DISCUSSION PAPER SERIES

DP14273

INFERRING COMPLEMENTARITY FROM CORRELATIONS RATHER THAN STRUCTURAL ESTIMATION

Alessandro Iaria and Ao WANG

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

INFERRING COMPLEMENTARITY FROM CORRELATIONS RATHER THAN STRUCTURAL ESTIMATION

Alessandro Iaria and Ao WANG

Discussion Paper DP14273
Published 05 January 2020
Submitted 03 January 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre’s research programmes:

- Industrial Organization

INFERRING COMPLEMENTARITY FROM CORRELATIONS RATHER THAN STRUCTURAL ESTIMATION

Abstract

According to the Hicksian criterion, two products are complements if their (compensated) cross-price elasticity is negative. While attractive in theory, the implementation of the Hicksian criterion can be hard: computing elasticities requires the estimation of structural models allowing for both complementarity and substitutability. Here, we instead investigate the correlation criterion, whose implementation only requires the comparison of observed market shares. We show that, in a large class of non-parametric models, the correlation criterion satisfies all the axioms by Manzini et al. (2018) and how, in mixed logit models, it can be used to learn about the Hicksian criterion.

JEL Classification: N/A

Keywords: Hicksian Complementarity, Substitutability, Correlation, demand elasticity, Demand estimation, Market Shares

Alessandro Iaria - iaria.alessandro@gmail.com
University of Bristol and CEPR

Ao WANG - ao.wang@ensae.fr
CREST

Acknowledgements
Iaria gratefully acknowledges funding from the British Academy/Leverhulme SRG no. 170142. We thank Paola Manzini and Marco Mariotti for their helpful insights.

Inferring Complementarity from Correlations rather than Structural Estimation

Alessandro Iaria and Ao Wang*

December 28, 2019

Abstract

According to the Hicksian criterion, two products are complements if their (compensated) cross-price elasticity is negative. While attractive in theory, the implementation of the Hicksian criterion can be hard: computing elasticities requires the estimation of structural models allowing for both complementarity and substitutability. Here, we instead investigate the correlation criterion, whose implementation only requires the comparison of observed market shares. We show that, in a large class of non-parametric models, the correlation criterion satisfies all the axioms by Manzini et al. (2018) and how, in mixed logit models, it can be used to learn about the Hicksian criterion.

1 Introduction

Understanding whether products, services, or more abstract options are either substitutes or complements has important economic repercussions. For example, mergers between firms selling complementary products can be welfare enhancing (Cournot (1838) and Song et al. (2017)), organizational change can be extremely difficult when there are complementarities among existing practices (Milgrom et al. (1990) and Brynjolfsson and Milgrom (2013)), the adoption of one technology may boost the adoption of others if they are complementary (Augereau et al. (2006) and Kretschmer et al. (2012)), and the optimal pricing of multi-product firms may drastically change with complementarities (Armstrong (2016) and Thomassen et al. (2017)).

Since the Hicks-Allen revolution in demand theory, the Hicksian criterion of negative (compensated) cross-price elasticity has represented the standard notion of complementarity among products (Samuelson (1974) and Chambers et al. (2010)). While this definition has theoretical

*University of Bristol and CEPR IO (alessandro.iaria@bristol.ac.uk) and CREST (ao.wang@ensae.fr). Iaria gratefully acknowledges funding from the British Academy/Leverhulme SRG no. 170142. We thank Paola Manzini and Marco Mariotti for their helpful insights.

advantages (e.g., symmetry), its practical implementation as a test for complementarity is not always straightforward (Gentzkow (2007) and Dubé (2019)): demand elasticities are not directly observed and need to be inferred from structural models whose estimation is notoriously challenging (Berry et al. (2014), Fox and Lazzati (2017), and Iaria and Wang (2019)).

As a practical alternative to the Hicksian criterion, we investigate the correlation criterion, according to which two products are complements whenever their purchases are positively correlated (Brynjolfsson and Milgrom (2013) and Manzini et al. (2018)). Conveniently, the implementation of the correlation criterion only requires comparisons of observed market shares (rather than the estimation of structural models). However, little is known about its micro-foundations and its relationship to the Hicksian criterion. As a notable exception, Manzini et al. (2018) show that—without restrictions to the data generating process—the correlation criterion violates the axiom of monotonicity or revealed preferences: complements can turn into substitutes following an increase in their joint purchases.[1]

In this paper, we present two novel results about the correlation criterion. First, we demonstrate that, given plausible restrictions to the data generating process, the correlation criterion satisfies all the axioms suggested by Manzini et al. (2018)—including monotonicity. The restrictions we require are satisfied by a large class of non-parametric models along the lines of Gentzkow (2007)'s, where the average utility of a bundle equals the sum of the average utilities of the single products plus a term capturing the extra utility associated with their joint purchase. We refer to these extra utility terms simply as *demand synergies*. In this context, violations of the monotonicity axiom by the correlation criterion boil down to the assumptions one is willing to make about unobserved heterogeneity in the demand synergies.

Second, even if little can be learned about the Hicksian criterion only by comparing observed market shares, we show that the correlation criterion can still be informative about Hicksian complementarity and substitutability when individual-level purchase data are available. This result relies on three ingredients: the availability of individual-level purchase data, the data to be generated by a mixed logit version of Gentzkow (2007)'s model of demand for bundles, and a simple non-parametric estimator of the *sign* of the covariance between the individual-level purchases of different products. The proposed estimator is simple to implement and only requires the computation of frequency counters, while its sign consistency hinges on some of the properties of mixed logit models of demand for bundles from Iaria and Wang (2019).

This paper relates to a small but growing literature that investigates how to test for complementarity given standard consumer purchase data (for a recent review, see Dubé (2019)). Gentzkow (2007) proposes a way of empirically implementing the Hicksian criterion by extending standard discrete choice models to allow for both complementarity and substitutability. Chambers et al. (2010) characterize the testable implications of "gross" complementarity, i.e.,

[1] Even though the correlation criterion in general violates monotonicity, it satisfies the other two desirable axioms proposed by Manzini et al. (2018): duality and responsiveness. See Manzini et al. (2018) for details.

a negative *un*compensated cross-price elasticitity. Fox and Lazzati (2017), Allen and Rehbeck (2018), and Iaria and Wang (2019) build on Gentzkow (2007) to study the identification and estimation of demand models useful to empirically implement the Hicksian criterion.

All of these focus on variants of the Hicksian criterion which require the estimation of structural models sometimes difficult to implement in practice.[2] For those cases in which the Hicksian criterion is impractical, Manzini et al. (2018) propose the use of convenient criteria for complementarity that only require the comparison of observed market shares. They then demonstrate a striking fact: none of the proposed criteria can—in general—satisfy all of their desirable axioms. In particular, the correlation criterion will in general violate the axiom of monotonicity or revealed preferences. We add to Manzini et al. (2018) by proposing two positive results about the correlation criterion. First, mild restrictions to the data generating process are sufficient for the correlation criterion to satisfy also monotonicity. Second, when individual-level purchase data are available, the correlation criterion can be used to learn about Hicksian complementarity and substitutability in mixed logit versions of Gentzkow (2007)'s model.

The rest of the paper is organized as follows. Section 2 introduces relevant notation and definitions. Section 3 presents sufficient restrictions on the data generating process to prevent violations of the monotonicity axiom. Section 4 proposes a simple method to learn about the Hicksian criterion on the basis of the correlation criterion with individual-level purchase data. Section 5 concludes with some final remarks and Appendix section 6 contains all the proofs.

2 Notation and Definitions

Suppose the econometrician observes data on the bundle-level purchase probabilities of two products, x and y. Each unit of observation can be thought of as a market and the purchase probabilities as bundle-level market shares. Markets can be interpreted in a broad sense as, for example: different time periods for the same geographical area, different geographical areas at a certain point in time, or a combination of both. We define the sampling space of bundle-level market shares for any given market as $\mathbf{T} \equiv \{(P_{(x,y)}, P_x, P_y, P_0) | P_k \in (0,1), \forall k; \sum P_k = 1\}$, where $P_{(x,y)}$ is the market share of the joint purchase of products x and y, P_x (respectively P_y) is the market share of purchase of x but not y (respectively y but not x), and P_0 is the market share of the outside option—the choice of not purchasing any of the two products. $\mathbf{T}$ is the space of possible values $(P_{(x,y)}, P_x, P_y, P_0)$ can take in any given market. For simplicity, we rule out the "zero market share" problem (Gandhi et al. (2019)) and assume that $P_k \in (0,1)$ for

[2] In the model-free environment of Chambers et al. (2010), the proposed revealed-preference procedure does not require structural estimation, but it does hinge on the availability of price variation, which in some cases is not available—as in Gentzkow (2007). Differently, structural models along the lines of Gentzkow (2007)'s can be used to implement variants of the Hicksian criterion in the absence of price or income changes. In these variants, the demand elasticities are computed with respect to changes in non-price product-specific attributes (Allen and Rehbeck (2018)).

$k \in \{(x,y), x, y, 0\}$ and that $P_{(x,y)} + P_x + P_y + P_0 = 1$.

We now define the correlation criterion and the monotonicity axiom, and briefly illustrate why—without further restrictions on the data generating process—the correlation criterion will violate monotonicity.

Correlation Criterion. The correlation criterion partitions the sampling space $\mathbf{T}$ into the three subsets $\mathbf{C}$, $\mathbf{I}$, and $\mathbf{S}$: the collections of possible values of $(P_{(x,y)}, P_x, P_y, P_0)$ that exhibit complementarity, independence, and substitutability. $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$ if and only if $\frac{P_{(x,y)}}{P_{(x,y)}+P_y} > P_{(x,y)} + P_x$. Note that this condition is equivalent to $P_{(x,y)}P_0 > P_xP_y$. Similarly, $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{I}$ if and only if $P_{(x,y)}P_0 = P_xP_y$ and $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{S}$ if and only if $P_{(x,y)}P_0 < P_xP_y$. According to the correlation criterion, $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$ (respectively $\mathbf{S}$) if and only if knowledge that one product is purchased increases (respectively decreases) the probability that the other product is purchased as well.

Monotonicity Axiom. This axiom embodies the coherence of the principle of revealed preferences. Monotonicity requires that if $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$, $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{T}$ with $P'_{(x,y)} \geq P_{(x,y)}$, $P'_x \leq P_x$, and $P'_y \leq P_y$, then $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{C}$. Symmetrically, if $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{S}$, $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{T}$ with $P'_{(x,y)} \leq P_{(x,y)}$, $P'_x \geq P_x$, and $P'_y \geq P_y$, then $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{S}$. The monotonicity axiom states that, if two products are complements for some $(P_{(x,y)}, P_x, P_y, P_0)$, then they cannot be substitutes or independent for any $(P'_{(x,y)}, P'_x, P'_y, P'_0)$ with larger joint purchases and smaller or equal individual purchases of each product, and vice versa (i.e., substitutes cannot turn into complements following a decrease in their joint purchases).

Correlation Criterion and Failure of Monotonicity. In general, the correlation criterion does not satisfy the monotonicity axiom. To see this, we report a counter-example due to Manzini et al. (2018). Suppose that $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$ and consider another possible value from $\mathbf{T}$ such that $P'_{(x,y)} = P_{(x,y)} + \epsilon$, $P'_x = P_x$, $P'_y = P_y$, $P'_0 = P_0 - \epsilon$ for a small enough ϵ. It then follows that $P'_{(x,y)}P'_0 = (P_{(x,y)m} + \epsilon)(P_0 - \epsilon)$ and for $\epsilon \to P_0$, $P'_{(x,y)}P'_0 \to 0$. This implies that, for $\epsilon \to P_0$, $P'_{(x,y)}P'_0 < P'_xP'_y$ and $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{S}$, violating monotonicity.

As discussed by Manzini et al. (2018), the correlation criterion will *in general* satisfy the other two desirable axioms of duality and responsiveness. The restrictions we propose to the data generating process will not affect this: when we claim that monotonicity holds, it is implied that also duality and responsiveness hold.

3 Gentzkow (2007)'s Model and Monotonicity

Manzini et al. (2018) resolve the inconsistency between the correlation criterion and the monotonicity axiom by proposing a weaker version of the axiom while maintaining an unrestricted data generating process. Differently, we add structure to the data generating process and resolve the inconsistency within the more specific context of Gentzkow (2007)'s model. To provide some intuition, we first illustrate in a simple model how violations of the monotonicity axiom boil down to the assumptions one is willing to make about unobserved heterogeneity in demand synergies. We then generalize these insights by characterizing violations of the monotonicity axiom in the context of a non-parametric class of models that includes the mixed logit and the probit as special cases.

3.1 Some Intuition: The Multinomial Logit Model

Assume that the indirect utilities of individual i in any given market are:

$$\begin{aligned} U_{ix} &= \delta_x + \epsilon_{ix} \\ U_{iy} &= \delta_y + \epsilon_{iy} \\ U_{i(x,y)} &= \delta_{(x,y)} + \epsilon_{i(x,y)} \\ &= \delta_x + \delta_y + \Gamma + \epsilon_{i(x,y)} \\ U_{i0} &= \epsilon_{i0}, \end{aligned} \tag{1}$$

where δ_x, δ_y, and $\delta_{(x,y)}$ are the market-level average utilities associated respectively to purchasing only product x, only product y, and the bundle (x, y). Following Gentzkow (2007), we specify $\delta_{(x,y)} = \delta_x + \delta_y + \Gamma$, where Γ is the market-level extra portion of average utility associated to the joint purchase of x and y. For example, if $\Gamma > 0$, individuals in this market obtain a higher average utility from the joint rather than the separate purchase of x and y. We refer to Γ simply as *demand synergy parameter*. $(\epsilon_{i(x,y)}, \epsilon_{ix}, \epsilon_{iy}, \epsilon_{i0})$ are unobserved and assumed to be i.i.d. Gumbel, giving rise to a multinomial logit (MNL) model.

According to the correlation criterion, $(P_{(x,y)}, P_x, P_y, P_0)$ belongs to the set of values exhibiting complementarity, $\mathbf{C}$, if and only if $P_{(x,y)}P_0 > P_xP_y$. In the context of MNL model (1), this happens if and only if $\exp(\delta_x + \delta_y + \Gamma) > \exp(\delta_x + \delta_y)$, or equivalently if and only if $\Gamma > 0$. Clearly, when the demand synergy parameter can take different values for any $(P_{(x,y)}, P_x, P_y, P_0) \neq (P'_{(x,y)}, P'_x, P'_y, P'_0)$, i.e. $\Gamma \neq \Gamma'$, knowing that $\Gamma > 0$—and so that $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$—carries no information about the membership of $(P'_{(x,y)}, P'_x, P'_y, P'_0)$ to $\mathbf{C}$, $\mathbf{I}$, or $\mathbf{S}$. This leads violations of the monotonicity axiom.

A way to prevent violations of the monotonicity axiom in MNL models is to restrict the variation of the demand synergy parameter in $\mathbf{T}$: $\Gamma' = \Gamma$ for any $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{T}$. Note that this is the original assumption made by Gentzkow (2007). In this case, if there exists a $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{T}$ that belongs to $\mathbf{C}$, implying that $\Gamma > 0$, then $\mathbf{T} = \mathbf{C}$. Similarly, if $\Gamma = 0$, then $\mathbf{T} = \mathbf{I}$ and if $\Gamma < 0$, then $\mathbf{T} = \mathbf{S}$. As a consequence, the correlation criterion does not violate the monotonicity axiom when the data generating process is MNL model (1) with $\Gamma' = \Gamma$ for any $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{T}$.

3.2 More in General

By building on the insights from MNL model (1), we propose sufficient restrictions to the data generating process that prevent violations of the monotonicity axiom in a large class of non-parametric models of demand for bundles.

Assumption 1: Data Generating Process.

(i) $(P_{(x,y)}(\delta), P_x(\delta), P_y(\delta), P_0(\delta))$ is a function of the product-specific average utilities $\delta = (\delta_x, \delta_y) \in \mathbb{R}^2$, with $P_k(\delta) \in (0,1)$, $k \in \{(x,y), x, y, 0\}$, and $\sum_k P_k(\delta) = 1$. Accordingly, the sampling space is $\mathbf{T} \equiv \{(P_{(x,y)}, P_x, P_y, P_0) \big| P_k = P_k(\delta), \forall k; \delta = (\delta_x, \delta_y) \in \mathbb{R}^2\}$.

(ii) $P_z(\delta_x, \delta_y)$ is strictly increasing in δ_z and strictly decreasing in $\delta_{z'}$ with $z, z' \in \{x, y\}$ and $z \neq z'$.

(iii) $P_{(x,y)}(\delta_x, \delta_y)$ is strictly increasing in both δ_x and in δ_y.

(iv) $(P_{(x,y)}(\delta), P_x(\delta), P_y(\delta), P_0(\delta))$ is C^1 with respect to $\delta = (\delta_x, \delta_y)$.

Intuitively, condition (i) requires that the only source of variation among different values of $(P_{(x,y)}, P_x, P_y, P_0)$ in $\mathbf{T}$ is the vector of product-specific average utilities $\delta = (\delta_x, \delta_y)$. This encapsulates the requirement of a constant demand synergy parameter Γ in MNL model (1). Condition (ii) requires the market share function of each product to be increasing in its own average utility and decreasing in that of the other product, while condition (iii) requires the market share function of the joint purchase to be instead increasing in the average utility of each of the products. Finally, condition (iv) requires that the market share functions are continuously differentiable in $\delta = (\delta_x, \delta_y)$.

Example: Mixed Logit Model. A more general example than MNL model (1) which is compatible with Assumption 1 is a mixed logit model with unobserved heterogeneity both in the alternative-specific constants $\delta_i = (\delta_{xi}, \delta_{yi})$ and in the demand synergy parameter Γ_i. This is an extension of Gentzkow (2007)'s original model, which assumes Γ to be constant across individuals. Note that heterogeneity in Γ_i means that some individuals can perceive products x and y as complements (with $\Gamma_i > 0$), while others can see them as substitutes (with $\Gamma_i < 0$). Importantly, Assumption 1 does not require any restriction on the joint distribution F of (δ_i, Γ_i), so that:

$$\begin{aligned} \text{den}(\delta, \delta_i, \Gamma_i) = & \; 1 + \exp(\delta_x + \delta_{xi}) + \exp(\delta_y + \delta_{yi}) \\ & + \exp(\delta_x + \delta_{xi} + \delta_y + \delta_{yi} + \Gamma_i) \\ P_0 = P_0(\delta) = & \int \frac{1}{\text{den}(\delta, \delta_i, \Gamma_i)} dF(\delta_i, \Gamma_i) \\ P_z = P_z(\delta) = & \int \frac{\exp(\delta_z + \delta_{zi})}{\text{den}(\delta, \delta_i, \Gamma_i)} dF(\delta_i, \Gamma_i); \text{ for } z = x,\, y \\ P_{(x,y)} = P_{(x,y)}(\delta) = & \int \frac{\exp(\delta_x + \delta_y + \delta_{xi} + \delta_{yi} + \Gamma_i)}{\text{den}(\delta, \delta_i, \Gamma_i)} dF(\delta_i, \Gamma_i), \end{aligned} \tag{2}$$

where $\delta = (\delta_x, \delta_y)$, $\delta_{(x,y)i} = \delta_x + \delta_y + \delta_{xi} + \delta_{yi} + \Gamma_i$, and the distribution of (δ_i, Γ_i) is the same across different values in $\mathbf{T}$: $F' = F$ for any $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{T}$. This last requirement essentially implies condition (i) from Assumption 1 and is the mixed logit analogue to a constant demand synergy Γ in MNL model (1).

Given any reference point $\delta = (\delta_x, \delta_y)$ whose corresponding $(P_{(x,y)}, P_x, P_y, P_0)$ belong to $\mathbf{C}$, applicability of the monotonicity axiom requires the existence of at least *another* $\delta' = (\delta'_x, \delta'_y) \neq \delta$ such that $P'_{(x,y)} \geq P_{(x,y)}$ and $P'_z \leq P_z$, for $z = x, y$. In what follows, we refer to any such point as a *feasible point* for δ. Note that feasibility is a pre-requisite for the monotonicity axiom to be violated: only among the feasible points for δ one can challenge the monotonicity axiom.

In the next Theorem, we show that there exist feasible points in a neighborhood of $\delta = (\delta_x, \delta_y)$—so that the monotonicity axiom applies—if and only if the Jacobian matrix of the mapping $(P_x(\cdot), P_y(\cdot))$ evaluated at $\delta = (\delta_x, \delta_y)$ has a negative determinant.[3]

Theorem 1. *Suppose Assumption 1 holds, $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$, and $Det\left(\frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)}\right) \neq 0$. Then, the following two statements are equivalent:*

1. *$Det\left(\frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)}\right) < 0$.*
2. *For any neighborhood $\mathbf{N}$ of $\delta = (\delta_x, \delta_y)$, there exists a feasible point $\delta' = (\delta'_x, \delta'_y) \in \mathbf{N}$ for δ such that $\delta' \neq \delta$.*

Proof. See Appendix 6.2. □

[3]For given matrix A, we denote its transpose by A^{T} and its determinant by $Det(A)$.

Symmetric arguments lead to the result that for any $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{S}$, there exists a feasible point in any neighborhood of δ if and only if Det $\left(\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\right) > 0$. The next Corollary summarizes our main conclusion about the monotonicity axiom.

Corollary 1. *Suppose Assumption 1 holds. Then, the correlation criterion does not locally violate the monotonicity axiom.*

Proof. Denote the set of δ's corresponding to the collection of possible values of $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{T}$ that exhibit complementarity as $\mathbf{C}_\delta \equiv \left\{(\delta_x, \delta_y)\middle|P_{(x,y)}(\delta)P_0(\delta) > P_x(\delta)P_y(\delta)\right\}$, and similarly the sets $\mathbf{S}_\delta$ and $\mathbf{I}_\delta$. We first prove that the correlation criterion doest not locally violate the monotonicity axiom in the space of $\delta = (\delta_x, \delta_y)$, and then show that this is enough for the argument to hold locally also in the space of $(P_x(\delta), P_y(\delta), P_{(x,y)}(\delta), P_0(\delta))$. Because of continuity of the market share function, $\mathbf{I}_\delta$ is a one-dimensional curve that defines the boundary between $\mathbf{C}_\delta$ and $\mathbf{S}_\delta$. In other words, $\mathbf{I}_\delta$ divides $\mathbb{R}^2 = \{(\delta_x, \delta_y)\big|(\delta_x, \delta_y) \in \mathbb{R} \times \mathbb{R}\}$ into $\mathbf{C}_\delta$ and $\mathbf{S}_\delta$. Because both $\mathbf{C}$ and $\mathbf{S}$ are open and the market share function is continuous, then both $\mathbf{C}_\delta$ and $\mathbf{S}_\delta$ are also open.[4] Similarly, Det $\left(\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\right) = 0$ divides $\mathbb{R}^2$ into O_+, the set of points associated to a positive determinant, and O_-, the set of points associated to a negative determinant. Then, for any reference point $\delta \in \mathbf{C}_\delta \cap O_-$ or $\delta \in \mathbf{S}_\delta \cap O_+$, according to Theorem 1, we can always find (infinitely many) feasible points in any neighbourhood of $\delta = (\delta_x, \delta_y)$. In addition, given that both $\mathbf{C}_\delta$ and O_- (respectively $\mathbf{S}_\delta$ and O_+) are open, $\mathbf{C}_\delta \cap O_-$ (respectively $\mathbf{S}_\delta \cap O_+$) is also open and all such feasible points still belong to $\mathbf{C}_\delta \cap O_-$ (respectively $\mathbf{S}_\delta \cap O_+$) and hence to $\mathbf{C}_\delta$ (respectively $\mathbf{S}_\delta$). Hence, the correlation criterion does not locally violate the monotonicity axiom for all the reference points in $\mathbf{C}_\delta \cap O_-$ and in $\mathbf{S}_\delta \cap O_+$. Moreover, for any $\delta \in \mathbf{C}_\delta \cap O_+$ or $\delta \in \mathbf{S}_\delta \cap O_-$, as shown by Theorem 1, there exists a neighborhood of δ such that there are no feasible points. Consequently, both in $\mathbf{C}_\delta \cap O_+$ and in $\mathbf{S}_\delta \cap O_-$, the monotonicity axiom is not locally applicable and, therefore, cannot be locally violated. Finally, note that in any of the four cases ($\delta \in \mathbf{C}_\delta \cap O_-$, $\delta \in \mathbf{S}_\delta \cap O_+$, $\delta \in \mathbf{C}_\delta \cap O_+$, or $\delta \in \mathbf{S}_\delta \cap O_-$), Det $\left(\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\right) \neq 0$. Then, according to the inverse function theorem, (P_x, P_y) is locally invertible. Therefore, all arguments that hold locally for $\delta = (\delta_x, \delta_y)$ also hold locally for $(P_x(\delta), P_y(\delta))$ and consequently for $(P_x(\delta), P_y(\delta), P_{(x,y)}(\delta), P_0(\delta))$.[5] □

To get some intuition about Corollary 1, in Figure 1 we illustrate the main forces at play. The sampling space $\mathbf{T}$ is partitioned into three subsets $\mathbf{C}$, $\mathbf{S}$, and $\mathbf{I}$ by the correlation criterion. Then, relying on the market share function from Assumption 1, we can map the collections of market shares $\mathbf{C}$, $\mathbf{S}$, and $\mathbf{I}$ onto the corresponding collections of δ's $\mathbf{C}_\delta$, $\mathbf{S}_\delta$, and $\mathbf{I}_\delta$. Figure 1 depicts a possible partition of $\mathbb{R}^2 = \{(\delta_x, \delta_y)\big|(\delta_x, \delta_y) \in \mathbb{R} \times \mathbb{R}\}$. These sets are such that $\mathbf{I}_\delta$ always separates $\mathbf{C}_\delta$ from $\mathbf{S}_\delta$. Similarly, $\mathbb{R}^2$ can be partitioned into three subsets on the basis

[4] The inverse of an open set under a continuous mapping is still open.

[5] Note that for $(P_x(\delta), P_y(\delta), P_{(x,y)}(\delta), P_0(\delta))$, the meaning of "locally" is with respect to the relative topology defined at $(P_x(\delta), P_y(\delta), P_{(x,y)}(\delta), P_0(\delta))$.

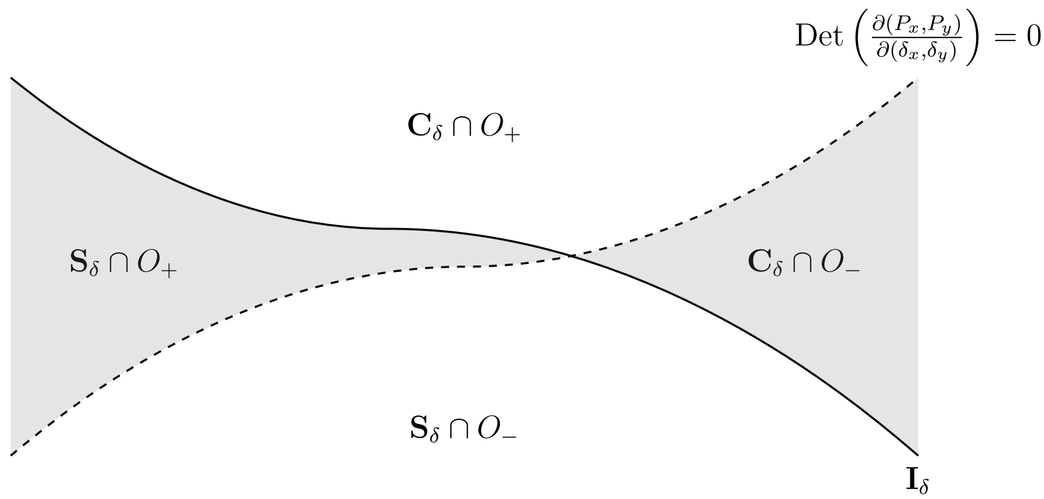

Figure 1: Illustration of Corollary 1

of Det $\left(\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\right)$: the set of δ's corresponding to a zero determinant will separate the set for which the determinant is positive, O_+, from the one for which it is negative, O_-. As in Figure 1, by intersecting these partitions we end up dividing $\mathbb{R}^2$ into four mutually exclusive regions. Because of Theorem 1, we know that in $\mathbf{C}_\delta \cap O_+$ and $\mathbf{S}_\delta \cap O_-$ (the white areas in Figure 1), the monotonicity axiom is not locally applicable—and therefore cannot be locally violated. In the remaining $\mathbf{C}_\delta \cap O_-$ and $\mathbf{S}_\delta \cap O_+$ (the grey areas in Figure 1), as shown in the proof of Corollary 1, the monotonicity axiom instead holds locally—again preventing any local violation.

4 Relationship to Hicksian Complementarity

In this section, we describe the relationship between correlation and Hicksian criteria in the context of two models discussed above: MNL model (1) and mixed logit model (2).

MNL model (1). As shown by Manzini et al. (2018), in the MNL model the two criteria coincide. As we saw above, the correlation criterion classifies x and y as complements if and only if $\Gamma > 0$. Symmetrically, as shown by Gentzkow (2007), the demand elasticity of x with respect to an increase in δ_y is positive (i.e., Hicksian complementarity) if and only if $\Gamma > 0$. This equivalence is however specific to the MNL, a model that rules out any correlation in the unobserved preferences of x and y *beyond* Γ.

Mixed Logit Model (2). Given the greater flexibility of this model, the relationship between correlation and Hicksian criteria is remarkably more complex than in the case of the MNL model.

In particular, the practical usefulness of this more complex relationship depends on the quality of the available data. Even though little can be learned about the Hicksian criterion uniquely from observed market shares, we show that the correlation criterion can be used to infer either Hicksian complementarity or substitutability when individual-level purchase data are available.

Define the individual-level *marginal* purchase probability of product $z \in \{x, y\}$ as:

$$P_{z.}(\delta, \delta_i, \Gamma_i) = \frac{\exp(\delta_z + \delta_{zi}) + \exp(\delta_x + \delta_y + \delta_{xi} + \delta_{yi} + \Gamma_i)}{\text{den}(\delta, \delta_i, \Gamma_i)} \tag{3}$$

and its market-level counterpart, the *marginal* market share of $z \in \{x, y\}$, as:

$$P_{z.} = \int P_{z.}(\delta, \delta_i, \Gamma_i) dF(\delta_i, \Gamma_i). \tag{4}$$

In mixed logit model (2), x and y are Hicksian substitutes if and only if $\partial P_{x.}/\partial \delta_y < 0$, or equivalently:

$$\begin{aligned} P_{(x,y)} &< \int P_{x.}(\delta, \delta_i, \Gamma_i) P_{y.}(\delta, \delta_i, \Gamma_i) dF(\delta_i, \Gamma_i) \\ &= \mathbb{E}[P_{x.}(\delta, \delta_i, \Gamma_i) P_{y.}(\delta, \delta_i, \Gamma_i)]. \end{aligned} \tag{5}$$

They are instead substitutes according to the correlation criterion if and only if

$$\begin{aligned} P_{(x,y)} &< (P_{(x,y)} + P_x)(P_{(x,y)} + P_y) \\ &= P_{x.} P_{y.} \\ &= \mathbb{E}[P_{x.}(\delta, \delta_i, \Gamma_i)] \mathbb{E}[P_{y.}(\delta, \delta_i, \Gamma_i)]. \end{aligned} \tag{6}$$

We refer to the difference between the right-hand sides of (5) and of (6) as $R_{xy}(\delta, F)$, and note that $R_{xy}(\delta, F)$ is the covariance between $P_{x.}(\delta, \delta_i, \Gamma_i)$ and $P_{y.}(\delta, \delta_i, \Gamma_i)$. Denote by $\mathbf{S}_\delta^H$ and by $\mathbf{S}_\delta$ the collections of δ's satisfying, respectively, (5) and (6) (and similarly $\mathbf{C}_\delta^H$ and $\mathbf{C}_\delta$). It then follows that:

(a) If $R_{xy}(\delta, F) > 0$, then $\delta \in \mathbf{S}_\delta$ implies $\delta \in \mathbf{S}_\delta^H$.

(b) If $R_{xy}(\delta, F) < 0$, then $\delta \in \mathbf{C}_\delta$ implies $\delta \in \mathbf{C}_\delta^H$.

Learning that $\delta \in \mathbf{S}_\delta$ (respectively $\delta \in \mathbf{C}_\delta$) from $(P_x, P_y, P_{(x,y)})$ is sufficient to conclude that $\delta \in \mathbf{S}_\delta^H$ (respectively $\delta \in \mathbf{C}_\delta^H$) *as long as* $R_{xy}(\delta, F) > 0$ (respectively $R_{xy}(\delta, F) < 0$). Unfortunately, without strong restrictions on F, the sign of the covariance $R_{xy}(\delta, F)$ is not identified given only market share data $(P_x, P_y, P_{(x,y)})$. In turn, without knowing the sign of $R_{xy}(\delta, F)$, it cannot be determined whether case (a) or case (b) applies, and consequently whether $\mathbf{S}_\delta \subset \mathbf{S}_\delta^H$ or $\mathbf{C}_\delta \subset \mathbf{C}_\delta^H$. Despite this complication, we now show that given some standard individual-level purchase data, it is possible to infer both the sign of $R_{xy}(\delta, F)$ and whether $\delta \in \mathbf{S}_\delta$ or $\delta \in \mathbf{C}_\delta$ without requiring any restriction on F or expensive structural estimation.

Suppose that there are $i = 1, ..., I$ individuals, each making a purchase from $\{(x, y), x, y, 0\}$ in each market $t = 1, ..., T$ and that the econometrician can observe these individual-level purchases over the T markets. Consistent with mixed logit model (2), suppose that the data generating process of individual i's purchase in market t, w_{it}, is a MNL model:

$$\begin{aligned} \Pr(w_{it} = z) &= \frac{\exp(\delta_z + \delta_{zi})}{\mathrm{den}(\delta, \delta_i, \Gamma_i)}, \; z = x, y \\ \Pr(w_{it} = (x, y)) &= \frac{\exp(\delta_x + \delta_y + \delta_{xi} + \delta_{yi} + \Gamma_i)}{\mathrm{den}(\delta, \delta_i, \Gamma_i)}. \end{aligned} \tag{7}$$

The random coefficients (δ_i, Γ_i) are assumed to be i.i.d. according to F within each market, and F to be the same across markets. Define $P_t = P(\delta_t) = (P_{x.}(\delta_t), P_{y.}(\delta_t))$ for $t = 1, ..., T$. We then propose to infer the sign of $R_{txy} = R_{xy}(\delta_t, F)$ on the basis of the sign of the following statistic:

$$\hat{R}_{txy}(\epsilon) = \frac{1}{I}\sum_{i=1}^{I} \hat{P}_{itx.}(\epsilon)\hat{P}_{ity.}(\epsilon) - \frac{1}{I}\sum_{i=1}^{I} \hat{P}_{itx.}(\epsilon)\frac{1}{I}\sum_{i=1}^{I} \hat{P}_{ity.}(\epsilon) \tag{8}$$

with

$$\hat{P}_{itz.}(\epsilon) = \frac{1}{\sum_{t'=1}^{T} \mathbf{1}_{|P_{t'} - P_t| < \epsilon}} \sum_{t'=1}^{T} \mathbf{1}_{|P_{t'} - P_t| < \epsilon}[\mathbf{1}_{w_{it'} = z} + \mathbf{1}_{w_{it'} = (x,y)}], z = x, y$$

where $|\cdot|$ denotes the Euclidean distance, $\mathbf{1}_C$ the indicator function for condition C, and ϵ the radius of a neighborhood of marginal market shares around P_t. Note that, for any given radius ϵ, $\hat{R}_{txy}(\epsilon)$ is a finite-sample approximation of the covariance R_{txy} that only requires the computation of frequency counters. Intuitively, the size of the radius ϵ determines the "quality" of the approximation (with smaller being better). However, given our interest only in the *sign* of R_{txy}, we will not need ϵ to shrink asymptotically for the sign consistency of the estimator.

In spite of $\hat{R}_{txy}(\epsilon)$'s practical simplicity, its sign consistency presents a technical difficulty: $P_{itz.}(\epsilon) \equiv \lim_{T\to\infty} \hat{P}_{itz.}(\epsilon)$ is not guaranteed to approximate $P_{itz.}$ "well enough." More specifically, one faces two challenges. First, for given i, one needs to quantify the *un*observed distance between $\delta_{t'}$ and δ_t on the basis of the observed distance between $P_{t'}$ and P_t. Second, one needs to control uniformly the approximation of $P_{itz.}(\epsilon)$ to $P_{itz.}$ for all $i = 1, ..., I$ as I increases. In the next Theorem, we address these complications and show $\hat{R}_{txy}(\epsilon)$'s sign consistency by relying on properties of the mixed logit model of demand for bundles from Iaria and Wang (2019).

Theorem 2. *Suppose that $R_{txy} \neq 0$ and that $\delta_{t'}$ is i.i.d. across markets according to some unknown distribution function in $\mathbb{R}^2$. Then, there exists $\epsilon_0 > 0$ such that, for any $\epsilon < \epsilon_0$,*

$$\lim_{I\to\infty}\lim_{T\to\infty} \Pr[sign(\hat{R}_{txy}(\epsilon)) = sign(R_{txy})] = 1.$$

Proof. See Appendix 6.3. □

5 Discussion

We investigate some features of the correlation criterion, an intuitively and practically appealing way of determining complementarity between products uniquely on the basis of their observed market shares. In those cases in which one can afford the estimation of a structural model of demand, the Hicksian criterion is clearly to be preferred. Among the main advantages, a structural model allows to measure the "intensity" of the complementarity or the substitutability (i.e., the magnitude of the estimated cross-price elasticities) and to simulate counterfactuals. However, when the estimation of appropriate structural models is hard, simpler options such as the correlation criterion may be necessary. In addition, even when the estimation of complex structural models is possible, tests based on the comparison of observed market shares can be helpful at the model selection stage. For example, ruling out complementarity can help motivate the estimation of simpler demand models: models allowing also for complementarity are remarkably more complex than models allowing only for substitutability (Berry et al. (2014), Dubé (2019), and Iaria and Wang (2019)).

6 Appendix

6.1 Preliminary Results

Here we present a simple result useful in proving Theorem 1 and Corollary 1.[6]

Lemma 1. *Suppose Assumption 1 holds and* $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$*. If there is* $(P'_{(x,y)}, P'_x, P'_y, P'_0) \in \mathbf{T}$*,* $(P'_{(x,y)}, P'_x, P'_y, P'_0) \neq (P_{(x,y)}, P_x, P_y, P_0)$*, such that* $P'_{(x,y)} \geq P_{(x,y)}$ *and* $P'_z \leq P_z$*, for* $z = x, y$*, then* $\delta'_z \geq \delta_z$*, for* $z = x, y$*.*

Proof. We prove this by contradiction. Without loss of generality, suppose that $\delta'_x < \delta_x$. Because $P_y(\delta)$ is a strictly decreasing function of δ_x, then $P'_y \leq P_y$ requires $\delta'_y < \delta_y$. Note that $P_{(x,y)}(\delta)$ is a strictly increasing function of both δ_x and δ_y. Consequently, $P'_{(x,y)} < P_{(x,y)}$, which contradicts $P'_{(x,y)} \geq P_{(x,y)}$. □

6.2 Proof of Theorem 1

By applying a Taylor expansion to $(P_x(\delta'_x, \delta'_y), P_y(\delta'_x, \delta'_y))$ around δ, we obtain:

$$\begin{bmatrix} P_x(\delta'_x, \delta'_y) - P_x(\delta_x, \delta_y) \\ P_y(\delta'_x, \delta'_y) - P_y(\delta_x, \delta_y) \end{bmatrix} = \frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)} \begin{bmatrix} \delta'_x - \delta_x \\ \delta'_y - \delta_y \end{bmatrix} + o(|(\delta'_x - \delta_x, \delta'_y - \delta_y)|).$$

[6] For brevity, we limit our discussion of feasibility to the case of $(P_{(x,y)}, P_x, P_y, P_0) \in \mathbf{C}$, but symmetric arguments hold also for observations in $\mathbf{S}$.

Then, by dividing both sides by $d' = |(\delta_{xm'} - \delta_{xm}, \delta_{ym'} - \delta_{ym})|$, we obtain:

$$\begin{bmatrix} P_x(\delta'_x, \delta'_y) - P_x(\delta_x, \delta_y) \\ P_y(\delta'_x, \delta'_y) - P_y(\delta_x, \delta_y) \end{bmatrix} \frac{1}{d'} = \frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)} \begin{bmatrix} \delta'_x - \delta_x \\ \delta'_y - \delta_y \end{bmatrix} \frac{1}{d'} + o(1).$$

Note that

$$\tilde{\delta}' = \begin{bmatrix} \delta'_x - \delta_x \\ \delta'_y - \delta_y \end{bmatrix} \frac{1}{d'} \tag{9}$$

is a unit vector. Consequently, for d' small enough, the sign of

$$\begin{bmatrix} P_x(\delta'_x, \delta'_y) - P_x(\delta_x, \delta_y) \\ P_y(\delta'_x, \delta'_y) - P_y(\delta_x, \delta_y) \end{bmatrix}$$

is determined by $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\tilde{\delta}'$. Note that the diagonal elements of $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}$, $\frac{\partial P_x}{\partial \delta_x}$, $\frac{\partial P_y}{\partial \delta_y}$, are positive, while the off-diagonal elements, $\frac{\partial P_x}{\partial \delta_y}$, $\frac{\partial P_y}{\partial \delta_x}$, are negative. Then

$$v_{(1,0)} = \frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial P_x}{\partial \delta_x} \\ \frac{\partial P_x}{\partial \delta_y} \end{bmatrix}$$

lies in the bottom-right part of the plane relative to $(0, 0)$, and

$$v_{(0,1)} = \frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\partial P_y}{\partial \delta_x} \\ \frac{\partial P_y}{\partial \delta_y} \end{bmatrix}$$

lies in the top-left part of the plane relative to $(0, 0)$.

(1) $\implies$ (2). It suffices to prove the existence of a $\tilde{\delta}' > 0$ such that $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\tilde{\delta}' < 0$. Because the determinant of $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}$ is negative, then the orientation of $v_{(1,0)}$ and $v_{(0,1)}$ is opposite to that of $(1, 0)$ and $(0, 1)$. Consequently, there exist $\lambda_1 > 0$ and $\lambda_2 > 0$, such that $\lambda_1 v_{(1,0)} + \lambda_2 v_{(0,1)}$ lies in the bottom-left part of the plane relative to $(0, 0)$: $\lambda_1 v_{(1,0)} + \lambda_2 v_{(0,1)} < 0$. Then, by choosing $\tilde{\delta}' = \frac{1}{\sqrt{\lambda_1^2+\lambda_2^2}}(\lambda_1, \lambda_2)^{\mathrm{T}}$ (or equivalently, $\delta'_x - \delta_x = \lambda_1$ and $\delta'_y - \delta_y = \lambda_2$), we obtain

$$\frac{\partial(P_x, P_y)}{\partial(\delta_x, \delta_y)}\tilde{\delta}' = \frac{\lambda_1}{\sqrt{\lambda_1^2 + \lambda_2^2}} v_{(1,0)} + \frac{\lambda_2}{\sqrt{\lambda_1^2 + \lambda_2^2}} \lambda_2 v_{(0,1)} < 0.$$

(2) $\implies$ (1). As shown in Lemma 1, any feasible point $\delta' = (\delta'_x, \delta'_y)$ for $\delta = (\delta_x, \delta_y)$ must satisfy $\delta'_x \geq \delta_x$ and $\delta'_y \geq \delta_y$. Then, according to (2), we have a sequence of δ' such that $\delta' \to \delta$, $\delta' \geq \delta$, $\delta' \neq \delta$, and $P_z(\delta') \leq P_z$ for $z = x, y$. As in (9), denote by $\tilde{\delta}'$ the unit vector equal to $(\delta'_x - \delta_x, \delta'_y - \delta_y)^{\mathrm{T}}$ divided by d'. Note that $\tilde{\delta}' \geq 0$ and $\tilde{\delta}' \neq 0$.

We prove the result by contradiction. Suppose that the determinant of $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}$ is positive. Then, according to the Inverse Function Theorem, there is a local bijection between $(P_x(\delta'),P_y(\delta'))$ and δ' in a neighborhood of δ. Therefore, $(P_x(\delta')-P_x, P_y(\delta')-P_y)\neq(0,0)$. Moreover, due to the positive determinant of $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}$, the orientation of $v_{(1,0)}$ and $v_{(0,1)}$ remains the same as $(0,1)$ and $(1,0)$. Consequently, for any $\lambda_1,\lambda_2\geq 0$ and $(\lambda_1,\lambda_2)\neq(0,0)$, $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}(\lambda_1,\lambda_2)^{\mathrm{T}}$ cannot lie in the bottom-left part of the plane relative to $(0,0)$. Let $\tilde{\delta}'=\frac{1}{\sqrt{\lambda_1^2+\lambda_2^2}}(\lambda_1,\lambda_2)^{\mathrm{T}}$ (or equivalently, $\delta'_x-\delta_x=\lambda_1$ and $\delta'_y-\delta_y=\lambda_2$). Then, at least one element of $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\tilde{\delta}'$ must be positive. In addition, we know that the sign of $(P_x(\delta')-P_x,P_y(\delta')-P_y)$ is determined by $\frac{\partial(P_x,P_y)}{\partial(\delta_x,\delta_y)}\tilde{\delta}'$. Consequently, at least one between $P_x(\delta')-P_x$ and $P_y(\delta')-P_y$ must be positive, which contradicts δ' being a feasible point for δ.

6.3 Proof of Theorem 2

Given ϵ, we first calculate $P_{itz.}(\epsilon)\equiv\lim_{T\to\infty}\hat{P}_{itz.}(\epsilon)$. Because $\delta_{t'}$ is i.i.d according to some distribution function $G(\cdot)$, for $z=x,y,(x,y)$, it follows that $\mathbf{1}_{|P_{t'}-P_t|<\epsilon}\mathbf{1}_{w_{it'}=z}$ is i.i.d. According to the law of large numbers, as $T\to\infty$

$$\begin{aligned}\hat{P}_{itz.}(\epsilon) &= \frac{1}{\sum_{t'=1}^{T}\mathbf{1}_{|P_{t'}-P_t|<\epsilon}}\sum_{t'=1}^{T}\mathbf{1}_{|P_{t'}-P_t|<\epsilon}[\mathbf{1}_{w_{it'}=z}+\mathbf{1}_{w_{it'}=(x,y)}]\\ &\to P_{itz.}(\epsilon)\\ &= \frac{\mathbb{E}[\mathbf{1}_{|P(\delta_{t'})-P_t|<\epsilon}[\mathbf{1}_{w_{it'}=z}+\mathbf{1}_{w_{it'}=(x,y)}]]}{\Pr[|P(\delta_{t'})-P_t|<\epsilon]}\\ &= \frac{\int_{|P(\delta_{t'})-P_t|<\epsilon}P_{it'z.}dG(\delta_{t'})}{\Pr[|P(\delta_{t'})-P_t|<\epsilon]}.\end{aligned}\tag{10}$$

Note that:

$$|P_{itz.}(\epsilon)-P_{itz.}|\leq\frac{\int_{|P(\delta_{t'})-P_t|<\epsilon}|P_{it'z.}-P_{itz.}|dG(\delta_{t'})}{\Pr[|P(\delta_{t'})-P_t|<\epsilon]}.\tag{11}$$

Moreover,

$$|P_{it'z.}-P_{itz.}|=|P_{z.}(\delta_{t'},\delta_{zi},\Gamma_i)-P_{z.}(\delta_t,\delta_{zi},\Gamma_i)|\leq\sup_{\delta\in\mathbb{R}^2}\left|\frac{\partial P_{z.}(\delta,\delta_{zi},\Gamma_i)}{\partial\delta}\right||\delta_{t'}-\delta_t|\,.$$

According to Iaria and Wang (2019) (see proof of Theorem 2, the "real analytic property"), $\sup_{\delta\in\mathbb{R}^2}\left|\frac{\partial P_{z.}(\delta,\delta_{zi},\Gamma_i)}{\partial\delta}\right|=A$, where A is finite. Moreover, Iaria and Wang (2019) (see Theorem 5, the "demand inverse") show that there is a global bijection between $P(\delta)$ and δ with everywhere positive-definite Jacobian. Consequently, $|P(\delta_{t'})-P_t|<\epsilon$ implies $|\delta_{t'}-\delta_t|\leq\eta(\epsilon,t)$, where

$\eta_t(\epsilon) \to 0$ as $\epsilon \to 0$. Then, we obtain:

$$|P_{it'z.} - P_{itz.}| = |P_{z.}(\delta_{t'}, \delta_{zi}, \Gamma_i) - P_{z.}(\delta_t, \delta_{zi}, \Gamma_i)| \le A\eta_t(\epsilon).$$

By plugging this into the right-hand side of (11), we obtain $|P_{itz.}(\epsilon) - P_{itz.}| \le A\eta_t(\epsilon)$. Define $\hat{R}_{txy} = \frac{1}{I}\sum_{i=1}^{I} P_{itx.}P_{ity.} - \frac{1}{I}\sum_{i=1}^{I} P_{itx.}\frac{1}{I}\sum_{i=1}^{I} P_{ity.}$. According to the law of large numbers, as $I \to \infty$, $\hat{R}_{txy}$ converges to R_{txy} in probability. Then, for any given I:

$$\begin{aligned}
\left|\lim_{T\to\infty} \hat{R}_{txy}(\epsilon) - R_{txy}\right| &= \left|\frac{1}{I}\sum_{i=1}^{I} P_{itx.}(\epsilon)P_{ity.}(\epsilon) - \frac{1}{I}\sum_{i=1}^{I} P_{itx.}(\epsilon)\frac{1}{I}\sum_{i=1}^{I} P_{ity.}(\epsilon) - R_{txy}\right| \\
&= \Big|\frac{1}{I}\sum_{i=1}^{I}(P_{itx.}(\epsilon) - P_{itx.} + P_{itx.})(P_{ity.}(\epsilon) - P_{ity.} + P_{ity.}) \\
&\quad - \frac{1}{I}\sum_{i=1}^{I}(P_{itx.}(\epsilon) - P_{itx.} + P_{itx.})\frac{1}{I}\sum_{i=1}^{I}(P_{ity.}(\epsilon) - P_{ity.} + P_{ity.}) - R_{txy}\Big| \\
&\le |\hat{R}_{txy} - R_{txy}| + 4A\eta_t(\epsilon).
\end{aligned} \tag{12}$$

Note that $\lim_{I\to\infty} |\hat{R}_{txy} - R_{txy}| \to 0$ as $I \to \infty$. Then, we can choose ϵ_0 such that $4A\eta_t(\epsilon_0) = 0.5|R_{txy}|$. Finally, for any $\epsilon < \epsilon_0$:

$$\limsup_{I\to\infty}\left|\lim_{T\to\infty} \hat{R}_{txy}(\epsilon) - R_{txy}\right| \le 0.5|R_{txy}|.$$

In other words, the distance between $\lim_{T\to\infty} \hat{R}_{txy}(\epsilon)$ and R_{txy} is bounded by $0.5|R_{txy}| < |R_{txy}|$ with probability one and therefore the sign of $\lim_{T\to\infty} \hat{R}_{txy}(\epsilon)$ is asymptotically equal to that of R_{txy} with probability one.

References

Allen, R. and Rehbeck, J. (2018). Hicksian complementarity and perturbed utility models. *Available at SSRN 2801345.*

Armstrong, M. (2016). Nonlinear pricing. *Annual Review of Economics*, 8:583–614.

Augereau, A., Greenstein, S., and Rysman, M. (2006). Coordination versus differentiation in a standards war: 56k modems. *The RAND Journal of Economics*, 37(4):887–909.

Berry, S., Khwaja, A., Kumar, V., Musalem, A., Wilbur, K. C., Allenby, G., Anand, B., Chintagunta, P., Hanemann, W. M., Jeziorski, P., et al. (2014). Structural models of complementary choices. *Marketing Letters*, 25(3):245–256.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Brynjolfsson, E. and Milgrom, P. (2013). Complementarity in organizations. *The handbook of organizational economics*, pages 11–55.

Chambers, C. P., Echenique, F., and Shmaya, E. (2010). On behavioral complementarity and its implications. *Journal of Economic Theory*, 145(6):2332–2355.

Cournot, A. (1838). *Researches into the Mathematical Principles of the Theory of Wealth.* Macmillan Co.

Dubé, J.-P. H. (2019). Microeconometric models of consumer demand.

Fox, J. T. and Lazzati, N. (2017). A note on identification of discrete choice models for bundles and binary games. *Quantitative Economics*, 8(3):1021–1036.

Gandhi, A., Lu, Z., and Shi, X. (2019). Estimating demand for differentiated products with zeroes in market share data.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *The American Economic Review*, 97(3):713–744.

Iaria, A. and Wang, A. (2019). Identification and estimation of demand for bundles. *Available at SSRN 3458543.*

Kretschmer, T., Miravete, E. J., and Pernías, J. C. (2012). Competitive pressure and the adoption of complementary innovations. *American Economic Review*, 102(4):1540–70.

Manzini, P., Mariotti, M., and Ülkü, L. (2018). Stochastic complementarity. *The Economic Journal.*

Milgrom, P., Roberts, J., et al. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *American economic review*, 80(3):511–528.

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):1255–1289.

Song, M., Nicholson, S., and Lucarelli, C. (2017). Mergers with interfirm bundling: a case of pharmaceutical cocktails. *The RAND Journal of Economics*, 48(3):810–834.

Thomassen, Ø., Smith, H., Seiler, S., and Schiraldi, P. (2017). Multi-category competition and market power: a model of supermarket pricing. *American Economic Review*, 107(8):2308–51.