Direct Complementarity

Jonathan Weinstein*

Washington University in St. Louis

First version: April 2017; This version: August 2017

Preliminary Draft, not for circulation

Abstract

We point out that the standard definition of complementarity between two goods (a negative compensated cross-price effect) is sensitive to the definitions of other goods. For instance, in a portfolio choice problem, two assets 1 and 2, which are substitutes, may become complements if the asset 3 is replaced by a mutual fund mixing all three assets, even though the prices of all portfolios are unchanged. We create an alternate definition of complementarity, called *direct complementarity*, which is robust to such changes. We also define a robust notion of *relative complementarity* which ranks pairs of goods according to how well they complement each other. Mathematically, the key is that our concepts are relations on bundle-space, while classic concepts are relations on its dual, price-space.

*I thank attendees at the 2017 North American Summer Meetings of the Econometric Society and the 2017 Games and Economic Behavior conference in honor of Ehud Kalai, as well as John Nachbar and Brian Rogers, for useful feedback. This version is preliminary and lacks citations.

1 Introduction

The notions of complementarity and substitutability are central to our understanding of preference and choice in multidimensional environments. The simplest formal context for these notions is a preference over a product $X \times Y$ of ordered sets, represented by a utility function u. In this context, we ordinarily say that quantities x and y are complements if the function u is *supermodular*: that is, if for every pair $x_1 \geq x_2 \in X$, $u(x_1, y) - u(x_2, y)$ is weakly increasing in y. From this it would follow that $x^*(y)$, the optimal x for fixed y, if it exists uniquely, is weakly increasing, and likewise for $y^*(x)$, a crucial conclusion in applications to supermodular games. If the domain is $\mathbb{R}^2$ and u is twice-differentiable, then supermodularity of u is equivalent to non-negativity of the cross-partial derivative: $u_{xy} \geq 0$. Under the reverse of this condition, we would say that x and y are substitutes.

In the theory of consumer choice, we use rather different notions of complement and substitute. We customarily assume that the consumer's preference $\succeq$ on $\mathbb{R}^n$ (the space of bundles of n goods) is represented by a twice-differentiable function $u : \mathbb{R}^n \to \mathbb{R}$. Denote partial derivatives by u_i, u_{ij}, etc. Assume, as is common, that $u_i(x) > 0$ for all x and i. The usual notion of supermodularity of u is *not* useful here, because u is only defined up to monotone transformation, and monotone transformations do *not* preserve supermodularity. In fact, under the given conditions any preference has both a supermodular and a submodular representation. This is closely related to the fact that u has no maximum, either globally or locally, so that without further constraints there are no meaningful comparative statics. Accordingly, in consumer theory we define complements and substitutes by looking at the choices a consumer makes under linear budget constraints. If the compensated cross-price effect is negative – that is, if a price increase in good i, along with a change in income which keeps utility constant, decreases the demand for good j – we say that goods i and j are *Hicksian complements*. It is shown in standard textbooks that (unlike gross complementarity, which looks at uncompensated cross-price effects) this definition is symmetric. But it has a weakness which has apparently gone unnoticed in the literature, as illustrated in our key example:

Example: Alice, at time 0, is forming a portfolio from one safe asset (asset 0) and three risky assets (1, 2 and 3.) There are no short-sale constraints. Her preferences are given by

$$u(x_0, x_1, x_2, x_3) = x_0 + x_1 + x_2 + x_3 - \frac{x_1^2}{2} - \frac{x_2^2}{2} - \frac{x_3^2}{2}$$

For motivation, note that preferences of this form will arise if every asset's time-1 average value is identical, the risky assets are identically, independently and normally distributed, and Alice has CARA preferences. Utility has been formulated as the certainty equivalent of her portfolio.

At prices $(p_0, ..., p_3)$, demand for each risky asset is given by $x_i^* = 1 - p_i/p_0$, with remaining wealth invested in the safe asset. The three risky assets hence have no cross-price effects and are neither complements nor substitutes. Notice also that because utility is quasi-linear in Asset 0, income effects apply only to Asset 0, so that when evaluating effects on Assets 1, 2 and 3 we do not need to know Alice's wealth or worry about compensated vs. uncompensated cross-price effects. For convenience we henceforth fix $p_0 = 1$, i.e. there is a risk-free interest rate of 0.

Now suppose Asset 3 is replaced by a mutual fund, M, containing one-third of a share of each of assets 1, 2, and 3, so Alice will be forming her portfolio from assets 0, 1, 2, and M. Prices $q_0 = 1, q_1, q_2, q_M$ will be such that Alice faces the very same optimization problem; the change of coordinates is given by

$$q_1 = p_1, q_2 = p_2, q_M = \frac{p_1 + p_2 + p_3}{3}$$

The quantities z_i Alice must purchase in order to form an equivalent portfolio are given by the linear transformation

$$z_0 = x_0, z_1 = x_1 - x_3, z_2 = x_2 - x_3, z_M = 3x_3$$

Rewriting Alice's demand in the new coordinate system, i.e. using z and q, gives

$$z_1 = -2q_1 - q_2 + 3q_3, z_2 = -q_1 - 2q_2 + 3q_3, z_M = 3q_1 + 3q_2 - 9q_3 + 3$$

So in the new coordinate system, Assets 1 and 2 have a negative cross-price effect and are Hicksian complements, though they are defined identically as before and the full set of available bundles is the same as before. Why does this happen? One key point is that in the new coordinate system, the meaning of an increase in the price of Asset 1 has changed. When we talk about a price increase, it is always implicit that the price of other goods is held fixed. While Asset 1 is the same in both problems, the change in the definition of Asset 3, to Asset M, changes the meaning of a price increase in Asset 1. When we increase the price of Asset 1 with Asset M fixed, the meaning in the original coordinates is that p_1 increases, p_2 is fixed, and p_3 decreases. Still in terms of the original coordinates, this causes a decrease x_1 and increase in x_3, with x_2 held fixed. So why do we see a cross-price effect on Asset 2? Because though Asset 2 is the same asset, $z_2 \neq x_2$, and holding x_2 fixed is not the same as holding z_2 fixed. To increase x_3 with x_2 fixed, in the z-coordinates we must increase z_M while decreasing z_2.

To help gain further insight, let's rewrite Alice's utility in terms of the z_i:

$$v(z_0, z_1, z_2, z_3) = z_0 + z_1 + z_2 + z_M - \frac{z_1^2}{2} - \frac{z_2^2}{2} - \frac{z_M^2}{6} - \frac{z_1 z_M}{3} - \frac{z_2 z_M}{3}$$

In this formulation, it is intuitive that Asset 1 and Asset 2 are each substitutes for Asset M, since the corresponding cross-partials v_{13}, v_{23} are negative – and of course from a portfolio-management point of view, this makes sense because Asset M is positively correlated with each of Assets 1 and 2. However, $v_{12} = 0$, just as $u_{12} = 0$, corresponding to the independent returns of Assets 1 and 2, so there is an intuitive sense in which Assets 1 and 2 are neither complements nor substitutes, in contrast with their status as Hicksian complements. In fact, the cross-partial $u_{12} = 0$ is preserved by any redefinition of assets other than 1 and 2. The cross-price effect, in the $z - q$ coordinate system, between Assets 1 and 2 can be thought of as "indirect." An increase in the price q_1 of Asset 1 causes an increase in demand for its substitute, Asset M, which is in turn a substitute for Asset 2, so there is a decrease in demand for Asset 2. This motivates us to form a definition of "direct" complements and substitutes, under which Assets 1 and 2 are neutrals, regardless of how the remaining assets are coordinatized. A difficulty is that, as mentioned earlier, u_{12}, though not sensitive to the definition of Asset 3, *is* sensitive to monotone transformations of u. Accordingly, our first definition will focus on the case of a quasi-linear utility function:

Definition 1. *Let preferences be quasi-linear, represented by $u(x) = x_0 + f(x_1, \ldots, x_k)$. Then, at a bundle x, we say that two goods $i, j \neq 0$ are* direct complements *if $u_{ij}(x) > 0$ and* direct substitutes *if $u_{ij}(x) < 0$.*

Later, we will define direct complements and substitutes for general utility functions. The definition above extends naturally because locally, near any point, any smooth utility function looks quasi-linear (up to a second-order approximation) under an appropriate change of coordinates, where the direction of the income effect takes the role of Good 0. We will also exhibit alternate equivalent definitions to further motivate the notion of direct complements and substitutes. Throughout, direct complements will maintain the key property of insensitivity to coordinate changes of the kind in the example.

To further explicate the distinction between direct complements and Hicksian complements, recall the characterization of compensated cross-price effects as cross-partials of the expenditure function:

$$\frac{\partial h_i(p,u)}{\partial p_j} = \frac{\partial^2 E(p,u)}{\partial p_i \partial p_j}$$

where h_i is Hicksian demand and E is the expenditure function (the minimum cost of achieving utility u at prices p.) In this formulation, it is clear that Hicksian complementarity is a binary relation on price vectors, *not* on goods – more specifically, on the infinitesimal changes in price represented by ∂p_i and ∂p_j. Crucially, price vectors p do *not* lie in the space of bundles of goods; rather, the set of price vectors is the *dual* space to the set of bundles, i.e. the set of linear functions from bundles to $\mathbb{R}$. When, as customary, we speak of the relation of Hicksian complementarity as if it is a relation on goods, we are implicitly making use of an isomorphism between bundle-space and its dual, price-space. As should be familiar from linear algebra, this isomorphism is non-canonical, i.e. basis-dependent. The change ∂p_i in price-space is defined by increasing the price of i with the price of other goods *in the basis* held fixed. In the example, when Asset 3 is replaced by Asset M, changing *one* basis element in bundle-space, *all* basis vectors in price-space change meanings: As discussed earlier, ∂q_1 represents an increase in the price of Asset 1 with the prices of Assets 2 and M held fixed, while ∂p_1 of course assumed the prices of Assets 2 and 3 were held fixed. On the other hand, the value of u_{12} depends only on the definitions of Assets 1 and 2.

The vital point here is that, fundamentally, direct complementarity is a symmetric relation between *goods*, while Hicksian complementarity is a symmetric relation between *price changes*. Standard textbook language, by treating Hicksian complementarity as a relation between *goods*, confounds a vector space with its dual.

The next sections are a little technical; for more intuition based on the example, a reader might benefit from looking at Section 4 next.

2 Definitions and Preliminary Results

Let V be a finite-dimensional real vector space; we write V and not $\mathbb{R}^n$ so as not to designate a favored basis. We use the standard notation V^* for the dual space, i.e. the set of linear functions from V to $\mathbb{R}$. Let $u : V \to \mathbb{R}$ be a C^∞ utility function. We write $Du(x) \in V^*$ for the "Fréchet derivative" of u at x, meaning that $Du(x)(v)$ is the first-order approximation around x of $u(x+v) - u(x)$. For a coordinate vector e_i, $Du(x)(e_i)$ is the usual partial derivative in coordinate i. Write $u \sim \hat{u}$ if $\hat{u} = f \circ u$ for a C^∞ function $f : \mathbb{R} \to \mathbb{R}$ with $f' > 0$ everywhere.

Definition 2. *For any C^∞ utility function $u : V \to \mathbb{R}$, the equivalence class $[u] = \{\hat{u} : \hat{u} \sim u\}$ is an* ordinal C^∞ function.

We will define the derivative of $[u]$, $D[u]$, as the set of derivatives $\{D\hat{u} : \hat{u} \sim u\}$. The chain rule says $D\hat{u}(x) = f'(u(x))Du(x)$, so

$$D[u](x) = \{\alpha Du(x) : \alpha \in (0, \infty)\}$$

i.e. the derivative is defined up to positive scalar. Then $D[u](x)$ corresponds to a half-space in V, namely the directions corresponding to increased utility, or "goods," $G(x) = \{v \in V : Du(x)(v) > 0\}$, bounded by the "indifference plane," $I(x) = \mathrm{Ker}(Du(x))$. By standard linear algebra, all functionals which are positive on the half-space G differ only by a positive scalar, hence are in $D[u](x)$. So G carries all the first-order information available from the ordinal function $[u]$. Note that, given a choice of coordinates, this information includes all marginal rates of substitution (MRS) between any two goods i and j; the MRS is the λ such that $e_i - \lambda e_j \in I$.

At each $x \in V$, $D^2u(x)$ can he viewed as a (symmetric) bilinear form $D^2u(x) : V \times V \to \mathbb{R}$, where in a coordinate system, $D^2u(x)(e_i, e_j)$ would be the usual cross-partial in coordinates (i, j). An equivalent formulation we will find useful is to write $D^2u(x) \in (V \otimes V)^*$, i.e. a linear functional on

2-tensors over V. Again, let $\hat{u} = f \circ u$, then

$$D^2\hat{u}(x)(v, w) = f'(u(x))D^2u(x)(v, w) + f''(u(x))Du(x)(v)Du(x)(w)$$

$$D^2\hat{u}(x) = f'(u(x))D^2u(x) + f''(u(x))[Du(x) \otimes Du(x)]$$

where the second line follows by definition of the form $[Du(x) \otimes Du(x)]$. Then we can write

$$D^2[u] = \{\alpha D^2u + \beta(Du \otimes Du) : \alpha \in (0, \infty), \beta \in \mathbb{R}\}$$

(We sometimes suppress x in our notation.) The second degree of freedom, parametrized by β, motivates us to define the "first-order-indifferent tensors,"

$$I^2 := \mathrm{Ker}(Du \otimes Du) = \mathrm{Span}(I \otimes V \cup V \otimes I) \subseteq (V \otimes V)$$

a space of codimension one in $(V \otimes V)$. When restricted to I^2, $D^2[u]$ is well-defined up to a positive scalar. Conversely, any form which, when restricted to I^2, is a positive scalar multiple of D^2u, is an element of $D^2[u]$, i.e.

$$D^2[u] = \{\nu \in (V \otimes V)^* : \exists\alpha \in \mathbb{R}^+ : \forall t \in I^2 : \nu(t) = \alpha D^2u(t)\}$$

The reverse inclusion holds because of the standard fact that $Du(x) \otimes Du(x)$ spans the forms which annihilate its kernel, I^2. So $D^2[u]$ determines, and is determined by, a half-space in I^2 which we call the "complementary tensors," $C(x) := \{t \in I^2 : D^2u(x)(t) > 0\}$, bounded by the "neutral tensors," $N(x) := I^2 \cap \mathrm{Ker}(D^2u)$, and we call the other half of I^2 the "substitutive tensors" $S = -C$. Note that $D^2[u]$ provides no information in its behavior outside of I^2 – such behavior is constrained only by linearity and is otherwise completely representation-dependent. In particular, if we are working with a basis where every coordinate vector is a "good," i.e. $e_i \in G$, then for large positive f'' we have $D^2u(e_i, e_j) > 0$ for all i, j, and for very negative f'' the reverse inequality. That is, as asserted in the introduction, $[u]$ contains both supermodular and submodular functions.

To summarize, around a point x, the first-order information about an ordinal function $[u]$ can be characterized by the half-space of goods, G, bounded by the indifference plane I. The first derivative also determines the first-order-indifferent tensors I^2. Then, the second derivative partitions I^2 into C, N and S. To make the interpretation of this partition more clear, let $v_1 \in I$, $v_2 \in G$, so $v_1 \otimes v_2 \in I^2$. Then, what does it mean to have $v_1 \otimes v_2 \in C$? It means that a small movement in the v_2 direction changes v_1 from a neutral to a good, i.e. $Du(x + \varepsilon v_2)(v_1) > 0$ for small enough positive ε. To be even more concrete, let e_1, e_2, e_3 be basis vectors which are all in G, and let

$$v_{12} := \frac{e_1}{Du(x)(e_1)} - \frac{e_2}{Du(x)(e_2)} \in I$$

be the vector, unique up to a scalar, which is in both the indifference plane and the $e_1 - e_2$ plane. Then $v_{12} \otimes e_3 \in C$ means that motion in the e_3 direction makes v_{12} a good – which is the same as saying that it increases the relative value of Good 1 to Good 2, i.e. the marginal rate of substitution between the goods. We interpret this to mean that Good 3 complements Good 1 *better* than it complements Good 2. This *relative* complementarity – equivalent to complementarity between Good 3 and the "neutral" v_{12} – is robust to changes in representation.

It is interesting to note that when preferences are locally strictly convex around x, $D^2u(x)(v, v) < 0$ for all $v \in I$, i.e. D^2u is negative-definite on I, so $-D^2u$ is an inner product on I, unique up to scalar. A pair $v_1, v_2 \in I$ are substitutes in our sense if $-D^2u(v_1, v_2) > 0$, i.e. if the "angle" between them according to this inner product is less than $\pi/2$, and complements if the angle is greater than $\pi/2$.

2.1 Link strength and Relative Complementarity

Given utility function u and $v_1, v_2 \in G(x)$, define $L_{u,x}[v_1, v_2]$, the *link strength* between v_1 and v_2 at point x, by

$$L_{u,x}[v_1, v_2] = \frac{D^2u(x)(v_1, v_2)}{(Du(x)(v_1))(Du(x)(v_2))}$$

Now, if $\hat{u} = f \circ u$ is another utility representing the same preferences, we have

$$L_{\hat{u},x}[v_1, v_2] = \frac{f''(u(x))}{f'(u(x))^2} + \frac{L_{u,x}[v_1, v_2]}{f'(u(x)}$$

so changing representations effects a positive affine transformation on Ls. Therefore, while link strength is not invariant, *relative* link strength is. Another quick way to check this, using the language from the last section, is to note that $L_{u,x}[v_1, v_2] > L_{u,x}[v_3, v_4]$ if and only if the tensor

$$\left(\frac{v_1 \otimes v_2}{(Du(x)(v_1))(Du(x)(v_2))} - \frac{v_3 \otimes v_4}{(Du(x)(v_3))(Du(x)(v_4))}\right)$$

which is by construction in $I^2(x)$, is a complementary tensor, i.e. in $C(x)$. Just as tensors in $I^2(x)$ can be classified as complementary without representation-dependence, tensors whose difference is in I^2 can be classified according to relative complementarity, according to whether their difference is in $C(x)$. So, the role played by the denominator in the definition of L is to allow us to compare tensors of the form

$$t_{v_1,v_2} = \frac{v_1 \otimes v_2}{(Du(x)(v_1))(Du(x)(v_2))}$$

which satisfy $Du(x) \otimes Du(x)(t) = 1$, so that their differences are all in $I^2 = \mathrm{Ker}(Du \otimes Du)$.

The comparison of $L[v_1, v_2]$ with $L[v_1, v_3]$ has a nice equivalent characterization, connecting with the discussion of relative complementarity at the end of the last section. Let $M_{v_3 v_2} = Du(v_2)/Du(v_3)$ be the marginal rate of substitution of good v_3 for good v_2. Then

$$Du(v_1)(L[v_1, v_2] - L_[v_1, v_3]) = \frac{\partial(\ln M_{v_3 v_2})}{\partial v_1}$$

which we summarize by saying that if v_1 is more strongly linked to v_2 than to v_3, movement in the v_1 direction increases the relative value of v_2 to v_3.

3 Direct Complementarity

At a point x, let v_x^* (the direction of the income effect) be such that motion in the v_x^* direction does not change any marginal rate of substitution; equivalently, for each $w \in I(x)$, $D^2u(v_x^*, w) = 0$. A dimension-counting argument shows that v_x^* must exist, and generically will be unique up to scalar; we assume in this section that this uniqueness holds.[1] The condition on v_x^* is equivalent to $L_{u,x}[v_x^*, v'] = L_{u,x}[v_x^*, v'']$ for all v', v''.

Once we identify v_x^*, we can construct a "locally quasilinear" representation u^q such that $D^2u^q(v_x^*, v) = 0$ for all $v \in V$, by letting $u^q = f \circ u$ for any f satisfying:

$$\frac{f''(u(x))}{f'(u(x))} = -\frac{D^2u(x)(v_x^*, v_x^*)}{(Du(x)(v_x^*))^2}$$

To prove the claimed property, simply check it for $v \in I(x)$ and for $v = v_x^*$, and note that these span V and the property extends linearly.

Then, by analogy with the quasilinear case mentioned earlier, we say:

[1]The dimension-counting argument: D^2u induces a map from V to I^*, defined by $v \mapsto (w \mapsto D^2u(v, w))$, which is well-defined on $[u]$ up to a scalar multiple. Since V has dimension n while I^* has dimension $n-1$, the map has non-trivial kernel, and any non-zero element of the kernel has the desired property.

Definition 3. *w, z are* direct complements *at x if*

$$D^2u^q(x)(w, z) > 0 \Leftrightarrow L_{u,x}[w, z] > L_{u,x}[w, v_x^*]$$

(The equivalence of the definitions follows from the fact that ranking of Ls is invariant to representation; this also establishes that the choice among locally quasilinear u^q could not matter.)

That is, w is a direct complement to z if it complements z better than it complements v_x^*, which is neutral with respect to all goods. This relation is symmetric, and would not be if v_x^* was replaced by any other good. An interesting special case is that for $n = 2$, two goods are direct complements if both are normal, but direct substitutes if one is inferior. (In contrast, the Hicksian notion is trivial when $n = 2$; the two goods are always substitutes.)

Another equivalent definition can be created as follows: Any bundle w can be decomposed as

$$w = \lambda_w v_x^* + w^n$$

where $w^n \in I$ is a neutral. Conceptually, bundles are composed of **nutrients** (utility-rich at first-order, second-order-neutral) and *flavor* (first-order-neutral, with second-order impact). Direct complementarity of (w, z) is equivalent to complementarity of the flavors, $D^2u(x)(w^n, z^n) > 0$, a condition which is representation-invariant as mentioned earlier. We could have also used the pairs (w^n, z) or (w, z^n). Conceptually, it is the first-order-neutral flavors which are complements (or substitutes), while v_x^* is simply a flavorless gruel of utility, complementing any bundle equally well.

4 Relationship between Direct and Hicksian Complementarity

For simplicity, we'll fix here a basis $e_0, ..., e_k$ for V (letting $k = n - 1$) and a quasi-linear utility function $u(x) = x_0 + f(x_1, \ldots, x_k)$. Let W be the span of

$e_1, ..., e_k$. We can view $D^2 f(x)$ as a bilinear form on W, or equivalently as a function from W to W^*. Represent $D^2 f(x)$ as a $k \times k$ matrix H, using the dual basis for W^*. This is nothing but the ordinary Hessian matrix. A little thought shows that cross-price effects are given by the matrix H^{-1}, so

- Goods i, j are direct complements iff $H_{ij} > 0$.
- Goods i, j are Hicksian complements iff $H^{-1}_{ij} < 0$.

This helps us understand the example from the introduction. The Hessian in the original coordinate system is just $H = -I$ where I is the identity, so that both H and H^{-1} are diagonal, and no goods are complements or substitutes in any sense. To translate to the new coordinates, we define a change-of-coordinate matrix

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

and write

$$\hat{H} = CHC^T = \begin{pmatrix} -1 & 0 & -\frac{1}{3} \\ 0 & -1 & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

The change in coordinates affects only the third row and column of H, leaving $\hat{H}_{12} = H_{12} = 0$ unchanged; this would be true for any H. It *does*, though, affect H^{-1}_{12}, which becomes $\hat{H}^{-1}_{12} = -1$, so that as calculated earlier, Goods 1 and 2 are now Hicksian complements. Please note that the inverse is continuous, so by setting $H_{12} = -\varepsilon$ we would easily obtain an example where Goods 1 and 2 change from (Hicksian and direct) substitutes to Hicksian complements (but still direct substitutes) under the change of coordinates.

As one more advertisement for direct complementarity, please note that, for quasi-concave utility, our definition always states that a good is a direct substitute for itself, as it should. In contrast, applying the Hicksian definition would declare a good to be a complement to itself, a point that textbooks wisely dodge to avoid confusion. My suggestion to view the Hicksian notion as a relation on price-changes resolves the paradox: in its effect on increasing the expenditure function, a price increase indeed complements itself.