Correction for Craword, Griffith and Iaria (2021) "A Survey of Preference Estimation with Unobserved Choice Set Heterogeneity" *Journal of Econometrics* 222:1 4-43*

Gregory S. Crawford† Rachel Griffith‡ Alessandro Iaria§

July 26, 2022

Correction

The Sufficient Set Logit (SSL) proposed by Crawford, Griffith, and Iaria (2021) in equation (3.5) is incorrect under Condition 1 (p. 12). The mistake can be fully corrected by a modification of the definition of sufficient set in Condition 1.[1] In summary, by requiring the definition of sufficient set to be "symmetric" among its elements, in a sense to be clarified below, the source of inconsistency is removed and the Conditional Maximum Likelihood Estimator (CMLE) of the resulting SSL in equation (3.5) will have desirable asymptotic properties on the basis of McFadden (1978). In the rest of this note, we discuss the source of inconsistency in SSL (3.5) under Condition 1, we propose an amended version of Condition 1, and then modify the definitions of FPH (equivalently IP) and PPH so that they satisfy the amended version of Condition 1. We then show that the results of our empirical illustration remain essentially unchanged when performed on the basis of this new definition.

The source of inconsistency of the CMLE of SSL (3.5) under Condition 1 can be readily seen as a violation of the "uniform conditioning property" by McFadden (1978). McFadden (1978) shows that when a researcher intends to estimate a Multinomial Logit (MNL) model with choice set $\mathscr{CS}$ from a subset $\mathscr{D} \subset \mathscr{CS}$, then she must augment the systematic utility of each alternative j, say

*We would like to thank Xavier D'Haultfœuille and Ao Wang for the insightful discussions and feedback on this note.
†Department of Economics, University of Zurich and CEPR, gregory.crawford@econ.uzh.ch
‡Department of Economics, University of Manchester and IFS, rgriffith@ifs.org.uk
§Corresponding author. Department of Economics, University of Bristol and CEPR, alessandro.iaria@bristol.ac.uk
[1] Guevara (2022) points out this inconsistency and proposes additional assumptions to overcome it.

V_j, by the probability with which the restricted choice set $\mathscr{D}$ is "sampled" conditional on alternative j being chosen, $V_j + \ln(\Pr[\mathscr{D}|j])$. The uniform conditioning property holds if for any $j, k \in \mathscr{D}$, $\Pr[\mathscr{D}|j] = \Pr[\mathscr{D}|k]$, so that all the terms $\ln(\Pr[\mathscr{D}|j]), j \in \mathscr{D}$, cancel out from the MNL expression based on the restricted $\mathscr{D}$, and the corresponding CMLE will have desirable properties (Andersen, 1970). This will, for example, be the case if $\mathscr{D}$ is a "fixed" subset of $\mathscr{CS}$ as in the classic specification test by Hausman and McFadden (1984), or if $\mathscr{D}$ is sampled by the researcher from the possible subsets of $\mathscr{CS}$ at random (uniformly). However, when the uniform conditioning property does not hold, for the CMLE of the MNL based on the restricted $\mathscr{D}$ to have desirable properties, the researcher will have to appropriately account for the additional terms $\ln(\Pr[\mathscr{D}|j])$, $j \in \mathscr{D}$.[2]

When $\mathscr{CS}_i^\star$ is the true but unobserved set of choice sequences of individual i, Crawford, Griffith, and Iaria (2021) define a "sufficient set" f as any correspondence satisfying the following property.

Condition 1. Given any choice sequence $Y_i \in \mathscr{CS}_i^\star$, the correspondence f is such that: (i) $Y_i \in f(Y_i)$ and (ii) $f(Y_i) \subseteq \mathscr{CS}_i^\star$.

Given Condition 1, for every individual i and choice sequence $Y_i = j$ such that $f(j) = r$, Crawford, Griffith, and Iaria (2021) then refer to the MNL conditional on $f(Y) = r$ in equation (3.5) as the SSL:

$$\Pr[Y_i = j | f(Y_i) = r, \theta] = \frac{\prod_{t=1}^{T} \exp(V(X_{ij_t t}, \theta))}{\sum_{k \in f(Y_i)=r} \prod_{t=1}^{T} \exp(V(X_{ik_t t}, \theta))} \tag{3.5}$$

and claim that "θ can be consistently estimated by the CMLE derived from $\Pr[Y_i = j|f(Y_i) = r, \theta]$ on the basis of McFadden (1978)." However, under Condition 1 the equality sign in (3.5) is incorrect, because in such case the right-hand side of (3.5) would only be a lower bound for $\Pr[Y_i = j|f(Y_i) = r, \theta]$. Hence, while it is true that "θ can be consistently estimated by the CMLE derived from $\Pr[Y_i = j|f(Y_i) = r, \theta]$ on the basis of McFadden (1978)," Crawford, Griffith, and Iaria (2021) do not provide the correct expression for $\Pr[Y_i = j|f(Y_i) = r, \theta]$ under Condition 1.[3]

The mistake arises in the second equality of (B.1) in Appendix B (p. 31). Under Condition 1, it is generally *not* the case that Y_i s.t. $f(Y_i) = r \iff Y_i \in r$, but only that Y_i s.t. $f(Y_i) = r \implies Y_i \in r$ (from Condition 1-(i)). Indeed, Condition 1 does not rule out cases of f with some $Y_i' \in f(Y_i) = r$

[2] Provided, of course, that $\Pr[\mathscr{D}|j] > 0$, $j \in \mathscr{D}$; what McFadden (1978) calls the "positive conditioning property."

[3] The correct expression for $\Pr[Y_i = j|f(Y_i) = r, \theta]$ under Condition 1 would replace "$k \in f(Y_i) = r$" in the denominator of (3.5) with "k s.t. $f(k) = r$." (We thank Xavier D'Haultfœuille for pointing this out in a private conversation.) An alternative route to the one we propose would then be to leave Condition 1 as it is and to instead amend equation (3.5). However, we find this option less appealing from a practical point of view, in that the sufficient set $f(Y_i)$ defined by the researcher would not necessarily correspond to the denominator of the SSL in (3.5), and this would considerably complicate implementation. We instead propose to amend the definition of sufficient set in Condition 1 so that equation (3.5) will remain unaltered and any (valid) sufficient set will automatically determine the denominator of the SSL.

such that $f(Y_i') \neq r$. This means that, under Condition 1, the second equality of (B.1) should instead be substituted by $\geq$ and that, in turn, the equality in (3.5) should be replaced by $\geq$. While under Condition 1 it is still possible to come up with examples of sufficient sets that satisfy (3.5) with an equality, such as the CP sufficient set proposed by Chamberlain (1980), this inequality would not in general be particularly useful. However, the equality in (3.5) holds by focusing on the following, more restrictive, class of sufficient sets.

Condition 1C. Given any choice sequence $Y_i \in \mathscr{CS}_i^\star$, the correspondence f is such that: (i) $Y_i \in f(Y_i)$, (ii) $f(Y_i) \subseteq \mathscr{CS}_i^\star$, and (iii) for any $Y_i' \neq Y_i$ such that $Y_i' \in f(Y_i)$, we have $f(Y_i') = f(Y_i)$.

Condition 1C augments Condition 1 by requirement (iii), which further restricts the class of sufficient sets to those that are "symmetric" in all their component sequences. The sufficient sets that satisfy Condition 1C are such that Y_i s.t. $f(Y_i) = r \iff Y_i \in r$, where it is clear that Condition 1C-(iii) specifically rules out the problematic cases with some $Y_i' \in f(Y_i) = r$ such that $f(Y_i') \neq r$. Condition 1C-(iii) avoids violations of the uniform conditioning property. In fact, under Condition 1, it is simple to find examples of sufficient sets for which $j, k \in f(j) = r$ but $\Pr[f(Y_i) = r | Y_i = j] = 1 \neq \Pr[f(Y_i) = r | Y_i = k] = 0$, a violation of the uniform conditioning property. Suppose that i's chosen sequence is $Y_i = (1, 2)$. Now consider the corresponding FPH sufficient set $f_{FPH}(1, 2) = \{(1, 1), (1, 2), (2, 1), (2, 2)\} = r$. It is then clear that the uniform conditioning property does not hold: while $\Pr[f_{FPH}(Y_i) = r | Y_i = (1, 2)] = \Pr[f_{FPH}(Y_i) = r | Y_i = (2, 1)] = 1$, unfortunately $\Pr[f_{FPH}(Y_i) = r | Y_i = (1, 1)] = \Pr[f_{FPH}(Y_i) = r | Y_i = (2, 2)] = 0$. Consequently, under Condition 1 the CMLE derived from (3.5) cannot in general be consistent. For the uniform conditioning property to hold in this example, the sequences $(1, 1)$ and $(2, 2)$ should be removed from the summation in the denominator of (3.5), giving rise to the CP sufficient set proposed by Chamberlain (1980): $f_{CP}(1, 2) = \{(1, 2), (2, 1)\}$. Condition 1C guarantees that the only sufficient sets to be considered are those that satisfy the uniform conditioning property, in that for any $j, k \in f(Y_i) = r$, $\Pr[f(Y_i) = r | Y_i = j] = \Pr[f(Y_i) = r | Y_i = k] = 1$.[4]

[4]For completeness, also Condition 3 in Appendix G (p. 40) in Crawford, Griffith, and Iaria (2021) should be amended along the same lines of Condition 1C.

Correct FPH and PPH Sufficient Sets

The FPH (equivalently the IP) and the PPH sufficient sets as defined in Crawford, Griffith, and Iaria (2021) (pp. 20-21) do not satisfy Condition 1C.[5] We propose more conservative definitions of these sufficient sets that satisfy Condition 1C and then illustrate their use with some examples. For the definitions that follow, suppose that i's observed choice sequence is $Y_i = (j_1, ..., j_t, ..., j_T) \in \mathscr{CS}_i^\star$.

Full Purchase History. An amended version of the FPH (equivalently the IP) sufficient set that satisfies Condition 1C is defined as:

$$f_{FPH}^C(j_1, ..., j_T) = \{(Y_1, ..., Y_T) : \forall t, Y_t \in \{j_1, ..., j_T\} \text{ and } |\{Y_1, ..., Y_T\}| = |\{j_1, ..., j_T\}|\},$$

where, to distinguish this version of FPH from the incorrect f_{FPH} reported in Crawford, Griffith, and Iaria (2021), we add a superscript C to its notation. Differently from f_{FPH}, f_{FPH}^C cannot be expressed as a cartesian product $\times_{t=1}^T f_{FPH,t}^C$ and consequently (3.5) will not simplify to (3.6). When all alternatives in the observed choice sequence appear only once (i.e., there is no repeated choice of any alternative across choice situations), so that $|\{j_1, ..., j_T\}| = T$, then $f_{FPH}^C = f_{CP}$. However, when $|\{j_1, ..., j_T\}| < T$, it is possible to express f_{FPH}^C as the union of CP sufficient sets $f_{FPH}^C = \bigcup_s f_{CP,s}$, where each $f_{CP,s}$ is simple to compute in practice. From this, it follows that $f_{CP} \subseteq f_{FPH}^C$. We illustrate this in practice in a few examples below.

Past Purchase History. An amended version of the PPH sufficient set that satisfies Condition 1C and preserves the cartesian structure $f_{PPH}^C = \times_{t=1}^T f_{PPH,t}^C$ is defined as:

$$f_{PPH,t}^C(j_1, ..., j_T) = \begin{cases} \{Y_t : Y_t = j_t\} & \text{if } t = 1 \\ \{Y_t : Y_t = j_t\} & \text{if } t > 1 \text{ and } j_t \notin \{j_1, ..., j_{t-1}\} \\ \{Y_t : Y_t \in \{j_1, ..., j_{t-1}\}\} & \text{if } t > 1 \text{ and } j_t \in \{j_1, ..., j_{t-1}\}. \end{cases}$$

In words, $\{j_1, ..., j_{t-1}\}$ is the collection of different alternatives chosen by i up to choice situation $t-1$. Therefore, $j_t \notin \{j_1, ..., j_{t-1}\}$ means that the alternative chosen in t is "new," in that i did not choose it in the previous choice situations. In practice, any choice situation t in which i is observed to choose an alternative j_t that she did not choose in the past will correspond to a singleton $f_{PPH,t}^C = \{j_t\}$, and

[5] As mentioned above, the CP sufficient set instead satisfies Condition 1C without modifications.

it will drop out of i's likelihood function. As a consequence, the choice situations used to construct i's likelihood function are those in which i chooses something that they have chosen in the past, and in those choice situations their sufficient set will correspond to everything they have ever chosen before, $f^C_{PPH,t} = \{j_1, ..., j_{t-1}\}$. Note that for f^C_{PPH} to contain at least two sequences, it is necessary that $T \geq 3$ and that the observed choice sequence is made of fewer than T different alternatives, $|\{j_1, ..., j_T\}| < T$. Similar to f_{CP}, also $f^C_{PPH} \subseteq f^C_{FPH}$.

Example 1: $Y_i = (1, 2, 3)$. In this case, $f^C_{FPH}(1, 2, 3) = f_{CP}(1, 2, 3) = \{(1, 3, 2), (1, 2, 3), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$ and $f^C_{PPH}(1, 2, 3) = \{(1, 2, 3)\}$. ■

Example 2: $Y_i = (1, 2, 2)$. Here $f_{CP}(1, 2, 2) = \{(1, 2, 2), (2, 1, 2), (2, 2, 1)\}$ and $f^C_{FPH}(1, 2, 2) = f_{CP}(1, 2, 2) \bigcup f_{CP}(1, 1, 2) = \{(1, 2, 2), (2, 1, 2), (2, 2, 1), (1, 1, 2), (1, 2, 1), (2, 1, 1)\}$. Moreover, $f^C_{PPH}(1, 2, 2) = \{1\} \times \{2\} \times \{1, 2\} = \{(1, 2, 1), (1, 2, 2)\}$. ■

Example 3: $Y_i = (3, 5, 5, 4)$. This is the example in Appendix F.1.2 (pp. 38-39) in Crawford, Griffith, and Iaria (2021), with $f_{CP}(3, 5, 5, 4) = \{(3, 5, 5, 4), (5, 3, 5, 4), (5, 5, 3, 4), (5, 5, 4, 3), (4, 3, 5, 5), (3, 4, 5, 5), (3, 5, 4, 5), (5, 3, 4, 5), (5, 4, 3, 5), (5, 4, 5, 3), (4, 5, 3, 5), (4, 5, 5, 3)\}$. In this case, $f^C_{FPH}(3, 5, 5, 4) = f_{CP}(3, 5, 5, 4) \bigcup f_{CP}(3, 3, 5, 4) \bigcup f_{CP}(4, 4, 5, 3)$, which we do not enumerate for brevity but that includes 36 choice sequences, and $f^C_{PPH}(3, 5, 5, 4) = \{3\} \times \{5\} \times \{3, 5\} \times \{4\} = \{(3, 5, 3, 4), (3, 5, 5, 4)\}$. ■

Example 4: $Y_i = (3, 5, 5, 3)$. Here $f_{CP}(3, 5, 5, 3) = \{(3, 5, 5, 3), (5, 3, 5, 3), (5, 5, 3, 3), (3, 3, 5, 5), (3, 5, 3, 5), (5, 3, 3, 5)\}$, $f^C_{FPH}(3, 5, 5, 3) = f_{CP}(3, 5, 5, 3) \bigcup f_{CP}(3, 5, 5, 5) \bigcup f_{CP}(5, 3, 3, 3)$, and $f^C_{PPH}(3, 5, 5, 3) = \{3\} \times \{5\} \times \{3, 5\} \times \{3, 5\} = \{(3, 5, 3, 3), (3, 5, 3, 5), (3, 5, 5, 3), (3, 5, 5, 5)\}$. ■

Corrected Empirical Example

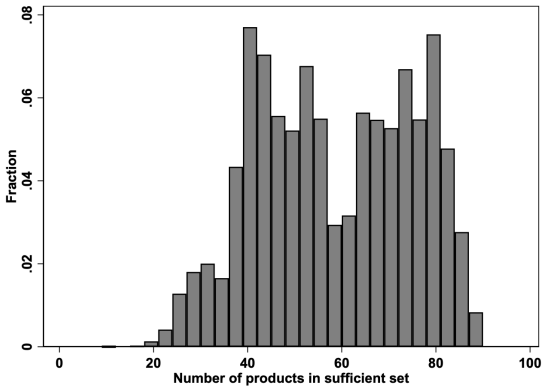
Section 5 of Crawford, Griffith, and Iaria (2021) includes an illustrative example of the use of the Past Purchase History sufficient set, f_{PPH}, to estimate parameters of a model of demand for chocolate purchased outside the home. We repeat the analysis using the correct definition of the sufficient set, f^C_{PPH}, and report here the figures and text that changes as a result. Overall the analysis remains essentially unchanged.

Figure 1 shows the distribution of sizes of the four sufficient sets used in estimation. Consistent with the more conservative definition, the number of products in the PPH sufficient set is smaller, as

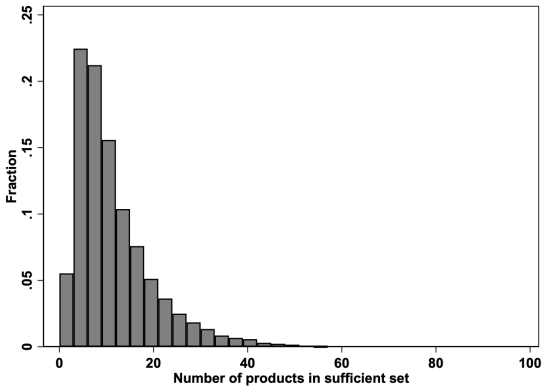
shown in the corrected version of Figure 1.

Figure 1: Number of Products in Sufficient Sets

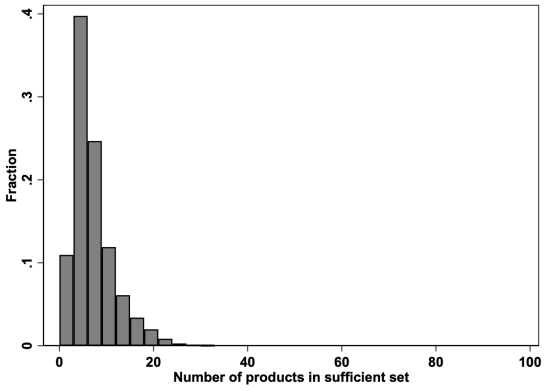
(a) Complete



(b) Past Purchase History



(c) Past Purchase History, last 12 months



(d) Past Purchase History, last 11 months

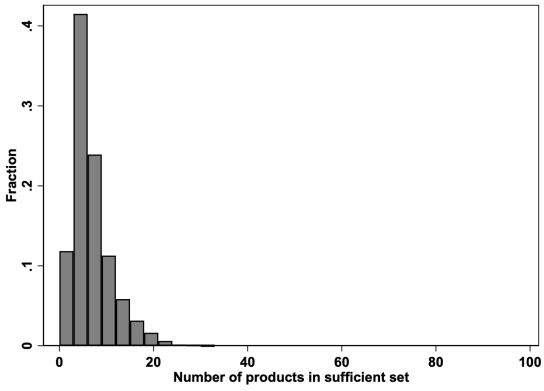


Table 3 presents the mean and standard deviation of the estimated price and advertising coefficients using each of the four sufficient sets. The estimates change somewhat, but the overall conclusion that the mean of the coefficient on price reduces substantially from the Complete sufficient set to the Past Purchase History, and reduces again when we use only information on purchases made in the year prior to the current purchase occasion still holds, as does the statement that the standard deviation of the individual estimates is smaller for the estimates using any of the Past Purchase History sufficient sets. Similarly, for the advertising coefficients, the mean of the estimates is higher when using the Complete sufficient set than when using any of the Past Purchase History sufficient sets.

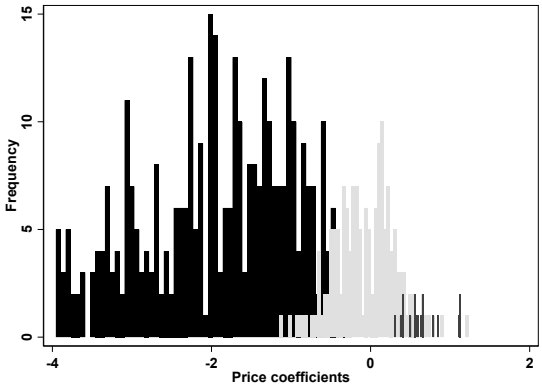
Table 3: Coefficient Estimates

		Complete	PPH	PPH 12 months	PPH 11 months
Price	Mean	-1.740	-0.775	-0.600	-0.578
	Std Dev	5.665	2.844	2.596	2.925
Advertising	Mean	0.165	0.047	0.029	0.027
	Std Dev	0.054	0.050	0.038	0.036
Product Effects		yes	yes	yes	yes
Time Effects		yes	yes	yes	yes

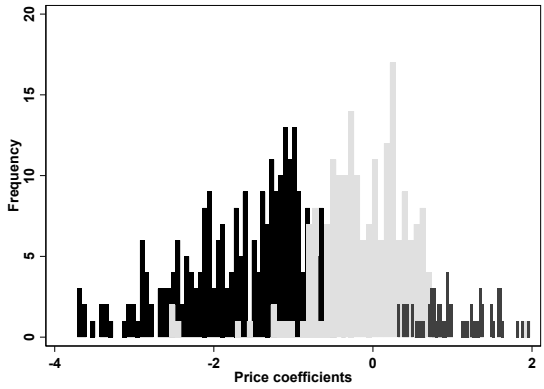
Figures 2 shows the distribution of the individual estimated price coefficients across the four sufficient sets; these look broadly similar.

Figure 2: Distributions of Estimated Price Coefficients

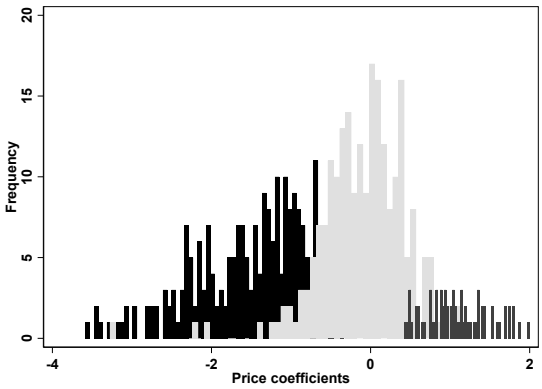
(a) Complete



(b) Past Purchase History



(c) Past Purchase History, last 12 months



(d) Past Purchase History, last 11 months

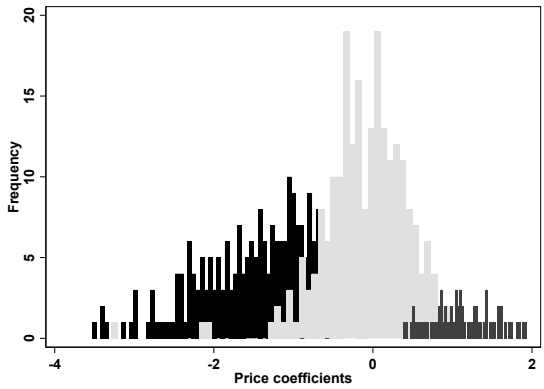


Table 4 reports some of the Hausman tests discussed in section 4.3.1 of Crawford, Griffith, and Iaria

Table 4: Hausman Tests

	% of sample with p-value on individual Hausman test			
	>0.1	0.05-0.1	0.01-0.05	<0.01
Price coefficients				
Complete v PPH	21.6	4.2	8.4	65.8
PPH v PPH 1 year	33.6	6.6	10.8	49.0
PPH 1 year v PPH 11 months	65.8	7.3	11.2	15.7
Advertising coefficients				
Complete v PPH	0.0	0.0	0.0	100.0
PPH v PPH 1 year	4.6	16.8	0.0	78.6
PPH 1 year v PPH 11 months	63.3	0.0	18.3	18.5

(2021). For both the price and the advertising coefficients, we find stronger evidence that individuals consider at least the chocolate bars they bought in the previous year. For example, for a larger share of the sample, 65.8% versus 37% in Crawford, Griffith, and Iaria (2021), the Hausman tests on the price coefficients are not rejected at the 10% when comparing PPH 12 months versus PPH 11 months.

In Table 5 in Crawford, Griffith, and Iaria (2021), we considered the complementary value of advertising following the ideas in Becker and Murphy (1993). The corrected values are reported in the corrected version of Table 5, and lead to essentially the same conclusions.

Table 5: Complementary Value of Advertising

	Complete	PPH	PPH 12 months	PPH 11 months
Mean	0.761	0.510	0.375	0.404
Std Dev	0.188	0.339	0.311	0.314

In the text in Crawford, Griffith, and Iaria (2021), we discussed how the estimates show that different assumptions about sufficient sets may have important practical consequences and lead to very different economic implications. The text remains essentially unchanged, but we include the corrected values here for completeness. At mean advertising, a one-standard deviation increase in the log advertising stock, $\ln(a_{ibt})$, equal to 0.69 (or 69%), implies an increase in the valuation of a product of 52.5 pence

when using the Complete sufficient set.[6] As the average price of a chocolate product is 58 pence, this is a 90% increase. By contrast, the estimates obtained using the PPH 12 months sufficient set suggest a one-standard deviation increase in the log advertising stock increases the value of a product by 25.9 pence, or a 45% increase.[7]

Concluding Remarks

While the FPH and PPH sufficient sets proposed by Crawford, Griffith, and Iaria (2021) on the basis of Condition 1 are theoretically incorrect (in the sense discussed above), the simulation results in Appendix D, Tables 6 and 7 (pp. 34-35), suggest that—in practice—the SSL (3.5) appears to be robust to this type of misspecification. In fact, the results in Table 6 provide no indication that the misspecified FPH SSL and PPH SSL (both incompatible with Condition 1C) perform any worse than the correctly specified CP SSL. Similarly, the results in Table 7 do not highlight any particular estimation bias for the misspecified PPH SSL. Importantly, this is not to say that the inconsistency discussed in this note does not matter, but only that the CMLE of SSL (3.5) shows some robustness against misspecifications of sufficient sets according to Condition 1C but still consistent with Condition 1. Along the same lines, the estimation results of the empirical illustration in Section 5 (p. 25) in Crawford, Griffith, and Iaria (2021) remain essentially unchanged when performed on the basis of the corrected version of the PPH SSL.

References

Andersen, E. B. (1970): "Asymptotic properties of conditional maximum-likelihood estimators," *Journal of the Royal Statistical Society: Series B (Methodological)*, 32(2), 283–301.

Becker, G. S., and K. M. Murphy (1993): "A simple theory of advertising as a good or bad," *The Quarterly Journal of Economics*, 108(4), 941–964.

Chamberlain, G. (1980): "Analysis of Covariance with Qualitative Data," *The Review of Economic Studies*, 47, 225–238.

Crawford, G. S., R. Griffith, and A. Iaria (2021): "A survey of preference estimation with unobserved choice set heterogeneity," *Journal of Econometrics*, 222(1), 4–43.

Guevara, A. (2022): "A Note on A survey of preference estimation with unobserved choice set heterogeneity by Gregory S. Crawford, Rachel Griffith, and Alessandro Iaria," *arXiv preprint arXiv:2205.00852*.

Hausman, J., and D. McFadden (1984): "Specification tests for the multinomial logit model," *Econometrica: Journal of the econometric society*, pp. 1219–1240.

McFadden, D. (1978): "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Residential Location*, ed. by A. K. et. al. North Holland Pub. Co.

[6] $(0.69 \times 0.761) = 0.525$, where 0.761 is the mean complementary value of advertising from Table 5 using the Complete sufficient set.

[7] $(0.69 \times 0.375) = 0.259$, where 0.0375 is the mean complementary value of advertising from Table 5 using the Past Purchase History 12 months.