Stochastic Complementarity*

Paola Manzini† Marco Mariotti‡ Levent Ülkü§

This version: December 2014

Abstract

Classical definitions of complementarity are based on cross price elasticities, and so they do not apply, for example, when goods are free. This context includes many relevant cases such as online newspapers and public attractions. We look for a complementarity notion that does not rely on price variation and that is: *behavioural* (based only on observable choice data); and *model-free* (valid whether the agent is rational or not). We uncover a conflict between properties that complementarity should intuitively possess. We discuss three ways out of the impossibility.

J.E.L. codes: D0.

Keywords: Complements and substitutes; Correlation; Stochastic choice.

*We thank Matthew Gentzkow, Faruk Gül, Rich McLean and Luigi Pistaferri for useful comments.

†School of Economics and Finance, University of St. Andrews, and IZA - Institute for Labor Economics, Bonn; paola.manzini@st-andrews.ac.uk

‡School of Economics and Finance, Queen Mary University of London; m.mariotti@qmul.ac.uk

§Department of Economics and CIE, ITAM; levent.ulku@itam.mx

1 Introduction

Suppose a local government wants to know whether two free public attractions, say a museum and a park, are complements or substitutes. The authority would be surprised on hearing that this situation does not fit the standard definitions of complementarity: these are based on price elasticities, while in this case both prices are fixed at zero. Indeed, price variations are just a *tool* to check complementarity: the notion of complementarity itself is not *intrinsically* related to price variations.

Similar observations apply in many other situations. Consider online newspapers, reviews/advice (e.g. financial) on social networks, public radio broadcasts, file sharing: all of these goods are often free. Another leading example is that of complementarity in business practices, such as training the workforce and allowing it more decisional discretion (Brynjolfson and Milgrom [4]). More abstractly, the 'goods' may be characteristics embodied in the objects of choice, so that any price variation is perfectly correlated between the goods. Of course there may be no prices at all: is beauty a complement or a substitute of wealth in a partner?

In this paper we look for definitions of complementarity (and substitutability) between two goods that (1) are not based on price variations; and (2) are *behavioural* and *model free*, in the sense that they just use choice data as inputs: they do not commit to any 'psychological' variable (such as utility) as the driver of behaviour, and they do not rely on assumptions about the specific choice procedure (such as 'utility maximisation plus errors') that generates individual consumption data. We uncover a conflict between properties that complementarity should intuitively possess, and we propose three possible ways out of this conflict.

Feature (2) above distinguishes our approach both from the 'supermodularity' approach (see section 6) and from the recent econometric approaches that deal with the zero price problem in the behavioural study of complementarity, pioneered by Gentzkow [8]. The assume that a specific decision model underlies choice data, typically a Random Utility Model (RUM) such as the multinomial logit and variations thereof. While these models may be in practice good approximations of actual behaviour, the recent wave of abstract works on stochastic choice (Brady and Rhebeck [3], Echenique, Saito and Tserenjigmid [7], Gül, Natenzon and Pesendorfer [9], Manzini and Mariotti [11]) has highlighted a wide variety of possible 'choice errors' and choice procedures, and so a number of reasons why agents' behaviour may fail to be described by a logit model, and indeed even by the much larger class of RUMs. It is therefore interesting to complement those analyses with model-free ones.

Complementarity in general is such a central concept in Economics (you had no trouble grasping the meaning of our initial paragraph) that its study hardly needs to be motivated. On the one hand, complementarity has deeply engaged some of the giants of the profession (see the historical overview in Samuelson [15]). On the other hand, knowing whether goods are complements or substitutes (or neither) is of major practical importance in disparate areas: for example, suppliers must have information about complementarity when introducing new products or when pricing existing products; so do regulators to evaluate the competitiveness of a market; businesses may be reluctant to change a practice because of its complementarity with another; and so on.

Let us now formulate the problem more precisely. Consider two goods, the online and the print versions of a newspaper. Available data are in the form $(p_{OP}, p_O, p_P, p_\varnothing)$, where p_O and p_P denote the consumption frequency of the online version only and of the print version only, respectively, p_{OP} denotes the frequency of joint consumption, and $p_\varnothing$ denotes the frequency with which neither version is read. When could one say, on the basis of these data alone, that the print and online versions are complementary?

One natural answer would seem to be 'whenever they are positively correlated': that is, when the posterior probability of reading one version conditional on reading the other version, $\frac{p_{OP}}{p_{OP}+p_P}$, is greater than the prior probability, $p_{OP} + p_O$. This interpretation is often assumed. For example, Brynjolfson and Milgrom [4] write about the measurement of complementarities in business practices:

> Measuring the correlation, or clustering, of practices is perhaps the most common approach to testing for complementarities (p. 33)

and in the mentioned study by Gentzkow, he writes

> The basic fact in the raw data is that a consumer who reads any one paper is on average more likely to have also read a second paper. If all heterogeneity in utilities were uncorrelated across papers, this would be strong evidence that all three are complements.[1]

[1] He considers two print papers and one online version.

The difficulty to be solved in this latter case is that the observed correlation may partly reflect correlated unobservable tastes for the goods, rather than 'true' complementarity: for instance, a news junkie may consume both paper and online versions even when there is no 'true' complementarity, which *in a model of utility driven consumer* means a positive difference between the value of joint consumption and the sum of the values of single good consumptions.[2]

In our approach we do not and cannot distinguish between 'true' complementarity/correlation and 'taste' correlation - we just take consumption data at face value. This allows us to deal also with consumers who, for example, choose on the basis of a boundedly rational procedure. These consumers may not maximise a utility function. They may not even 'have' a utility function.

However, we maintain that even so - adopting a purely behavioural, model-free view - the identification of complementarity with pure correlation cannot be taken for granted. Statistical association in general captures the idea that more of one variable 'goes together' with more of the other. This is an intuitive property of complementarity. But, equally intuitively, we would also expect complementarity to have a different property, namely that increases in joint consumption not compensated by increases in single good consumption should be evidence of complementarity (and similarly, that increases of single good consumption not compensated by increases in joint consumption should be evidence of substitutability). As it turns out, these two intuitions are in conflict. Suppose that the data are given by the following table

	Read Print	Did not read print
Read Online	0.3	0.2
Did not read online	0.2	0.3

that is $(p_{OP}, p_O, p_P, p_\varnothing) = (0.3, 0.2, 0.2, 0.3)$. Then the data indicate a positive correlation $(\frac{p_{OP}}{p_{OP}+p_P} = 0.6 > 0.5 = p_{OP} + p_O)$. Suppose now that joint consumption rises to $p'_{OP} = 0.55$ while single good consumptions stay the same. Then the correlation turns negative $(\frac{p'_{OP}}{p'_{OP}+p'_P} = 0.73 < 0.75 = p'_{OP} + p'_O)$. An increase in joint consumption has transformed the goods from complementary to substitutes!

Our first main contribution is to show that this simple example illustrates a deeper

[2] As Gentzkow shows, in the two good model this is equivalent to a positive compensated cross price elasticity of demand.

conflict between two seemingly natural principles that criteria for complementarity should satisfy. One principle is *monotonicity*: an increase joint consumption accompanied by decreases in single good consumption should strengthen an existing complementarity.

The second principle is *duality*: if in a dataset O and P are complementary, then they are substitutes in the 'opposite' dataset in which the instances of consumption of P are switched with the instances of non-consumption of P (holding fixed the consumption/non-consumption of O). Duality is evidently satisfied by all common measures of correlation and association. This is the dual table of the previous one:

	Read print	Did not read print
Read online	0.2	0.3
Did not read online	0.3	0.2

theorem 1 and its corollary 1 show that any concept of complementarity that satisfies monotonicity and duality must be also *unresponsive*, in the sense that the level of non-consumption $p_{\varnothing}$ *alone* determines whether the goods are complements or substitutes, irrespective of the distribution between single and joint consumption. A second impossibility result (theorem 2) shows that duality and monotonicity are in outright conflict if it is also assumed that the frontier between complementarity and independence is thin, as is the case for the standard elasticity-based criterion.

We then look for ways out of the impossibility. We first show that correlation is the only symmetric criterion of complementarity that satisfies both duality and a modified monotonicity condition. Next, we examine a monotonic criterion that is economically intuitive if the numbers p_{OP}, p_O and p_P are taken as expressing the values of the respective options: O and P are complementary (resp., substitutes) if $p_{OP} > p_O + p_P$ (resp., $p_{OP} < p_O + p_P$). This criterion satisfies a different notion of duality, based on exchanging joint consumption with total single good consumption. We finally consider a third criterion for complementarity, which is also monotonic. It says that O and P are complementary (resp., substitutes) if $p_{OP} > \max\{p_O, p_P\}$ (resp., $p_{OP} < \min\{p_O, p_P\}$). This criterion satisfies yet a different notion of duality, based on exchanging joint consumption with one type of single good consumption. We consider these as the three main candidate criteria of complementarity.[3]

[3] As a matter of fact, while various commentators (including the authors) have different preferences over these three criteria, and any of the three gets some support, none has been suggested outside of the

Having defined complementarity in these ways, we finally go back, with some examples, to the question that is usually the starting point of the analysis: How do tastes or cognitive variables affect complementarity? For example, if choice behaviour is at least in part guided by preferences, what aspect of preferences makes two good complementary or substitutes? In order to answer such questions we need to postulate specific models of the process leading to choice. We look at two models in particular. The first is the classical Luce (or multinomial logit) model of stochastic choice. The second is (a simplified version of) the more recent 'stochastic consideration set' model of Manzini and Mariotti [11] and Brady and Rehbeck [3]. We discover that in both cases the correlation criterion on data reflects supermodularity types of condition on preferences.

2 Preliminaries

There are two goods, x and y. A *datapoint* is an ordered four-tuple $p = (p_{xy}, p_x, p_y, p_\varnothing)$ with $p_k \in (0,1)$ for $k \in \{xy, x, y, \varnothing\}$ and $\sum_{k\in\{xy,x,y,\varnothing\}} p_k = 1$. The interpretation is that p_{xy} denotes the probability (or frequency) of joint consumption of x and y, p_x and p_y denote the probabilities of consumption of x but not y and of y but not x, respectively, and $p_\varnothing$ denotes the probability of consuming neither x nor y.

We consider the partitions of the space

$$T = \left\{(a, b, c, d) \in (0,1)^4 : a + b + c + d = 1\right\}$$

in three regions: the *complementarity region* C, the *substitution region* S and the *independence region* I. If $p \in C$ (resp. $p \in S$, resp. $p \in I$) we say that x and y are complements (resp. substitutes, resp. independent) at p. We call any such partition a *criterion*.

Here are some examples of criteria:

Example 1 *(correlation):*

$$\begin{aligned} C &= \left\{(p_{xy}, p_x, p_y, p_\varnothing) \in T : \frac{p_{xy}}{p_{xy} + p_y} > p_{xy} + p_x\right\} \\ S &= \left\{(p_{xy}, p_x, p_y, p_\varnothing) \in T : \frac{p_{xy}}{p_{xy} + p_y} < p_{xy} + p_x\right\} \end{aligned}$$

three.

According to the correlation criterion a datapoint is in C (resp., S) if and only if the information that one of the goods is consumed increases (resp., decreases) the probability the other good is also consumed.

Example 2 *(additivity)*

$$
\begin{aligned}
C &= \{(p_{xy}, p_x, p_y, p_\varnothing) \in T : p_{xy} > p_x + p_y\} \\
S &= \{(p_{xy}, p_x, p_y, p_\varnothing) \in T : p_{xy} < p_x + p_y\}
\end{aligned}
$$

The additivity criterion is natural whenever one thinks of the probabilities as expressing 'values' (as is the case in the logit model). Then it says that x and y are complements whenever the value of joint consumption is greater than the sums of the values of the goods when consumed singly. This is in fact the notion of complementarity used in many applications, e.g. the literature on bundling (e.g. Armstrong [1]).

Example 3 *(maxmin)*

$$
\begin{aligned}
C &= \{(p_{xy}, p_x, p_y, p_\varnothing) \in T : p_{xy} > \max\{p_x, p_y\}\} \\
S &= \{(p_{xy}, p_x, p_y, p_\varnothing) \in T : p_{xy} < \min\{p_x, p_y\}\}
\end{aligned}
$$

The maxmin criterion fits, for instance, the situation in which one good is an 'accessory' and only the 'dominant' single good consumption is relevant in comparison with joint consumption to declare complementarity. To check whether steak and pepper are complementary you may want to compare the probability of consumption of steak with that of steak and pepper, rather than with that of pepper alone. Substitution is declared symmetrically.

Example 4 *(supermodularity)*

$$
\begin{aligned}
C &= \{(p_{xy}, p_x, p_y, p_\varnothing) \in T : p_{xy} + p_\varnothing > p_x + p_y\} \\
S &= \{(p_{xy}, p_x, p_y, p_\varnothing) \in T : p_{xy} + p_\varnothing < p_x + p_y\}
\end{aligned}
$$

Here the the goods are delcared complementary if a supermodularity condition on p is satisfied (with p seen as a function defined on the set of consumption bundles $\{xy, x, y, \varnothing\}$). Supermodularity-type conditions capture complementarity when imposed on an objective function to be maximised (Topkis [16], Milgrom and Roberts [13], Milgrom and Shannon [14]). Note that the condition is equivalent to $p_{xy} + p_\varnothing > \frac{1}{2}$.

For illustration, consider table 1, calculated on the basis of the Gentzkow [8] data on 5-day readership of the online and print version of the Washington Post:

	Read print	Did not read print
Read online	0.137	0.043
Did not read online	0.447	0.373

Table 1: Washington Post, 5-day readership of online and print version (Gentzkow [8]).

In this case the above example criteria are in deep conflict: the correlation and supermodularity criteria indicate that the two versions are complementary, the additivity criterion indicates that they are substitutes, and the maximin criterion indicates that they are independent. Therefore, in order to assess the different criteria, we propose an axiomatic analysis, looking for natural properties that criteria should possess given the interpretation.

3 Impossibilities

In this section we study the core conflict between properties.

Symmetry:
1) If $(a, b, c, d) \in C$ then $(a, c, b, d) \in C$.
2) If $(a, b, c, d) \in S$ then $(a, c, b, d) \in S$.

Symmetry says that exchanging the amounts of single good consumptions is immaterial for the purpose of classifying goods into complementary or substitutes. Samuelson [15] considers its symmetry as one the two major improvements of the Slutsky-Hicks-Allen-Schultz 'compensated' definitions compared to the 'uncompensated' one.

Note that the two symmetry conditions imply an analogous property for I: if $(a, b, c, d) \in I$, then $(a, c, b, d) \in I$. For if $(a, c, b, d) \notin I$, then one of the two conditions would yield $(a, b, c, d) \notin I$.

As explained in the introduction, we view duality as the 'soul' of all association-based definitions of complementarity and substitution:

Duality

1) If $(a,b,c,d) \in C$ then $(b,a,d,c) \in S$.
2) If $(a,b,c,d) \in S$ then $(b,a,d,c) \in C$.

Suppose that you have two datapoints p and q. Suppose that, whether x is consumed or not, y is consumed at q with the same frequency with which it is not consumed at p. If a datapoint were presented in table form, as in the introduction, q would be obtained from q by switching the rows. For example, q could be obtained when y is consumed only in weekends while p is obtained when y is consumed only in weekdays (assuming for simplicity that y's consumption pattern is the same whether x is consumed or not). In this sense q expresses a behaviour that is the 'opposite' of the behaviour at p. Then duality says that x and y are complements at p only if they are substitutes at q, and vice-versa.

Note that, as for Symmetry, the two duality conditions imply a third one, that if $(a,b,c,d) \in I$, then $(b,a,d,c) \in I$.

The second main principle we have discussed is:

Monotonicity
1) If $(a,b,c,d) \in C$, $(a',b',c',d') \in T$, $a' \geq a$, $b \geq b'$ and $c \geq c'$ then $(a',b',c',d') \in C$.
2) If $(a,b,c,d) \in S$, $(a',b',c',d') \in T$, $a \geq a'$, $b' \geq b$ and $c' \geq c$ then $(a',b',c',d') \in S$.

Monotonicity says that, if goods are complements, then an increase in joint consumption without an increase in single consumption cannot transform them into substitutes or render them independent, and vice-versa.

There do exist criteria that satisfy Symmetry, Duality and Monotonicity: for example, the supermodularity criterion above. However, this criterion seems highly unsatisfactory, because it declares the goods complementary at any datapoint for which $p_{\varnothing} > \frac{1}{2}$, for *all* possible values of p_{xy}, p_x and p_y. Instead, it is desirable that $p_{\varnothing}$ should not be decisive *by itself* to declare either complementarity or substitution: it should also respond to variations the frequencies of joint and single consumptions.

The following property captures an even weaker version of this idea. Essentially, it just excludes the (bizarre) claim that is implicit in a complementarity criterion such as $p_{\varnothing} > \frac{1}{2}$: 'these goods are clearly complementary: they are rarely consumed together'. While it allows in principle $p_{\varnothing}$ to be decisive, it should not be 'high' non-consumption to indicate complementarity.

Responsiveness: There exists $\tau \in (0,1)$ such that, for all $d \in (\tau, 1)$, $(a,b,c,d) \in S$ for some $(a,b,c) \in (0,1)^3$.

For example, the correlation criterion satisfies Responsiveness. It is also satisfies Symmetry and Duality, but, as noted in the introduction, it is not monotonic. The additivity criterion satisfies all properties except part (2) of Duality. The maxmin criterion fails only Duality.

It turns out that all possible symmetric criteria that are correlation-based in the sense of satisfying duality must either fail Monotonicity or Responsiveness:

Theorem 1 *There exists no criterion that satisfies Symmetry, Monotonicity, Duality and Responsiveness.*

Proof: We start by proving:

Claim: Let (C, I, S) be a criterion that satisfies Symmetry and Duality. If $(a,b,c,d) \in C$ then $(d,b,c,a) \in C$.

To see this, suppose $(a,b,c,d) \in C$. By Symmetry $(a,c,b,d) \in C$. By Duality $(c,a,d,b) \in S$. By Symmetry $(c,d,a,b) \in S$. By Duality $(d,c,b,a) \in C$. Finally, by Symmetry $(d,b,c,a) \in C$.

Returning to the proof of the main result, suppose that a criterion (C, I, S) satisfies Symmetry, Monotonicity, Duality and Responsiveness. By Responsiveness and Duality[4] there exists a $p = (a,b,c,d) \in C$. Let $\theta = \min\{a,b,c,d\}$, and note in particular that it must be $d < 1 - \theta$.

We will now show that for all $q = (a', b', c', d') \in T$, if $d' > 1 - \theta$ then $q \in C$. This contradicts Responsiveness and thus proves the impossibility. Take such a q, and let $r = (d', b', c', a')$. Note that $b' < \theta$ (otherwise, if $b' \geq \theta$, then $d' > 1 - b'$ and thus $b' + d' > 1$), and similarly $c' < \theta$. Then $d' > 1 - \theta > a$, $d' < \theta \leq a$ and $c' < \theta \leq c$. By Monotonicity, $r \in C$ and by the Claim above, we conclude that $q \in C$. ■

Symmetry is not implied by the other three axioms. For example the criterion given by $C = \{(a,b,c,d) : a > b\}$ and $S = \{(a,b,c,d) : a < b\}$ satisfies all of them but not

[4] To see this, recall that by Responsiveness there exists $\tau \in (0,1)$ such that, for all $d \in (\tau, 1)$, $(a,b,c,d) \in S$ for some $(a,b,c) \in (0,1)^3$. Then by (2) of Duality $(b,a,d,c) \in C$.

Symmetry. This - together with the other examples given previously - shows that the impossibility result of theorem 1 is tight.

To clarify the role played by Symmetry in the impossibility, consider the following strengthening of Duality.

Duality*

1) If $(a, b, c, d) \in C$ then $(b, a, d, c) \in S$ and $(c, d, a, b) \in S$.

2) If $(a, b, c, d) \in S$ then $(b, a, d, c) \in C$ and $(c, d, a, b) \in C$.

Duality* adds to Duality the requirement that switching columns in a table leads to the same effect as switching rows. In the presence of Symmetry, Duality and Duality* are equivalent. However Duality* alone does not imply Symmetry. For example the criterion defined by $C = \{(a, b, c, d) : b > a \text{ and } c > d\}$ and $S = \{(a, b, c, d) : b < a \text{ and } c < d\}$ satisfies Duality* but fails Symmetry: $(0.3, 0.31, 0.2, 0.19) \in C$ yet $(0.3, 0.2, 0.31, 0.19) \in I$. Note that this criterion also fails Monotonicity, which should not be surprising, given the following result.

Lemma 1 *If a criterion satisfies Duality* and Monotonicity, then it satisfies Symmetry.*

Proof. Suppose that a criterion (C, I, S) satisfy Duality* and Monotonicity but fails Symmetry. We consider four cases.

Proof. We consider four cases.

Case 1: $(a, b, c, d) \in C$ but $(a, c, b, d) \in S$. By Duality* $(b, a, d, c) \in S$ and $(c, a, d, b) \in C$. By Monotonicity $c > b$. Applying Duality* to the first two datapoints we get $(c, d, a, b) \in S$ and $(b, d, a, c) \in C$, and then by Monotonicity $b > c$, a contradiction.

Case 2: $(a, b, c, d) \in C$ but $(a, c, b, d) \in I$. By Duality* $(b, a, d, c) \in S$ and $(c, a, d, b) \in I$. By Monotonicity $c > b$. Applying Duality* to the first two datapoints we get $(c, d, a, b) \in S$ and $(b, d, a, c) \in I$, and then by Monotonicity $b > c$, a contradiction.

Cases 3 and 4 where $(a, b, c, d) \in S$ and $(a, c, b, d) \notin S$ are similar. ■

Hence,

Corollary 1 *There exists no criterion that satisfies Duality*, Monotonicity and Responsiveness.*

Finally, the incompatibility between correlation and monotonicity properties can also be observed from a different angle. Consider:

$I-$**Monotonicity**
1) If $(a,b,c,d) \in I$, $(a',b',c',d') \in T$, and $a' \geq a$, $b' \leq b$ and $c' \leq c$, with at least one inequality strict, then $(a',b',c',d') \in C$.
2) If $(a,b,c,d) \in I$, $(a',b',c',d') \in T$, and $a' \leq a$, $b' \geq b$ and $c' \geq c$, with at least one inequality strict, then $(a',b',c',d') \in S$.

Loosely, $I-$Monotonicity says that, if the goods are independent, then increasing joint consumption while decreasing single good consumption makes them complementary. This monotonicity property incorporates a responsiveness requirement: essentially, it implies that the Independence area is thin, as is the case for all standard definitions of complementarity/substitutability.

Theorem 2 *There exists no criterion that satisfies Symmetry, $I-$Monotonicity, and Duality.*

Proof: Suppose that (C,I,S) satisfies the axioms. Take $p=(a,a,b,b)$ with $b>a$. It cannot be $(a,a,b,b)\in S$, for then by Duality $(a,a,b,b)\in C$, a contradiction. Similarly, it cannot be $(a,a,b,b)\in C$. Then $(a,a,b,b)\in I$. By Symmetry, $(a,b,a,b)\in I$. By Duality $(b,a,b,a)\in I$. But this contradicts $I-$Monotonicity. ■

4 Possibilities

We now turn to 'resolutions' of the conflicts. We analyse three plausible criteria, correlation, additivity and maxmin. Correlation is obtained by preserving Duality and appropriately modifying the monotonicity properties. For the other two criteria, we retain the monotonicity properties but vary the notion of duality. A duality operation produces the 'opposite' behaviour to a given one, and a duality property in our context asserts, loosely, that if a datapoint is classified in a certain way, then its dual is classified in the opposite way. This is an intuitive requirement but, as we will see, there are other reasonable ways to interpret the concept of 'opposite' behaviour, hence other reasonable versions of duality.

We will focus on criteria that satisfy the following regularity condition: If $(a, b, c, d) \in I$, $(b, a, c, d) \in I$ and $(c, b, a, d) \in I$, then $a = b = c$. It is easy to check that all three criteria that we shall characterise satisfy it, and that so does the supermodularity criterion.

4.1 Correlation

Recall that according to the correlation criterion two goods are complements (substitutes) if their consumption is positively (negatively) correlated. While, as we have seen, the criterion fails Monotonicity, it satisfies a different monotonicity condition. Let us write, for any vector $q \in \Re^4_{++}$,

$$q^* = \frac{1}{\sum q_i} q$$

so that $q^* \in T$.

Scale Monotonicity
1) If $(a, b, c, d) \in C$ and $m \geq n > 0$, then $(ma, nb, c, d)^* \in C$.
2) If $(a, b, c, d) \in S$ and $n \geq m > 0$, then $(ma, nb, c, d)^* \in S$.
3) If $(a, b, c, d) \in I$, then $(ma, nb, c, d)^* \in I$ ($\in C$, $\in S$) if $m = n$ ($> n$, $< n$).

Suppose that the total time spent reading the online version (alone or together with the print version) changes, but the time spent reading the online version alone decreases (resp., increases) as a proportion of the time spent reading both versions. Suppose also that the time left is allocated exactly in the same proportion as before between reading the print version and not reading either version. Parts (1) and (2) of Scale Monotonicity say that if the initial consumption pattern indicated complementarity (resp., substitutability) , then the new consumption pattern should also indicate complementarity (resp., substitutability). Part (3) of the axioms states a similar idea based on I-Monotonicity.

Theorem 3 *A criterion satisfies Symmetry, Duality and Scale Monotonicity if and only if it is the correlation criterion.*

Proof: That the three axioms are necessarily satisfied by the correlation definition is trivial. Suppose that a criterion (C, I, S) satisfies the three axioms. Begin by noting that

$$\begin{aligned}(p_{xy}, p_x, p_y, p_\varnothing) \quad &\in \quad C \Leftrightarrow \frac{p_{xy}}{p_{xy}+p_y} > p_{xy} + p_x \\ &\Leftrightarrow \quad p_{xy}\left(1 - p_{xy} - p_x - p_y\right) > p_x p_y \\ &\Leftrightarrow \quad p_{xy} p_\varnothing > p_x p_y\end{aligned}$$

and similarly $(p_{xy}, p_x, p_y, p_\varnothing) \in S \Leftrightarrow p_{xy}p_\varnothing < p_x p_y$. Then, since C, I and S form a partition, the result follows from the following three claims.

Claim 1: $C \subseteq \{(a, b, c, d) \in T : ad > bc\}$. Take $(a, b, c, d) \in C$ and suppose towards a contradiction that $ad \leq bc$. It follows that $\min\{a, d\} \leq \max\{b, c\}$. Symmetry and Duality imply that w.l.o.g. we can assume $d \leq a$ and $b \leq c$ so that $d \leq c$.

We will show that $b < a$. First note that $(d, c, b, a) \in C$ by Symmetry and Duality. By Scale Monotonicity (recall $d \leq c$) $(\frac{c}{d}d, \frac{d}{c}c, b, a)^* = (c, d, b, a) \in C$. Now Using Symmetry and Duality again we get $(a, b, d, c) \in C$. Duality gives $(b, a, c, d) \in S$. Now if $b \geq a$, applying Scale Monotonicity $(\frac{a}{b}b, \frac{b}{a}a, c, d)^* = (a, b, c, d) \in S$, a contradiction. Hence $b < a$ as we wanted to show.

We have $(a, \frac{ad}{bc}b, c, d)^* = (a, \frac{ad}{c}, c, d)^* \in C$ by Scale Monotonicity since $\frac{ad}{bc} \leq 1$. By Symmetry and Duality $(d, \frac{ad}{c}, c, a) \in C$. Applying Scale Monotonicity again, $(\frac{c}{d}d, \frac{c}{d}\frac{ad}{c}, c, a)^* = (c, a, c, a)^* \in C$. Apply Symmetry and Duality again and we get $(a, c, a, c)^* \in C$. By SM $(\frac{b}{a}a, \frac{d}{c}c, a, c)^* = (b, d, a, c) \in C$ since $ad \leq bc$ gives $\frac{d}{c} \leq \frac{b}{a}$. Finally this implies, by Symmetry and Duality, that $(a, b, c, d) \in S$, contradiction.

Claim 2: $S \subseteq \{ad < bc\}$. Take $(a, b, c, d) \in C$ and suppose towards a contradiction that $ad \geq bc$. It follows that $\min\{b, c\} \leq \max\{a, d\}$. Symmetry and Duality say that w.l.o.g. we can assume $a \leq d$ and $c \leq b$ so that $c \leq d$.

We will show that $a < b$. First note that $(d, c, b, a) \in S$ by Symmetry and Duality. By Scale Monotonicity (recall $c \leq d$) $(\frac{c}{d}d, \frac{d}{c}c, b, a)^* = (c, d, b, a) \in S$. Now Using Symmetry and Duality again we get $(a, b, d, c) \in S$. Duality gives $(b, a, c, d) \in C$. Now if $a \geq b$, applying Scale Monotonicity $(\frac{a}{b}b, \frac{b}{a}a, c, d)^* = (a, b, c, d) \in C$, a contradiction. Hence $a < b$ as we wanted to show.

We have $(a, \frac{ad}{bc}b, c, d)^* = (a, \frac{ad}{c}, c, d)^* \in S$ by Scale Monotonicity since $\frac{ad}{bc} \geq 1$. By Symmetry and Duality $(d, \frac{ad}{c}, c, a) \in S$. Applying Scale Monotonicity again, $(\frac{c}{d}d, \frac{c}{d}\frac{ad}{c}, c, a)^* = (c, a, c, a)^* \in S$. Apply Symmetry* and Symmetry to get $(a, c, a, c)^* \in S$. By Scale Monotonicity $(\frac{b}{a}a, \frac{d}{c}c, a, c)^* = (b, d, a, c) \in S$ since $ad \geq bc$ gives $\frac{d}{c} \geq \frac{b}{a}$. Finally this implies, by Symmetry and Duality, that $(a, b, c, d) \in C$, contradiction.

Claim 3: $I \subseteq \{ad = bc\}$. Take $(a, b, c, d) \in I$ and suppose towards a contradiction that $ad < bc$. Then set w.l.o.g. $d < c$ and consequently, using part (3) of Scale Monotonicity in an exact adaptation of Claim 1, $a > b$. The rest of the argument mirrors that in Claim 1. Similarly follow, with the obvious necessary modifications, the proof of Claim 2 if $ad > bc$. ■

4.2 Additivity

As noted before, the additivity criterion given in Example 2 is symmetric and monotonic. It also satisfies the notion of duality based on the operation illustrated below:

	y	$\sim y$
x	a	c
$\sim x$	b	d

$\longrightarrow$

	y	$\sim y$
x	$b+c$	$a\left(\frac{c}{b+c}\right)$
$\sim x$	$a\left(\frac{b}{b+c}\right)$	d

.

The operation consists of exchanging *Total* single good consumption with *Joint* consumption (with the joint consumption allocated to the two goods in proportion to the amounts that were consumed singly).

(T,J)-Duality For $\alpha = \frac{b}{b+c}$:
1) If $(a, b, c, d) \in C$, then $(b + c, \alpha a, (1 - \alpha) a, d) \in S$.
2) If $(a, b, c, d) \in S$, then $(b + c, \alpha a, (1 - \alpha) a, d) \in C$.

(T,J)-Duality says that the duality operation above transforms complementarity into substitution and viceversa. For example, if online and print newspapers are complements for a consumer who reads both versions two thirds of the time and a single version (either print or online) one fourth of the time, then they must be substitutes for a consumer who reads both versions one fourth of the time and the single versions two thirds of the time.

Note that if (a', b', c', d') is a (T,J)-dual to (a, b, c, d) in the sense of this axiom, i.e., if $a' = b+c$, $b' = ab/(b+c)$, $c' = ac/(b+c)$ and $d' = d$, then (a, b, c, d) is dual to (a', b', c', d') in the same way as well. Consequently, (T,J)-Duality implies: if $(a, b, c, d) \in I$, then $\left(a + b, \frac{ab}{a+b}, \frac{ac}{a+b}, d\right) \in I$.

Theorem 4 *A criterion satisfies Monotonicity, $I-$Monotonicity and (T,J)-Duality if and only if it is the additivity criterion.*

Proof. It is straightforward to show that the additivity criterion satisfies the three axioms. Suppose that (C, I, S) satisfies the three axioms. The result follows from the following three claims.

Claim 1: $C \subseteq \{(a, b, c, d) \in T : a > b + c\}$. Suppose towards a contradiction that $(a, b, c, d) \in C$ and $a \leq b + c$. By (T,J)-Duality, $(b + c, ab/(b + c), ac/(b + c), d) \in S$. Since $a/(b + c) \leq 1$ this contradicts Monotonicity.

Claim 2: $S \subseteq \{(a, b, c, d) \in T : a < b + c\}$. The proof is symmetric to that of Claim 1.

Claim 3: $I \subseteq \{(a, b, c, d) \in T : a = b + c\}$. Suppose that $(a, b, c, d) \in I$ but $a < b + c$. (T,J)-Duality yields $(b + c, ab/(b + c), ac/(b + c), d) \in I$, which contradicts I-Monotonicity. Similarly if $a > b + c$. ■

4.3 Maxmin

The Maxmin criterion is a monotonic criterion that differs structurally from the other two because it has a thick independence region (so that it will not satisfy $I-$Monotonicity). It expresses yet a different notion of duality, based on the operation illustrated below:

	y	$\sim y$
x	a	c
$\sim x$	b	d

$\rightarrow$

	y	$\sim y$
x	b	c
$\sim x$	a	d

Here, the behaviour 'opposite' to a given one is defined by exchanging joint consumption with *only one* of the single good consumptions. Ideally, we would like to impose a property of the following type. Suppose that online and print newspapers are complements for a consumer who, when he reads the print version, also reads the online version $\alpha\%$ of the time; then, they must be substitutes for a consumer who, when he reads the print version, also reads the online version $(1 - \alpha)\%$ of the time (and analogously starting from substitutability). It is a consequence of our characterisation below that this type of duality together with Symmetry and Monotonicity leads to another impossibility. So we use a weakened version of the property, which requires a full switch between complementarity and substitutability only for (at least) one of the two goods, and settles for merely switching out of the initial region (possibly to independence) for both goods.

(S,J)-Duality

1) If $(a, b, c, d) \in C$ then $q = (b, a, c, d) \notin C$ and $q' = (c, b, a, d) \notin C$, and at least one of $q \in S$ and $q' \in S$.
2) If $(a, b, c, d) \in S$ then $q = (b, a, c, d) \notin S$ and $q' = (c, b, a, d) \notin S$, and at least one of $q \in C$ and $q' \in C$.

Theorem 5 *A criterion (C, I, S) satisfies Symmetry, (S,J)-Duality and Monotonicity if and only if is the maxmin criterion.*

Proof: Necessity is easily checked. For sufficiency, we show first that $C \subseteq \{p : p_{xy} > \max\{p_x, p_y\}\}$. Suppose by contradiction that $p = (a, b, c, d) \in C$ but $a \leq \max\{b, c\}$. W.l.o.g. set $b \geq c$. Then by Monotonicity $q = (b, a, c, d) \in C$. On the other hand, (1) of (S,J)-Duality implies that $q \notin C$, a contradiction.

Next, we show $\{p : p_{xy} > \max\{p_x, p_y\}\} \subseteq C$. Suppose not, and let $a > \max\{b, c\}$ with $p = (a, b, c, d) \notin C$. Suppose first that $p \in S$. Then by Monotonicity $q = (b, a, c, d) \in S$ but by (2) of (S,J)-Duality $q \notin S$, contradiction. Suppose then that $p \in I$. Let $q = (b, a, c, d)$ and $q' = (c, b, a, d)$, and assume w.l.o.g. (by Symmetry) that $b \geq c$. If $q \in S$, then by Symmetry $(b, c, a, d) \in S$, so that by Monotonicity $q' \in S$; but then by (2) of (S,J)-Duality we should have either $p \in C$ or $q \in C$, a contradiction. If $q \in C$, then by Monotonicity (since $a > b$) it must be $p \in C$, a contradiction. Therefore $q \in I$ and $p \in I$. If $q' \in C$ we contradict Monotonicity, and if $q' \in S$, we contradict (2) of (S,J)-Duality. Therefore $q' \in I$. By the regularity condition this implies $a = b = c$, contradicting the assumption $a > \max\{b, c\}$.

The proof for S is analogous, and since (C, I, S) is a partition the result follows. ■

5 From behaviour to psychology: two examples with the correlation criterion

So far we have followed a rigorously behavioural approach, eschewing any hypothesis on the choice process that generates the data. Sometimes, though, one may entertain a hypothesis on the decision process that has generated the data. Even so, the previous analysis can be useful. We can ask what the psychological primitives must look like in a model for behavioural complementarity to be observed. In this way, we can obtain non-obvious complementarity conditions expressed, e.g., in terms of preferences, but justified by purely behavioural properties.

To perform this exercise we need to postulate some decision models: we study two 'polar' representatives. The first is the logit model, in which preferences are random and applied to a deterministic set. The second model is a simplification of the stochastic choice model in Manzini and Mariotti [11] and Brady and Rehbeck [3], in which preferences are deterministic but there is randomness in the subset of alternatives that are actively considered by the agent. For reasons of space, we perform the analysis only for the correlation criterion, which is the case yielding the most intriguing answers. As we shall see, in both polar cases this criterion implies supermodularity-style conditions on the psychological primitives.

In the *logit model*, we assume that each bundle $\sigma \in \{xy, x, y, \varnothing\}$ has a 'systematic utility' $u : \{xy, x, y, \varnothing\} \to \mathcal{R}_{++}$, and that σ is chosen with logit probability, namely

$$p_{\text{logit}}^{(u,\lambda)}(\sigma) = \frac{\exp\left(\frac{u_\sigma}{\lambda}\right)}{\sum_{\tau \in \{xy,x,y,\varnothing\}} \exp\left(\frac{u_\tau}{\lambda}\right)} \tag{1}$$

where $\lambda > 0$ is a scaling factor (measuring the variance of the underlying Gumbel errors, see McFadden [12]). In this specification, purely random behaviour (i.e. uniform distribution on $\{xy, x, y, \varnothing\}$) is obtained in the limit as λ tends to infinity and rational deterministic behaviour is obtained for $\lambda = 0$.

In the *stochastic consideration set model* the agent has a preference relation $\succ$ on $\{xy, x, y\}$. The agent considers each nonempty bundle σ with a probability $\alpha \in (0, 1)$ independent of σ. The agent chooses a bundle by maximising $\succ$ on the set of bundles that he considers. In the event that the agent does not consider any bundle, the agent picks the empty bundle. Note that unlike in the multinomial logit case we are only given ordinal preference information.[5] Therefore σ is chosen with probability

$$p_{\text{cons}}^{(\succ,\alpha)}(\sigma) = \begin{cases} \alpha(1-\alpha)^{\beta(\sigma)} & \text{if } \sigma \in \{xy, x, y\} \\ (1-\alpha)^3 & \text{if } \sigma = \varnothing \end{cases} \tag{2}$$

where

$$\beta(\sigma) = |\{\tau \in \{xy, x, y\} : \tau \succ \sigma\}|.$$

[5]In Manzini and Mariotti [11] and Brady and Rehbeck [3] the consideration coefficients depend, respectively, on the individual alternatives and on the menu.

Then in the logit model the correlation criterion yields:

$$
\begin{aligned}
p_{\text{logit}}^{(u,\lambda)} &\in C \\
&\Leftrightarrow \frac{\dfrac{e^{u_{xy}/\lambda}}{e^{u_{xy}/\lambda}+e^{u_x/\lambda}+e^{u_y/\lambda}+e^{u_\varnothing/\lambda}}}{\dfrac{e^{u_{xy}/\lambda}+e^{u_x/\lambda}}{e^{u_{xy}/\lambda}+e^{u_x/\lambda}+e^{u_y/\lambda}+e^{u_\varnothing/\lambda}}} > \frac{e^{u_{xy}/\lambda}+e^{u_y/\lambda}}{e^{u_{xy}/\lambda}+e^{u_x/\lambda}+e^{u_y/\lambda}+e^{u_\varnothing/\lambda}} \\
&\Leftrightarrow e^{u_{xy}/\lambda}e^{u_\varnothing/\lambda} > e^{u_x/\lambda}e^{u_y/\lambda} \\
&\Leftrightarrow u_{xy}+u_\varnothing > u_x+u_y
\end{aligned}
$$

and similarly

$$p_{\text{logit}}^{(u,\lambda)} \in S \Leftrightarrow u_{xy}+u_\varnothing < u_x+u_y$$

That is:

- in the two-good logit model x and y are complementary according to the correlation criterion if and only if the systematic utility u is *strictly supermodular* on Σ, and substitutes if and only if u is strictly submodular.

Remarkably, this holds *independently of the scaling factor* λ. As we shall see, this scale independence property is lost as soon as we consider a multi-good case. Note also that this complementarity condition is *not* invariant to monotonic transformations of utility.

Turning to the stochastic consideration set model, simple calculations show the following:[6] provided that α is greater than a threshold value $\alpha^* \in (0,1)$,

$$
\begin{aligned}
p_{\text{cons}}^{(\succ,\alpha)} &\in S \Leftrightarrow x \succ y \succ xy \text{ or } y \succ x \succ xy \\
p_{\text{cons}}^{(\succ,\alpha)} &\in C \text{ for any other preference ordering}
\end{aligned}
$$

That is:

[6] For example, if $x \succ y \succ xy$, then x and y are complementary with the correlation criterion iff

$$\frac{(1-\alpha)^2\alpha}{(1-\alpha)^2\alpha+\alpha} > (1-\alpha)^2\alpha + (1-\alpha)\alpha$$

Thus it is easy to check that there exists a unique $\alpha^* \in (0,1)$ such that x and y are complementary for $\alpha < \alpha^*$, substitutes for $\alpha > \alpha^*$ and independent for $\alpha = \alpha^*$.

- In the two good stochastic consideration set model and for α sufficiently high, x and y are complementary according to the correlation criterion if and only if joint consumption is not at the bottom of the preference ordering, and they are substitutes otherwise.

Since the model uses ordinal information on preferences, unlike in the logit case, we have obtained a purely ordinal preference condition for complementarity/substitutability.

However, the two-good case is in some respects very specific. We now proceed to consider an extension of the correlation criterion to the multi-good case, and show that then complementarity in both models can be expressed as a supermodularity condition on preferences, one cardinal an the other ordinal.

5.1 The multi-good case

Let $X = \{x, y, ...\}$ be a finite set of goods, and let Σ be the power set of X. We are interested as usual in the complementarity between x and y. A datapoint is a probability distribution p on Σ, with $p(\sigma)$ denoting the probability of bundle $\sigma \in \Sigma$. Define the following sets:

$$\begin{aligned} XY &= \{\sigma \in \Sigma : x \in \sigma, y \in \sigma\} \\ \bar{X}\bar{Y} &= \{\sigma \in \Sigma : x \notin \sigma, y \notin \sigma\} \\ \bar{X}Y &= \{\sigma \in \Sigma : x \notin \sigma, y \in \sigma\} \\ X\bar{Y} &= \{\sigma \in \Sigma : x \in \sigma, y \notin \sigma\} \end{aligned}$$

We study the following generalised correlation criterion:

$$\begin{aligned} p \in C &\Leftrightarrow \frac{\sum_{\sigma \in XY} p(\sigma)}{\sum_{\sigma \in XY \cup X\bar{Y}} p(\sigma)} > \sum_{\sigma \in XY \cup \bar{X}Y} p(\sigma) \\ &\Leftrightarrow \sum_{\sigma \in XY} p(\sigma) \left(1 - \sum_{\sigma \in XY} p(\sigma) - \sum_{\sigma \in \bar{X}Y} p(\sigma) - \sum_{\sigma \in X\bar{Y}} p(\sigma)\right) > \left(\sum_{\sigma \in X\bar{Y}} p(\sigma) \sum_{\sigma \in \bar{X}Y} p(\sigma)\right) \\ &\Leftrightarrow \sum_{\sigma \in XY} p(\sigma) \sum_{\sigma \in \bar{X}\bar{Y}} p(\sigma) > \sum_{\sigma \in X\bar{Y}} p(\sigma) \sum_{\sigma \in \bar{X}Y} p(\sigma) \end{aligned}$$

and similarly

$$p \in S \Leftrightarrow \sum_{\sigma \in XY} p(\sigma) \sum_{\sigma \in \bar{X}\bar{Y}} p(\sigma) < \sum_{\sigma \in X\bar{Y}} p(\sigma) \sum_{\sigma \in \bar{X}Y} p(\sigma)$$

In this definition, 'joint consumption' of x and y is implicitly interpreted as meaning all instances of bundles where both x and y are consumed, and single consumption of x (resp., y) is interpreted as all instances of bundles where x (resp., y) is consumed but not y (resp., x). Obviously, other interpretations are possible. It could be the case, for example, that x is consumed together with z but not y because x is complementary to z but y is not. In this case the consumption of x without y would not express in a clean way the fact that x and y are substitutes. For example, you may consume 50% of the time coffee and milk and 50% of the time tea and milk, so that according to the criteria we have studied coffee would be independent from milk, while in an intuitive sense they are complementary. Here we ignore for simplicity this problem.[7]

5.1.1 Logit

In the logit model (1) generalises to

$$p_{\text{logit}}^{(u,\lambda)}(\sigma) = \frac{\exp\left(\frac{u_\sigma}{\lambda}\right)}{\sum_{\tau\in\Sigma}\exp\left(\frac{u_\tau}{\lambda}\right)} \text{ for } \sigma \in \Sigma$$

with $u : \Sigma \to \mathcal{R}_{++}$, so that with the correlation criterion we have

$$\begin{aligned} p_{\text{logit}}^{(u,\lambda)} &\in C \Leftrightarrow \sum_{\sigma\in XY, \tau\in\bar{X}\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right) > \sum_{\sigma\in\bar{X}Y, \tau\in X\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right) \\ p_{\text{logit}}^{(u,\lambda)} &\in S \Leftrightarrow \sum_{\sigma\in XY, \tau\in\bar{X}\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right) < \sum_{\sigma\in\bar{X}Y, \tau\in X\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right) \end{aligned}$$

Unlike in the two-good case, the scaling factor λ now becomes important to assess complementarity. We study the limiting behaviour of the complementarity conditions for λ tending to zero, which captures the case of small errors with respect to utility maximisa-

[7] As Samuleson [15] discusses, similar conceptual problems in the multi-good case arise also for the standard elasticity-based definitions. If milk is complementary to coffee but it is even more complementary to tea, a rise in the price of coffee, leading to a substitution of tea for coffee, will also generate an increase in the consumption of milk, making milk look like a substitute of coffee.

tion.[8]

Theorem 6 *In the limit for* $\lambda \to 0$*, according to the correlation criterion*

$$\begin{aligned} p_{logit}^{(u,\lambda)} &\in C \Leftrightarrow \max_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} (u_\sigma + u_\tau) > \max_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} (u_\sigma + u_\tau) \\ p_{logit}^{(u,\lambda)} &\in S \Leftrightarrow \max_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} (u_\sigma + u_\tau) < \max_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} (u_\sigma + u_\tau) \end{aligned}$$

Proof: Note first that, for all $\lambda > 0$,

$$\begin{aligned} \sum_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right) &> \sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right) \Leftrightarrow \\ \lambda \ln\left(\sum_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right)\right) &> \lambda \ln\left(\sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right)\right) \end{aligned}$$

Next, define

$$k(\lambda) = \max_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \frac{u_\sigma + u_\tau}{\lambda}$$

[8]In the case of large errors (behaviour is almost purely random) it is easy to see that the goods are always approximately independent according to the correlation criterion:

$$\begin{aligned} \lim_{\lambda\to\infty}\left(\sum_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right)\right) &= |XY|\cdot|\bar{X}\bar{Y}| \\ &= 2^{n-2}\cdot 2^{n-2} \\ &= |\bar{X}Y|\cdot|X\bar{Y}| \\ &= \lim_{\lambda\to\infty}\left(\sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right)\right) \end{aligned}$$

so that neither the complementarity nor the substitutability condition can hold in the limit.

Then we have:

$$\begin{aligned}
& \lim_{\lambda\to 0}\left(\lambda \ln\left(\sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right)\right)\right) \\
= & \lim_{\lambda\to 0}\left(\lambda \ln\left(\sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma+u_\tau}{\lambda}\right)\frac{\exp k(\lambda)}{\exp k(\lambda)}\right)\right) \\
= & \lim_{\lambda\to 0}\left(\lambda k(\lambda) + \lambda \ln \sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma+u_\tau}{\lambda} - k(\lambda)\right)\right) \\
= & \max_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} (u_\sigma+u_\tau) + \lim_{\lambda\to 0}\left(\lambda \ln \sum_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} \exp\left(\frac{u_\sigma+u_\tau}{\lambda} - k(\lambda)\right)\right) \\
= & \max_{\sigma\in\bar{X}Y,\tau\in X\bar{Y}} (u_\sigma+u_\tau)
\end{aligned}$$

To see that the last equality holds, note that each term in the summation is of the form $\exp\left(\frac{A}{\lambda}\right)$, where $A \leq 0$, and thus is either constant and equal to one or tends to zero as λ tends to zero. Moreover, at least one term is equal to one, so that the logarithm of the sum remains finite in the limit.

An analogous calculation yields

$$\lim_{\lambda\to 0}\left(\lambda \ln\left(\sum_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} \exp\left(\frac{u_\sigma}{\lambda}\right)\exp\left(\frac{u_\tau}{\lambda}\right)\right)\right) = \max_{\sigma\in XY,\tau\in\bar{X}\bar{Y}} (u_\sigma+u_\tau)$$

from which the result follows. ■

This result generalises the supermodularity condition we found in the two-good case. In the multi-good case the highest utility elements are taken as 'representatives' of the various classes to be used in the two-good supermodularity formula.

5.1.2 Consideration sets

Turning to the stochastic consideration set model, the probability of choosing bundle $\sigma \in \Sigma$ is defined as

$$p_{\text{cons}}^{(\succ,\alpha)}(\sigma) = \begin{cases} \alpha(1-\alpha)^{\beta(\sigma)} & \text{if } \sigma\in\Sigma\backslash\varnothing \\ (1-\alpha)^{|\Sigma\backslash\varnothing|} & \text{if } \sigma=\varnothing \end{cases} \tag{3}$$

where $\succ$ is defined on $\Sigma\backslash\varnothing$ and for all $\sigma \in \Sigma\backslash\varnothing$

$$\beta(\sigma) = |\{\tau \in \Sigma\backslash\varnothing : \tau \succ \sigma\}|$$

Therefore with the correlation criterion, after simplifying:

$$p_{\text{cons}}^{(\succ,\alpha)} \in C \Leftrightarrow \frac{\sum_{\sigma\in XY}(1-\alpha)^{\beta(\sigma)}}{\sum_{\sigma\in XY\cup X\bar{Y}}\alpha(1-\alpha)^{\beta(\sigma)}} > \sum_{\sigma\in XY\cup\bar{X}Y}(1-\alpha)^{\beta(\sigma)}$$

$$p_{\text{cons}}^{(\succ,\alpha)} \in S \Leftrightarrow \frac{\sum_{\sigma\in XY}(1-\alpha)^{\beta(\sigma)}}{\sum_{\sigma\in XY\cup X\bar{Y}}\alpha(1-\alpha)^{\beta(\sigma)}} < \sum_{\sigma\in XY\cup\bar{X}Y}(1-\alpha)^{\beta(\sigma)}$$

We study the limiting case for α tending to one: this expresses small deviations from the 'rationality' case in which all alternatives are considered.

For $\sigma \in \Sigma\backslash\varnothing$, define the function $u_\succ$ given by

$$u_\succ(\sigma) = -\beta(\sigma),$$

which is a representation of the preference $\succ$.

Theorem 7 *In the limit for $\alpha \to 1$, according to the correlation criterion*

$$p_{cons}^{(\succ,\alpha)} \in C \Leftrightarrow \max_{\sigma\in XY,\sigma'\in\bar{X}\bar{Y}}(u_\succ(\sigma)+u_\succ(\sigma')) > \max_{\sigma\in X\bar{Y},\sigma'\in\bar{X}Y}(u_\succ(\sigma)+u_\succ(\sigma'))$$

$$p_{cons}^{(\succ,\alpha)} \in S \Leftrightarrow \max_{\sigma\in XY,\sigma'\in\bar{X}\bar{Y}}(u_\succ(\sigma)+u_\succ(\sigma')) < \max_{\sigma\in X\bar{Y},\sigma'\in\bar{X}Y}(u_\succ(\sigma)+u_\succ(\sigma'))$$

Proof: Fixing the preference $\succ$, let σ^* denote the best bundle. Note that all terms $\alpha(1-\alpha)^{\beta(\sigma)}$ tend to zero as $\alpha \to 1$ except the one corresponding to σ^* (since $\beta(\sigma^*) = 0$). We consider four cases, depending on whether σ^* is in XY, $\bar{X}Y$, $\bar{X}\bar{Y}$ or $X\bar{Y}$.

Simple calculations[9] show that the complementarity condition can be written as

$$\sum_{\sigma\in XY} p(\sigma) > \frac{\sum_{\sigma\in X\bar{Y}} p(\sigma) \sum_{\sigma\in \bar{X}Y} p(\sigma)}{\sum_{\sigma\in \bar{X}\bar{Y}} p(\sigma)} \tag{4}$$

Suppose $\sigma^* \in XY$. The LHS in (4) tends to one as $\alpha \to 1$ (all terms in the sum tend to zero except $p(\sigma^*)$). In the RHS all terms in the sums tend to zero, so that the limit of the RHS depends on a comparison between the minimum powers of $(1-\alpha)$ that appear in the numerator and in the the denominator, respectively. More precisely, the RHS can be written as

$$\frac{\sum_{\sigma\in X\bar{Y},\sigma'\in \bar{X}Y} \alpha(1-\alpha)^{\beta(\sigma)} \alpha(1-\alpha)^{\beta(\sigma')}}{\sum_{\sigma\in \bar{X}\bar{Y}\backslash\varnothing} \alpha(1-\alpha)^{\beta(\sigma)} + (1-\alpha)^{|\Sigma\backslash\varnothing|}} = \frac{\alpha^2 \sum_{\sigma\in X\bar{Y},\sigma'\in \bar{X}Y} (1-\alpha)^{\beta(\sigma)+\beta(\sigma')}}{\alpha\sum_{\sigma\in \bar{X}\bar{Y}\backslash\varnothing} (1-\alpha)^{\beta(\sigma)} + (1-\alpha)^{|\Sigma\backslash\varnothing|}}$$

So, given that $|\Sigma\backslash\varnothing| > \beta(\sigma)$ for all $\sigma \in \Sigma\backslash\varnothing$, if

$$\min_{\sigma\in X\bar{Y},\sigma'\in \bar{X}Y} (\beta(\sigma)+\beta(\sigma')) > \min_{\sigma\in \bar{X}\bar{Y}\backslash\varnothing} \beta(\sigma)$$

then the RHS tends to zero and therefore the complementarity condition holds in the limit, whereas if the reverse inequality holds then the RHS tends to infinity and the goods are substitutes for α large enough. In summary:

Fact 1: Let $\sigma^* \in XY$. Then, for $\alpha \to 1$, $p_{\text{cons}}^{(\succ,\alpha)} \in C$ ($\in S$) according to the correlation criterion if there are strictly fewer (strictly more) than $\min_{\sigma\in X\bar{Y},\sigma'\in \bar{X}Y} (\beta(\sigma)+\beta(\sigma'))$ bundles that are preferred to the best bundle in $\bar{X}\bar{Y}$.

A similar analysis solves the other cases:

---

[9] To see this observe the following:

$$\begin{aligned}
\frac{\sum_{\sigma\in XY} p(\sigma)}{\sum_{\sigma\in XY\cup X\bar{Y}} p(\sigma)} &> \sum_{\sigma\in XY\cup\bar{X}Y} p(\sigma) \Leftrightarrow \sum_{\sigma\in XY} p(\sigma) > \sum_{\sigma\in XY\cup X\bar{Y}} p(\sigma) \sum_{\sigma\in XY\cup\bar{X}Y} p(\sigma) = \\
&= \sum_{\sigma\in XY} (p(\sigma))^2 + \sum_{\sigma\in XY} p(\sigma) \sum_{\sigma\in\bar{X}Y} p(\sigma) + \sum_{\sigma\in X\bar{Y}} p(\sigma) \sum_{\sigma\in XY} p(\sigma) + \sum_{\sigma\in X\bar{Y}} p(\sigma) \sum_{\sigma\in\bar{X}Y} p(\sigma) \\
&\Leftrightarrow \sum_{\sigma\in XY} p(\sigma) \left(1 - \sum_{\sigma\in XY} p(\sigma) - \sum_{\sigma\in\bar{X}Y} p(\sigma) - \sum_{\sigma\in X\bar{Y}} p(\sigma)\right) > \sum_{\sigma\in X\bar{Y}} p(\sigma) \sum_{\sigma\in\bar{X}Y} p(\sigma) \\
&\Leftrightarrow \sum_{\sigma\in XY} p(\sigma) \sum_{\sigma\in\bar{X}\bar{Y}} p(\sigma) > \sum_{\sigma\in X\bar{Y}} p(\sigma) \sum_{\sigma\in\bar{X}Y} p(\sigma) \Leftrightarrow \sum_{\sigma\in XY} p(\sigma) > \frac{\sum_{\sigma\in X\bar{Y}} p(\sigma) \sum_{\sigma\in\bar{X}Y} p(\sigma)}{\sum_{\sigma\in\bar{X}\bar{Y}} p(\sigma)}
\end{aligned}$$

Fact 2: Let $\sigma^* \in \bar{X}\bar{Y}$. Then, for $\alpha \to 1$, $p_{\text{cons}}^{(\succ,\alpha)} \in C$ $(\in S)$ according to the correlation criterion if there are strictly fewer (strictly more) than $\min_{\sigma\in X\bar{Y},\sigma'\in\bar{X}Y}\left(\beta(\sigma)+\beta(\sigma')\right)$ bundles that are preferred to the best bundle in XY.

Fact 3: Let $\sigma^* \in \bar{X}Y$. Then, for $\alpha \to 1$, $p_{\text{cons}}^{(\succ,\alpha)} \in C$ $(\in S)$ according to the correlation criterion if there are strictly fewer (strictly more) than $\min_{\sigma\in XY,\sigma'\in\bar{X}\bar{Y}}\left(\beta(\sigma)+\beta(\sigma')\right)$ bundles that are preferred to the best bundle in $X\bar{Y}$.

Fact 4: Let $\sigma^* \in \bar{X}Y$. Then, for $\alpha \to 1$, $p_{\text{cons}}^{(\succ,\alpha)} \in C$ $(\in S)$ according to the correlation criterion if there are strictly fewer (strictly more) than $\min_{\sigma\in XY,\sigma'\in\bar{X}\bar{Y}}\left(\beta(\sigma)+\beta(\sigma')\right)$ bundles that are preferred to the best bundle in $\bar{X}Y$.

Now we can summarise these facts into a single statement by noting that, by definition, if $R \in \left\{XY, \bar{X}Y, \bar{X}\bar{Y}, X\bar{Y}\right\}$ is such that $\sigma^* \in R$, then

$$\min_{\sigma\in R}\beta(\sigma) = 0$$

Therefore by inspection of the four conditions we can conclude that, in the limit for $\alpha \to 1$, according to the correlation criterion

$$\begin{aligned}
p_{\text{cons}}^{(\succ,\alpha)} &\in C \Leftrightarrow \min_{\sigma\in XY,\sigma'\in\bar{X}\bar{Y}}\left(\beta(\sigma)+\beta(\sigma')\right) < \min_{\sigma\in X\bar{Y},\sigma'\in\bar{X}Y}\left(\beta(\sigma)+\beta(\sigma')\right) \\
p_{\text{cons}}^{(\succ,\alpha)} &\in S \Leftrightarrow \min_{\sigma\in XY,\sigma'\in\bar{X}\bar{Y}}\left(\beta(\sigma)+\beta(\sigma')\right) > \min_{\sigma\in X\bar{Y},\sigma'\in\bar{X}Y}\left(\beta(\sigma)+\beta(\sigma')\right)
\end{aligned}$$

which proves the result. ■

Strikingly, the condition in the statement is an exact ordinal analog of the condition obtained for the 'rationality limit' of the logit case, in which the utility of a bundle σ is measured by (the opposite of) the number of bundles better than σ. One way of understanding this analogy is to think that both models are special cases of the RUM family, and that both models are based on an underlying preference. When the parameters of these models converge to 'rationality', the deterministic preference effect (as opposed to the stochastic effect) dominates. The correlation definition of complementarity captures the complementarity information contained in these 'unbiased' preference. However, in the multinomial logit case, preference is cardinal, while in the mood model preference is ordinal: the numerical analogy between the conditions of the two models holds for one utility representation of $\succ$ but it may not hold for other, ordinally equivalent, representations.

6 Related literature

While referring to Samuelson [15] for an erudite discussion of the literature on complementarity up to the 70's, we mention here some notable more recent works. It is a surprising fact that the full behavioural implications of the classical definitions of complementarity and substitutability, based on cross price elasticities, have only recently been uncovered, in two papers by Chambers, Echenique and Shmaya ([5] and [6]). The key difference between their work and ours is that their hypothetical data include observations of consumption decisions for different prices (as the classical definition requires), whereas ours are based on consumption decisions alone.

The work by Gentzkow [8] we have already mentioned asks the question of whether online and print versions of a newspaper are complements or substitutes. Within a random utility model, he analyses the identification of complementarity (as opposed to taste correlation) in the data by using exogenous variations in factors that do not interact with preferences. This requires the development of an innovative econometric identification technique which, however, is meaningful only within the random utility model. Our approach, in contrast, is to investigate whether complementarity or substitution can be identified in a model-free fashion.

A large literature exists in which supermodularity of a utility function gives, by definition, a complementarity relationship between the goods, and likewise submodularity is equivalent to substitutability (see for example Bikhchandani and Mamer [2] and Gül and Stachetti [10] for applications of submodularity and related notions of substitutability in general equilibrium models with indivisibilities.) These definitions are cardinal. Complementarity in the form of supermodularity is also the bread and butter of modern monotone comparative static techniques as surveyed by Topkis [16]. Our approach is a complement to this line of work, in that our axiomatic analysis starts with the data, rather than with the underlying preference.

7 Conclusions

While correlation in consumption or usage data is usually taken as a behavioural indicator of complementarity, we have shown that, in general, criteria for complementarity based on statistical association conflict with a basic monotonicity requirement. Our axiomatic analysis suggests that if monotonicity is considered important, then different criteria (additivity and maxmin) may be preferable.

We have illustrated that the theoretical distinction between criteria is also relevant in practice, since correlation, additivity and maxmin give strongly contrasting indications using the data found in a leading application (Gentzkow [8]).

The conclusions from the study of two specific RUM models are intriguing. In that context, we have shown that identifying complementarity with correlation is equivalent, in terms of primitives, to identifying complementarity with a form of supermodularity of the utility function. Because supermodularity is a natural and accepted criterion for complementarity when expressed in terms of utility, this might be taken as a validation of the correlation criterion, at least for the case where the mechanism of choice is described by one of the models examined. But our findings also have an alternative interpretation. They could be taken as suggesting that, for the case of random choice over bundles, supermodularity in utility in itself is a poor or at least insufficient descriptor of complementarity. One reason for this interpretation in, for example, the logit model is the following. The 'cross-partial' of utility $(u_{xy} - u_x) - (u_y - u_\varnothing)$ may be positive while at the same time $u_\varnothing > u_{xy}$. In this case, correlation in the data must be driven by the *lack* of joint consumption (that is, $p_\varnothing$ is high) rather than by joint consumption, which is unintuitive. Consider again the data in table 1. There, the proportion of people who read both versions is only about one third of the proportion of people that did not read either version. The positive correlation is mostly driven by the latter proportion. In fact, the additivity criterion picks this up and declares the goods substitutes rather than complements. Also, a positive cross-partial of utility is compatible with $u_{xy} < \max\{u_x, u_y\}$. Then once again positive correlation can only be driven by a high proportion of non-consumers. In the table, the proportion of people who read both versions is less than one third of the proportion of people that read only the print version. The maxmin criterion picks this up by failing to declare complementarity.

References

[1] Armstrong, M. (2013) "A More General Theory of Commodity Bundling", *Journal of Economic Theory* 148(2): 448-72.

[2] Bikhchandani, Sushil and John W. Mamer (1997) "Competitive Equilibrium in an Exchange Economy with Indivisibilities", *Journal of Economic Theory* 75: 385-413.

[3] Brady, Richard and John Rehbeck (2014) "Menu-Dependent Stochastic Consideration", mimeo, University of California at San Diego

[4] Brynjolfson, E. and P. Milgrom (2013) "Complementarity in Organizations", ch. 1 in R. Gibbons and J. Roberts *Handbook of Industrial Organization*, Princeton University Press.

[5] Chambers, Chris, Echenique, Federico and Eran Shmaya (2010) "Testable Implications of Gross Substitutes in Demand for Two Goods", *American Economic Journal: Microeconomics*, 3 (1): 129-36.

[6] Chambers, Chris, Echenique, Federico and Eran Shmaya (2011) "On Behavioral Complementarity and its Implications", *Journal of Economic Theory*, 145 (6): 2332-2355.

[7] Echenique, Federico, Kota Saito and Gerelt Tserenjigmid (2013) "The Perception-Adjusted Luce Model", mimeo, California Institute of Techonology.

[8] Gentzkow, M. (2007) "Valuing New Goods in a Model with Complementarity: Online Newspapers", *American Economic Review*: 713-44.

[9] Gül, Faruk, Paulo Natenzon and Wolfgang Pesendorfer (2014) "Random Choice as Behavioral Optimization", *Econometrica*, 82 (5): 1873– 1912.

[10] Gül, Faruk and Ennio Stachetti (1999) "Walrasian Equilibrium with Gross Substitutes", *Journal of Economic Theory*, 87 (1): 95–124.

[11] Manzini, Paola, and Marco Mariotti (2014) "Stochastic Choice and Consideration Sets", *Econometrica*, 82 (3): 1153-76.

[12] McFadden, D.L. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka (ed.), Frontiers in Econometrics, 105-142, Academic Press: New York.

[13] Milgrom, P. and J. Roberts (1992) *Economics, Organizations and Management*, Upper Saddle River, NJ, Prentice Hall.

[14] Milgrom, P. and C. Shannon (1994) "Monotonone Comparative Statics", *Econometrica* 62: 157-180.

[15] Samuelson, Paul A. (1974) "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory", *Journal of Economic Literature* 12(4): 1255-89.

[16] Topkis, Donald M. (1998) *Supermodularity and Complementarity*, Princeton University Press, Princeton.