No. DP19-2

RCESR Discussion Paper Series

Do Large-scale Point-of-sale Data Satisfy the Generalized Axiom of Revealed Preference in Aggregation Using Representative Price Indexes?: A Case Involving Processed Food and Beverages

March 2020

Hideyasu Sato, Hitotsubashi University

RCESR

The Research Center for Economic and Social Risks
Institute of Economic Research
Hitotsubashi University

2-1 Naka, Kunitachi, Tokyo, 186-8603 JAPAN
http://risk.ier.hit-u.ac.jp/

Do Large-scale Point-of-sale Data Satisfy the Generalized Axiom of Revealed Preference in Aggregation Using Representative Price Indexes?: A Case Involving Processed Food and Beverages[1]

Hideyasu Sato[2]
Hitotsubashi University

March 31, 2020

Abstract

The necessary conditions for data to be rationalized by weakly separable utility functions are verified by aggregation using representative price indexes. For processed food and beverages, the generalized axiom of revealed preference (GARP) is tested using large-scale product-level point-of-sale data. If GARP is not satisfied, the Afriat efficiency index (AEI) is introduced to assess the degree of optimization error. We find that the larger the number of observations in the time series direction, the less likely GARP is to be satisfied. However, the maximum level of AEI is, at most, more than 99.6%, indicating that the degree of the optimization error is small.

Keywords: Aggregation, Revealed preference, Weak separability, POS data
JEL Classification: C43, D12, Q11

---

[1] This work was supported by JSPS KAKENHI, 19H01467.
[2] Corresponding author address: Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo 186-8603, Japan; email: sato.hideyasu@gmail.com

1. Introduction

In the empirical analysis of consumer demand, it is standard to aggregate goods. This is done to save a degree of freedom in the parameter estimation of the demand structure model, avoid the problem of multicollinearity, and reduce the calculation load of computers (Caps and Love, 2002: p. 807). In addition, there is a case where consumption data that is already aggregated must be used from the viewpoint of availability. In this case, the aggregation method uses a typical index formula. For example, in food demand analysis using Japanese public data, it is common to obtain aggregated food prices from the Consumer Price Index (aggregated by the Laspeyre-type method) and to construct quantity indexes using the corresponding expenditure values from the *Survey of Household Economy*.

However, whether there exists a well-behaved utility function that rationalizes given aggregate demand and price data is an empirical question. This verification is usually conducted by testing whether the data can be rationalized by weakly separable utility functions. If there exists no weakly separable utility function through which it is possible to rationalize the data, the demand estimation carried out using aggregate data is erroneous. Eales and Unnevehr (1988), Nayga and Caps (1994), and Sellen and Goddard (1997) examined the weak separability of utility functions using econometric methods. In these studies, a parametric method was adopted after specifying a utility function form. In such a case, the problem arises that it is impossible to distinguish whether the rejection of weak separability is caused by an error in the function specification or whether the hypothesis of true weak separability is rejected[1].

Proposed by Varian (1983), one of the main approaches to avoid this problem is the use of nonparametric tests. The author shows necessary and sufficient conditions for the data to be rationalized by weakly separable utility functions and presents an algorithm for verification using the generalized axiom of revealed preference (GARP)[2]. The algorithm comprises three steps: (1) first, GARP is tested for data before aggregation; (2) if GARP is satisfied, then GARP is tested for each commodity group after aggregation; (3) if GARP is also satisfied in the second step, GARP is also tested for aggregation data. In the second step, parameters called "Afriat numbers" that satisfy GARP in this step are chosen

[1] These studies are based on Goldman and Uzawa (1964), and are based on the fact that the Slutsky substitution matrix term among goods in different goods groups is proportional to the income effect. In recent years, parametric tests using the methods of Blackorby et al. (1991) and Moschini et al. (1994), which are generalized versions of Goldman and Uzawa (1964), have been studied (Dhar et al., 2003; Frumouzi et al., 2012; Lakkakula et al., 2016).

[2] This research relies on the studies of data rationalization by Afriat (1967), Diwert (1973), and Varian (1982).

accordingly. Afriat numbers correspond to the inverse of the price indexes and quantity indexes of the commodity group after aggregation. Even if GARP is rejected in the third step, it may pass the GARP test in this step under new Afriat numbers that pass the GARP test in the second step. Therefore, Varian's (1983) algorithm is a sufficient condition for weak separability.

We focus on verifying data rationalizability under the assumption of aggregation using representative price indexes, as is generally done in the aggregation of public data. Corresponding to Varian's (1983) algorithm, after the GARP test of the first step, the Afriat numbers are calculated in advance using the representative indexes and used to conduct the GARP tests of the second and third steps. In this study, the first and third steps of the GARP tests are carried out to verify the necessary conditions for data rationalization by aggregation using representative price indexes.

As an approach different from the test of weak separability, Lewbel (1996) proposed a method using the test of the generalized composite commodity theorem (GCCT). The GCCT showed that the demand function for composite goods satisfies some desirable properties of the demand function if the logarithmic price ratio of both composite and individual goods is independent of the logarithmic price of composite goods. Since Lewbel (1996), there has been an increasing number of studies testing the GCCT for the rationalizability of food and beverage demand data. According to the reviews by Shumway and Davis (2001: p.167) up to 2000, the rationalization of food and beverage data tends not to be rejected by the GCCT rather than by separability[3]. Subsequent studies validating the GCCT, such as Reed et al. (2005), Xie and Myrland (2011), Schulz et al. (2012), and Heng et al. (2018), have provided some acceptance for the rationalization of aggregate food and beverage demand data.

On the other hand, these previous studies were conducted to examine the rationalization of data by aggregating goods already aggregated into larger groups of goods. Indeed, the demand theory is based on the rational consumption behavior for individual goods. Therefore, verification using data at the product level should be carried out. In addition, research using microdata, such as scanner data, targets limited goods groups, such as "carbonated drink" or "ground meat". This implicitly assumes weak separability for other large amounts of food and beverages.

In this study, the necessary condition for the weak separability of the utility function is examined nonparametrically using large-scale point-of-sale (POS) data that record purchase information of processed food and beverages at the product level. The aims of

[3] However, it should be noted that the review by Shumway and Davis (2001) includes homothetic separability and verification on the production side.

this study are as follows: (1) to avoid specification problems using a non-parametric test; (2) to ensure consistency with consumer theory by using non-aggregate data at the product level; and (3) to cover a wide range of goods and commodity groups for food and beverage demand. Verification covering a wide range of food and beverage products characterizes household consumption behavior at the food and beverage expenditure stage based on data.

The remainder of this paper is organized as follows. Section 2 confirms the main theorems of data rationalization, and explains the analytical procedures of this study. Section 3 describes the data used in the analysis and Section 4 describes the results of the analysis. Finally, Section 5 presents the conclusion and scope for future work.

2. Confirmation of Theorems on Weak Separability Test

2. 1. Data Rationalization

First, we confirm the theorems on the rationalizability of data by Afriat (1967), Diewert (1973), and Varian (1982).

Definition 1: *Define* $D = \{\mathbf{p}_{\mathrm{t}}; \mathbf{x}_{\mathrm{t}}\}_{t\in T}$ *as the data set consisting of a strictly positive price vector* $\mathbf{p}_{\mathrm{t}} \in \mathbb{R}^n_{++}$ *and a nonnegative consumption bundle* $\mathbf{x}_{\mathrm{t}} \in \mathbb{R}^n_{+}$ *at observation* $t \in T$ *respectively. The data set* D *is rationalizable if there exists a well-behaved utility function* $u: \mathbb{R}^n_+ \to \mathbb{R}$ *that satisfies the following condition* (1)[4]*:*

$$\forall t \in T\ \mathbf{x}_t \in \arg\max_{\mathbf{x}} u(\mathbf{x}) \ \text{ s.t.}\, \mathbf{p}_t\mathbf{x} \leq \mathbf{p}_t\mathbf{x}_t. \tag{1}$$

Definition 2: *We say that* $\mathbf{x}_t$ *is directly revealed preferred to* $\mathbf{x}$ *at observation* $t \in T$ *if* $\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}$ *and is expressed as* $\mathbf{x}_t R^D \mathbf{x}$. *We say that* $\mathbf{x}_t$ *is strictly directly revealed preferred to* $\mathbf{x}$ *if* $\mathbf{p}_t\mathbf{x}_t > \mathbf{p}_t\mathbf{x}$ *and is expressed as* $\mathbf{x}_t P^D \mathbf{x}$. *We say that* $\mathbf{x}_t$ *is revealed preferred to* $\mathbf{x}$ *if there exists a sequence* $\mathbf{x}_t R^D \mathbf{x}_j, \mathbf{x}_j R^D \mathbf{x}_k, \dots, \mathbf{x}_m R^D \mathbf{x}$ *and is expressed as* $\mathbf{x}_t R \mathbf{x}$. *The relation* R *is called transitive closure of* R^D.

Definition 3: *We say that the data set* $D = \{\mathbf{p}_t; \mathbf{x}_t\}_{t\in T}$ *satisfies GARP if* $\mathbf{x}_t R \mathbf{x}_s$ *does not imply* $\mathbf{x}_s P^D \mathbf{x}_t$, *that is,* $\mathbf{x}_t R \mathbf{x}_s \Rightarrow \mathbf{p}_s\mathbf{x}_s \leq \mathbf{p}_s\mathbf{x}_t$.

[4] Well-behaved utility functions are those that satisfy monotonicity, continuity, and concavity. $\mathbf{px}$ in Eq.(1) means the inner product $\sum_{i=1}^{N} p_i x_i$, where $i = 1, \dots, N$ indicates the individual goods numbers.

Theorem 1: *The following conditions* (i) – (iv) *are equivalent:*

(i) *there exists a non-satiated utility function that rationalizes the data set* $D = \{\mathbf{p}_t; \mathbf{x}_t\}_{t \in T}$;

(ii) *the data set* $D = \{\mathbf{p}_t; \mathbf{x}_t\}_{t \in T}$ *satisfies GARP*;

(iii) *there exist numbers* $U_t, \lambda_t > 0\ \forall t \in T$ *that satisfy the following inequality* (2):

$$\forall t, s \in T\ U_t \leq U_s + \lambda_s \mathbf{p}_s(\mathbf{x}_t - \mathbf{x}_s). \tag{2}$$

(iv) *there exists a well-behaved utility function that rationalizes the data set* $D = \{\mathbf{p}_t; \mathbf{x}_t\}_{t \in T}$.

Satisfying GARP for the data set D is a necessary and sufficient condition for the existence of a well-behaved utility function that rationalizes the data set. Equation (2) is Afriat's inequality, where U_t and U_s are interpreted as utility levels and λ_s as the marginal utility of income.

2. 2. Rationalization of aggregate data

Next, we confirm the theorems on rationalizability of aggregate data by Varian (1983).

Definition 4: The *data set* D *is divided into two commodity groups and rewritten as* $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$. $\mathbf{q}_t$ *is a price vector corresponding to the consumption bundle* $\mathbf{y}_t$. *The data set* D *is rationalizable by a weakly separable utility function, which means that there exists a well-behaved macro-utility function* u *and a sub-utility function* v *that satisfy the following condition* (3):

$$(\mathbf{x}_t, \mathbf{y}_t) \in \arg\max_{\mathbf{x},\mathbf{y}} u(v(\mathbf{x}), \mathbf{y})\ \ \mathrm{s.t.}\, \mathbf{p}_t\mathbf{x} + \mathbf{q}_t\mathbf{y} \leq \mathbf{p}_t\mathbf{x}_t + \mathbf{q}_t\mathbf{y}_t. \tag{3}$$

Theorem 2: *The following conditions* (i) - (iv) *are equivalent:*

(i) *there exists a non-satiated utility function that rationalizes the data set* $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t \in T}$;

(ii) *there exist numbers* $U_t, V_t, \lambda_t > 0, \mu_t > 0\ \forall t \in T$ *that satisfy the following inequalities* (4) *and* (5):

$$\forall t, s \in T\ U_t \leq U_s + \lambda_s \left[\mathbf{q}_s(\mathbf{y}_t - \mathbf{y}_s) + \frac{1}{\mu_s}(V_t - V_s)\right], \tag{4}$$

$$\forall t, s \in T \; V_t \leq V_s + \mu_s \mathbf{p}_s(\mathbf{x}_t - \mathbf{x}_s); \tag{5}$$

(iii) *the data sets* $\{\mathbf{p}_t; \mathbf{x}_t\}_{t\in T}$ *and* $\{1/\mu_t, \mathbf{q}_t; V_t, \mathbf{y}_t\}_{t\in T}$ *satisfy GARP under specific numbers* (V_t, μ_t);

(iv) *there exists a well-behaved utility function that rationalizes the data set* $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t\in T}$;

The numbers $1/\mu_t$ and V_t in (iii) are interpreted as the price and quantity of the goods group $\mathbf{x}_t$ aggregated by the sub-utility function v, respectively, and the numbers (V_t, μ_t) are called "Afriat numbers."

Here, the correspondence with Varian's (1983) algorithm mentioned in Section 1 is confirmed once again. First, we test GARP of the data set $D = \{\mathbf{p}_t, \mathbf{q}_t; \mathbf{x}_t, \mathbf{y}_t\}_{t\in T}$. If GARP is rejected at this stage, weak separability is rejected; if accepted, we proceed to the next step. Next, GARP of the data set $\{\mathbf{p}_t; \mathbf{x}_t\}_{t\in T}$ is tested. At this stage, if GARP is rejected, weak separability is rejected; if accepted, we proceed to the next step. Finally, GARP of the data set $\{1/\mu_t^*, \mathbf{q}_t; V_t^*, \mathbf{y}_t\}_{t\in T}$ is tested, where (V_t^*, μ_t^*) are the specific Afriat numbers that pass the GARP test in the second step. If GARP is passed at this final stage, weak separability is accepted. On the other hand, if GARP is rejected, weak separability cannot be rejected. Because the GARP test at the final stage depends on the value of $(V_t^* \mu_t^*)$, Afriat numbers other than those accepted $(V_t^* \mu_t^*)$ may still pass the GARP test at the final stage. Therefore, Varian's (1983) algorithm is recognized as a sufficient condition for weak separability[5].

Diewert and Parkan (1985) generalize the assumption of data aggregation in Varian (1983) and show the necessary and sufficient conditions for data rationalization. Definition 5 below is based on the study by Diewert and Parkan (1985) and describes the premise of this study's analysis:

Definition 5: *The data set* D *is divided into goods groups of* $h = 1, \dots, H$ *and rewritten as* $D = \{\mathbf{p}_t^1, \dots, \mathbf{p}_t^H; \mathbf{x}_t^1, \dots, \mathbf{x}_t^H\}_{t\in T}$. *The data set* D *can be rationalized by a weakly separable utility function, which means that there exists a well-behaved macro-utility function* u *and a sub-utility function* v *that satisfy condition* (6) *below*[6]:

[5] Barnett and Choi (1989), Fleissig and Whitney (2003), and Hjertstrand (2009) conducted some simulations to confirm the sufficient conditions.

[6] In Diewert and Parkan (1985: pp.133-142), the arguments of the macro utility function take the subutility function of $h = 1, \dots, H$ and one nonaggregate good. In this sense, the setting in this study is more specific than that of Diewert and Parkan (1985).

$$(\mathbf{x}_t^1, \dots, \mathbf{x}_t^H) \in \arg\max_{\mathbf{x}^1,\dots,\mathbf{x}^H} u\big(v_1(\mathbf{x}^1), \dots, v_H(\mathbf{x}^H)\big)$$

$$\text{s.t.} \sum_{h=1}^{H} \mathbf{p}_t^h \mathbf{x}^h \leq \sum_{h=1}^{H} \mathbf{p}_t^h \mathbf{x}_t^h. \tag{6}$$

Theorem 3: *The following conditions* (i) - (iv) *are equivalent*:

(i) *there exists a non-satiated utility function that rationalizes the data set* $D = \{\mathbf{p}_t^1, \dots, \mathbf{p}_t^H; \mathbf{x}_t^1, \dots, \mathbf{x}_t^H\}_{t\in T}$;

(ii) *there exist numbers* $U_t, V_t^h, \lambda_t > 0, \mu_t^h > 0\ \forall t \in T, h = 1, \dots, H$ *that satisfy the following inequalities* (7) *and* (8):

$$\forall t, s \in T\ U_t \leq U_s + \lambda_s \left[\sum_{h=1}^{H} \frac{1}{\mu_s^h}\left(V_t^h - V_s^h\right)\right], \tag{7}$$

$$\forall t, s \in T, h = 1, \dots, H\ V_t^h \leq V_s^h + \mu_s^h \mathbf{p}_s^h\left(\mathbf{x}_t^h - \mathbf{x}_s^h\right). \tag{8}$$

(iii) *the data sets* $\{\mathbf{p}_t^h; \mathbf{x}_t^h\}_{t\in T}^{h=1,\dots,H}$ *and* $\{1/\mu_t^1, \dots, 1/\mu_t^H; V_t^1, \dots, V_t^H\}_{t\in T}$ *satisfy GARP under specific numbers* $(1/\mu_t^1, \dots, 1/\mu_t^H, V_t^1, \dots, V_t^H)$;

(iv) *there exists a well-behaved utility function that rationalizes the data set* $D = \{\mathbf{p}_t^1, \dots, \mathbf{p}_t^H; \mathbf{x}_t^1, \dots, \mathbf{x}_t^H\}_{t\in T}$.

2. 3. Analytical procedure used in this study

When the data set is divided into $h = 1, \dots, H$, and the existence of well-behaved sub-utility functions that rationalize the data is considered in each goods group, the number of aggregation patterns is as many as 2^H. Considering the large number of goods in POS data, the combination of sub-utility functions $\{v_1(\mathbf{x}^1), \dots, v_H(\mathbf{x}^H)\}$ is used only to aggregate the divided data set $D = \{\mathbf{p}_t^1, \dots, \mathbf{p}_t^H; \mathbf{x}_t^1, \dots, \mathbf{x}_t^H\}_{t\in T}$ under the setting in Definition 5 to verify weak separability. Since quantities and prices are collected and used in one country as a whole for each commodity, the utility maximization behavior of one representative individual is assumed.

In the analysis procedure, the nonparametric GARP test is carried out for the data set before aggregation[7]. We introduce the Afriat efficiency index (AEI) $e \in (0,1]$ to search

[7] The GARP tests are performed using the package "revealedPrefs" of statistical software R. See Boelaert (2019) for more information.

for the maximum efficiency level at which GARP is satisfied, although the data set is not rationalized by the weakly separable utility function unless it passes the GARP test at this stage[8]. The following is the definition of GARP with the AEI:

Definition 6: *We say that* $\mathbf{x}_t$ *is directly revealed preferred to* $\mathbf{x}$ *at efficiency level* $e \in (0,1]$ *if* $e^t\mathbf{p}_t\mathbf{x}_t \geq \mathbf{p}_t\mathbf{x}$ *at observation* $t \in T$ *and is expressed as* $\mathbf{x}_t R^D(e)\mathbf{x}$. *Let the transitive closure of* $R^D(e)$ *be* $R(e)$. *We say that the data set meets GARP at efficiency level* e *if* $\mathbf{x}_t R(e)\mathbf{x}_s \Rightarrow e\mathbf{p}_s\mathbf{x}_s \leq \mathbf{p}_s\mathbf{x}_t$ *is satisfied.*

$e = 1$ is equal to the ordinary GARP inequality. $1 - e$ is interpreted as the percentage of expenditure wasted by consumer optimization errors. If GARP is not satisfied for the data set before aggregation, the level of e is lowered and GARP is tested again, and this is repeated until the data set is rationalized. The degree of the optimization error of the representative consumer is confirmed by obtaining the maximum efficiency level $e = e^*$ that passes the GARP test.

Next, the data set is divided. The division is based on the classification of goods by INTAGE Inc., which provides POS data for analysis[9]. GARP is tested on the data set $\{\mathbf{p}_t^h; \mathbf{x}_t^h\}_{t\in T}^{h=1,\dots,H}$ by the goods group, and the degree of the optimization error is confirmed by introducing the AEI and the test of the data set before aggregation when the test is not passed.

The last is the GARP test of the dataset $\{1/\mu_t^1, \dots, 1/\mu_t^H; V_t^1, \dots, V_t^H\}_{t\in T}$. It is necessary to obtain the unknown Afriat numbers $(1/\mu_t^1, \dots, 1/\mu_t^H, V_t^1, \dots, V_t^H)$. In this study, representative price indexes and quantity indexes obtained by dividing the value by the price indexes are created according to the goods group classification and used as Afriat number[10]. Laspeyres-type PL, Paasche-type PP, Fisher-type PF, and Töurnqvist-type PT are adopted for the price index. For all $t \in T$ and $h = 1, \dots, H$, the respective index formulas are as in Equations (9) – (12) below[11]:

---

[8] The name "AEI" follows Varian (1994).

[9] Theoretically, the economic model should be constructed in the general equilibrium framework using information from the supply side, and the goods group that can be aggregated should be discussed. On the other hand, since this is a large-scale method for estimating the entire target supply and demand system, this study carries out the verification by giving the commodity group classification exogenously from the viewpoint of feasibility.

[10] It should be noted that even if the GARP test based on the chosen Afriat numbers is rejected, there is still a possibility of rationalizing the data using other Afriat numbers. In this study, it is assumed that the aggregation of goods (price) is based on a typical index formula.

[11] Fisher-type and Töurnqvist-type price indexes can be interpreted as the cost-of-living index from the economic approach in the index theory. The cost-of-living index is defined by the ratio of the expenditure function when the utility level is fixed, the Fisher-type is assumed to be a specific homothetic utility

$$PL_t^h = \frac{\mathbf{p}_t^h \mathbf{x}_0^h}{\mathbf{p}_t^h \mathbf{x}_0^h}, \tag{9}$$

$$PP_t^h = \frac{\mathbf{p}_t^h \mathbf{x}_t^h}{\mathbf{p}_t^h \mathbf{x}_{\boldsymbol{t}}^h}, \tag{10}$$

$$PF_t^h = \sqrt{PL_t^h * PP_t^h} \tag{11}$$

$$\ln PT_t^h = 0.5 * \left[\mathbf{w}_0^h\left(\ln \mathbf{p}_t^h - \ln \mathbf{p}_0^h\right) + \mathbf{w}_t^h\left(\ln \mathbf{p}_t^h - \ln \mathbf{p}_0^h\right)\right],$$

$$\mathbf{w}_0^h = \left\{\frac{p_0^{h,i} x_0^{h,i}}{\mathbf{p}_0^h \mathbf{x}_0^h}\right\}^{i \in I(h)}, \mathbf{w}_{\boldsymbol{t}}^h = \left\{\frac{p_t^{h,i} x_t^{h,i}}{\mathbf{p}_t^h \mathbf{x}_t^h}\right\}^{i \in I(h)}. \tag{12}$$

In the final stage, the degree of optimization error is confirmed by introducing the AEI.

3. Data

This study uses the Nationwide Retail Store Panel Survey (SRI) by INTAGE Inc., which includes large-scale POS data collected from about 4,000 supermarkets, convenience stores, discount stores, and drug stores in Japan. SRI records the values of product sales and the sales quantities of food and daily necessities by stores and goods.

The sales values and quantities across stores selling the same product are totaled at the national level, and the price of individual goods is calculated as the unit price obtained by dividing the sales value by the sales quantity. The goods to be analyzed comprises 538,704 products including items such as rice, bread, noodles, dairy products, frozen foods, processed meat and fishery products, canned foods, confectioneries, seasonings, and non-alcoholic drinks. Product data for which no purchases are observed during the analysis period are excluded[12].

The weekly data comprise a total of 626 weeks from the first week of January 2006 to the last week of December 2017, and are tabulated at the monthly and annual levels considering the frequency of purchase and storage of goods. The month (or year) to which the first day of the week belongs is treated as data for that month (or year) on both counts[13].

---

function, and the Töurnqvist-type is assumed to be a translog expenditure function that does not assume homotheticity.

[12] It should be noted that the selection bias is caused by missing observations. The remedy for this will be in future work.

[13] Since the number of stores surveyed increases to over 200 as of September 2014, the verification that divided the data into two in September 2014 is also shown in the Appendix.

Monthly data cover 144 months, while annual data cover 12 years. In addition, to confirm the robustness of the long and short analysis periods, the data of the oldest month (or year) are successively erased by month (or year), and the GARP test is repeated. Therefore, the base time for calculating each price index is the last month (or year) of the data so that it does not change with the length of the data in the time series direction.

Figure 1 is a line graph showing the number of deleted months plus one on the horizontal axis and the coverage ratio of purchased products on the vertical axis. “Coverage ratio of purchased products” means “the ratio of the number of products purchased continuously to the number of products purchased at least once during the analysis period.” For example, when the value of the horizontal axis is 1, the value of the vertical axis is approximately 0.027. This indicates that, using monthly data for the entire 144 months, of the 538,704 products that are observed to be purchased more than once, 14,672 (about 2.7%) are purchased continuously. The coverage ratio of purchased products increases monotonously as the period becomes shorter, and the coverage ratio in the case of using two months, November and December, 2017, is approximately 21%. The longer the period, the lower the likelihood of continuous purchase, and new products will drop out. The coverage ratio of annual data is illustrated in Figure 2. The percentage of missing observations is slightly lower for annual than monthly totals.

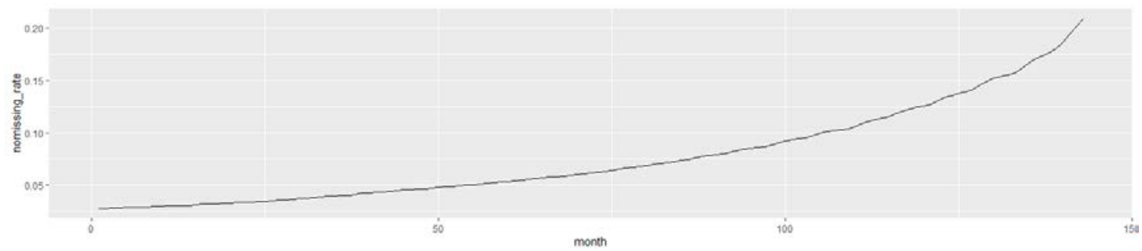

Figure 1: Removal Months and Coverage Ratio of Purchased Products

Note: The horizontal axis represents the number of months deleted from the first month of the total 144 months plus one, while the vertical axis represents the ratio of the number of products continuously purchased out of the products purchased at least once during the analysis period.

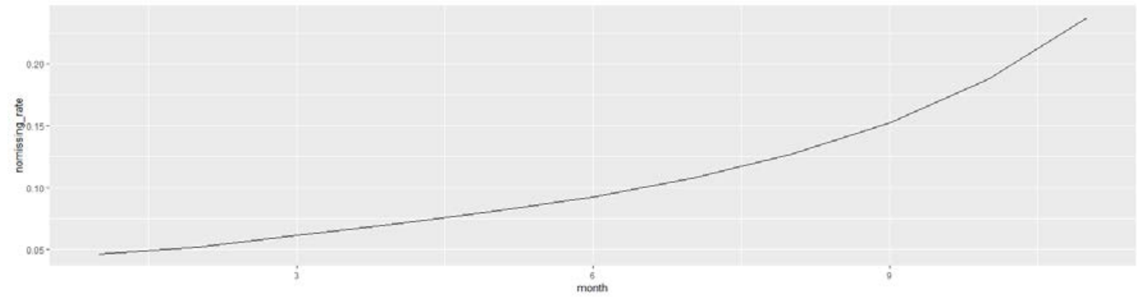

Figure 2: Removal Years and Coverage Ratio of Purchased Products

Note: The horizontal axis represents the number of years deleted from the first year of the total 12 years plus one, while the vertical axis represents the ratio of the number of products continuously purchased out of the products purchased at least once during the analysis period.

The data are divided and aggregated based on the commodity group classification by INTAGE Inc., which comprises Category 1 (major division) and Category 2 (subdivision). Category 1 includes 140 goods groups, while Category 2 includes 1,109 goods groups nested within Category 1[14].

4. Analysis results

4. 1. The GARP test of pre-aggregation data

The GARP test of the product level data set is carried out before the goods are aggregated. The data set analysis period is 144 months for monthly data and 12 years for annual data. The results of the GARP test show that the data set is rationalized using both monthly and annual data. In addition, the results obtained are robust to the analysis period. In other words, the GARP test consistently passes, even though the GARP test is repeated by deleting the time series observation of one month (or year) at a time from the oldest month (or year) in the analysis period. Therefore, we conclude that the large-scale POS data used are rationalized by a well-behaved utility function.

4. 2. The GARP tests of aggregate data

The GARP tests are performed with data constructed by goods aggregation. Goods are aggregated into either Category 1 or 2. In either case, by defining the commodity group as $h = 1, \dots, H$, the data set is represented as $D = \{1/\mu_t^1, \dots, 1/\mu_t^H; V_t^1, \dots, V_t^H\}_{t \in T}$. The values $\left(1/\mu_t^h, V_t^h\right), h = 1, \dots, H\ \forall t \in T$ expressed by Afriat numbers are price indexes

[14] Depending on the analysis period, the number of products may be less than or equal to the aforementioned number of product groups.

and quantity indexes. The price indexes are determined in advance by the representative aggregation method, and the corresponding quantity indexes are calculated by dividing the sales values by the price indexes.

The results of the GARP tests using monthly data are illustrated in Figures 3 and 4, which are based on data sets using Categories 1 and 2, respectively. The horizontal axis in each figure represents the number of months deleted from the first month of the monthly data of all 144 months plus one, and the vertical axis represents the maximum AEI level e^* that passed the GARP test. For example, when the horizontal axis value is 1, it represents e^* in the case of using data for 144 months from January 2006, and when it is 2, it represents e^* in the case of using data for 143 months from February 2006.

First, let us look at the movement of e^* using the Laspeyres price index shown in Figure. 3, represented by solid lines. The movement of e^* fluctuates until the months from 2006 to 2007 are removed; some of them pass the GARP test, while others show the lowest e^* value in the graph. When the number of deleted months exceeds about two years, the movement of e^* becomes stable and always takes a value above approximately 0.9990. In particular, the GARP test is often passed if the data used cover less than 50 months, and the value of e^* does not fall below 0.9995 even if the data does not pass the GARP test.

Next, let us focus on the movement of e^* when using the Paasche price index, represented by a dotted line in Figure. 3. As with the Laspeyres price index, e^* moves strongly until the months from 2006 to 2007 are removed. However, the value of e^* is often less than or equal to approximately 0.9990, and therefore, the degree of optimization error is small compared with the results obtained using other price indexes. When the number of deleted months exceeds about two years, e^* becomes 1 in most cases and the tendency to pass the GARP test is widely observed.

The results using the Fisher and the Töurnqvist price indexes are indicated in Figure 3 by dashed and long-dashed lines, respectively. As with the Laspeyres and Paasche price indexes, these results show a large movement of e^* until the months around 2006 to 2007 are removed. The results using these two indexes appear relatively similar.

Whichever index formula is used, there is a common point that the movement of e^* is large until about the first 25 months are deleted. As the number of deleted months decreases, the value of e^* increases and shows a stable tendency. The GARP test tends to be passed for monthly data of about four years. The most important point is that even if the GARP test is not passed, the value of $1 - e^*$ does not exceed 0.004, and therefore, the degree of optimization error is minor.

We now proceed to discuss Figure 4, which shows the result of using the Category 2 commodity group classification. The results of all index formulas show that the movement of e^* is very similar to that of the Category 1 commodity group. Although the reasons for this cannot be clearly stated, one factor may be that Category 2 is nested within the goods group Category 1. However, the level of e^* tends to be lower than that shown in Figure. 3. Therefore, the degree of optimization error is small even in commodity group Category 2[15].

When annual data are used, GARP is satisfied in both Categories 1 and 2. This result is robust for long and short analysis periods. The annual data comprise the observation number of 12 units in the time series direction. When the number of observations in the time series decreases, it tends to pass the GARP test easily.

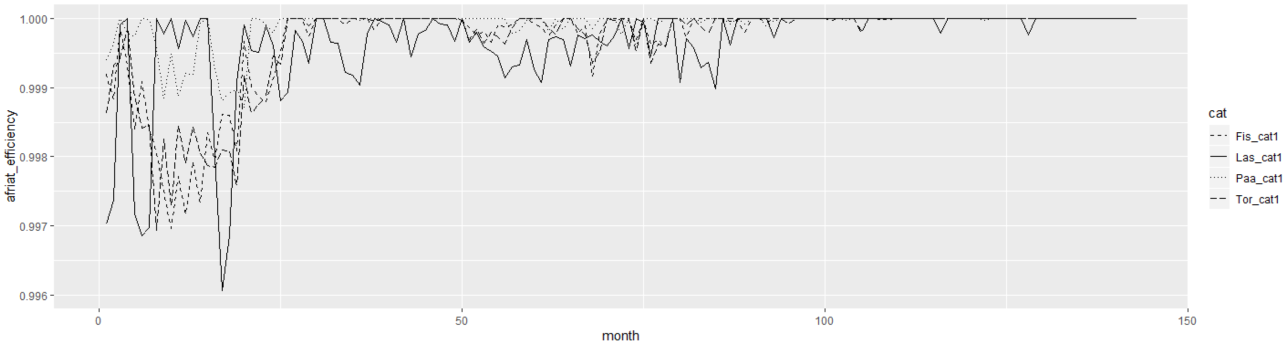

Figure 3: Removal Months and the Maximum AEI Level That Passes the GARP Test (Category 1)

Note: The horizontal axis represents the number of months deleted from the first month of the total 144 months plus one, while the vertical axis shows the maximum AEI level that passes the GARP test. As for the caption, Fis_cat1, Las_cat1, Paa_cat1, and Tor_cat1 represent the aggregation in the Category 1 commodity group for the Fisher, Laspeyres, Paacshe, and Töurnqvist price indexes, respectively.

[15] The number of stores surveyed has increased since September 2014 (Month 106), but the impact cannot be seen clearly. In the Appendix, data are divided into before and after September 2014 for analysis, but the results do not change significantly.

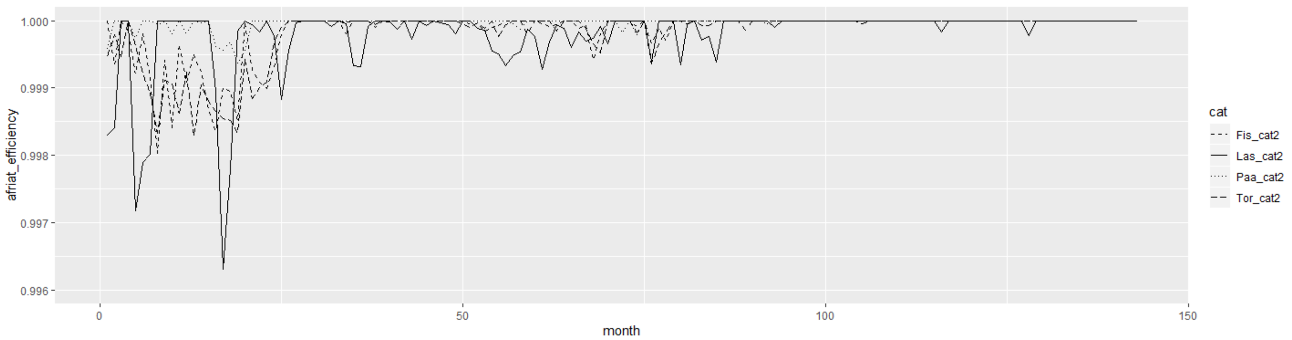

Figure 4: Removal Months and the Maximum AEI Level That Passes the GARP Test (Category 2)

Note: The horizontal axis represents the number of months deleted from the first month of the total 144 months plus one, while the vertical axis shows the maximum AEI level that passes the GARP test. As for the caption, Fis_cat2, Las_cat2, Paa_cat2, and Tor_cat2 represent the aggregation in the Category 2 commodity group for the Fisher, Laspeyres, Paacshe, and Töurnqvist price indexes, respectively.

5. Conclusion

The results reveal that the large-scale POS data for processed foods and beverages can be rationalized by a well-behaved utility function. This result is also robust for the length of the analysis period. On the other hand, data aggregated using representative price indexes, such as Laspeyres, Paacshe, Fisher, and Töurnqvist, tend not to pass the GARP test when the analysis period is long; however, the degree of optimization error is minor. Although it is not exactly clear whether this is because the observation number in the time series direction is small, a robust result is obtained that passes the GARP test when the aggregated data at the annual level are used.

Two methods (Categories 1 and 2) were adopted for the commodity group division of aggregation. The former covers 140 commodity groups and the latter covers 1109 commodity groups. Although the number of commodity groups is quite different, the movement of the largest AEI that passes the GARP test is generally similar. It is necessary to examine what results can be obtained from aggregation at a higher level because, in the analysis of food and beverage demand, aggregation may be conducted by a coarser classification of goods.

The validation of this paper is a necessary condition for the existence of weakly separable utility functions that rationalize aggregate data using representative price indexes. Under the Afriat numbers calculated by the representative price indexes, it is necessary to verify whether the data for *each* commodity group satisfy GARP.

In addition, efficient algorithms have been developed since Varian (1983)[16]. The application and extension of these algorithms will provide further insight into the verification of weak separability using large product-level data.

[16] For example, verification by nonlinear programming (Swofford and Whitney, 1994; Elger and Jones, 2008; Fleissig and Whitney, 2008) and verification by mixed integer programming (Cherchye et al., 2015). Some studies have examined how many types of utility functions exist to rationalize data (Crawford and Pendakur, 2013).

Appendix

Results of the GARP tests when data are divided after September 2014

The number of stores surveyed has increased since September 2014. Therefore, when the GARP test is carried out, an increase in the purchase quantity and total expenditure will be observed after September 2014. Beatty and Crawford (2011) noted that the GARP test is more likely to be passed if the change in total expenditures relative to the change in relative prices is large. In this section, considering this point, the data are divided around September 2014 and the GARP test is carried out for each division. Within the data of all 144 months, the divided first half of the data becomes 105 months, and the latter half of the data becomes 39 months.

Using the monthly data before aggregation, the divided data passed the GARP test as well as the results of not dividing around September 2014. This result is robust even when the analysis period is shortened by sequentially deleting from the first month.

Figures A.1 to A.4 show the results of the GARP test for representative price indexes. Figures A.1 and A.2 show the results of using the first half of the data and the second half of the data when aggregated into Category 1, respectively. The first half of the data reveals that the value of e^* is stable around 1 when the data over 25 months from the start month are excluded. This demonstrates the same tendency as the result in Section 4. The results of the Laspeyres price index, which is sensitive to the time deviation from the base month, tend to lower the e^* level compared to the results in Section 4. In Figure A.2, as in Section 4, the value of e^* is stable around 1. Figures A.3 and A.4 show the results of using the first half of the data and the second half of the data when aggregated into Category 2, respectively. In comparison with Section 4, the value of e^* shows fairly stable movement and the level approaches 1.

Based on the above results, it is concluded that the GARP test results are not significantly affected by the increase in the number of surveyed stores at a particular point in time, and the main factor influencing the results obtained in this section is considered the decrease in the number of observations in the time series direction due to data division.

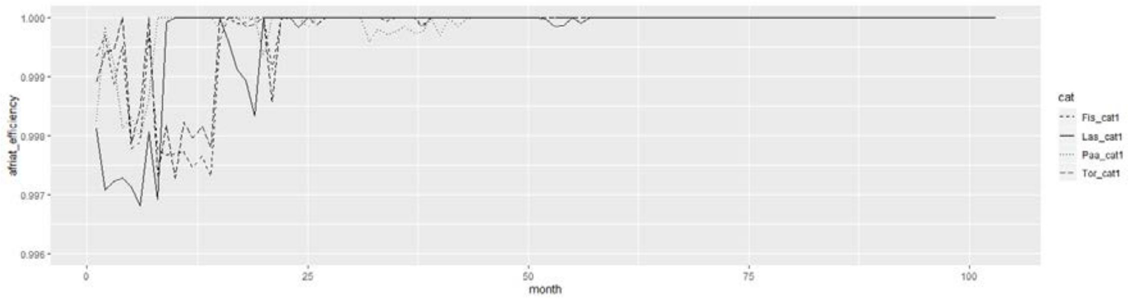

Figure A.1: Removal Months and Maximum AEI Level That Passes the GARP Test (Category 1, First-half of the Data)

Note: The horizontal axis represents the number of months deleted from the first month of the total 105 months plus one, while the vertical axis shows the maximum AEI level that passes the GARP test. As for the caption, Fis_cat1, Las_cat1, Paa_cat1, and Tor_cat1 represent the aggregation in the Category 1 commodity group for the Fisher, Laspeyres, Paasche, and Töurnqvist price indexes, respectively.

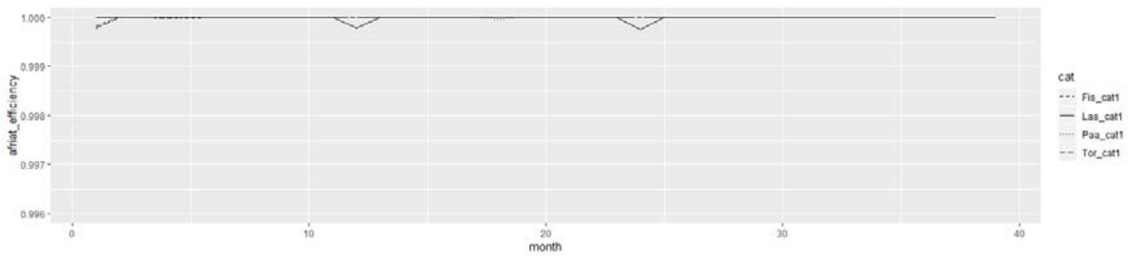

Figure A.2: Removal Months and Maximum AEI Level That Passes the GARP Test (Category 1, Second-half of the Data)

Note: The horizontal axis represents the number of months deleted from the first month of the total 39 months plus one, while the vertical axis shows the maximum AEI level that passes the GARP test. As for the caption, Fis_cat1, Las_cat1, Paa_cat1, and Tor_cat1 represent the aggregation in the Category 1 commodity group for the Fisher, Laspeyres, Paasche, and Töurnqvist price indexes, respectively.

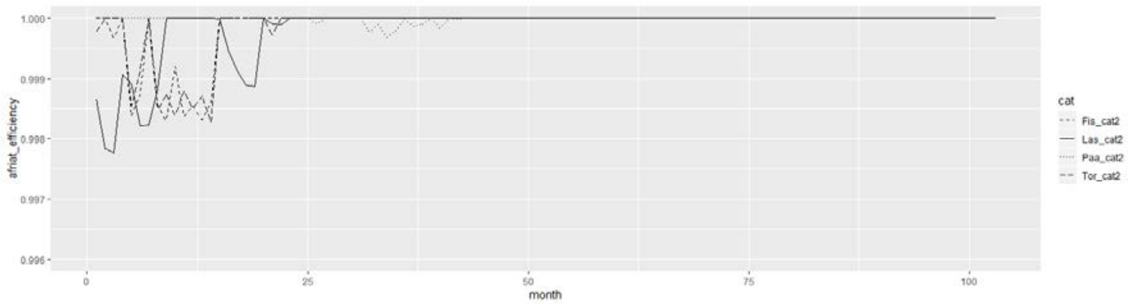

Figure A.3: Removal Months and Maximum AEI Level That Passes the GARP Test (Category 2, First-half of the Data)

Note: The horizontal axis represents the number of months deleted from the first month of the total 105 months + 1, while the vertical axis shows the maximum AEI level that passes GARP. As for the caption, Fis_cat2, Las_cat2, Paa_cat2, and Tor_cat2 represent the aggregation in the Category 2 commodity group for the Fisher, Laspeyres, Paasche, and Töurnqvist price indexes, respectively.

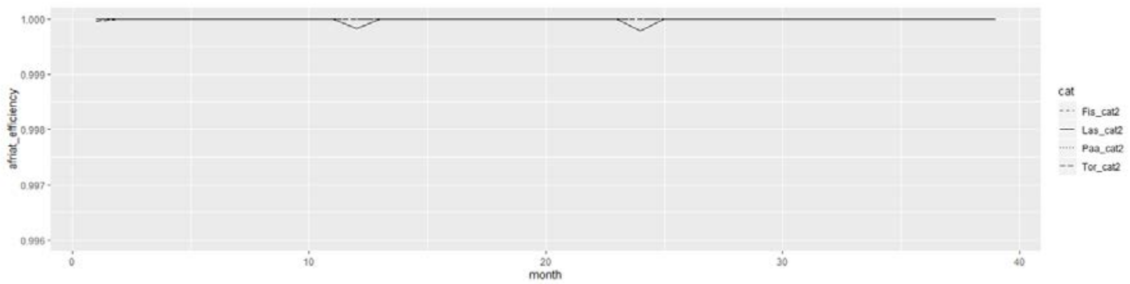

Figure A.4: Removal Months and Maximum AEI Level That Passes the GARP Test (Category 2, Second-half of the Data)

Note: The horizontal axis represents the number of months deleted from the first month of the total 39 months + 1, while the vertical axis shows the maximum AEI level that passes the GARP test. As for the caption, Fis_cat2, Las_cat2, Paa_cat2, and Tor_cat2 represent the aggregation in the Category 2 commodity group for the Fisher, Laspeyres, Paasche, and Töurnqvist price indexes, respectively.

References

Afriat, S. N.. 1967. The Construction of a Utility Function from Expenditure Data. *International Economic Review* 8 1: 67–77.

Barnett, W.A., Choi, S.. 1989. A Monte Carlo Study of Tests of Blockwise Weak Separability. *Journal of Business & Economic Statistics* 7 3: 363–77.

Beatty, T. K. M., Crawford, I. A.. 2011. How Demanding Is the Revealed Preference Approach to Demand?. *American Economic Review* 101 6: 2782–95.

Blackorby, C., Davidson, R., Schworm, W.. 1991. Implicit Separability: Characterisation and Implications for Consumer Demands. *Journal of Economic Theory* 55 2: 364–99.

Boelaert, J.. 2019. revealedPrefs: Revealed Preferences and Microeconomic Rationality. R package version 0.4.1.

Capps, O., Love, H.. 2002. Econometric Considerations in the Use of Electronic Scanner Data to Conduct Consumer Demand Analysis. *American Journal of Agricultural Economics* 84 3: 807–16.

Cherchye, L., Demuynck, T., De Rock, B., Hjertstrand, P.. 2015. Revealed Preference Tests for Weak Separability: An Integer Programming Approach. *Journal of Econometrics* 186 1: 129–41.

Crawford, I., Pendakur, K.. 2013. How Many Types Are There?. *Economic Journal* 123 567: 77–95.

Diewert, E. W.. 1973. Afriat and Revealed Preference Theory. *Review of Economic Studies* 40 3: 419–25.

Diewert, W. E., Parkan, C.. 1985. Tests for the Consistency of Consumer Data. *Journal of Econometrics* 30 1–2: 127–47.

Dhar, T., Chavas, J., Gould, B. W.. 2003. An Empirical Assessment of Endogeneity Issues in Demand Analysis for Differentiated Products. *American Journal of Agricultural Economics* 85 3: 605–17.

Eales, J., Unnevehr, L.. 1988. Demand for Beef and Chicken Products: Separability and Structural Change. *American Journal of Agricultural Economics* 70 3: 521–32.

Elger, T., Jones, B. E.. 2008. Can Rejections of Non-separability Be Attributed to Random Measurement Errors in the Data?. *Economics Letters* 99 1: 44–7.

Fleissig, A., Whitney, G.A.. 2003. A New PC-based Test for Varian's Weak Separability Condition. *Journal of Business & Economic Statistics* 21 1: 133–45.

Fleissig, A. R., Whitney, G. A.. 2008. A Nonparametric Test of Weak Separability and Consumer Preferences. *Journal of Econometrics* 147 2: 275–81.

Fourmouzi, V., Genius, M., Midmore, P.. 2012. The Demand for Organic and

Conventional Produce In London, UK: A System Approach. *Journal of Agricultural Economics* 63.3: 677–93.

Goldman, S. M., Uzawa, H.. 1964. A Note on Separability in Demand Analysis. *Econometrica* 32 3: 387–98.

Heng, Y., House, L. A., Kim, H.. 2018. The Competition of Beverage Products in Current Market: A Composite Demand Analysis. *Agricultural and Resource Economics Review* 47 1: 118–31.

Hjertstrand, P.. 2009. A Monte Carlo Study of the Necessary and Sufficient Conditions for Weak Separabillity, 151–82. In *Measurement Error: Consequences, Applications and Solutions*, ed. Binner, J., Edgerton, D., Elger, T., Bingley: Emerald Group Publishing Limited.

Lakkakula, P., Schmitz, A., Ripplinger, D.. 2016. U.S. Sweetener Demand Analysis: A QUAIDS Model Application. *Journal of Agricultural and Resource Economics* 41 3: 533–48.

Lewbel, A.. 1996. Aggregation Without Separability: A Generalized Composite Commodity Theorem. *American Economic Review* 86 3: 524–43.

Moschini, G., Moro, D., Green, R. D.. 1994. Maintaining and Testing Separability in Demand Systems. *American Journal of Agricultural Economics* 76 1: 61–73.

Nayga, R., Capps, O.. 1994. Tests of Weak Separability in Disaggregated Meat Products. *American Journal of Agricultural Economics* 76 4: 800–8.

Reed, A. J., Levedahl, J. W., Hallahan, C.. 2005. The Generalized Composite Commodity Theorem and Food Demand Estimation. *American Journal of Agricultural Economics* 87 1: 28–37.

Schulz, L. L., Schroeder, T. C., Xiac, T.. 2012. Studying Composite Demand Using Scanner Data: The Case of Ground Beef in the US. *Agricultural Economics* 43 s1: 49–57.

Sellen, D., Goddard, E.. 1997. Weak Separability in Coffee Demand Systems. *European Review of Agricultural Economics* 24 1: 133–44.

Shumway, C. R., Davis, G. C.. 2001. Does Consistent Aggregation Really Matter?. *Australian Journal of Agricultural and Resource Economics* 45 2: 1–34.

Swofford, J. L., Whitney, G. A.. 1994. A Revealed Preference Test for Weakly Separable Utility Maximization with Incomplete Adjustment. *Journal of Econometrics* 60 1–2: 235–49.

Varian, H. R.. 1982. The Nonparametric Approach to Demand Analysis. *Econometrica* 50 4: 945–73.

Varian, H. R.. 1983. Non-Parametric Tests of Consumer Behaviour. *Review of Economic*

Studies 50 1: 99–110.

Varian, H. R.. 1994. Goodness-of-Fit for Revealed Preference Tests. Econometrics 9401001. Germany: University Library of Munich.

Xie, J., Myrland, Ø.. 2011. Consistent Aggregation in Fish Demand: A Study of French Salmon Demand. *Marine Resource Economics* 26 4: 267–80.