoy effect.

For the study of a potential link between the decoy effect and the Nested Logit we need to relax one formal aspect of the problem. While the decoy effect is concerned about the observed attributes of the alternatives, the Nested Logit conveys instead assumptions about the unobserved error structure. Having said that, a natural nesting structure for this problem would involve having O and D in the same nest N because those alternatives are expected to be more similar among them, compared to C. The implicit assumption is that the error terms of the alternatives will be correlated if the observed parts are similar.

Assuming the Objective-Decoy nesting structure as the DGP, the probability of choosing O when the three alternatives are available would have the following form

$$P_n(O|C_n = \{O, C, D\}) = P_n(O|C_n = \{O, D\})P_n(\{O, D\}|C_n = \{O, C, D\})$$

$$P_n(O|C_n = \{O, C, D\}) = \frac{e^{\mu_N V_{On}}}{e^{\mu_N V_{On}} + e^{\mu_N V_{Dn}}} \frac{e^{\frac{\mu}{\mu_N}\ln\left(e^{\mu_N V_{On}} + e^{\mu_N V_{Dn}}\right)}}{e^{\mu V_C} + e^{\frac{\mu}{\mu_N}\ln\left(e^{\mu_N V_{On}} + e^{\mu_N V_{D^n}}\right)}},$$

where μ is the scale of the root and μ_N is the scale of the nest N.

The condition needed for the Nested Logit to be consistent with utility maximization is that the scales of the nests become larger as going down the tree. The intuition is that alternatives within a nest should be more similar among them than compared to alternatives outside the nest. A formal demonstration for this condition is provided by McFadden (1978).

In this example, consistency with utility maximization would imply that $\mu/\mu_N \leq 1$. Under such a setting, the Nested Logit would work in the opposite direction of the decoy effect because it was designed to account for what is called the similarity principle (see. e.g. Luce, 1977). Thus, a new alternative should reclaim a larger share from existing alternatives that are more similar to it, than from those that are more dissimilar to it. Hence, equally to the Logit, a diagrammatic analysis for the Nested Logit model that satisfies the utility maximization principle (i.e. $\mu/\mu_N \leq 1$) would show that the area of the attribute space for which $\theta_r^{NL} > 1$ would be the empty set.

Instead, when $\mu/\mu_N > 1$ there will be an area in which the decoy may increase the choice probability of O in the Nested Logit model. This decoy region is shown in the diagram depicted in Fig. 5. This diagram was built considering two attributes (tt_r, tc_r) with domain from 50 to 65, and from 0 to 15, respectively, both valued negatively with coefficients $\beta_{tt} = \beta_{tt} = -1$. The ratio of the scales was set to$\mu/\mu_N = 5$. The Objective (O) and the competitor (C) were located over the trade-off line at points(55, 10) and (60, 5) respectively. The diagram depicts in grey all the locations of the decoy that result in $\theta_r^{NL} > 1$. The strength of the decoy is depicted with the color scale shown at the right of Fig. 5.

Fig. 5 shows that for the non-RUM NL model ($\mu/\mu_N > 1$), the ratioθ_r^{NL} may become larger than 1 just below the trade-off line. Then, it quickly grows up to about 1.2 to finally fall slowly to one. The degree in which $\theta_r^{NL} > 1$ will depend on the parameters ($\beta_{tc}, \beta_{tt}, \frac{\mu}{\mu_N}$)used, but the shape described will be qualitatively the same.

It should be evident that the decoy region implied by Fig. 5 is completely unrelated to the decoy effects that have been detected empirically in previous literature, which are summarized in Fig. 3. Hence, it can be asserted that the non-RUM NL ($\mu/\mu_N > 1$) model seems not to be a congruent representation of the DGP that is behind of what has been described as the decoy effect, in the same sense that the scale shift bizarre condition was not valid for the Logit.

The analysis of other Multivariate Extreme Value (MEV) models, such as the Cross-Nested Logit, is difficult with the diagrammatic approach. However, it is expected that, just like the Nested Logit, the fulfillment of McFaddenś (1978) utility maximization conditions would prevent the occurrence of the decoy effect because those models are intended to replicate the similarity effect, which is opposite to the decoy. It cannot be discarded however that some non-RUM MEV models could be able to replicate the decoy effect to some extent. This analysis is left for further research.

We now analyze four practical implementations of the context-RUM models described in Section 2. Consider first the weigh-change (WC) model. As a direct extension of the Logit model, a context-RUM version of the weigh-change model would imply that the coefficients β_k in $V_{jn} = \sum_k \beta_k x_{kjn}$somehow change because of the inclusion of the decoy. Since the change in βcould take any desired form, there is no need to develop a diagrammatic analysis for this model because it will be able to fit any desired shift that might be caused by a decoy. For example, if the data implies that for AD decoys in Fig. 3 $\theta_r^{WC} = 1 + \delta$, with δ any positive scalar, an ad-hoc set of $\tilde{\beta}$ will always exist such that

$$\theta_r^{WC} = 1 + \delta = \frac{P_n^{WC}\left(O|C_n = \{O, C, D_r\}; \tilde{\beta}\right)}{P_n^{WC}(O|C_n = \{O, C\}; \beta)}.$$

In practice, β and $\tilde{\beta}$ could be estimated from the data, provided the existence of different scenarios in which alternatives are once favored and then not favored by the decoy, and that those relations can be clearly identified by the researcher.

However, the fact that the WC context-RUM model has the flexibility to accommodate any desired decoy outcome does not mean that it is the best model to account for it. From a modeling perspective, WC could be seen as an ad-hoc specification that (i) may over-fit the data, (ii) requires preconceived identification of the decoy alternatives and (iii) may not be used to describe the true underlying behavior that causes the decoy. In a real application it would be unclear how to interpret the differences between $\tilde{\beta}$ and β, and how they would behave outside the estimation sample. In Section 5 we will assess the WC context-RUM model in estimation and cross-validation, shedding some light into these issues.

Consider now the case of the emergent value (EV) model. In a context-RUM context, EV would imply the addition of a term J_{jn} to the systematic utility of each alternative, such that it becomes $V_{jn} = \sum_k \beta_k x_{kjn} + J_{jn}$. The term J_{jn} would have to depend on the relative dominance or the position of D compared O and C.

J_{jn} could be defined in countless ways. The most straightforward form would be to consider that $J_{jn} = \beta_D 1_{Dj}$, where β_D is a coefficient to be estimated and 1_{Dj} is a dummy variable that takes value 1 if j is being favored by the decoy, and zero otherwise. Another possibility, applicable however only to capture the compromise effect, is the one proposed by Chorus and Bierlaire (2013), who considered an approach in which 1_{Dj} is replaced by the number of variables in j that are not extreme in the given choice context[2]. A third example is the work of Rooderkerk’s et al. (2011), who proposed a model that considered an ad-hoc combination of what they defined as the positioning vector and the preference vector. For the analysis in this article we will use the most straightforward dummy variable approach for modeling the emergent value principle, since it calls for the lowest level of cognizance from the modeler side and facilitates a full comparative analysis with other methods.

Equivalent to what occurs with the WC model, the dummy EV model would be able to fit any desired decoy effect. Indeed, if we call β_D the coefficient of the EV dummy, and the data implies, for example, that for AD decoys $\theta_r^{EV} = 1 + \delta$, with δ any positive scalar, an ad-hoc set β_D will always exist such that

$$\theta_r^{EV} = 1 + \delta = \frac{P_n^{EV}(O|C_n = \{O, C, D_r\}; \beta_D)}{P_n^{EV}(O|C_n = \{O, C\})}.$$

As with the WC, this flexibility is not necessarily an advantage from a modeling perspective. In Section 5 we will assess the EV context-RUM model in estimation and cross-validation, shedding some light into this issue.

The case of the value-shift (VS) model is similar to that of the WC and the EV models, with the additional detriment that it is unclear how to transform, non-arbitrarily, this conceptual cognitive model into a practical context-RUM model. In a context-RUM setting, the value shift principle would imply that the x_{kjn} in $V_{jn} = \sum_k \beta_k x_{kjn}$ changes because of the inclusion of the decoy. However, even though the cognitive behavioral assumption behind the VS model is clearly illustrated by the Müller-Lyer’s (1889) illusion (Fig. 2), it is unclear how to devise experiments to detect this effect and to differentiate it from, for example, a WC effect. Consequently, we will not analyze VS later in Section 5.

The final context-RUM model aimed to account for the decoy effect is prospect theory. To account for this cognitive principle in a context-RUM model, we build on the work developed by Chorus et al. (2008) and Chorus (2010).

Based on the work of Quiggin (1994), Chorus et al. (2008) suggested a practical prospect-theory choice model termed Random Regret Minimization (RRM) model. In this early RRM model (which we will call e-RRM), the paired regret R_{ijn} of an alternative i with respect to alternative j, is calculated as a sum of the losses for each attribute k, weighed by its respective β_k.

$$R_{ijn} = \sum_k \max\left[0, \beta_k\left(x_{kjn} - x_{kin}\right)\right]. \tag{1}$$

[2] Chorus and Bierlaire’s (2013) model fully coincides with the simple dummy variable approach for the SP experiment considered in Section 5 since it consists of compromise alternatives that are non-extreme in all attributes for all cases.

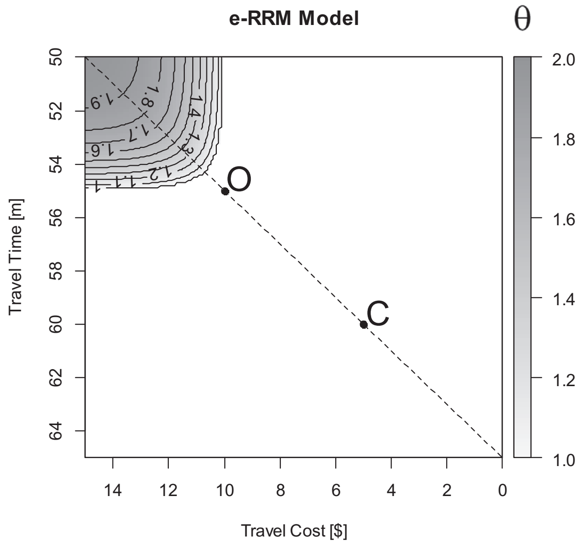

Fig. 6. Diagrammatic analysis of the decoy region for the e-RRM.

Then, the overall regret of an alternative A is calculated as the maximum paired regret across all available alternatives.

$$R_{in} = \max_p \left\{ R_{ipn}; \forall p \in C_n \right\}$$

Chorus's et al. (2008) e-RRM model complies with the reference-dependence principle of prospect-theory because in it the incumbent alternative gets compared with the best alternative available in the choice-set. It also fulfills the loss aversion principle because only the losses are accounted for in the calculation of the relative regret.

Before developing a diagrammatic analysis, it can be shown that Chorus's et al. (2008) e-RRM model could be behind the DGP that leads to the Compromise (CP) decoy. This occurs because, for example, a CP Decoy as the one shown in Fig. 3 would imply a larger regret in travel time for C, relative to O, favoring then the later. For the same reason, this model could be as well behind the Phantom (P) decoys, assuming that the regrets attached to the O and C alternatives remain attached to them after the non-availability of P is revealed to the choice-maker.

However, Chorus et al. (2008) e-RRM model cannot be behind the DGP that leads to Asymmetrically Dominated (AD) decoys. If an AD Decoy is incorporated to the choice-set, it would have no impact on the regret of the dominant alternative (O), because only the losses are accounted, and there would be no losses for O in such a case. Something similar occurs with the competitor (C), but for a different reason. Although C is inferior to D in some attributes, the largest paired regret will still be the one compared to O, leaving the overall regret of C unchanged as well. This is the result of the imposition by Chorus et al. (2008) of a stronger version of Quigginś (1994) assumption of Irrelevance of Statewise Dominated Alternatives (ISDA). Quigginś (1994) ISDA states that the addition of overall dominated alternatives (inferior to every other option) should have no impact in the choices. In Chorusś et al. (2008) e-RRM this condition holds not only for overall dominated alternatives, but also for asymmetrically dominated alternatives.

The diagrammatic analysis of the decoy region for the e-RRM model is depicted in Fig. 6, which was constructed using the same parameters, besides the scale, used to build Fig. 5. It can be noted that the decoy effect occurs only when the decoy is close the CP type. The fact that the decoy manifests above the trade-off line could be explained if, instead of the trade-off line, the indifference curve has a more pronounced convex shape.

Aimed at addressing numerical problems resulting from the calculation of the maximums within the likelihood function, Chorus (2010) proposed a modification of the e-RRM in two aspects. The first modification consists in that the relative regret is replaced by the following expression

$$R_{ijn} = \sum_k \ln\left(1 + \exp\left(\beta_k\left[x_{kjn} - x_{kin}\right]\right)\right),$$

which can either be seen as an approximation of Eq. (1), or as the result of maximizing an expected regret function with Extreme Value I errors[3]. Second, instead of assuming that the overall regret depended only on the alternative with the

[3] Under the latter assumption, a scale term would appear resulting in what was defined as the μRRM model by van Cranenburgh et al (2015). In the more general case, the relative regret would have the following form $R_{ijn} = \sum_k \frac{1}{\mu_k} \ln(1 + \exp(\mu_k \beta_k [x_{kjn} - x_{kin}]))$.

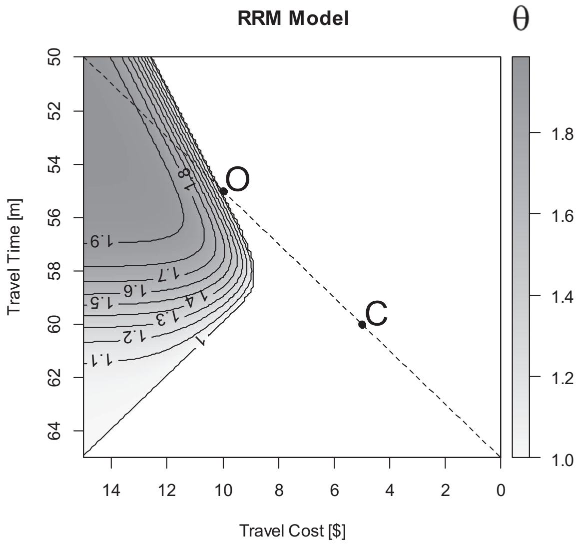

Fig. 7. Diagrammatic analysis of the decoy region for the RRM model.

largest regret, the author considered now for this the sum of the regrets over all alternatives

$$R_{in} = \sum_{p \in C_n} R_{ipn}. \tag{2}$$

This second modification makes the difference that allows RRM to be able to account not only for CP decoys (as it was implicitly acknowledged before by Chorus et al, 2008 and Chorus and Bierlaire, 2013), but also for AD decoys. This can be shown with the diagram depicted in Fig. 7, which was constructed using the same parameters, besides the scale, used to build Fig. 5.

Comparing Figs. 3 and 7 it can be noted that Chorus's (2010) RRM model may be the DGP behind all types of decoys that have been described in the literature. Besides, differently from the EV and WC models, this RRM model is more parsimonious, as it reveals the decoy outcomes from an overall cognitive assumption (prospect theory) with fewer parameters and no ad-hoc dummies, reducing by this the risk of over fitting the data.

The RRM model is not only able to replicate the decoy outcomes but also suggests different strengths of the decoy types. Fig. 7 shows that, if RRM is behind the decoy effect, the strength of the decoys types should be, from lower to higher, F, RF, R and CP. If such relative strength is observed in reality, it would be additional evidence suggesting that the RRM, or something similar, is indeed behind the DGP resulting in decoy outcomes.

The relative impact of the decoy types reported in the literature seems to agree, to some extent, with the shape suggested by Fig. 7. First, studying AD decoys, Huber et al (1982) found that the lowest impact was achieved by F, followed RF and R. Likewise, Weddell (1991) concluded that AD decoys that were too much worse than O, would have no impact on the share of O. On the contrary, Pettibone and Wedell (2000) states, but without providing conclusive empirical evidence, that AD decoys "tend to be" larger than ND decoys, what seems to contradict Fig. 7. In Section 5 we will explore this issue using a common SP database and formal statistical tests.

Similar to what we considered for the Logit, the analysis depicted in Fig. 7 does not account for the potential change of scale that may occur when individuals face choice-sets of different sizes (2, without D, and 3 with D). van Cranenburgh et al. (2016) suggest that the scale of the regret in the RRM model decreases inversely proportional to the choice-set size, in a degree that has to be estimated empirically. This would imply for Fig. 7 that the impact of the decoy will be diminished, but the shape of it would not be affected and, therefore, the qualitative result inferred from it would be unaffected. Consequently, we will ignore such effect in the analysis.

With the purpose of obtaining a decoy region that is closer to the area that has been reported in the literature, we propose a variation of the RRM model that we denominate Regret By Aspect (RBA). The base of this behavioral model consists in assuming that individuals care only about the attribute (aspect) of the alternative that has the largest expected accumulated regret for the available choice-set.

Assuming that the relative expected regret of alternative i with respect to alternative j and for attribute k has the following form

$$R_{ijkn} = \frac{1}{\mu_{ijk}} \ln\left(1 + \exp\left(\mu_{ijk}\beta_k\left[x_{kjn} - x_{kin}\right]\right)\right),$$

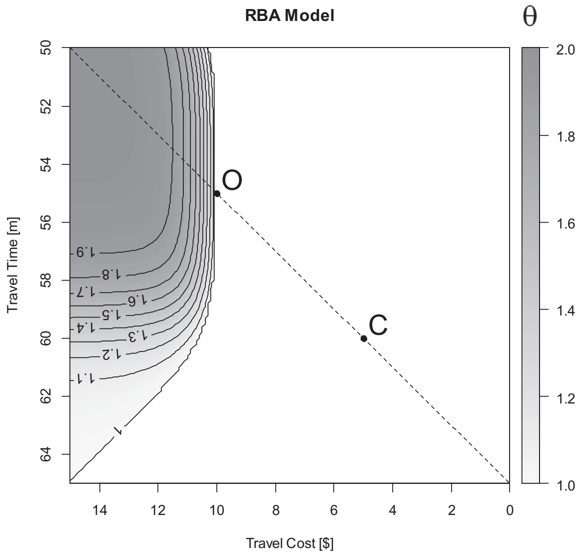

Fig. 8. Diagrammatic analysis of the decoy region for the RBA model.

the accumulated regret of alternative i and attribute k will be

$$R_{ikn} = \sum_j \frac{1}{\mu_{ijk}} \ln\left(1 + \exp\left(\mu_{ijk}\beta_k\left[x_{kjn} - x_{kin}\right]\right)\right).$$

Then, if individuals only care about the worst performing attribute, and Extreme Value 1 errors are assumed, the regret for alternative i becomes

$$R_{in} = \frac{1}{\mu_i} \ln\left(\sum_k \exp\left(\sum_j \frac{\mu_i}{\mu_{jk}} \ln\left(1 + \exp\left(\mu_{jk}\beta_k\left[x_{kjn} - x_{kin}\right]\right)\right)\right)\right). \tag{3}$$

The diagram of the decoy region for the proposed RBA model is depicted in Fig. 8, which was constructed considering the same parameters used to build Fig. 5, assuming all the scale parameters μ to be equal to one.

It can be noted that, to some extent, the decoy region of RBA can be described as the sum of the regions that are obtained by the e-RRM model and the AD decoys. Alternatively, it can be described as the region obtained with RRM, but turned to the right. Similar to the RRM, it can be affirmed that the RBA is able to replicate all the decoy types that have been described in the literature, but with a much steeper boundary for the attribute in which O is outperformed by C, which in this illustration corresponds to travel cost. In Section 5 we will assess the empirical suitability of both RRM and RBA for replicating decoy outcomes in an application with real data.

Finally, we consider the Contextual Concavity Model (CCM), which was proposed by Kivetz et al. (2004). This model can be classified among the prospect theory class and has been studied, among others, by Leong and Hensher (2012) and Chorus and Bierlaire (2013). CCM considers that the utility of an alternative i is constructed by comparing the level of each of its attributes k to the respective worst level $\breve{x}_{kn}$ in the choice context, as shown in Eq. (4). This comparison is assumed to be non-linear with a concavity defined by an estimated parameter ϕ (hence the model's name) and it calls for a modeler a-priori assumption for sign of β_k.

$$V_{jn} = \sum_k \left(\beta_k x_{kjn} - \beta_k \breve{x}_{kn}\right)^{\phi_k}. \tag{4}$$

When considering a convex relation ($\phi > 1$), the diagram of the decoy region for the CCM model resulted in a decoy that favored the competitor instead of the objective, in direct contradiction with what was reported in empirical literature. The results were much better when considering a concave relation ($\phi < 1$). The diagram of the decoy region for the CCM model is depicted in Fig. 9, which was constructed using $\phi = 0.5$ for both time and cost, and the same β_k used for the previous models.

It can be noted that this version of the CCM recovered, to some extent, not only the compromise effect, but also AD decoys, a fact that was already recognized as a possibility by Kivetz et al. (2004). However, the shape of the decoy types implied by the CCM model seems not fully coherent with what was depicted in Fig. 3. The most remarkable issue is that the strength of the decoy effect seems to grow, instead of decrease, when moving further away from the trade-off line, contradicting previous empirical evidence reported in the literature. The suitability of CCM, and other models, will be further analyzed in Section 5 using SP data.

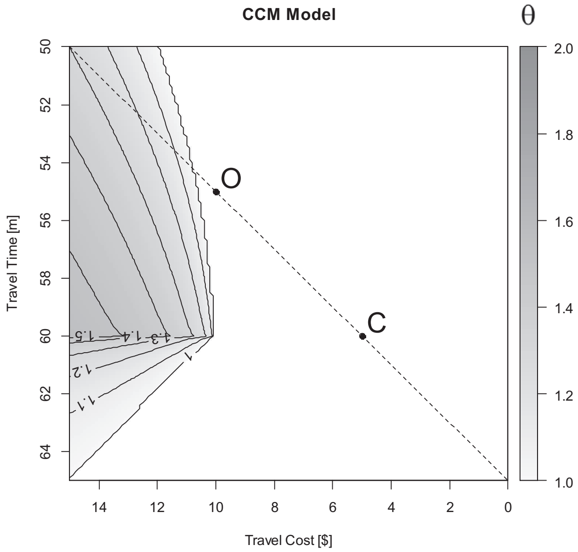

Fig. 9. Diagrammatic analysis of the decoy region for the CCM model.

"Suppose that you go shopping by car on a Saturday at 4 PM and have the following 3 routes available"

1. Ch$1670 29 min
2. Ch$1750 25 min
3. Ch$1590 29 min

Which route would you choose?____

Fig. 10. Example of choice task 3 of 10.

4. Detection and characterization of the decoy effect in a route choice SP choice case study

In this section we begin describing the Stated Preference (SP) survey on route choice that we used as a case study for the analysis of the decoy effect. We use this data first to detect the decoy effect and then to estimate the relative strength of its different types, contrasting the findings with what was reported in previous literature and with what would be attained if different context-RUM choice models would have been behind the DGP.

4.1. A tailored route choice SP choice case study to detect the decoy effect

The SP survey was applied via internet to a convenience valid sample of 134 individuals, mainly (92%) undergraduate students from the same university in Santiago de Chile. Further details of the survey can be found in Fukushi (2015). The database is available to the interested reader, upon request.

Besides some basic socioeconomic questions, the core of the survey consisted in 10 trinomial route choices to a hypothetical shopping trip by car performed on a Saturday at 4 PM. The alternatives in the choice-task were only described by their travel cost (in Chilean pesos) and travel time (in minutes), as it is illustrated in Fig. 10.

The first two of 10 choice tasks depicted a trade-off between three pre-defined non-dominated alternatives. Then, the responses to these choice tasks were used to calculate a rough estimator of the value of time VT_n of each individual, their trade-off between travel time and travel cost. The reader is referred to Fukushi (2015) for further details on how this rough estimator was calculated.

VT_n was then used to tailor the next 8 choice tasks presented to each individual. For each case, two of the three alternatives were put over the trade-off line implied by the respective VT_n. Under this setting, one of the alternatives was cheap but slow (A) and the other was fast but expensive (B). The third alternative of each choice task was defined to

Table 1
Contingency table for various decoy types applied to the SP route choice survey.

Decoy		Frequency (F)		Range-frequency (RF)		Range (R)		Compromise (CP)		Total	
		B		B		B		B		B	
	Choice	A	B	A	B	A	B	A	B	A	B
A	A	51	29	41	27	50	47	46	18	188	121
	B	12	38	10	46	16	21	2	1	40	106
Total		130		124		134		67		455	

A: cheapest; B: fastest. Wide rows and columns depict the alternative favored by the decoy, for each decoy type. Narrow rows and columns depict the choice. Each cell shows the number of individuals, among the valid sample, making the respective choice. Choices of the Decoy are excluded

emulate four decoy types: F, RF, R and CP. Each decoy setting was put once favoring A and once favoring B, to complete the eight choice tasks of this part of the experiment.

The size of the attributes, the order of the alternatives, and the order of the type of decoy represented, were shuffled across choice-tasks to reduce the chances that the individual would notice the purpose of the experiment and change his/her choice behavior. In particular, the order presented to all individuals was: Range favoring A (RA), Range Frequency favoring B (RFB), Frequency favoring A (FA), Compromise favoring B (CB), Range Frequency favoring A (RFA), Range favoring B (RB), Compromise favoring A (CA) and Frequency favoring B (FB). Thus, the example in Fig. 10 corresponds to choice task 3, an F decoy (route 1) favoring alternative A (route 3, the cheapest). Choice task 10 would be its complement, an F decoy favoring alternative B (the fastest).

We proceed now to use the data from this case study to analyze the presence of the decoy effect. Fukushi (2015) compared different statistical tests to detect the decoy effect, concluding that McNemmar's (1947) test for marginal homogeneity had the largest power among the set analyzed. Therefore, we will use that test to detect the presence of the decoy effect in the route choice SP data. The statistic of McNemar's test is constructed from a contingency table that records the number of individuals that reacted to the experimental stimulus, and those that were indifferent to it.

The contingency table for the SP route experiment is shown in Table 1. An observation in this experiment corresponds to the response of the individual to the decoy stimulus. The rows show the number of individuals that made each choice when the decoy favored alternative A (the cheapest) and the columns show the respective choice when the decoy favored alternative B (the fastest) for each decoy type (F, RF, R and CP). The individuals who chose the decoy alternative are excluded from the analysis reported in Table 1.

Consider for example the Range (R) decoy, summarized at the middle of Table 1. In this case nobody chose the decoy D, so all 134 individuals are included in the analysis. The N_{AA}^{R} =50 in the upper left implies that 50 of 134 individuals chose alternative A, both when the R decoy favored A and when it favored B. Equivalently, N_{BB}^{R} =21 people chose B, independent of the position of the R decoy. The choice among A and B for the other 63 individuals was affected to some extent by the depiction of the R decoy. N_{AB}^{R} =47 individuals followed the alternative favored by the R decoy, and $N_{BA}^{R} = 16$ reacted contrary to it.

The null hypothesis of McNemar's test is that there is no decoy effect, case in which $H_0 : N_{BA}^{dt} = N_{AB}^{dt}$, where dt corresponds to the decoy type under analysis. The alternate hypothesis is then that $H_1 : N_{BA}^{dt} \neq N_{AB}^{dt}$. A one sided test is also possible. The statistic of McNemarś test of partial homogeneity has the following form

$$M^{dt} = \frac{\left(N_{AB}^{dt} - N_{BA}^{dt}\right)^2}{N_{AB}^{dt} + N_{BA}^{dt}},$$

and follows a χ^2distribution with one degree of freedom.

M^{dt} statistic for the different decoy types (F, RF, R, CP) took, respectively, values 7.05, 7.08, 15.3, 12.8, all of which imply a p-value below 1%. Consequently, we reject the null hypothesis that there is no decoy effect is this SP experiment of route choice, for all the decoy types investigated.

McNemar's test assumes independence between observations. This assumption holds for the tests applied to each decoy type but it would fail if we would like to test the hypothesis for all the observations together, using the information under the column "Total" in Table 1. The test in this case would need a multilevel (or "cluster") correction that accounts for the correlation among choices made by the same individual across the four decoy types. For this, we consider Durkalskiś et al. (2003) test, which assumes constant within-cluster correlation. The statistic of Durkalski test is

$$D^{Total} = \frac{\left(\sum_{n=1}^{N} \frac{N_{ABn}^{Total} - N_{BAn}^{Total}}{4}\right)^2}{\sum_{n=1}^{N} \left(\frac{N_{ABn}^{Total} - N_{BAn}^{Total}}{4}\right)^2},$$

which follows a χ^2distribution with one degree of freedom and where the index n corresponds to each individual in the sample. The value of this statistic is 31.4, which implies a p-value below 1%, suggesting again the rejection of the null hypothesis that there is no decoy effect.

Table 2
Analysis of strength by decoy type.

	Include all valid observations				Eliminate IDs choosing AD decoy				Eliminate IDs choosing any decoy			
	θ^{dt}	s.e.	N	p-value*	θ^{dt}	s.e.	N	p-value*	θ^{dt}	s.e.	N	p-value*
Frequency (F)	1.28	0.135	260	< 1%	1.31	0.154	242	< 1%	1.16	0.102	114	< 1%
Range-frequency (RF)	1.24	0.144	248	< 1%	1.25	0.150	242	< 1%	1.32	0.213	114	< 1%
Range (R)	1.22	0.0672	268	< 1%	1.22	0.0572	242	< 1%	1.80	0.491	114	< 1%
Compromise (CP)	2.16	1.12	134		2.28	1.25	114		2.28	1.25	114	

* p-value of non-paired t-test between θ^{dt} for the incumbent decoy type and the next decoy type in the table.

4.1. Analysis of the strength of different decoy types

Most of the previous literature on the decoy effect focuses only on its detection in different contexts, which is already a controversial issue (see e. g. Huber et al., 1982; Doyle et al., 1999; Shafir et al., 2002; Latty and Beekman, 2011; Yang and Lynn, 2014; Frederick et al, 2014; Huber et al., 2014; Tsetsos et al, 2015; Trueblood et al., 2015). The focus of the present article is instead on modeling and characterizing it. Thus, the next step of the analysis corresponds to the characterization of the strength of the different types of decoys. Although this is only a reduced case study, if the relative impact of the different decoy types is in anyway close to the shape depicted in Figs. 6–9, that would suggest that, to some extent, the DGP behind the finding of decoy effects may be somehow related with the respective parsimonious models.

The characterization of the strength of the different types of decoys is performed by sampling enumeration using an auxiliary binary emergent value (EV) model estimated for each decoy type as the base for forecasting. The systematic utilities of the auxiliary EV model for each decoy type (dt) has the following form,

$$\begin{aligned} V_{An}^{dt} &= \beta_A^{dt} + \beta_{tt}^{dt} tt_{An}^{dt} + \beta_{tc}^{dt} tc_{An}^{dt} + \beta_d^{dt} d_{An}^{dt} \\ V_{Bn}^{dt} &= \beta_{tt}^{dt} tt_{Bn}^{dt} + \beta_{tc}^{dt} tc_{Bn}^{dt} + \beta_d^{dt} d_{Bn}^{dt}, \end{aligned} \tag{5}$$

where β^{dt} is a vector of parameters for each decoy type and d_{in}^{dt} is a dummy that takes value one if i is being favored by the decoy. It should be recalled that either A or B is always favored by the decoy.

For being able to estimate this auxiliary binary model, all the choices for the decoy alternative were excluded from the database. Besides, for avoiding artificial unbalance between alternatives, the respective pair of a decoy choice for the given individual and decoy type was also excluded. Additionally, no correction was considered to account for the correlation among responses from the same person, since this panel data effect has not impact in the consistency of the estimators of the model parameters (Daly and Hess, 2011).

After estimating the auxiliary EV model for each decoy type, the impact of the decoy for each observation is calculated as the ratio θ_n^{dt} of the simulated choice probability that each individual would choose the objective (either A or B, as it corresponds) when the decoy favors it (imposing $d_{On}^{dt} \equiv 1$) and when it does not (imposing $d_{On}^{dt} \equiv 0$).

$$\theta_n^{dt} = \frac{P_n^{dt}\left(O|d_{On}^{dt} \equiv 1\right)}{P_n^{dt}\left(O|d_{On}^{dt} \equiv 0\right)}$$

The estimated impact of the different decoy types θ^{dt} is then calculated as the average of θ_n^{dt} across the sample of individuals. The results are summarized in Table 2, where it is presented the average, the standard error, the number of observations and a non-paired t-test between the estimator of θ^{dt} for the incumbent decoy type, and that of the following decoy type in the table.

Three cases are analyzed in Table 2. At the left are reported the results obtained when all valid observations are considered. At the center are shown the results attained when eliminating all the observations from individuals that chose an Asymmetrically Dominated (AD) decoy at least once, since that may be an indication of poor attention to the choice-task. Finally, at the right are depicted the results attained when eliminating all the observations from individuals that chose any decoy, including Compromise (CP) decoy. Since CP decoys are not dominated, their choice does not imply poor attention to the choice-task and, therefore, discarding individuals that choose them can be seen as a way of selecting only those individuals that behaved in a way that was fully coherent with the compromise principle.

The first result that can be highlighted, valid for all the subsamples analyzed, is that the relative strength of the ND decoys (CP) is substantially larger than for the AD ones (F, RF and R). This is in line with what would result if RRM or RBA would have been behind the DGP (Figs. 7 and 8, respectively), but not if CCM would be playing that role (Fig. 9). This result contradicts the statement of Pettibone and Wedell (2000).

The relative strength of the decoys within the AD type depends on the subsample analyzed. Consider first the case where all observations are included in the analysis, at the left of Table 2. In this case the relative strength within the AD decoys is, from lowest to highest, R, RF and then F. This is precisely opposite to the order suggested by Figs. 7 and 8 for RRM and RBA, and has no relation with what is suggested by Fig. 9 for CCM. Although the differences are small, they are significant, according to the non-paired t-test that shows a p-value below 1% for all cases. The results are qualitatively the same when the individuals choosing the AD decoy were eliminated, as shown at the middle of Table 2. This suggests that the potential source for the discrepancy of the impact among AD decoy types compared to the one suggested by RRM, RBA and CCM (Figs. 7–9) is not related with a potential low attention to the choice-task.

The results among the AD decoys change when all the individuals choosing any type of decoy are eliminated. In this case, the order of the impact of the decoy types is F, RF and R, which is fully coherent with the order suggested by Figs. 7 and 8 for RRM and RBA. Furthermore, even the size of the impact is remarkably close to the ones suggested by Figs. 7 and 8. Regarding CCM, the results again suggest that this model seems to have no relation with the DGP behind the decoy effect.

We interpret these results as a suggestion that the parsimonious RRM and RBA models may indeed be, to some extent, behind the DGP that results in what has been described as the decoy effect, especially when comparing the impact between dominated and non-dominated decoys. Within dominated decoys, the order of magnitude of the decoys is fairly similar to what is suggested by RRM or RBA, but the relative strength fully seems to coincide only for a class of individuals that behave in a way that is coherent with the compromise principle. For other classes, RRM and RBA, although being able to replicate a decoy outcome, they may not be the full DGP at play. Regarding CCM, all the results analyzed strongly suggest that this model seems to have no relation with the DGP behind the decoy effect.

Besides the analysis of the relative strength of the decoy types, note that the standard error of θ^{dt}can be interpreted as measure of the robustness of the impact of the different decoy types. Under this point of view, the results in Table 2 suggest then that CP decoys are always the more volatile ones and R decoys are markedly more robust when considering all the observations and when the individuals choosing the AD decoys are fully excluded.

The analysis of the existence, strength and robustness of the decoy types inferred from Tables 1 and 2 may be affected, to some extent, by a plausible hypothesis that cannot be fully examined with the available data. Given the repeated nature of the choice-tasks, it is conceivable that the knowledge gained from previous choices may decrease the strength of the decoy effect in posterior choices. However, such an hypothesis cannot be investigated with this data because, despite the sequence of decoy types presented was shuffled, the same order (RA, RFB, FA, CB, RFA, RB, CA, FB) was used for all the individuals in the choice set, confounding the decoy type with the order of appearance in the experiment. Therefore, for example, if the mentioned hypothesis is true, the measurement of the relative strength of R decoys may be upwards biased, compared to CP decoys, because the formers were presented relatively earlier than the latter ones. To explore such an impact we would require a different type of experiment in which the order of the decoy types was fully randomized across individuals.

A final issue that may impact the analysis deployed in Tables 1 and 2 is selection bias because all the choices of the decoy alternative were dropped from the sample. This is likely not an issue for the case of AD experiments, because pruning those choices simply implies retaining observations from individuals that were paying full attention to the choice task. On the contrary, getting rid of the decoy choices for the CP experiment entails the implicit selection of individuals that tend to avoid extreme options, what may redound in a potential overestimation of the impact of the decoy effect. The analysis considered in Section 5, which does not require the elimination of the decoy choices, is one way to circumvent this potential limitation.

5. Assessment of models in terms of their estimation and forecasting performance

The final step of analysis corresponds to the assessment of the different models in terms of their capability for replicating the choice behavior observed in the SP survey of route choice. In this case the decoy alternative is included in the specification so that a trinomial choice model is considered. The analysis is performed first regarding estimation and then forecasting, using for the latter a 20% cross-validation sample by individual, with 100 repetitions. This sampling is performed by individual because they are the subjects of experimentation and because sampling by choices will often result in the partial pruning of complementary observations, potentially favoring spuriously one model over the other.

5.1. General setting

Six models are compared. The first model is a Logit with linear utility on travel time (with coefficient β_{tt}) and travel cost (with coefficient β_{tc}). This classical Logit model that ignores the context is used as a benchmark of the worst case in which the potential impact of the decoy is simply ignored. To control for lexicographic effects, all models consider two dummy variables, one for the alternative presented first β_F, and one for the alternative presented last β_L (see Fig. 9). These Alternative-Specific Constants (ASC) also allow replicating the sample share and, by being maintained in all the models, favor a fair comparison among them.

The second model estimated corresponds to an Emergent Value (EV) model of the same form of the one depicted in Eq. (5), but including the decoy as a third alternative and a respective set of lexicographic ASCs. The third model estimated is Weigh Change (WC), which is the same as the Logit but considering a different set of coefficients for travel time (β_{tt}^D) and travel cost (β_{tc}^D) for the alternative that was being favored by the decoy in each case. The fourth model corresponds to an RRM model as the one depicted in Eq (2), but also considering ASCs. The fifth model corresponds to the RBA model depicted in Eq (3) with ASCs and the final model corresponds to the CCM model in Eq. (4) with its respective set of ASCs to account for first and last lexicographic effects.

An additional correction was considered for the AD experiments. The decoy alternative in these experiments is dominated by the objective and, thus, if someone chooses the decoy under such circumstances it is an indication of poor

Table 3
Model estimation for asymmetrically dominated decoys.

	Logit		Emergent value (EV)		Weight change (WC)		RRM		RBA		CCM	
	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e
β_F	0.378	0.184	0.374	0.184	0.414	0.187	−0.360	0.187	−0.350	0.179	−0.101	0.270
β_L	0.180	0.157	0.184	0.157	0.247	0.173	−0.127	0.159	−0.072	0.150	−0.253	0.226
β_{tt}	−9.20	1.01	-9.71	1.02	−9.53	1.32	−7.15	0.725	−10.1	0.81	−11.8	2.24
β_{tc}	−2.92	0.279	−3.13	0.280	−3.10	0.348	−2.38	0.216	−3.63	0.259	−8.59	2.67
β_{tt}^D					−0.0136	0.205						
β_{tc}^D					−0.0684	0.176						
β_D			0.421	0.085								
θ_{order}	−1.65	0.171	−1.48	0.177	−1.50	0.185	1.65	0.170	1.78	0.163	−1.79	0.189
ϕ_{tt}											1.03	0.149
ϕ_{tc}											0.494	0.0736
N	804		804		804		804		804		804	
$\bar{\rho}^2$	0.344		0.360		0.355		0.348		0.350		0.356	

understanding or attention to the choice task. Under the hypothesis that understanding/attention grows with the repeated choices, the logarithm of the order of appearance of the choice task was added to the decoy alternative in all AD decoy experiments. A negative sign (positive for RRM and RBA) of the coefficient θ_{order} accompanying this variable would be a confirmation of this hypothesis.

As it was highlighted before, the consistency of the estimators is not affected by neglecting the panel data effect in model estimation induced by the repeated nature of the choices in the SP experiment. However, the standard errors of the estimators would be different from those obtained directly as the inverse of the Fisher information matrix. For this reason, in the results that we report we use the correction proposed by Daly and Hess (2011), which consists in considering robust or "bhhh" standard errors that are integrated at the level of the individual, rather than at the level of the observation. The standard errors obtained with this correction tended to be slightly larger than those obtained directly from the Fisher Information matrix, but the results inferred from them were qualitatively the same.

Given the differences in strength by decoy type detected in Section 4, the analysis in this section is separated between Asymmetrically Dominated (AD) decoys (F, RF and R) and Non-Dominated Decoys (CP). The study cannot be separated by the user classes considered in Section 4 because eliminating all choices of the decoy would make the model non identifiable. A different tool for class analysis could be based in the definition of latent classes, but the database at hand was too homogenous (almost only student of the same university) to apply such an approach.

5.2. Analysis of asymmetrically dominated (AD) decoys

The estimation results for the AD decoys are summarized in Table 3. In this case there are 804 observations available, which correspond to the 134 individuals times the 6 choice-tasks concerning the AD decoys.

Regarding model fit, Table 3 shows that the best model in terms of $\bar{\rho}^2$[4] is EV, closely followed by CCM and WC. In turn, both RRM and RBA are closer in fit to the Logit model, which is the benchmark for the worst case modeling scenario. All differences in $\bar{\rho}^2$ are statistically significant, according to a Horowitz test (Horowitz, 1983).

Regarding the estimators, first of all, the coefficients of travel time and travel cost are both negative, as they should be, for almost all models. The only remarkable case is WC, where the coefficients for time and cost of the decoy option are very small, both non-significant, with a p-value over 70%. This would mean that individuals do not care about travel time and travel cost when a decoy is presented, which is a highly questionable behavioral assumption.

Regarding the lexicographic effect, Table 3 shows a significant inclination toward choosing the first alternative presented (β_F) for all models, except for CCM. Furthermore, θ_{order} is statistically significant for all models (p-value below 1%) and with a sign that implies that understanding/attention to the choice task did grew with the repeated choices. Finally, the estimators of the exponents of the CCM imply a linear effect for time ($\phi_{tt} = 1$) and a concave effect for cost ($\phi_{tc} < 1$).

The specification of the models reported in Table 3 correspond to their best setting for the AD sub-sample. For the RRM, we explored the more general version μRRM that includes a scale factor, proposed by van Cranenburgh et al. (2015). For such a case, we faced estimation problems and, when estimation was feasible, the fit of μRRM was always equal or below the fit of the RRM model. In the case of the RBA model, we explored the estimation of different scales for time and cost, achieving a larger fit when imposing the constraint $\mu_{tt} = \mu_{tc} = 1$.

The analysis is complemented by Table 4, where cross validation is applied to a 20% outer sample with 100 repetitions. This means that a randomly selected 20% of the full sample of individuals was set apart and each model was estimated with the remaining 80%. Afterwards, using the parameters estimated with the 80%, the fit on the 20% outer-sample was

[4] Adjusted McFadden's $\bar{\rho}^2 = 1 - \frac{L(\hat{\beta})-K}{L(0)}$, where K is the number of coefficients

Table 4
Estimation and forecasting for asymmetrically dominated (AD) decoys.

	$\bar{\rho}^2$Estimation	$\bar{\rho}^2$Forecasting	Paired t-test w/r to Logit forecasting		Paired t-test w/r to EV forecasting	
			t-test	p-value	t-test	p-value
Logit	0.343	0.311				
EV	0.359	0.320	−7.89	< 1%		
WC	0.355	0.310	0.746	45.7%	24.0	< 1%
RRM	0.348	0.315	−8.43	< 1%	5.96	< 1%
RBA	0.350	0.318	−6.10	< 1%	3.73	< 1%
CCM	0.356	0.310	0.855	39.4%	9.45	< 1%

20% cross-validation sample. 100 repetitions.

Table 5
Model estimation for compromise decoy.

	Logit		Emergent value (EV)		Weight change (WC)		RRM		RBA		CCM	
	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e
β_F	−2.24	0.568	−0.833	0.602	−1.05	0.687	2.69	0.557	2.11	0.604	−1.01	NA
β_L	−3.61	0.440	−3.95	0.449	−1.74	0.783	3.89	0.435	3.60	0.401	−2.68	NA
β_{tt}	−12.7	2.21	−18.1	3.40	−21.24	3.94	−13.9	1.86	−18.8	2.38	−0.317	NA
β_{tc}	0.898	0.957	−2.54	1.09	−2.81	1.13	−0.0535	0.649	-4.02	1.30	−2.25E−07	NA
β_{tt}^D					−1.01	0.443						
β_{tc}^D					1.75	0.918						
β_D			2.64	0.319								
μ_{tt}									0.563	0.215		
ϕ_{tt}											0.775	NA
ϕ_{tc}											1.09	NA
N	268		268		268		268		268		268	
$\bar{\rho}^2$	0.401		0.512		0.519		0.448		0.479		0.240	

evaluated, using $\bar{\rho}^2$ as a measure. To avoid finite sample bias, this process was repeated 100 times to build a sampling distribution of the $\bar{\rho}^2$ measure. These 100 $\bar{\rho}^2$ values were then used to construct Table 4, where we report the average $\bar{\rho}^2$ across the repetitions, as an estimator of the population out-of-sample fit $\bar{\rho}^2$. We also use the whole sampling distribution of out-of sample $\bar{\rho}^2$ to perform some statistical tests.

The first column of Table 4 reports the average $\bar{\rho}^2$ across the 100 estimations on the 80% of the full sample, what is often denominated as the learning sample. The information that can be inferred from this column is qualitatively the same as the one that was inferred from Table 3.

The second column in Table 4 reports the average $\bar{\rho}^2$ in the 20% validation sample across the 100 repetitions. The next column reports a paired t-test comparing the forecasting fit of the models to that of the Logit and the final column does the same thing, but making the comparison with the EV model. Results show that, although WC and CCM have better estimation fit than the parsimonious RRM or RBA, only the forecasting fit of the latter ones is significantly superior to that of the Logit model. Besides, EV has better forecasting properties than the Logit and, according to the final column, it outperforms all other models.

Summarizing, the results provide strong evidence suggesting that the Weight Change (WC) and the Contextual Concavity (CCM) models are far from being behind the data generation process (DGP) behind the Asymmetrically Dominated (AD) decoys. Results are instead promising for the parsimonious Random Regret Minimization (RRM) and Regret by Aspects (RBA) models which show a forecasting performance significantly above the simple Logit model. However, the fact that the ad-hoc Emergent Value (EV) model shows better fit in estimation and in forecasting suggests that, although in Section 3 it was shown that the parsimonious RRM and RBA can replicate Asymmetrically Dominated (AD) decoys, their modeling capability seems to be limited. It is unlikely that neither RRM nor RBA are fully behind the data generation process (DGP) that results in AD decoy outcomes. The quest for a parsimonious choice model that can fully capture this phenomenon is still open.

5.3. Analysis of compromise (CP) decoys

The estimation results for the Compromise (CP) decoy are summarized in Table 5. In this case there are 268 observations available, which correspond to the 134 individuals times the 2 choice-tasks concerning the CP decoys.

Regarding the model fit based on $\bar{\rho}^2$, this table shows that the best model is WC, closely followed by EV. Both RRM and RBA are below WC and EV, but considerably above the Logit model. Again, all differences in $\bar{\rho}^2$are statistically significant, according to a Horowitz test (Horowitz, 1983).

Regarding the estimated coefficients, the coefficients of travel time and travel cost are negative for EV, RRM, RBA and CCM, but $\beta_{tc}$is positive (p-value 35%) for the Logit and $\beta_{tc}^D$is also positive (p-value 6%) for WC. This suggests that, although WC has a relatively high fit, it may not be capturing the true DGP, since that would imply that when individuals face the

Table 6
Estimation and forecasting for compromise decoys.

			Paired t-test w/r to Logit forecasting		Paired t-test w/r to EV forecasting	
	$\bar{\rho}^2$ Estimation	$\bar{\rho}^2$ Forecasting	t-test	p-value	t-test	p-value
Logit	0.399	0.336				
EV	0.510	0.410	−3.88	< 1%		
WC	0.516	0.398	−4.30	< 1%	1.68	9.58%
RRM	0.445	0.379	−20.5	< 1%	1.66	9.94%
RBA	0.476	0.395	−13.1	< 1%	0.807	42.2%
CCM	0.255	0.176	24.0	< 1%	11.6	< 1%

20% cross-validation sample. 100 repetitions. T-test regarding forecasting.

Decoy, somehow they would be willing to spend more cost travelling, a nonsense from a behavioral perspective. Finally, although the coefficient of travel cost is negative for the RRM, it is not significant (p-value 93%) suggesting a potential weakness of the model in this context.

Qualitatively different from the AD experiment, the lexicographic constants suggest in this case that it tends to be a negative and significant effect against the extreme options (first and last). This different outcome can have a behavioral or a statistical explanation. The behavioral explanation could be that, maybe because of stress, when individuals face a true trinomial choice on the CP experiment, they tend to favor the alternative presented in the middle, whereas when facing a seemingly binary choice on the AD experiment, they tend to favor the first alternative available. The statistical possible explanation is that, because the choice-sets were not randomized across individuals, the last alternative in this CP experiment always was the fastest (between the objective and the competitor) and the first was always the cheapest, confounding by this the lexicographic effects with a potential extremeness aversion.

Equivalent to the analysis reported before for the AD decoys, the specifications of the models reported in Table 5 correspond to their best setting for on the CP sub-sample. For the RRM, we again explored the μRRM version, discarding it for the same reasons as we did for the AD sub-sample. In the case of the RBA model, we explored the estimation of different scales for time and cost, achieving a larger fit when allowing $\mu_{tt} \neq 1$.

The analysis is completed by Table 6, where cross validation is applied to a 20% outer sample and 100 repetitions. The information inferred from the first column is qualitatively the same as the one that was inferred from Table 5. The results on the other columns suggest first that EV, WC, RRM and RBA are all superior to Logit since the null hypothesis for the t-test in the third column is rejected for all with p-value below 1%. CCM is worse than Logit. Additionally, this Table shows that, although WC has the best estimation fit, EV is the best in forecasting. Finally, while RBA, RRM and WC, pass a 5% test when being compared to EV in forecasting, only the first would pass a test of 10% significance.

Summarizing, Tables 5 and 6 provide strong evidence suggesting that the Contextual Concavity (CCM) model is far from being the data generation process (DGP) behind the Compromise (CP) decoys. Regarding the Weight Chance (WC) model, despite it shows the best estimation fit, and a forecasting fit above the simple Logit model and close to the Emergent Value (EV) model, its estimators reported in Table 5 have poor behavioral meaning, making it an unlikely DGP behind CP decoys. The best forecasting performance is attained by the ad-hoc EV model, but closely followed by the WC and the parsimonious Random Regret Minimization (RRM) and Regret by Aspects (RBA). Just as it occurred for the AD decoys, this suggest that RRM and RBA do play a role in the data generation process (DGP) that results in CP decoy outcomes but there is still room for improvement. The quest for a parsimonious choice model that can fully capture this phenomenon is still open.

The analysis reported in Tables 5 and 6 for CP decoys is equivalent to the one developed by Chorus and Bierlaire (2013, Tables 3 and 7), although they did not analyzed RBA neither WC. Contrary to our results, the authors report a better fit for CCM compared to EV and RRM. In turn, in line with what we report, they also mention “stability issues” for the CCM model. The different outcome is likely explained by the use of different data sources in each case study. Other differences are that Chorus and Bierlaire (2013) did not considered ASCs and that they used a single 33% cross-validation sample, instead of 100 repetitions, making their conclusions more prone to sampling bias.

6. Detecting and handling the decoy effect in revealed preference (RP) mode choice data with multiple variables and alternatives

We finally consider a Revealed Preference (RP) mode choice experiment with the twofold purpose of exploring the potential presence of the decoy effect in this context and to explore possible ways to handle this phenomenon in a real choice situation with multiple alternatives and variables.

The database used in this case was not specially designed for detecting neither for handling the decoy effect. It is a mode choice RP known as “Las Condes-Centro” that was collected by Ortuzar and Donoso (1983) and Ortúzar and Fernández (1985). The database consists of 697 individuals choosing commuting mode in Santiago de Chile, among a choice-set that includes public transportation (bus, metro, shared-taxi and combinations) and private modes (car driver and carpool). This data has been widely used previously by, among others, Munizaga and Daziano (2002) and Guevara (2016).

Table 7
Logit mode choice model from "Las Condes-Centro" Database with and without emergent value variables.

Coefficients	Logit model $\hat{\beta}$	s.e	Emergent value model $\hat{\beta}$	s.e.
1. Walking time	−0.161	0.0187	−0.145	0.0193
2. Waiting time	−0.236	0.105	−0.242	0.105
3. IV time	−0.0824	0.0174	−0.0682	0.0181
4. Cost/Income	−0.0245	0.00670	−0.0137	0.00778
6. Female	−0.295	0.217	−0.335	0.217
7. Licenses	2.36	0.408	2.35	0.411
8. EV_AD			0.444	0.144
9. EV_CP			0.108	0.131
Final log-likelihood	−949.135		−944.353	
Adjusted rho-square	0.224		0.226	
N	697		697	

Alternative specific constants by mode omitted from this summarized report; IV time: In-vehicle travel time; Source of data: Ortuzar and Fernandez (1985). Specification of Logit model replicates Munizaga and Daziano (2002). EV_AD: number of alternatives that are dominated in time and cost by each mode. EV_CP: 1 if mode is not extreme in time and cost.

The base model specification considered for this RP mode choice problem replicates the Logit model studied by Munizaga and Daziano (2002). The attributes of each mode are in-vehicle travel time (IV time), walking and waiting time, and cost divided by family net income. The model also includes for the car-driver mode the variable Licenses, which corresponds to the ratio between the number of cars and the number of licenses in the household. A gender variable (1 if female) is also included for carpool and shared taxi.

Following the emergent value principle, we added to the base model specification two variables constructed to account for the AD and CP decoys. Thus, the significance of the coefficients of these auxiliary variables could be interpreted as a test for the existence of the decoy effect in this RP experiment.

The first challenge in investigating the decoy effect in this non-tailored RP setting is related with the number of variables in the model, which are more than two, as considered in the SP model analyzed in Sections 4 and 5. This could be done in countless ways. We handle this problem by defining two aggregated critical variables for which we will analyze dominance and extremeness aversion. The two critical variables considered are total travel time (in vehicle plus walking and waiting) and travel cost.

The second challenge for this RP setting is in defining dominance and extremeness aversion with more than three alternatives. For this, an emergent value variable to account for asymmetrically dominated decoys (EV_AD) was built as the number of modes that, for each alternative, were dominated both in total travel time and cost. Besides, an emergent value variable to account for compromise decoys (EV_CP) was built as a dummy that took value one if the mode under analysis was not extreme, both in total travel time and travel cost, and zero otherwise.

The results of this analysis are reported on Table 7, where we contrast the estimators attained with a traditional Logit model (replicating the model reported by Munizaga and Daziano, 2002) with a model where the proposed emergent value variables are included.

Results in Table 7 show that the effect of both EV variables is positive, but only the variable accounting for asymmetric dominance has a statistically significant impact on choice. The *t*-test for EV_AD is 3.09, which implies a *p*-value below 1%. On the contrary the *t*-test for EV_CP is 0.823, which implies a *p*-value of 41%. Furthermore, using a Horowitz test, it can be concluded that the fit of the Emergent Value model is statistically better than that of the Logit model, suggesting that the decoy effect is present in this RP mode choice data, at least in its asymmetrically dominated form.

7. Conclusion

This article assesses the suitability of a set of context-RUM choice models for being the underlying choice behavior behind the Data Generation Process (DGP) that results on decoy effect outcomes that have been described in empirical literature. We begin developing a set of practical context-RUM model by adapting various conceptual cognitive models that had been suggested in previous literature as possible explanations for finding decoy outcomes. After showing that neither Logit nor the Nested Logit can be behind the decoy effect, we demonstrate that such a goal could be achieved instead by ad-hoc versions of the context-RUM emergent value (EV) and weight change (WC) models. We also show that this can be achieved as well by the much more parsimonious Random Regret Minimization (RRM) model and the variation of it proposed in this article, the Regret by Aspect (RBA) model. Results also show that, although the parsimonious Contextual Concavity Model (CCM) may capture all decoy types, the relative strength of the decoy types implied by it seems to be different from what has been described in the previous literature.

Then, we report a case study of a Stated Preference (SP) survey of route choice in which we first studied the existence of the decoy effect and then analyzed the empirical strength of the various decoy types. The results showed, for the first time to the best of our knowledge, that the decoy effect was present in this context and that the relative empirical strength between dominated and non-dominated decoys was coherent with the relation that would result if the parsimonious RRM or RBA were behind the DGP. In turn, the relative strength within Asymmetrically Dominated (AD) decoys only coincided with that of RRM and RBA for a group of individuals that behaved fully accordingly with the compromise principle, being reversed for other individuals. The shapes of the decoy strength encountered in this experiment had no relation with the shape implied by the CCM.

Besides, we used this real SP data to assess the different models under study in terms of estimation and forecasting. This SP data showed first strong evidence suggesting that EV was superior in estimation, but principally in forecasting, compared to the WC model. Among the parsimonious models only RRM and RBA were superior to the Logit in forecasting, but inferior to EV for the AD experiment. For the CP experiment, the performance of RRM and RBA was statistically as good as that of EV, considering a 5% test, while only the latter passed a 10% test. CCM was clearly inferior, even below the Logit, while also showing important estimation difficulties.

The better fit in estimation of the EV model suggests that, although the more parsimonious RRM and RBA may be able to explain, to some extent, the DGP behind decoy effect, there still seems to be something missing in them, which is captured by the ad-hoc dummies of the EV models.

We finally consider a Revealed Preference (RP) mode choice experiment with which we detect, for the first time to the best of our knowledge, an AD decoy effect in this choice setting. We also use this experiment to illustrate how to handle the decoy phenomena in a real context with various alternatives and variables.

Of course, it cannot be affirmed that the results attained for these particular case studies have general validity. Much more empirical analysis is needed to sustain such a statement. For such a goal, the discrete choice analysis deployed in this article offer a systematic analytical framework that could facilitate future research in this area.

Various lines of future research can be identified in this area. The first is the quest for better parsimonious models able to account for the decoy effect. The pursuing of such a goal could be based on the extension and adaptation of practical models based on choice behavior that has been described in judgmental decision making literature, as well as on a complete analysis of the Multivariate Extreme Value (MEV) family of models. It would also be relevant to collect a much more extended and heterogeneous data that would allow, among other things, detailed identification of various classes of choice behaviors that may arise when facing a decoy. Another relevant line for future research lies in the identification of the conditions under which the decoy effect may be present in a real transportation context. Preliminary evidence from pilot surveys developed under this research, suggest that the decoy effect was difficult to detect when the choice-maker was presented with commuting, instead of shopping options, and when the objective and the competitor were far from the trade-off line. Another research venue in this topic corresponds to the design of transportation public policies that may benefit from the existence of the decoy effect, beyond the pricing policy depicted in Figs. 1 and 2. Finally, a big challenge that emerges from this research is the potential need for a reformulation of the welfare analysis under a DGP that is not based on the classical random-utility maximization principle that ignores the choice context.

Acknowledgments

This publication was partially funded by CONICYT (FONDECYT 1150590), the Complex Engineering Systems Institute, **ISCI (ICM-FIC: P05-004-F, CONICYT: FB0816)**, and a Leverhulme's Visiting Professorship (VP1-2015-054) held at the University of Leeds and hosted by Professor Stephene Hess. We would also like to acknowledge the support of Universidad de los Andes, in Chile, where the data collection and the preliminary stages of this research were developed while the second author was pursuing a Master degree under the supervision of the first author. Besides, we would like to acknowledge the valuable comments provided by Stephane Hess, Caspar Chorus, Michel Bierlaire, Mogens Fosgerau, three anonymous referees and the attendants of the conferences hEART2014, ICMC2015, IATBR2015 and ITEA2016, where preliminary versions of this research were presented. Of course, all potential errors remain ours. All models and data analysis was developed using the open-source software R (R Development Core Team, 2008).

References

Ariely, D., 2008. Predictably Irrational. HarperCollins, New York.
Ariely, D., Wallsten, T.S., 1995. Seeking subjective dominance in multidimensional space: an explanation of the asymmetric dominance effect. Organ. Behav. Hum. Decis. Process. 63 (3), 223–232.
Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, Vol. 9. MIT press.
Chorus, C., 2010. A new model of random regret minimization. EJTIR 10 (2), 181–196.
Chorus, C.G., Bierlaire, M., 2013. An empirical comparison of travel choice models that capture preferences for compromise alternatives. Transportation 40 (3), 549–562.
Chorus, C.G., Arentze, T.A., Timmermans, H.J., 2008. A random regret-minimization model of travel choice. Transp. Res. Part B 42 (1), 1–18.
Daly, AJ, Hess, S, 2011. Simple approaches for random utility modelling with panel data. The 90th Annual Meeting of the Transportation Research Board Paper presented at.
Doyle, J., O'Connor, D., Reynolds, G., 1999. The robustness of the asymmetrically dominated effect: buying frames, phantom alternatives, and in-store purchases. Psychol. Mark. 16, 225–243.
Durkalski, V.L., Palesch, Y.Y., Lipsitz, S.R., Rust, P.F., 2003. Analysis of clustered matched-pair data. Stat. Med. 22 (15), 2417–2428.

Frederick, S., Lee, L., Baskin, E., 2014. The limits of attraction. J. Mark. Res. 51 (4), 487–507.
Fukushi, M., 2015. Detectando y Modelando El Efecto Decoy En Transporte. Universidad de los Andes, Santiago, Chile Master Thesis.
Guevara, C.A., 2016. Mode-valued differences of in-vehicle travel time savings. Transportation 1–21.
Herne, K., 1998. Testing the reference-dependent model: an experiment on asymmetrically dominated reference points. J. Behav. Decis. Mak. 11 (3), 181–192.
Highhouse, S., 1996. Context-dependent selection: the effects of decoy and phantom job candidates. Organ. Behav. Hum. Decis. Process. 65 (1), 68–76.
Horowitz, J.L., 1983. Statistical comparison of non-nested probabilistic discrete choice models. Transp. Sci. 17 (3), 319–350.
Huber, J., Payne, J.W., Puto, C.P., 2014. Let's be honest about the attraction effect. J. Mark. Res. 51 (4), 520–525.
Huber, J., Payne, J., Puto, C., 1982. Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. J Consum. Res. 9 (1), 90–98.
Huber, J., Puto, C., 1983. Market boundaries and product choice: Illustrating attraction and substitution effects. J. Consum. Res. 10 (1), 31–44.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 263–291.
Kivetz, R., Netzer, O., Srinivasan, V., 2004. Alternative models for capturing the compromise effect. J. Mark. Res. 41 (3), 237–257.
Latty, T., Beekman, M., 2011. Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. Proc. R. Soc. Lond. B 278 (1703), 307–312.
Leong, W., Hensher, D.A., 2012. Embedding multiple heuristics into choice models: an exploratory analysis. J. Choice Model. 5 (3), 131–144.
Luce, R.D., 1977. The choice axiom after twenty years. J. Math. Psychol. 15 (3), 215–233.
McFadden, D., 1978. Modelling the Choice of Residential Location. Institute of Transportation Studies, University of California, California, pp. 75–96.
McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12 (2), 153–157.
Müller-Lyer, F.C., 1889. Optische urteilstäuschungen. Arch. Für Anat. Physiol. Physiol. Abt. 2, 263–270.
Munizaga, M., Daziano, R., 2002. Evaluation of mixed logit as a practical modelling alternative. In: Proceedings European Transport Conference. Cambridge, U.K..
Ortúzar, J.deD., Donoso, P.C.F., 1983. Survey design, implementation, data coding and evaluation for the estimation of disaggregate choice models in Santiago, Chile. 2nd International Conference on Survey Methods in Transport, Septiembre.
Ortúzar, J.deD., Fernández, J.E., 1985. On the stability of discrete choice models in different environments. Transp. Plann. Technol. 10 (3), 209–218.
Parducci, A., 1974. Contextual effects: a range-frequency analysis. In: Handbook of Perception, 2, pp. 127–141.
Pechtl, H., 2009. Value structures in a decoy and compromise effect experiment. Psychol. Mark. 26 (8), 736–759.
Pettibone, J.C., Wedell, D.H., 2000. Examining models of nondominated decoy effects across judgment and choice. Organ. Behav. Hum. Decis. Process. 81 (2), 300–328.
Pratkanis, A.R., Farquhar, P.H., 1992. A brief history of research on phantom alternatives: evidence for seven empirical generalizations about phantoms. Basic Appl. Soc. Psychol. 13 (1), 103–122.
Quiggin, J., 1994. Regret theory with general choice sets. J. Risk Uncertain. 8 (2), 153–165.
R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Rooderkerk, R.P., Van Heerde, H.J., Bijmolt, T.H., 2011. Incorporating context effects into a choice model. J. Mark. Res. 48 (4), 767–780.
Shafir, S., Waite, T., Smith, B., 2002. Context-dependent violations of rational choice in honeybees (Apis mellifera) and gray jays (Perisoreus canadensis). Behav. Ecol. Sociobiol. 51 (2), 180–187.
Simonson, I., 1989. Choice of based on reasons: the case of attraction and compromise effects. J. Consum. Res. 16, 158–174.
Simonson, I., Tversky, A., 1992. Choice in context: Tradeoff contrast and extremeness aversion. J. Mark. Res. 29 (3), 281.
Trueblood, J.S., Brown, S.D., Heathcote, A., 2015. The fragile nature of contextual preference reversals: Reply to Tsetsos, Chater, and Usher. Physchol. Rev. 122 (4), 848–853 (2015).
van Cranenburgh, S., Guevara, C.A., Chorus, C.G., 2015. New insights on random regret minimization models. Transp. Res. Part A 74, 91–109.
Van Cranenburgh, S., Prato, C.G. & Chorus, C.G. 2016 "Accounting for variation in choice set size in random regret minimization models" (working paper).
Wedell, D.H., 1991. Distinguishing among models of contextually induced preference reversals. J. Exp. Psychol. 17, 767–778.
Wedell, D.H., Pettibone, J.C., 1996. Using judgments to understand decoy effects in choice. Organ. Behav. Hum. Decis. Process. 67 (3), 326–344.
Yang, S., Lynn, M., 2014. More evidence challenging the robustness and usefulness of the attraction effect. J. Mark. Res. 51 (4), 508–513.