Revealed Complementarity*

Paola Manzini[†] Marco Mariotti[‡] Levent Ülkü[§]

This version: November 2015

Abstract

The Hicksian definition of complementarity and substitutability may not apply in contexts in which agents are not utility maximisers or where prices, whether implicit or explicit, are not available. We look for tools to identify complementarity and substitutability satisfying the following criteria: they are *behavioural* (based only on observable choice data); *model-free* (valid whether the agent is rational or not); and they do not rely on price variation. We uncover a conflict between properties that any complementarity notion should intuitively possess. We discuss three different possible resolutions of the conflict.

J.E.L. codes: D0.

Keywords: Complements and substitutes; Correlation; Stochastic choice.

*We thank Miguel Angel Ballester, Jean-Pierre Dubé, Matthew Gentzkow, Faruk Gül, Patrick Harless, Rich McLean, Luigi Pistaferri, Tomas Sjöström and seminar audiences at Birkbeck College, ECARES, LSE, the Universitity of Manchester and Rutgers University for very useful comments. The usual disclaimers apply. A previous version was circulated with the title 'Stochastic Complementarity'.

[†]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

1.1 Motivation

In this paper we study a notion of complementarity (and substitutability) between goods that is:

(1) not based on price variations;

(2) behavioural: it just uses choice data as inputs;

(3) model free: it does not commit to any specific cognitive procedure underlying the choice data.

To motivate this program, suppose that 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 to hear that this situation does not fit directly the textbook criteria for complementarity, which are based on price elasticities while in this case both prices are fixed at zero. The authority may also be surprised to hear that the gold standard concept of Hicksian complementarity is based on the assumption that agents are utility maximisers.[1]

Of course, any well-bred economist could explain to the authority that it is always possible to retrieve 'implicit' prices. To illustrate, one could check the impact of having the park located at different distances from the museum, observe the resulting effect on the demand for the museum, impute a value to the time needed to move between the two attractions, and then calculate an elasticity based on such values. Because it may be somewhat impractical to experiment with moving the park at various distances, the impact would be estimated from the observations of analogous effects relating to other museums and other parks, and these effects would be adjusted for the various factors affecting demand in the different locations. But, while this standard type of methodology

[1] It merits noting in this respect that goverments are more and more advised by 'behavioural economics' units. These include the US Office of Information and Regulatory Affairs, which led to the institution of the US government's Social and Behavioral Sciences Team in July 2015; the Behaviour Insights team in the UK, established in 2010 and then spun off into a separate company in 2014; the Behavioural Insights Unit established in 2012 as part of the New South Wales Premier and Cabinet's office. The German Chancellery is in the process of setting up a similar unit. Several other governmental units and groups exist in various countries and in the European Union Directorates General.

has merits[2], it is fair to say that it is very indirect and thus its validity necessarily rests on several assumptions (regarding the imputed value of time, the comparisons across locations, the specification of the model that generates demand, etcetera). And the fact remains that Hicksian compensated elasticities are meaningful only within a utility maximisation framework. The recent econometric approaches that deal with the zero price problem, pioneered by Gentzkow [14] typically assume an additive Random Utility Model (RUM) such as the multinomial logit and variations thereof (see section 2 for a discussion). However, the recent wave of abstract works on stochastic choice (e.g. Brady and Rhebeck [9], Echenique, Saito and Tserenjigmid [13], Gül, Natenzon and Pesendorfer [15], Manzini and Mariotti [17] among others) has highlighted a wide variety of possible 'choice errors' and choice procedures, and so a number of reasons why agents' behaviour might fail to be described by a logit model, and indeed even by the much larger class of RUMs.

We take the view that price variations are just a *tool* to check complementarity: the notion of complementarity itself is not *intrinsically* related to price variations. Furthermore, we also take the view that complementarity is meaningful independently of whether people maximise utility or not. We take seriously the multiplicity of plausible models and the consequent difficulty of model selection. For these reasons we look for definitions that complement those of the standard approach.

The scope of interest of this program is quite vast, firstly because consumption may be 'non-rational' in many different ways, and secondly because zero prices are observed for many other goods beside public attractions: online newspapers, reviews/advice (e.g. financial) on social networks, public radio broadcasts, file sharing 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 [10]). More abstractly, the 'goods' may be characteristics embodied in the objects of choice, so that any price variation is perfectly correlated between the goods. In some cases, prices may be especially difficult to conceptualise: is beauty a complement or a substitute of wealth in a partner?

[2] A neat early example of the methodology is Becker and Murphy's [4] analysis of the complementarity between advertising and advertised goods based on the implicit price of commercials, which are shown on television without a price. If networks stopped showing commercials, the public would have to pay for television content. The change in the price of content is the implicit price of commercials.

1.2 The basic ideas

Complementarity between goods means intuitively that they 'go together'. How can this concept be made operational?

The first idea in our analysis is to use *stochastic* choice data as a primitive. Because we are not going to exploit responses to price variation, and more in general the variation of choices across menus, we lose some information compared to the classical approach. To obviate this, we consider instead a multiplicity of choices from a fixed menu, in the form of choice frequencies. The rich structure of this type of data is an alternative source of information about underlying complementarities.

Secondly, we suggest that two basic principles should be examined:

a) *Statistical Association*: if choice data come in the form of frequencies then the consumption of complementary goods should exhibit some kind of statistical association. Association is precisely what it means to 'go together' in statistical language. Indeed, in the literature about complementarities in business practices, the positive *correlation* (or clustering) of practices is the most common complementarity test[3].

b) '*Revealed preference':* Consuming the goods jointly should not in itself be evidence of substitutability; consuming the goods individually should not in itself be evidence of complementarity. This principle seem almost self-evident from an economic point of view if choices must convey information about complementarity. It is also fundamental from a welfare perspective. Interest in complementarity is often a consequence of a welfare question, such as "is it welfare enhancing to build a park next to the existing museum?" or "would introducing a new product be welfare enhancing for the consumers (so that he would be willing to pay more)?". We take the classical view that an agent's choices encode welfare information, and that, as articulated by Bernheim and Rangel [6], they do so whether the agent is rational or not. It would be odd if the local authority, having decided that the construction of the park is in the interest of the community, was dissuaded after learning that the joint consumption of park and museum in a similar location has increased.

[3] See e.g. Brynjolfson and Milgrom [10], p.33.

1.3 Preview of results

Consider two goods, say the online and the print versions of a newspaper. As in Gentzkow [14], the data come 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.

Let's consider the statistical principle discussed above. For simplicity, let's also focus in this section on association as correlation (this is just for concreteness: later on we will use a much more abstract concept). Then we would say the two versions are complementary 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, $p_{OP} + p_O$.

It is easily shown, however, that this property flatly contradicts the intuitive revealed preference view that increases in joint consumption do not constitute evidence of substitutability. 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 conflict between two seemingly natural properties that criteria for complementarity should satisfy. One is *monotonicity*, embodying the revealed preference principle: an increase in joint consumption accompanied by (weak) decreases in single good consumption should not overturn an existing complementarity (and analogously for substitutability).

The second property 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 arguably the 'soul' of statistical association, as it is evidently satisfied by all common measures of correlation and association. It is intrinsic to the nature of association that 'inverting' behaviour changes the sign of the association.

The conflict in general is a little more subtle than in the simple example above. Even for symmetric concepts of complementarity for which complementarity is always bidirectional (i.e., if one good complements another then vice versa) the two properties do not flatly contradict each other. However, theorem 1 and its corollary 1 show that any symmetric concept of complementarity that satisfies monotonicity and duality must be also *unresponsive*, in the sense that the level of non-consumption $p_{\varnothing}$ *on its own* determines whether the goods are complements or substitutes, irrespective of the distribution between single and joint consumption. This is a very undesirable feature, and for this reason we interpret the result as one of conflict between the the statistical association and the revealed preference principles. Furthermore, 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 criteria.

We then look for ways out of the impossibility (Section 5). We first show that correlation is the only symmetric criterion of complementarity that satisfies both duality and a modified monotonicity condition, which embodies only particular aspects of the revealed preference principle.

Next, we examine monotonic criteria that satisfy modified notions of duality, based on alternative interpretations of what constitutes the 'opposite' of a given behaviour. One criterion 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 duality property based on exchanging joint consumption with total single good consumption.

The third criterion for complementarity to be considered 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 a 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 model-free stochastic complementarity.[4]

[4]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 three.

Having defined complementarity in these ways, in Section 6 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 basic Luce (or multinomial logit) model of stochastic choice. The second is (a particular version of) the more recent 'stochastic consideration set' model of Manzini and Mariotti [17] and Brady and Rehbeck [9]. We discover that in both cases the correlation criterion on data reflects novel supermodularity types of condition on preferences.

In the concluding discussion we argue the Hicksian complementarity criterion is ill-suited for the type of data we are considering.

2 Related literature

Samuelson [21] contains an erudite discussion of the subtleties of the concept of complementarity, with an exhaustive review of the classical literature.

The work by Gentzkow [14] we have already mentioned pioneers the approach to the zero price problem. He asks the question of whether the online and print versions of a newspaper are complements or substitutes. The main difficulty to be solved in this case is that the observed correlation in consumption 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 (random) utility maximisation* means a positive difference between the value of joint consumption and the sum of the values of single good consumptions.[5] Gentzkow finds sufficient conditions under which correlation in choice data is indicative of Hicksian complementarity, and 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

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

within the random utility model. Our approach, in contrast, is to investigate whether complementarity or substitution can be identified in a model-free fashion. Our agents may not even 'have' a utility function.

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 ([11] and [12]). 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.

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 [7], Gül and Stachetti [16] for applications of submodularity and related notions of substitutability in general equilibrium models with indivisibilities.[6]) 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 [22]. Our approach differs from this line of work, in that our axiomatic analysis starts with the data, rather than with the underlying preference.

3 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 set

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

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

[6]See Baldwin and Klemperer [3] for an innovative approach (based on tools from tropical geometry) that yields complements/substitutes types of conditions for the existence of equlibria with discrete goods.

(resp. substitutes, resp. independent) at p. We call any such partition a *criterion*.

Here are some examples of criteria:

Example 1 *(correlation):*

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

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 [2]).

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)*

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

Here the goods are declared 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 [22], Milgrom and Roberts [19], Milgrom and Shannon [20]). Note that the condition is equivalent to $p_{xy} + p_\varnothing > \frac{1}{2}$.

For illustration, consider table 1, calculated on the basis of Gentzkow's [14] 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 [14]).

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, examining the properties that criteria should possess given the interpretation.

4 Impossibilities

In this section we uncover the core conflict between the association and the revealed preference principle discussed in the introduction.

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 [21] considers its symmetry as one of 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 p 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 concerning the independence region, namely that if $(a, b, c, d) \in I$, then $(b, a, d, c) \in I$.

The revealed preference principle is formalised as follows:

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 is 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. This looks 'wrong' on two counts. First, it is desirable that no individual component of (a, b, c, d), should be decisive *by itself* to declare either complementarity or substitutability: the criterion should also respond to variations in the other components. Secondly, even granting the possibility of one component dictating the criterion, it is hard to justify the fact that it is a high value of $p_\varnothing$ on its own to mandate complementarity.

The following property captures a logically even weaker version of these ideas. 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, a 'very high' non-consumption level should not by itself 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 satisfy Duality and Monotonicity must fail 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 prove the Claim, 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 Duality and Responsiveness there exists a $p = (a, b, c, d)$ such that $p \in C$. To see this, by Responsiveness there exists $\tau \in (0, 1)$ such that, for all $d \in (\tau, 1)$, $(a, b, c, d) \in S$ for some $(a, b, c) \in (0, 1)^3$, and then by (2) of Duality $(b, a, d, c) \in C$.

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 $d' > 1 - \theta > a$, $b' < \theta \leq b$ 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 Duality, Monotonicity and Responsiveness but not 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. This observation motivates the following results.

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.

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 clash between association 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. ■

5 Possibilities

We now turn our attention 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 instead the monotonicity properties but vary the notion of duality. A duality operation produces the 'opposite' behaviour to the one to which the operation is applied. 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.

5.1 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

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

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 $ad > bc$.

■

5.2 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 that $a' = b + c$, $b' = ab/(b + c)$, $c' = ac/(b + c)$ 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 $a > b + c$. ∎

5.3 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

$\longrightarrow$

	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 settles for merely switching out of the initial region after the duality operation.

(S,J)-Duality
1) If $(a,b,c,d) \in C$ then $(b,a,c,d) \notin C$ and $(c,b,a,d) \notin C$.
2) If $(a,b,c,d) \in S$ then $(b,a,c,d) \notin S$ and $(c,b,a,d) \notin S$.
3) If $(a,b,c,d) \in I$ and $a \neq b$ (resp. $a \neq c$) then $(b,a,c,d) \notin I$ (resp. $(c,b,a,d) \notin I$).

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. In the other direction, suppose that (C,I,S) satisfies these three axioms axioms. We will first show that $C \subseteq \{(a,b,c,d) : a > b, c\}$. Suppose, towards a contradiction, that $(a,b,c,d) \in C$ but $a \leq \max\{b,c\}$. By (S,J)-Duality $(b,a,c,d) \notin C$ and $(c,b,a,d) \notin C$, and this contradicts Monotonicity. Hence $a > b, c$. Similarly if $(a,b,c,d) \in S$ but $a \geq \min\{b,c\}$, then (S,J)-Duality yields $(b,a,c,d) \notin S$ and $(c,b,a,d) \notin S$, again contradicting Monotonicity. Hence $S \subseteq \{(a,b,c,d) : a < b, c\}$.

It remains to show that $I \subseteq \{\min\{b,c\} \leq a \leq \max\{b,c\}\}$ and the proof will follow the fact that (C,I,S) is a partition. To this end take some $(a,b,c,d) \in I$. There are three cases to consider regarding where the dual datapoint (b,a,c,d) lies.

Case 1: $(b,a,c,d) \in I$. Then by (S,J)-Duality (part 3), $a = b$ and therefore $\min\{b,c\} \leq a \leq \max\{b,c\}$.

Case 2: $(b,a,c,d) \in S$. By Monotonicity we must have $a > b$, giving $a \geq \min\{b,c\}$. By Symmetry on the other hand, $(a,c,b,d) \in I$. By (S,J)-Duality applied to (b,a,c,d), $(c,a,b,d) \notin S$. Now either $(c,a,b,d) \in I$, in which case $a = c$ by (S,J)-Duality (part 3), or $(c,a,b,d) \in C$, in which case $c > a$ by Monotonicity. Hence $a \leq c$ and therefore $a \leq \max\{b,c\}$.

Case 3: $(b,a,c,d) \in C$. By Monotonicity $b > a$ and therefore $a \leq \max\{b,c\}$. By Symmetry $(a,c,b,d) \in I$. By (S,J)-Duality $(c,a,b,d) \notin C$. Either $(c,a,b,d) \in I$, in which case $a = c$ by (S,J)-Duality (part 3), or $(c, a, b, d) \in S$, in which case $c < a$ by Monotonicity. Hence $a \geq c$ and therefore $a \geq \min\{b, c\}$. ■

6 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 simple '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 [17] and Brady and Rehbeck [9], 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 [18]). 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.[7] 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

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

Then in the logit model the correlation criterion yields:

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

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.

[7] In Manzini and Mariotti [17] and Brady and Rehbeck [9] the consideration coefficients depend, respectively, on the individual alternatives and on the menu.

Turning to the stochastic consideration set model, simple calculations show the following:[8] 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:

- 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.

6.1 The multi-good case

The two-good case is in some respects very specific, as some relevant general features of the complementarity conditions (notably the dependence on the 'psychological' parameters) cannot be understood from it. We now consider an extension of the correlation criterion to the multi-good case. We will show that complementarity in both models can be expressed as a supermodularity condition on preferences, one cardinal an the other ordinal.

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

[8] 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 $\alpha = \alpha^*$.

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.[9]

[9] As Samuleson [21] 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.

6.1.1 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 maximisation.[10]

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

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

[10] 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.

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

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

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

Next, define

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

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. This role of the max operator in the formula is interesting and not obvious a priori.

6.1.2 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:

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

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*

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

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[11] 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:

[11] 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 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:

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

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

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.

6.2 Do the results support the correlation criterion?

We have shown that, in both of the models we have examined, identifying complementarity with correlation is equivalent, in terms of primitives, to identifying complementarity with a special 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. But our findings also have an alternative interpretation. They could be taken as an indication that, for the case of random choice over bundles, supermodularity in utility *over bundles* is in itself 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, complementarity in the data must be driven by the *lack* of joint consumption (that is, $p_\varnothing$ is high) rather than by joint consumption, and an increase in joint consumption can cause a reversion to substitutability, 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, and single good consumption is much higher. The positive correlation is mostly driven by non-readers: should we really be forced to say that the versions are complementary *because* of the lack of joint consumption - with the correlation possibly lost if joint consumption increases? This would be a violation of the revealed preference principle discussed in the Introduction. 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.

In sum, even when preferences govern behaviour (an assumption we have sought to eschew in our general treatment), it seems that while some form of statistical association in consumption may be a natural ingredient of complementarity, statistical association *per se* ignores some 'revealed preference' information that is contained in the data and that may be relevant to assess complementarity. For example, guns and bullets can be safely considered complements, but in many countries people who own a gun are relatively rare: so $p_\varnothing$ is high and p_x and p_y are very low (lower than joint consumption) and only account, say, for the collectors of either of the single items. Suppose that after an increase in crime we observe a massive increase in demand for guns by the joint consumers who worry about personal safety, with a corresponding increase in the consumption of bullets, whereas the collectors' demand, which is unrelated to self-defense motives, stays unchanged. We should interpret this shift as an increase in 'joint revealed preference' for both guns and bullets, and thus confirm our evaluation of guns and bullets as complements. But, as we have seen, this cannot be guaranteed by the correlation criterion. This criterion treats the shift analogously to a shift from observing few cases of smokers with lung cancers to observing many such cases, with the frequencies of smokers with no cancer and of non-smokers with cancers unchanged. From a purely statistical point of view this may or may not be evidence of increased correlation, because both the prior probability and the posterior probability of getting cancer conditional on smoking increase with the shift. But an increase in joint gun/bullet consumption relative to only gun or only bullet consumption is an act of choice that reveals a relative preference shift in favour of joint consumption that is not meaningful in the clinical example: there cannot be any act of choice that reveals a preference for smoking and getting cancer over just smoking.[12]

7 Concluding remarks

Complementarity in general is such a central concept in Economics that its study hardly needs to be motivated. Complementarity has deeply engaged at the theoretical level some

[12] We should also note that the logit model, viewed as an additive RUM of the form $u(b) = \delta_b + \varepsilon_b$ (with $b \in \{xy, x, y, \varnothing\}$, δ_b the deterministic component of utility and ε_b the Gumbel distributed error term) assumes zero correlation in the error terms. Adding some correlation to the error terms would break the link between supermodularity in utility and the correlation criterion (details available from the authors). Our point, however, is that even when supermodularity in utility is equivalent to correlation in consumption, it is still possible that neither is a good criterion of complementarity.

of the giants of the profession (see the historical overview in Samuelson [21]). 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. Finally, the concept of complementarity can be imaginatively applied in non-obvious contexts: for example, Becker and Murphy [4] introduced the idea of assimilating the theory of advertising into the theory of complementarity.

While statistical association is intuitively part of what it means for goods to 'go together' (and sheer correlation in consumption or usage data is often taken as a behavioural indicator of complementarity), we have shown that, in general, criteria for complementarity based on statistical association alone conflict with a basic monotonicity requirement that captures the 'revealed preference' aspect of complementarity. Our axiomatic analysis suggests that if monotonicity is considered primarily 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 [14]).

7.1 'True complementarity' vs. taste correlation

One criticism that could be made of our approach runs along the following lines.[13] Suppose that we observe a daily series of online/print news consumptions for an individual at fixed prices. Then each of our definitions will declare whether or not the two versions are complements or substitutes. A definition of complementarity based on random utility, on the other hand, will distinguish between the case in which the individual derives more utility from joint consumption than from the sum of the utilities of single-version consumption ('true complementarity'), and the case in which whenever the individual wakes up in a mood for reading online news he is also likely to be in the mood for reading the print version - he may simply wake up sometimes in the mood for news and some other times not in the mood for news, independently of the form in which they come

[13] We thank Matthew Gentzkow for raising this important issue.

(‘correlation in taste’). Now, we could hold the choices constant and vary whether or not they indicate ‘true complementarity’ by changing the correlation in taste, whereas our definitions will fail to record these changes. But the key point to understand here is that the kind of complementarity we are trying to capture with our definitions is a related but separate concept from the random utility based concept of complementarity. The latter concept is tied to a specific assumption on the process that drives behaviour. But one could make different assumptions. Suppose for example that behaviour was driven instead by ‘random consideration set’ mechanisms of the type discussed in section 6. Then we should separate ‘true complementarity’ from correlation in consideration rather than from correlation in taste, as our individual’s mood is now expressed by a shock in consideration and not by a shock in taste: he wakes up sometimes considering both types of news media and sometimes not considering either, while always deriving utility from either in the same way. This leads to a different identification problem and likely to a different measurement of complementarity. If we only observe behaviour, which of the two measurements of ‘true complementarity’ should we regard as ‘truer’?[14]

Our behavioural, model-free approach is designed *precisely* to cut through this type of modelling dilemmas. It takes choice data at face value. It serves a different purpose from ‘standard’ definitions: it suits the researcher or user who wishes to be non-committal as to the mechanism that generates behaviour, which is treated as unknown and unknowable. We view this approach as a complement, rather than as a substitute, of the standard one.

7.2 If you could, should you use Hicks complementarity with stochastic choice data?

An interesting by-product of our approach is its implication that *if* prices were available and consumers were utility maximisers, and therefore the standard Hicks criterion of complementarity could be applied, it would conflict with Monotonicity.[15] This can be quickly seen through the following reasoning. Define Hicksian complementarity with stochastic demand in the standard way using expected demand. Gentzkow [14] shows that, for the simple two-good logit model, Hicksian complementarity thus defined is equivalent to the

[14] Note that some specifications of the consideration set model can be rewritten as RUMs while others cannot.

[15] We thank Jean-Pierre Dubé for pointing this out to us.

supermodularity of the utility function. On the other hand, we have shown that the correlation criterion fails Monotonicity. Further, we have shown that the correlation criterion is equivalent to the supermodularity of utility in the two-good logit model, and therefore we can conclude that this application of the Hicks criterion may violate Monotonicity.

Because Monotonicity seems such a fundamental principle, we argue that this reasoning shows that the Hicks criterion may be ill-applied to the context of random utility. That is, even when it *could* be applied because prices are available, the Hicks criterion *should not* be uncritically applied to stochastic choice data. To refer to table 1 again, suppose that the data changed from those in that table, to the following:

	Read print	Did not read print
Read online	0.487	0.043
Did not read online	0.447	0.023

This can mean for example that for every 1000 people, 350 of the non-readers converted to reading. Each of these new readers reads both versions and none of them reads just one version. Now the previously positive consumption correlation has turned negative ($\frac{0.487}{0.487+0.447} = 0.521 < 0.530 = 0.487 + 0.043$). On the assumption that consumption is generated by a logit random utility model, negative (resp., positive) consumption correlation is, as noted before, equivalent to Hicksian substitutability (resp., complementarity) using expected demand. But how can the conversion of over one third of the population to reading jointly the print and the online version be taken as diagnostic of a switch from complementarity to substitution in the nature of these goods? This seems a perverse conclusion.

The paradoxical behaviour of the Hicks criterion in this circumstance stems from the often neglected fact that a parameter (in this case the sign of the Hicksian elasticity) that is meaningful in the deterministic utility model may not carry the same meaning when applied to the perturbed utility version of the model. For an analogy, think of the fact that the Arrow-Pratt risk aversion coefficient of a deterministic utility cannot be taken as a measure of risk aversion in the random utility version of the model (Apesteguia and Ballester [1], Blavatsky [8], Wilcox [23]): in that case, an individual with a higher Arrow-Pratt coefficient of risk aversion may be more likely to accept risk than one with a lower coefficient. We hope that our analysis helps to reinforce the notion that the interpretation of certain features of a utility function (such as supermodularity) is not necessarily inherited by the stochastic version of that utility. Complementarity criteria for stochastic choice data should, even if based on utility maximisation, address directly the stochastic nature of the primitives.

References

[1] Apesteguia, J. and M.A. Ballester (2105) "Well-Defined Stochastic Choice Models: The Case of Risk and Time Preferences", mimeo.

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

[3] Baldwin, E. and P. Klemperer (2015) "Understanding Preferences: "Demand Types", and the Existence of Equilibrium with Indivisibilities", mimeo, http://elizabeth-baldwin.me.uk/papers/demandtypes.pdf

[4] Becker, G.S. and K.M. Murphy (1993) "A Simple Theory of Advertising as a Good or Bad"; *Quarterly Journal of Economics* 108(4): 941-64.

[5] Bernheim, B. D., 2009, "Behavioral Welfare Economics", *Journal of the European Economic Association*, 7: 267–319.

[6] Bernheim, B. D. and A. Rangel, 2009, "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics", *Quarterly Journal of Economics*, 124: 51-104.

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

[8] Blavatskyy, P.R. (2011) "Probabilistic Risk Aversion with an Arbitrary Outcome Set," *Economics Letters* 1112: 34–37.

[9] Brady, R. and J. Rehbeck (2014) "Menu-Dependent Stochastic Consideration", mimeo, University of California at San Diego

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

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

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

[13] Echenique, F., K. Saito and G. Tserenjigmid (2013) "The Perception-Adjusted Luce Model", mimeo, California Institute of Technology.

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

[15] Gül, F., P. Natenzon and W. Pesendorfer (2014) "Random Choice as Behavioral Optimization", *Econometrica*, 82 (5): 1873– 1912.

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

[17] Manzini, P., and M. Mariotti (2014) "Stochastic Choice and Consideration Sets", *Econometrica*, 82 (3): 1153-76.

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

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

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

[21] Samuelson, P. 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.

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

[23] Wilcox, N. T. (2011), "Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice Under Risk" *Journal of Econometrics* 162 (1): 89–104.