The Generalized Composite Commodity Theorem: Aggregation of Grocery Items at Firm Level*

Preliminary and Incomplete: Please Do Not Quote or Cite

Wanwiphang Manachotphong†

February 2007

Abstract

The large number of products and prices in multi-product firms causes great difficulty in analyzing consumers' choice among them. Prices of all products in each firm play some roles in consumers' decision-making process. However, accounting for all of them would be very difficult of not impossible. According to the Generalized Composite Commodity Theorem (GCCT) developed by Lewbel (1996), it is possible to obtain valid aggregation of commodities if some certain conditions are satisfied. A valid aggregation would maintain the four conditions of rational demand system–e.g. adding up, homogeneity, slutsky symmetry and negative semi-definiteness. Using data from the UK supermarket industry, this paper shows that it is valid to aggregate grocery items by their supermarket firms.

* I am grateful to Howard Smith and Takamitsu Kurita for their guidance and support. I am also grateful to Arthur Lewbel for very useful discussions and suggestions. For financial support, I am grateful to the UK's Milk Development Council who made obtaining this data set possible. Any remaining errors are my own.

† Department of Economics, University of Oxford. Email: wanwiphang.manachotphong@economics.ox.ac.uk.

1 Introduction

The large number of products and prices in multi-product firms causes great difficulty in analyzing consumers' choice among them. In year 2005, for example, Asda, Tesco, Sainsbury and Morrison–the four biggest supermarkets in the UK–each sold over 30,000 different product lines. Prices of all the products in each firm play some roles in consumers' decision-making process. However, accounting for all of them would be very difficult if not impossible. Recent empirical advances provide various remedies to this so-called dimensionality problem. Which remedy is best depends on the nature of the analysis. In our case, consumers make choices upon grocery stores which provide similar product selections. Prices of individual products are unlikely to be as influential to consumers' decisions as the overall "expensiveness" of stores (or firms). Therefore, our aim is to aggregate products at store- or firm-level while still retaining all properties of a valid demand system, i.e. adding-up, homogeneity, Slutsky symmetry and negative semi-definiteness. In this paper, we use the generalized composite commodity theorem (GCCT) developed by Lewbel (1996) to test for valid aggregation of grocery products at firm level.

The first composite commodity theorem was developed by John R. Hicks (1936) and Wassily Leontief (1936). Hicks and Leontief suggest that if all individual product prices are perfectly colinear[1], those products can be aggregated into the same group. While perfect colinearity provides a sensible aggregation rationale, prices of different goods are hardly found to behave in such way. As a result, tests of valid aggregation using this version of composite commodity theorem often fail. Lewbel (1996) suggests that valid aggregation can be obtained under a more empirically realistic condition. Under Lewbel's GCCT, valid aggregation can be obtained if all the deviations between individual product prices and their group price index are independent of income and all the price indices in the demand system.

Other than GCCT, existing remedies to the dimensionality problem are such as multi-stage budgeting approach (Gorman, 1959) (Deaton and Muellbauer, 1980) and characteristics approach (Pinkse and Slade, 2004). In a multi-stage budgeting framework, we assume that consumers first decide to allocate expenditure among pre-defined sets of goods. Then, given the allocated expenditure, they maximize their utility by choosing how much to spend on each individual product in the groups. Here, weak separability is required because the decision among individual goods in one group would

[1]Let p_i^t denote price of product i at time t, Hicks(1936) and Leontief(1936) requires $p_a^1/p_a^0 = p_b^1/p_b^0 = ... = \delta$ where δ is a fixed price ratio.

be independent of consumption level of other groups. Therefore, when using the multi-stage budgeting approach, we need to both impose weak separability on consumers' preferences and restrict their consumption pattern (Deaton and Muellbauer, 1980). The characteristics approach, on the other hand, neither impose budget allocation pattern nor require weak separability. For this approach, the cross-price elasticity of a given pair of products is estimated in terms of similarity of their characteristics. Pinkse and Slade (2004), for example, estimated brand-level cross-price elasticities of beers in the UK as functions of price, sales volume, alcohol content, brewer identity, etc. Therefore, rather than having to estimate all the cross-price elasticities directly, we only have to estimate parameters associated with each characteristic. The characteristics approach provides a convenient way to reduce the number of dimensions, however, it can only be used when product characteristics are observed.

Although GCCT is not subject to the above limitations, it can replace other methods only when we do not need to assess cross-price elasticities at individual product level. GCCT remedies the dimensionality problem by aggregating products into broader–and fewer–groups. As a result, only cross-price elasticities at group level, not individual product level, can be estimated. For the utility functional form, GCCT can be used with all homothetic utility functions, almost ideal demand system (AIDS), translog demand system, and any utility function in which goods are aggregated into two groups (Lewbel, 1996).

To date, GCCT has been used in many applications. Davis, Lin, and Shumway (2000) apply GCCT to aggregate US and Mexican agricultural outputs. They show that the theory provides support for aggregation into as few as two agricultural output groups in each country. Capp and Love (2002) analyze the demand system of different types and brands of fruit juice. They compare the bias of price elasticities obtained through multi-stage budgeting approach and through GCCT approach and found that GCCT gives more accurate results. Reed, Levedahl and Hallahan (2005) use GCCT to estimate demand elasticity of aggregated food products. They first use GCCT to test for valid aggregation, then estimate own- and cross-price elasticities of each product group. Davis (2003) use family-wise test of multiple hypotheses to provide a stronger empirical support to GCCT. More discussion on the family-wise test can be found in section 3.2.

This paper contributes to the composite commodity literature in two aspects. First, we test for valid aggregation of products by their sellers (firms) rather than by their similarity. Second, we propose using the cointegrated vector autoregressive model (cointegrated VAR) (Johansen, 1995) to perform a more comprehensive test of valid aggregation.

Using a data set obtained from the UK supermarket industry, we show that it is valid to aggregate grocery items by their supermarket chains. In this context, the price index for each supermarket chain is constructed using prices of the same set of grocery items. Therefore, it is less likely–than in a case where different price indices are constructed using different sets of products–for the aggregation to pass the independence test[2]. More details on this discussion can be found in section 3.1.

Potential benefits from valid firm-level aggregation is substantial. Particularly, in the industrial organization context where we focus on competition at firm level rather than at individual product level. The ability to aggregate products by firm enables us to sidestep the dimensionality issue caused by too many price parameters. In this paper, valid aggregation of grocery items by supermarket chains allows us to easily study consumers' supermarket choice, supermarket substitution pattern and their intensity of competition. These merits also applies to other multi-product firms cases.

The rest of the paper is organized as follows. Section 2 provides theoretical overview of GCCT. Section 3 lays out an empirical overview and discusses our testing strategy. Section 4 explores the data set as well as describes our empirical implementation. Section 5 discusses the empirical result, while section 6 concludes.

2 Theoretical Overview

The discussion on Generalized Composite Commodity Theorem (GCCT) is taken directly from Lewbel (1996). Further reference can be found in the original paper.

Following Lewbel (1996), let p_i denote price of individual product $i = 1, ..., n$ and P_j denotes price index of product group $I = 1, ..., J$. Then, define $r_i = \log(p_i)$, $R_j = \log(P_j)$, $\rho_i = \log(p_i/P_I)$, $\mathbf{r} = n$-vector of all individual product prices, $\mathbf{R} = J$-vector of all price indices and $\boldsymbol{\rho} = n$-vector of $\log(p_i/P_I)$ where product i belongs to group I. The term ρ_i is the deviation of log individual price from log price index of the product group it belongs to. ρ_i is also called *relative price* of product i.

Before moving on to the aggregated product demand system, Lewbel first explains the disaggregated product (individual product) case as a benchmark. For any individual product, its observed budget share can be expressed in terms of Marshallian demand and an error term. Suppose w_i denotes observed budget share of individual product i, $g_i(\mathbf{r}, z)$ denotes a Marshallian

[2]The GCCT requires all the deviations between individual product prices and their group price index to be independence of all the price indices in the system.

demand function and e_i denotes an error term with zero conditional mean, we can express the empirical representation of individual product i's demand as:

$$w_i = g_i(\mathbf{r}, z) + e_i.$$

Since the error term e_i has zero conditional mean ($E(e_i|\mathbf{r}, z) = 0$), the observed budget share is an unbiased estimator of demand:

$$E(w_i|\mathbf{r}, z) = g_i(\mathbf{r}, z).$$

The function $g_i(\mathbf{r}, z)$ is a valid Marshallian demand function because it satisfies the following conditions:

1) Adding up – $\Sigma_{i=1}^{n} g_i = 1$
2) Homogeneity – $g_i(r - k, z - k) = g_i(r, z)$
3) Slutsky symmetry – $(\partial g_i/\partial r_j) + (\partial g_i/\partial z)g_j = (\partial g_j/\partial r_i) + (\partial g_j/\partial z)g_i$ for all i and j
4) Negative semi-definiteness – matrix $\tilde{s}$ that is consist of elements $\tilde{s}_{ij} = s_{ij} + g_i(\mathbf{r},z)g_j(\mathbf{r},z)$ for $i \neq j$ and $\tilde{s}_{ii} = s_{ii} + g_i(\mathbf{r},z)^2 - g_i(\mathbf{r},z)$ is nevative semidefinite.

When a demand function has the first three properties, it satisfies the first-order conditions for utility maximizations and can be called *integrable.* When a demand function has all the four properties, it is arose from rational decision and can be called *rational.*

Although disaggregated product demand system are integrable and rational, it is not always true for the aggregated product demand system counterpart. Lewbel (1996) shows that aggregated product demand system could be integrable and rational under two assumptions.

1. The disaggregated demand functions $g_i(\mathbf{r}, z)$ for $i = 1, ..., n$ are rational - that is, the function satisfies adding up, homogeneity, Slutsky symmetry, and negative semi-definiteness.

2. The relative price of each individual product ρ_i for $i = 1, ..., n$ are independent of all the aggregated product price indices R_j for $j = 1, ..., J$ and income z in the demand system.

To show how these two assumptions come to play their parts, first let some products $i = 1, ..., M$ where $(M < n)$ be aggregated into product group j. Here, the observed budget share of group j can be written as $W_j = \Sigma_{i\in j}^{M} w_i$. Similar to the disaggregated product case, if $G_j(\mathbf{R}, z)$ is the budget share demand function for product group j and u_j is an error term with zero conditional mean, we can express the empirical representation of the budget share demand function of aggregated products as:

$$W_j = G_j(\mathbf{R}, z) + u_j.$$

Since the error term u_j has zero conditional mean ($E(u_j|\mathbf{R}, z) = 0$), the observed budget share of product group j is an unbiased estimator of its budget share demand counterpart:

$$E(W_j|\mathbf{R}, z) = G_j(\mathbf{R}, z).$$

To proof whether $G_j(\mathbf{R}, z)$ satisfies the four requirements of rational demand function, Lewbel started from establishing the term $G_j^*(\mathbf{r}, z) \equiv \Sigma_{i\in j}^{M} g_i(\mathbf{r}, z)$ which is group j demand expressed in terms of logged individual prices $\mathbf{r}$ and income z. Then, he defined $\mathbf{R}^* = \mathbf{r} - \boldsymbol{\rho}$ where $\mathbf{R}^*$ is n-vector of group prices with R_j in row i and for every row, $i \in j$. The demand functions $g_i(\mathbf{r}, z)$, $G_j(\mathbf{R}, z)$ and $G_j^*(\mathbf{r}, z)$ are related to as follows:

$$\begin{aligned}
G_j^*(\mathbf{r}, z) &= G_j^*(\mathbf{R}^* + \boldsymbol{\rho}, z) = \Sigma_{i\in j}^{M} g_i(\mathbf{r}, z) = \Sigma_{i\in j}^{M} w_i - \Sigma_{i\in j}^{M} e_i \\
G_j^*(\mathbf{R}^* + \boldsymbol{\rho}, z) &= W_j - \Sigma_{i\in j}^{M} e_i \\
E[W_j|\mathbf{R}, z] &= E[G_j^*(\mathbf{R}^* + \boldsymbol{\rho}, z)|\mathbf{R}, z] + E[\Sigma_{i\in j}^{M} e_i|\mathbf{R}, z] \\
G_j(\mathbf{R}, z) &= E[G_j^*(\mathbf{R}^* + \boldsymbol{\rho}, z)|\mathbf{R}, z] + 0
\end{aligned}$$

If $\boldsymbol{\rho}$ is independent of the price index vector $\mathbf{R}$ (and thus, $\mathbf{R}^*$) and income z, we can proceed to the following step:

$$G_j(\mathbf{R}, z) = \int G_j^*(\mathbf{R}^* + \boldsymbol{\rho}, z) dF(\boldsymbol{\rho}),$$

where $F(\boldsymbol{\rho})$ is the distribution function of $\boldsymbol{\rho}$. This equation says that the aggregate group budget share $G_j(\mathbf{R}, z)$ is equal to the conditional expected value of the sum of all product demand functions that belong to the group $G_j^*(\mathbf{r}, z)$. Lewbel (1996) shows that when the two assumptions are satisfied, the budget share demand functions $G_j(\mathbf{R}, z)$ for $j = 1, ..., J$ is a valid system of composite demand equations. This is because they satisfy adding-up, homogeneity, and (if not perfectly) nearly Slutsky symmetry. The demand elasticities of $G_j(\mathbf{R}, z)$ for $j = 1, ..., J$ are best unbiased estimates of within-group sums of individual product demand elasticities.

In general, the first assumption–$g_i(\mathbf{r}, z)$ for $i = 1, ..., n$ are rational–is satisfied if we assume utility maximization. The second assumption, however, requires empirical testing of whether the relative price ρ_i for $i = 1, ..., n$ are independent of all the price indices R_j and income z. The next section discusses empirical strategies used to test for these independence.

3 Empirical Overview

Recall from the previous section, Lewbel's independence assumption requires each individual relative price ρ_i to be independent of *all* the price indices in the demand system (R_j for $j = 1, ..., J$) and income z. Since independence is very difficult if not possible to test, all previous works based their inferences on correlation and cointegration tests (Lewbel, 1996), (Davis, Lin, and Shumway, 2002), (Capps and Love, 2002) and (Reed, Levedahl, and Hallahan, 2005). Whether correlation or cointegration test is appropriate depends on the stationarity property of variables being tested. If variables are stationary, a standard correlation test such as Spearman's rank correlation test (Mendenhall, Scheaffer, and Wacherly, 1990) and F-test of significant coefficients (Theil, 1971, chapter 11) can be used. If variables are non-stationary, a cointegration test (Johansen, 1995) is more appropriate.

It is worth noting here that we cannot draw a definite confirmation of valid aggregation even if we find no correlation or no cointegration between ρ_i and the vector R. This is because no correlation or no cointegration does not imply independence. For empirical feasibility, however, we resort to test of correlation and cointegration while being aware for their potential errors.

A correlation test between ρ_i and the price index vector $\mathbf{R}$ and income z can be performed by regressing ρ_i on all the price indices R_j (for $j = 1, ..., J$) and income z. Likewise, a cointegration test could be performed through multiple equations estimation. Since the number of parameters being estimated grows with the number of lag length and the number of equations, reliable estimates would be obtained only when the sample size is sufficiently large.

In practice, it is difficult to obtain a long time-series data on price. Most data sets are available on quarterly or annual basis and usually last less than 50 years. For example, Lewbel (1996) obtained his data set from the U.S. national income and product accounts (NIPA). The data was collected annually from 1954 to 1993; Davis, Lin and Schumway (2000) obtained their U.S. and Mexican agricultural output information from Ball (1996). The data was collected annually from 1949 to 1991. In both studies, the number of observations were less than 100. Given a small sample size, it is not possible to obtain consistent estimates through multiple equations estimation. In other words, it is not feasible perform correlation/cointegration test between the relative price ρ_i and all the price indices R_j (for $j = 1, ..., J$) simultaneously.

To remedy the small sample problem, many previous works perform a so-called "single hypothesis testing". Here, rather than performing system estimations, they perform correlation test or cointegration test only on relative price ρ_i and the price index of the group it belongs to (R_j, where $i \in j$).

If the test is passed, the aggregation is assumed to be valid.

Single hypothesis testing is acceptable in many cases where relative prices are most likely to be correlated with the price index of the groups they belong to. That is, if the relative price ρ_i is independent of the price index R_j where $i \in j$, it is very likely that ρ_i would be independent of all other price indices R_k for $i \notin j$. In our case, where products are aggregated by firms not by their similarity, it is less likely that the same logic can be applied. So, the test of independence between ρ_i and all other price indices R_k for $i \notin j$ should be conducted. Davis (2003) and Davis, Lin, and Shumway (2000) implement multiple hypotheses testing methods–also called family-wise test–to test for independence between ρ_i and all price indices R_j for $j = 1, ..., J$ in the demand system. We follow their analysis and will discuss about the family-wise test in more detail at the end of this section.

The test strategy proposed in this paper is different from those implemented in the majority of previous works, e.g. (Lewbel, 1996) (Davis, Lin and Shumway, 2000) (Capps and Love, 2002). In those studies, a non-stationarity test (unit root test) was performed on each individual variable–ρ_i, R_j and z. Then, based on the stationarity property of each ρ_i, R_j pair, they choose an appropriate test method. If both ρ_i and R_j are stationary, they perform a correlation test. If both ρ_i and R_j are non-stationary, they perform a cointegration test. If one of the variables is stationary and another is non-stationary, they assume that the variables are not correlated.

One problem with this testing strategy is the low power of the non-stationarity test (large type-1 error) and so it is biased towards finding non-stationarity (Lo and Mackinley, 1989). Johansen (1995) suggests that non-stationarity test can be performed more accurately through test of restrictions in cointegrated vector autoregressive models (cointegrated VAR). This is because potential relations among variables in the cointegrated vector help improve the power and accuracy of the test. This remedy is adopted in this paper. The following section discusses the procedure through which we perform the test of valid aggregation.

3.1 3-Step Correlation/Cointegration Test

This section describes our 3-step testing strategy. In brief, we first perform a cointegration test on each ρ_i and R_j pair regardless of their stationarity properties. Then, according to the cointegration test result, we verify stationarity properties of variables through the "test for restrictions in a cointegrated VAR model". Since the stationarity test here is nested inside a bivariate system, potential relations between variables make the test result more accurate than that of a single variable test (Johansen, 1995). Lastly, we apply a family-wise test method to confirm the absence of own- and cross-group correlation/cointegration between ρ_i and R_j.

- **Step 1**: Cointegration test and Spearman's rank Correlation test

  Without testing for non-stationarity of variables, we first perform a cointegration test (based on Johansen, 1995) on all combinations of relative price ρ_i for $i = 1, ..., n$ and price index R_j for $j = 1, ..., J$. The null hypothesis is no cointegrated relationship between ρ_i and R_j. In other words, ρ_i and R_j form a nonstationary process. If this null hypothesis cannot be rejected, we conclude that the given pair of relative price ρ_i and firm-level price index R_j is not cointegrated. If the null hypothesis is rejected and the process has full rank, it means that there are two stationary processes in the bivariate system. In this case, we perform Spearman's rank correlated test on the two variables. If the null hypothesis is rejected and the process is I(1), it means that there is one stationary process in the bivariate system. If this happens, we proceed to the next step

- **Step2**: Test of restrictions in cointegrated VAR model

  When an I(1) process is found in step 1, it could be due to one of the two following cases. First, relative price ρ_i and price index R_j are actually cointegrated. In this case, the aggregation would be invalid. Second, one variable is nonstationary while another is. Since we did not test for non-stationarity of variables prior to performing cointegration tests, the second case is possible. If this is true, ρ_i and R_j would not be correlated in the first place and the aggregation would be valid. In brief, step 2 performs restriction tests in cointegrated VAR model to justify whether stationarity finding is due to the first or the second case.

- **Step3**: Family-wise or multiple hypotheses test

  According to GCCT,valid aggregation requires each relative price ρ_i (for $i = 1, ..., n$) to be independent of *all* price indices in the demand system (vector $\mathbf{R}$) and income z. However, data limitation has prevented many previous works from performing multiple equations estimations to test for this independence. In stead of testing for independence between relative price ρ_i and the price index vector $\mathbf{R}$, they only test for independence between relative price ρ_i and the price index it belongs to R_j (where $i \epsilon j$). This practice is justified if ρ_i is more likely to be correlated with R_j (where $i \epsilon j$) than with R_k (where $k \notin j$).

By proofing the independence between ρ_i and R_j (where $i \epsilon j$), the independence between ρ_i and R_k (where $k \epsilon j$) follows.

In our case, however, it is not as obvious that the relative price ρ_i is more likely to be correlated with the price index of the group it belongs to. Our firm-level aggregation involves constructing firm-specific price indices from the same set of grocery items. For instance, R_{firm1} is constructed from eggs, bread, milk and meat sold by firm 1, R_{firm2} is constructed from eggs, bread, milk and meat sold by firm 2, etc. Since prices of the same items– e.g. eggs–are likely to be correlated across firms, each relative price (e.g. ρ_{store1_eggs}) is likely to be correlated with R_k where $k \notin j$ as much as with R_i where $i \epsilon j$.

Davis(2003) discusses family-wise test methods that can be used in this context. A family-wise hypothesis tests whether all the associated individual hypotheses are true. In our case, a family hypothesis is that ρ_i is independent of all the elements in $\mathbf{R}$ and income z. Individual hypotheses are 1) ρ_i is independent of R_1, 2) ρ_i is independent of R_2, etc. Let H_0 denote a family hypothesis and $H_1, ..., H_J$ denote all the associated individual hypotheses. A family-wise hypothesis can be expressed in terms of associated individual hypotheses as follows:

$$H_0 = \cap_{j=1}^{J} H_j.$$

With sufficient number of observations, there would be quite a few candidates to test the above family-wise hypothesis. For example, the F-test (Theil, 1971, chapter 11) can be used in case of stationary variables; and the multivariate cointegration test (Johansen, 1995) can be used in case of nonstationary variables. With insufficient number of observations, we need to use methods which accuracy does not depend on the asymptotic property– e.g. size of the data set. The Bonferroni procedure is one of the methods which satisfies this requirement. This paper uses a modified Bonferroni procedure called Hochberg procedure (Hochberg,1979) to implement family-wise hypothesis testing.

3.2 Modified Bonferroni Method

The reasoning behind the Bonferroni procedure is that when we test more than one hypotheses at the same time, the chance of making at least one false rejection (type1 error) would increase with the number of joint hypotheses we are testing. Consider the following individual and joint hypotheses:

$$\begin{aligned} H_1 &: \theta_1 = 0 \\ H_2 &: \theta_2 = 0 \\ &\text{and} \\ H_A &: \theta_1 = \theta_2 = 0 \end{aligned}$$

It is easy to see why we are more likely to make false rejection with H_A than H_1 or H_2. Then consider the following joint hypotheses:

$$\begin{aligned} H_A &: \theta_1 = \theta_2 = 0 \\ &\text{and} \\ H_B &: \theta_1 = \theta_2 = \theta_3 = \theta_4 = \theta_5 = \theta_6 = \theta_7 = 0. \end{aligned}$$

It is also obvious to see that we are more likely to make false rejection with H_B than H_A. The above two sets of examples allows us to make two claims. First of all, since it is easier to make false rejection with multiple (joint) than single hypothesis, the rejection level of joint hypothesis should be greater than that of a single hypothesis. Second of all, since it is easier to make false rejection when the number of single hypotheses increases, the rejection level of joint hypothesis should increase with the number of single hypotheses.

A collection of hypotheses being jointly tested is called a "family hypothesis". Its associated rejection level can be called the "family-wise error rate" (FWER). As mentioned previously, if H_0 denote a family hypothesis and $H_1, ..., H_J$ denote all the individual hypotheses. A family-wise hypothesis can be expressed in terms of associated individual hypotheses as:

$$H_0 = \cap_{j=1}^{J} H_j$$

And if α denotes the rejection level at which we test a single hypothesis, the FWER can be expressed as:

$$FWER \leq \alpha \tag{1}$$

To make the rejection level decrease with the number of individual hypotheses in the family, Bonferroni proposed that FWER $= \alpha/s$. If any individual p-value is less than α/s, the family hypothesis H_0 is rejected. Although the Bonferroni's FWER satisfies the requirement in 1, it leads to very small chance of rejecting the null hypothesis when s gets large. All successive modified Bonferroni Methods deals with this problem in various ways.

Here, we use the method developed by Hochberg(1988) due to its ease of use and high power[3].

Let α be the level we use for a single hypothesis testing and let $p_{(1)} ... \leq p_{(s)}$ be the increasing arrangement of p-value associated with individual null hypotheses $H_1, ..., H_s$. The Hochberg procedure rejects individual hypothesis H_j when $p_{(j)} \leq \alpha/(s - j + 1)$. The family hypothesis H_0 is retained only if *all* the individual hypotheses are not rejected.

4 Data and Empirical Implementation

4.1 The data

Our price data are from the TNS's Worldpanel survey. TNS randomly recruited a number of households in the UK to participate in their panel survey program. Each household was given a personal scanner which they used to record all their grocery shopping activities. The output data provided by TNS includes items bought, unit bought, price paid and outlet of purchase by 26,133 households over a 3-year time span (October 2002 to September 2005).

For our 3-year panel, we aggregate the information of each variable into 78 successive time periods. The reason why we choose a time period to be two-week long is due to consumer's shopping behavior. Two weeks is long enough for us to believe that the decisions on grocery shopping in each period is independent. On the other hand, it is short enough for us to believe that all items bought during the two-week period contribute to a single utility maximization not multiple utility maximizations. Manachotphong and Smith (2006) uses the same reasoning to justify the length of consumers' shopping period in their supermarket choice analysis.

Given the length of each time period, we calculate a Tornqvist price index of each firm using 55 grocery items. This comes from choosing five most popular items from 11 most popular product categories which are consistent across stores. The TNS categorizes grocery items into more general 268 distinct categories. These categories are such as milk, butter, vegetable, fruit juice, toilet paper, canned food, canned fish, frozen vegetable, flour, breakfast cereal and etc. Out of 268 categories, our 11 popular categories constitute of about 29.3 percent of the budget share–almost one third of consumer's spending on groceries. Table (1) shows the average budget share of each of

[3] According to Davis(2003), two other popular modified Bonferroni procedures are Holm procedure(Holm, 1979) and Simes procedure(Simes, 1986). Both are equally easy to implement, but Holm is less powerful than Hochberg and Simes.

Table 1: Product Category and Budget Share

Product Category	Budget share (per cent)
Vegetable	6.3
Fruit	4.7
Milk	3.3
Cheese	3.0
Cooked Meat	2.6
Fresh Poultry	2.2
Breakfast Cereal	2.2
Bread	1.9
Yoghurt	1.5
Instant Coffee	1.0
Eggs	0.7
Total	29.3

Based on all consumer's spending in May 2005.

the eleven product categories analyzed in this paper.

In 2007, the UK supermarket industry consists of 17 national-level chains where the four biggest chains hold about 74.6 per cent of market share[4]. These four biggest firms are Asda, Morrisons, Sainsbury and Tesco. They are also nicknamed "The Big4". For calculation tractability, we analyze only the price indices of these four prominent firms.

4.2 Empirical Implementation

To facilitate our analysis, some notations are modified. First of all, the relative price of an individual product would be denoted as ρ_i^k (where i is used to index product category and k is used to index firm) rather than ρ_i (where i is used to index individual product). Therefore, ρ_{eggs}^{Tesco}– eggs at Tesco– would be a different product from $\rho_{eggs}^{Sainsbury}$– eggs at Sainsbury. Second of all, price indices are calculated at firm level rather than at an aggregated product level. For example, R_{Tesco} represents logged price index of grocery items in Tesco and $R_{Sainsbury}$ represents logged price index of grocery items in Sainsbury. Therefore, we can write $\rho_i^k = \log(p_i^k - P_k)$ and $R_j = \log(P_j)$ where p_i^k is the Tornqvist price index of product category i sold by firm k and P_j is the Tornqvist price index of firm j.

Since we analyze 11 most popular product categories in four firms, there are 44 relative prices ρ_i^k and four price indices R_j in total. As for income, we

[4]According to Taylor Nelson Sofres plc. (TNS).

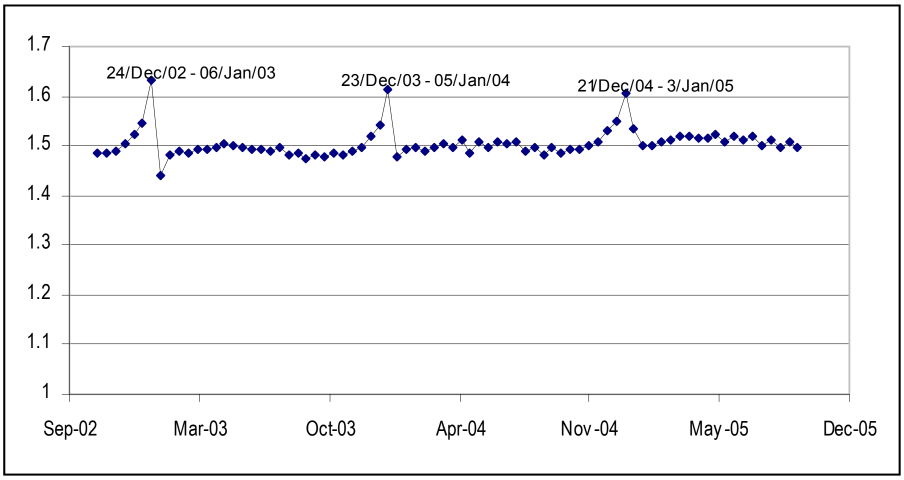

Figure 1: Log per-capita expenditure on grocery items

use per-capita expenditure on grocery items as a proxy.

Figure (1) plots log income (z) over time. It shows that, other than during the Christmas and New Year holidays, the per-capita spending on grocery is approximately stable. Figure (2), (3), (4) and (5) plot log price index of Asda, Morrisons, Sainsbury and Tesco respectively. These variables fluctuate more than expenditure. They also establish a steeper increasing trend over time.

Having calculated all the relative prices ρ_i^k, price indices R_j and income z, we now implement the 3-step test of correlation/cointegration.

4.2.1 3-Step Test: The Implementation

For step 1, we conduct a cointegration test (based on Johansen, 1995) on ρ_i^k and each of the $R_{Asda}, R_{Morrisons}, R_{Sainsbury}, R_{Tesco}$, and z. A time trend is added to reflect rising nature of price. Appropriate numbers of lags are added to each pair of ρ_i^k, R_j being tested. Exogenous dummy variables are also added to explain shocks to individual product prices and price indices. As firms usually charge premiums and give discounts according to their inventory situation and marketing strategy, it is justified to include those dummy variables into the model.

Prior to performing a cointegration test for each bivariate system, we conduct a diagnostic test of normality to ensure that the error terms are normally distributed. This is a fundamental assumption of cointegration analysis and needed to be satisfied. In addition, we account for potential

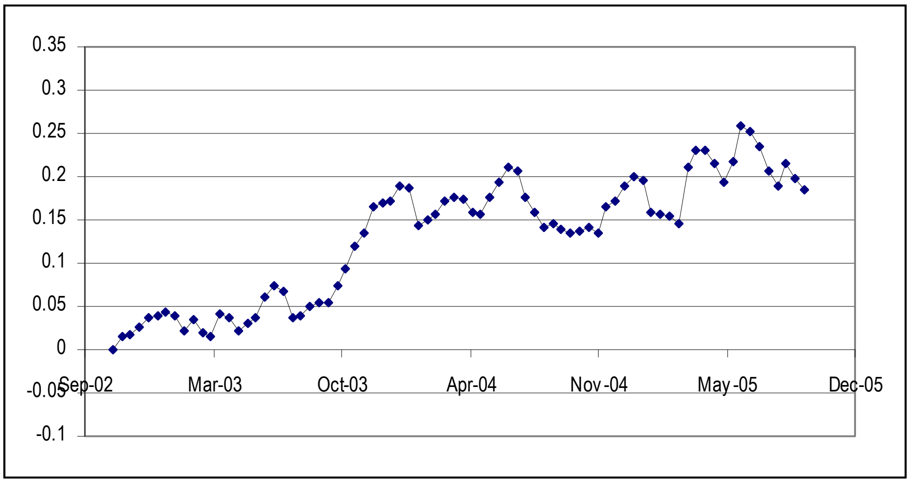

Figure 2: Log Asda Price Index (R_{Asda})

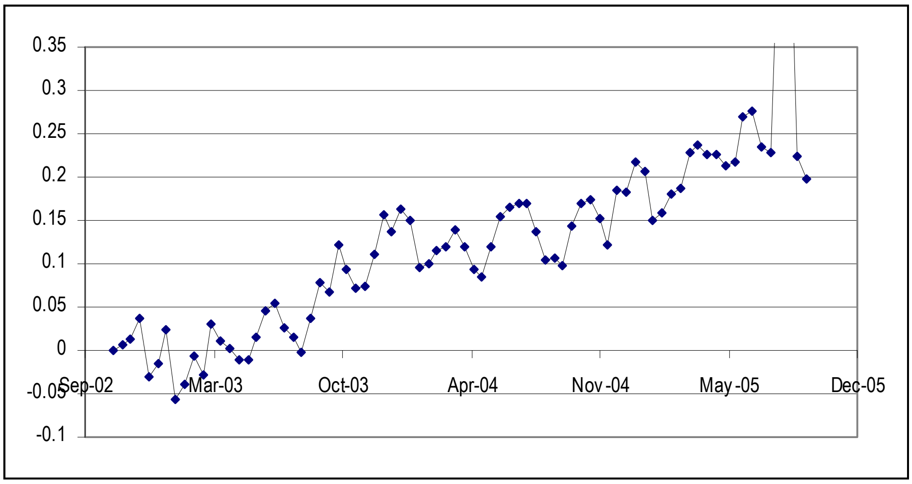

Figure 3: Log Morrisons Price Index ($R_{Morrisons}$)

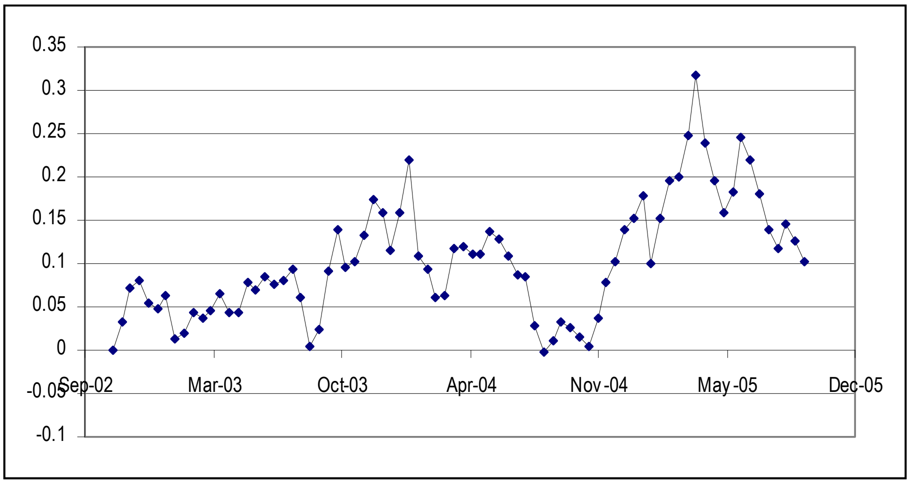

Figure 4: Log Sainsbury Price Index ($R_{Sainsbury}$)

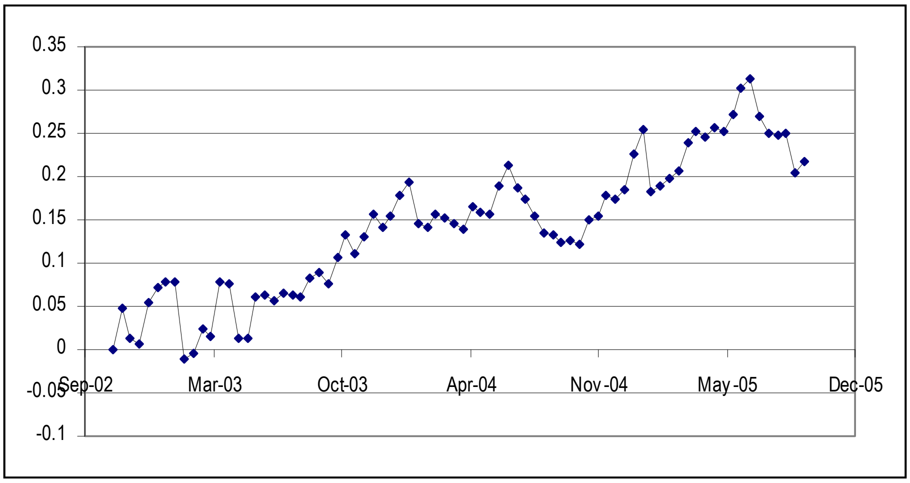

Figure 5: Log Tesco Price Index (R_{Tesco})

small sample bias by using Barrett correction and account for the low power for cointegration test by increasing the critical level α to 10 per cent. If we accept the null hypothesis of nonstationary process (two variables are *not* cointegrated), we report the p-value. If we reject the null hypothesis of nonstationary process (two variables are cointegrated), we proceed to the second step.

Step 2 helps verify whether the stationarity finding is because the two variables are actually cointegrated or because one is stationary while another is not. If the later case is true, the two variables would neither be correlated nor cointegrated (Granger and Hallman, 1989). The procedure used is called test of restrictions in cointegrated VAR model (Johansen, 1995).

If the test of restrictions shows that both variables are nonstationary, the two variables are likely to be cointegrated and we report the p-value. If the test shows that one variable is stationary while another is not, then we conclude that they are neither correlated nor cointegrated. If it appears that both variables are stationary, then we conduct a Spearman's rank correlation test and report the p-value.

In step 3, we use a modified Bonferroni method called the Hochberg procedure to perform family-wise hypothesis testing. Since GCCT requires each ρ_i^k to be independent of all price indices and income, a family-wise hypothesis test is required to confirm this fact. Let H_i^k denote the family hypothesis that ρ_i^k is neither correlated nor cointegrated with any of the price indices and income. Then, let H_j denote an individual hypothesis that ρ_i^k is neither correlated nor cointegrated with variable j. We can write

$$\begin{aligned} H_i^k &= H_{R(asda)} \cap H_{R(morrisons)} \cap H_{R(sainsbury)} \cap H_{R(tesco)} \cap z, \\ H_i^k &= \cap_j H_j. \end{aligned}$$

Table (2) shows the FWER of the original Bonferroni procedure and the Hochberg procedure. A critical level of 0.1 is chosen to account for the low power of cointegration test and small sample size (78 observations). We can see that when the number of associated individual hypothesis (s) gets large, the Bonferroni's FWER becomes very small $(FWER = \alpha/s)$. This makes it more unlikely for us to reject any individual null hypothesis. The Hochberg procedure, on the other hand, remedies this shortcoming by increasing the FWER for each individual hypothesis with the ordering of their p-value $(FWER_i = \alpha/(s - i + 1)$ where i is the ascending ordering of the p-value). The modification improves the power of the test without adding calculation burden.

Table 2: Comparison of the family-wise error rates (FWER)

Order of the p-value (i) (1 = smallest)	Bonferroni significance levels $p_{(i)} \leq \alpha/s$	Hochberg significance levels $p_{(i)} \leq \alpha/(s-i+1)$
1	0.025	0.020
2	0.025	0.025
3	0.025	0.033
4	0.025	0.050
5	0.025	0.100
Critical level for single hypothesis testing $\alpha = 0.1$		

5 Results

Table (3) reports aggregation test results. The first two columns define ρ_i^k. The next five columns report p-values from correlation or cointegration test between ρ_i^k and $R_{Asda,}R_{Morrisons}$, $R_{Sainsbury}$, R_{Tesco} and z respectively. If both ρ_i^k and R_j (or z) are nonstationary, we report the p-value from bivariate cointegration test with number of lags in brackets. If both ρ_i^k and R_j (or z) are stationary, we report the p-value from Spearman's rank correlation test. The last column reports the family-wise test result. Product i in firm k passes the test if all the p-values are higher than the Hochberg's FWER (see table(2)).

According to our 3-step test, all log price indices R_j are nonstationary, log income z is stationary, and most log relative prices ρ_i^k are nonstationary. Hence, all reported p-values between ρ_i^k and R_j are from cointegration test while, all reported p-values between ρ_i^k and z are from Spearman's rank correlation test. Blank cells mean that one variable is stationary and another is not. Therefore, no test is required. The two variable would neither be correlated nor cointegrated (Granger and Hallahan, 1989).

In terms of family-wise test, three out of 44 individual products fail. Those are cooked meats in Morrisons, cooked meat in Sainsbury, and instant coffee in Sainsbury. Thus, for Asda and Tesco, aggregation is valid for all product categories. For Morrisons, aggregation is valid for all categories except for cooked meats. For Sainsbury, aggregation is valid for all categories except for cooked meat and instant coffee.

It is worth noting that those three products failed the test because of one common reason–they are correlated with income. Here, we need to keep in mind that income is proxied by per-capita expenditure on grocery. Since prices of grocery items are more likely to be correlated with expenditure on grocery than income, the significance of correlations are likely to be overstated in our analysis. Unfortunately, actual income is not available on a bi-weekly basis. The best that we can do is to acknowledge the existence of this bias.

In sum, our results justify the aggregation of most grocery products at firm level. Only three out of 44 relative prices ρ_i^k failed the family-wise test because they are correlated with expenditure on grocery. We expect these correlations to weaken or even disappear had the actual income been used. It is, therefore, possible that all the ρ_i^k actually pass the family-wise test. In that best possible case, our proposed aggregation is completely valid.

6 Conclusion

This paper uses the generalized composite commodity theorem (GCCT) to test for valid aggregation of grocery items at firm level. The data is collected through homescan technology over a 78 bi-week time period (3 years). The analysis includes 11 most popular product categories in four biggest supermarket firms in the UK, namely Asda, Morrisons, Sainsbury and Tesco. This amounts to 44 distinct products and four product groups–one for each firm–in total. We propose a 3-step procedure to test Lewbel's assumptions for valid aggregation. Strong empirical support was found for 41 out of 44 products. All products in Asda and Tesco pass the test, ten out of eleven products in Morrisons pass the test, and nine out of eleven products in Sainsbury pass the test. Three products fail the test because they are correlated with expenditure on grocery, our proxy for income. Because relative prices are less likely to be correlated with income than with expenditure on grocery, it is possible that all those three products pass the test if the actual income is used. In that case, the GCCT would provide unanimously support for all aggregated product groups.

Table 3: Aggregation Test Results for Grocery Items by Firms

Relative Price (ρ_i^k)		Firm-Level Price Index (R_j)				Income	Hochberg test
k=	i=	Asda	Morrisons	Sainsbury	Tesco	z	$H_i^k = \bigcap_j H_j$
Asda	Milk	0.786(3)	0.559(2)	0.828(3)	0.859(3)	-	pass
	Vegetables	0.626(2)	0.191(2)	0.057(4)	0.465(3)	-	pass
	Fruits		stationary			0.835	pass
	Cheese	0.134(3)	0.197(3)	0.778(3)	0.181(2)	-	pass
	Cooked Meats	0.419(2)	0.227(2)	0.565(3)	0.271(3)	-	pass
	Fresh Poultry	0.455(2)	0.985(2)	0.458(2)	0.688(2)	-	pass
	Cereal	0.243(3)	0.559(2)	0.288(3)	0.533(3)	-	pass
	Bread	0.306(4)	0.903(2)	0.259(2)	0.541(2)	-	pass
	Yoghurt	0.995(3)	0.729(1)	0.228(3)	0.882(3)	-	pass
	Instant Coffee	0.212(2)	0.386(3)	0.075(3)	0.097(3)	-	pass
	Eggs	0.720(2)	0.737(2)	0.958(3)	0.945(3)	-	pass
Morrisons	Milk	0.155(2)	0.236(2)	0.277(3)	0.043(1)	-	pass
	Vegetables	0.463(2)	0.394(3)	0.343(2)	0.117(2)	-	pass
	Fruits	0.491(1)	0.069(1)	0.302(1)	0.095(1)	-	pass
	Cheese	0.083(2)	0.285(1)	0.436(2)	0.086(2)	-	pass
	Cooked Meats		stationary			0.015	fail
	Fresh Poultry	0.268(2)	0.313(1)	0.150(2)	0.107(1)	-	pass
	Cereal		stationary			0.765	pass
	Bread	0.097(2)	0.090(1)	0.043(1)	0.381(3)	-	pass
	Yoghurt	0.686(1)	0.238(1)	0.356(1)	0.118(1)	-	pass
	Instant Coffee		stationary			0.233	pass
	Eggs	0.537(3)	0.179(1)	0.045(1)	0.023(1)	-	pass
Sainsbury	Milk	0.622(3)	0.108(2)	0.523(3)	0.293(3)	-	pass
	Vegetables	0.491(2)	0.610(2)	0.032(2)	0.142(2)	-	pass
	Fruits		stationary			0.263	pass
	Cheese		stationary			0.078	pass
	Cooked Meats		stationary			0.003	fail
	Fresh Poultry	0.555(2)	0.612(2)	0.921(3)	0.139(2)	-	pass
	Cereal		stationary			0.089	pass
	Bread	0.598(1)	0.153(1)	0.393(1)	0.118(1)	-	pass
	Yoghurt	0.213(3)	0.301(2)	0.211(4)	0.176(4)	-	pass
	Instant Coffee		stationary			0.003	fail
	Eggs	0.094(2)	0.353(2)	0.954(3)	0.081(3)	-	pass
Tesco	Milk	0.684(2)	0.172(2)	0.796(3)	0.351(2)	-	pass
	Vegetables	0.809(2)	0.895(2)	0.102(2)	0.350(2)	-	pass
	Fruits	0.446(2)	0.210(2)	0.673(3)	0.548(2)	-	pass
	Cheese	0.749(3)	0.599(2)	0.403(1)	0.465(2)	-	pass
	Cooked Meats	0.525(4)	0.518(3)	0.172(4)	0.596(5)	-	pass
	Fresh Poultry	0.558(2)	0.397(2)	0.203(2)	0.216(3)	-	pass
	Cereal	0.079(2)	0.194(2)	0.065(2)	0.105(2)	-	pass
	Bread		stationary			0.166	pass
	Yoghurt	0.163(2)	0.453(2)	0.129(2)	0.199(2)	-	pass
	Instant coffee	0.290(3)	0.017(4)	0.369(3)	0.213(3)	-	pass
	Eggs	0.441(2)	0.777(2)	0.861(3)	0.388(2)	-	pass

References

[1] Ball, V. E. (1996). U.S. Agricultural Data Set. U. S. D. o. Agriculture, U.S. Department of Agriculture, Washington DC.

[2] Capp, O. J. and A. H. Love (2002). "Econometric Considerations in the Use of Electronic Scanner Data to Conduct Consumer Demand Analysis " American Journal of Agricultural Economic 84(3): 807-816.

[3] Davis, G. C. (2003). "The Generalized Composite Commodity Theorem: Stronger Support in the Presence of Data Limitations." The Review of Economics and Statistics 85(2): 476-480.

[4] Davis, G. C., N. Lin, et al. (2000). "Aggregation without Separability: Tests of the United States and Mexican Agricultural Production Data." American Journal of Agricultural Economic 82(1): 214-230.

[5] Doornik, J. A. and D. Hendry (2001). Modelling Dynamic Systems using PcGive: Volume II. London, Timberlake Consultants Ltd.

[6] Doornik, J. A. and D. Hendry (2001). Modelling Dynamic Systems using PcGive: VolumeI. London, Timberlake Consultants Ltd.

[7] Granger, C. W. J. and J. Hallman (1989). The Algebra of I(1). Finance and Economics Discussion Series, Paper 45, Division of Research and Statistics, Federal Reserve Board, Washington DC.

[8] Hicks, J. R. (1936). Value and Capital. Oxford, Oxford University Press.

[9] Hochberg, Y. (1988). "A Sharper Bonferroni Procedure for Multiple Tests of Significance." Biometrika 75: 800-802.

[10] Johansen, S. (1995). Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models. Oxford, Oxford University Press.

[11] Lehmann, E. L. and J. P. Romano (2006). Testing Statistical Hypotheses. New York, Springer.

[12] Leontief, W. (1936). "Composite Commodities and the Problem of Index Numbers." Econometrica 4: 39-59.

[13] Lewbel, A. (1996). "Aggregation Without Separability: A Generalized Composite Commodity Theorem." The American Economic Review 86(3): 524-543.

[14] Lo, A. and C. Mackinley (1989). "The Size and Power of the Variance Ratio Test in Finite Samples: a Monte Carlo Investigation." Journal of Econometrics 40: 203-238.

[15] Mendenhall, W., R. L. Scheaffer, et al. (1990). Mathematical Statistics with Applications. Boston, PWS-Kent Publishing Company.

[16] Reed, A. J., J. W. Levedahl, et al. (2005). "The Generalized Composite Commodity Theorem and Food Demand Estimation." American Journal of Agricultural Economic 87(1): 28-37.

[17] Theil, H. (1971). Principles of Econometrics. New York, Wiley.