RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

IFN Working Paper No. 1308, 2019

Index Numbers and Revealed Preference Rankings

Per Hjertstrand, James L. Swofford and Gerald A. Whitney

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Index Numbers and Revealed Preference Rankings[*]

Per Hjertstrand
Research Institute of Industrial Economics
per.hjertstrand@ifn.se

James L. Swofford
Department of Economics and Finance, University of South Alabama
jswoffor@southalabama.edu

Gerald A. Whitney
Department of Economics and Finance, University of New Orleans
gwhitney@uno.edu

ABSTRACT

For previously identified weakly separable blockings of goods and assets, we construct aggregates using four superlative index numbers, the Fisher, Sato-Vartia, Törnqvist and Walsh, two non-superlative indexes, the Laspeyres and Paasche and the atheoretical simple summation. We conduct several tests to examine how well each of these aggregates "fit" the data. These tests are how close the aggregates come to solving the revealed preference conditions for weak separability, how often each aggregate gets the direction of change correct and how well the aggregates mimic the preference ranking from revealed preference tests. We find that, as the number of goods and assets being aggregated increases, the problems with simple summation manifest.

JEL Code: C43

Keywords: Superlative index numbers; Aggregation; Preference orderings; Weak separability

[*] We thank two referees, Bill Barnett, and participants at the Society for Economic Measurement conference in Thessaloniki in July 2016 and the 3rd HenU/InferWorkshop on AppliedMacroeconomics inMarch 2017 in Kaifeng for very helpful comments. Hjertstrand would like to thank Jan Wallander och Tom Hedelius stiftelse, Marcus och Marianne Wallenberg stiftelse, and Johan och Jakob Söderberg stiftelse for financial support.

1. INTRODUCTION

Samuelson (1983, p. xx) discussed the relationship of index number theory to revealed preference analysis saying "Index number theory is shown to be merely an aspect of the theory of revealed preference. [...] this is the point of revealed preference–knowledge of but two (P, Q) situations (or of a limited number of situations) can at best put bounds on each one of our sought-for ratios."

Nonparametric revealed preference analysis is often used to test whether a data set composed of prices and quantities of goods can be rationalized by a well behaved utility function that is weakly separable in some subset of goods. Weakly separable utility, as its name implies, allows for a separation of a subset of goods into a sub-utility or aggregator function. Thus, weak separability is a necessary condition for the existence of an economic aggregate. The existence of such an economic aggregate is the principal justification for the use of superlative or other index numbers to aggregate over goods. A superlative index is exact when it can provide a second order approximation to a particular aggregator function.

Varian's (1983) revealed preference conditions for weakly separable utility maximization has the advantage of not relying on any particular functional form for the utility function. The solution values for the conditions can be interpreted as representing levels of utility for the utility and sub-utility functions. These solution values, while not necessarily unique, are consistent with the preferences revealed by the data. If period t is preferred, directly or indirectly, to period s, then the utility level assigned to t is greater than or equal to that assigned to s. However, since indexes, superlative or otherwise, need not mirror preferences, this property is not guaranteed to hold for index numbers.

This paper examines how well such aggregates "fit" the data. That is, we investigate how close these aggregates come to solving the revealed preference conditions for weak separability, when it is known from disaggregated data that a solution exists. We use the solution values from the revealed preference conditions to check how well the aggregates reflect the preference orderings in the data using three tests. First, we calculate how much each index needs to be perturbed in order to satisfy the revealed preference conditions for weak separability. Second, we investigate how often the aggregates get the revealed preference direction of change right between adjacent time periods. Third, we compare non-adjacent periods by calculating how often the aggregates correctly rank a bundle to be revealed preferred to any other bundle.

Barnett and Choi (2008) suggested a definition that spans all index numbers that are superlative. We consider aggregates based on four of the most well-known superlative index numbers, Fisher, Sato-Vartia, Törnqvist and Walsh.[1] We include the Fisher index because of its superior axiomatic and economic justifications (Diewert, 1993), and the Sato-Vartia index for its "consistency in aggregation" property.[2] The Fisher and Sato-Vartia indexes satisfy more of the axioms describing a desirable index than any other superlative index.[3] The Törnqvist index is best from a stochastic view point and widely used in monetary economics. Since some of our aggregates are from monetary data, we include it in our analysis. We include the Walsh index because, as discussed in IMF (2005), it is best from the view point of a pure index number.[4]

We also consider the Paasche and Laspeyres indexes, which are not superlative, but exact for a first order or linear approximation to an aggregator function. These two indexes are commonly used by many statistical agencies to calculate GDP and construct GDP deflators. A version of the Laspeyres index known as the Lowe index is commonly used by agencies to create consumer price indexes. Finally, because of its wide-spread use by central banks to construct monetary aggregates, we also include simple summation in our analysis.[5]

We find little difference among index numbers and simple summation when aggregating over a small number of goods. However, as the level of aggregation increases, the problems with simple summation emerge. This is consistent with the critique of simple summation indexes by Fisher (1922) and the Barnett (1980) critique of simple sum monetary aggregates.[6]

We review the economics of aggregation over goods in the following section. We also discuss revealed preference testing and the test procedure used to generate the results reported in this paper.

2. AGGREGATION OVER GOODS AND WEAK SEPARABILITY

Economic models typically start with a representative consumer faced with maximizing utility over a lifetime.[7] To make empirical models more tractable, empirical studies often model and examine the representative agent's decision as involving a small number of aggregated goods and assets.

Leontief (1936) showed that a set of goods and assets can be aggregated into a single composite good when the quantity vector in each time period is proportional to a base quantity vector.[8] In such cases, the long term bilateral quantity index equals the coefficient of proportionality at each time period (Diewert, 1993). However, when relative quantities are not proportional, aggregation becomes a more complex issue.

Suppose that there are K goods and assets. Let $\boldsymbol{x} = (x_1, \dots, x_K)$ denote the quantity-vector with the corresponding price-vector $\boldsymbol{p} = (p_1, \dots, p_K)$. Suppose that the goods and assets are split into two sub-groups $\boldsymbol{y} = (y_1, \dots, y_H) = (x_1, \dots, x_H)$ and $\boldsymbol{w} = (w_1, \dots, w_{K-H}) = (x_{H+1}, \dots, x_K)$. Let $\boldsymbol{r} = (r_1, \dots, r_H) = (p_1, \dots, p_H)$ denote the prices of the $\boldsymbol{y}$-goods and $\boldsymbol{z} = (z_1, \dots, z_{K-H}) =$

$(p_{H+1}, \dots, p_K)$ denote the prices of the goods and assets in $\boldsymbol{w}$. We assume that prices and quantities are observed at T time periods indexed by $\mathbb{T} = \{1, \dots, T\}$. An observation of the prices and quantities at time $t \in \mathbb{T}$ is denoted $\boldsymbol{p}_t = \{\boldsymbol{r}_t, \boldsymbol{z}_t\}$ and $\boldsymbol{x}_t = \{\boldsymbol{y}_t, \boldsymbol{w}_t\}$, respectively. We write $\mathbb{D} = \{\boldsymbol{r}_{t,}, \boldsymbol{z}_t; \boldsymbol{y}_t, \boldsymbol{w}_t\}_{t \in \mathbb{T}}$ to signify all price-quantity observations and refer to $\mathbb{D}$ as "the data".

A utility function $U(\boldsymbol{x})$ is said to be weakly separable in the $\boldsymbol{y}$ block of goods if there exists a macro function u and a sub-utility function V such that U can be written as $U(\boldsymbol{x}) = U(\boldsymbol{y}, \boldsymbol{w}) = u(V(\boldsymbol{y}), \boldsymbol{w})$. Weak separability is a necessary condition to construct an aggregate of goods and assets when relative quantities are *not* proportional.

We say that the data $\mathbb{D}$ can be *rationalized* by a weakly separable utility function if there exist well-behaved, i.e. continuous, concave and strictly increasing, functions u and V, such that $\{\boldsymbol{y}_t; \boldsymbol{w}_t\}_{t \in \mathbb{T}}$ solves the utility maximizing problem, i.e.,

$$\{\boldsymbol{y}_t; \boldsymbol{w}_t\}_{t \in \mathbb{T}} \quad solves \quad \max_{\{\boldsymbol{y}, \boldsymbol{w}\}} u(V(\boldsymbol{y}), \boldsymbol{w}) \quad s.t. \quad \boldsymbol{r}_t \boldsymbol{y} + \boldsymbol{z}_t \boldsymbol{w} \leq \boldsymbol{r}_t \boldsymbol{y}_t + \boldsymbol{z}_t \boldsymbol{w}_t.$$

Varian's (1983) necessary and sufficient non-parametric revealed preference conditions for weak separability are summarized in the following theorem.

Theorem 1 (Varian, 1983). Consider the data set $\mathbb{D} = \{\boldsymbol{r}_{t,}, \boldsymbol{z}_t; \boldsymbol{y}_t, \boldsymbol{w}_t\}_{t \in \mathbb{T}}$. *Conditions (a)-(d) are equivalent:*

(a) There exists a weakly separable (in the $\boldsymbol{y}$*-goods), concave, strictly increasing and continuous utility function rationalizing the data* $\mathbb{D}$.

(b) There exist numbers V_t *and* $\phi_t > 0$ *such that (for all* $s, t \in \mathbb{T}$*):*

$$V_t - V_s - \phi_t \boldsymbol{r}_t (\boldsymbol{y}_t - \boldsymbol{y}_s) \geq 0 \tag{1}$$

$$\{{\phi_t}^{-1}, \boldsymbol{z}_t; V_t, \boldsymbol{w}_t\}_{t \in \mathbb{T}} \text{ satisfies GARP}^{9} \tag{2}$$

(c) There exist numbers $V_t, U_t, \phi_t > 0$, *and* $\lambda_t > 0$ *such that (for all* $s, t \in \mathbb{T}$*):*

$$V_t - V_s - \phi_t \boldsymbol{r}_t(\boldsymbol{y}_t - \boldsymbol{y}_s) \geq 0 \quad (3)$$

$$U_t - U_s - \lambda_t \boldsymbol{z}_t(\boldsymbol{w}_t - \boldsymbol{w}_s) - \lambda_t(V_t - V_s)/\phi_t \geq 0 \quad (4)$$

(d) There exist numbers V_t, u_t and $\phi_t > 0$ such that (for all $s, t \in \mathbb{T}$):

$$V_t - V_s - \phi_t \boldsymbol{r}_t(\boldsymbol{y}_t - \boldsymbol{y}_s) \geq 0 \quad (5)$$

$$if\ \phi_t \boldsymbol{z}_t(\boldsymbol{w}_t - \boldsymbol{w}_s) + (V_t - V_s) \geq 0 \quad then\ u_t - u_s \geq 0 \quad (6)$$

$$if\ \phi_t \boldsymbol{z}_t(\boldsymbol{w}_t - \boldsymbol{w}_s) + (V_t - V_s) > 0 \quad then\ u_t - u_s > 0 \quad (7)$$

Conditions *(b), (c)* and *(d)* are testable conditions that can be used to check whether a data set can be rationalized by a well-behaved and weakly separable utility function.

Fleissig and Whitney (2003) suggested a sequential procedure to implement condition *(b)*.[10] The idea is, in a first step, to find the minimal perturbation of some superlative quantity index and its price index found from the adding up condition that satisfy (1), and in a second step, use these in place of V_t and ϕ_t, when testing whether (2) holds. Thus, this procedure can be used to analyze how “close” a superlative index is to satisfying the ordering of preferences discovered from revealed preference conditions. However as discussed in Hjertstrand (2009) because of its sequential nature, the procedure may over-reject weak separability.

In order to circumvent the problem of over-rejecting weak separability, several test procedures based on condition *(c)* have been proposed in the literature. See for examples, Diewert and Parkan (1985), Swofford and Whitney (1994), Fleissig and Whitney (2008) and Elger and Jones (2008).[11] However, any procedure based on *(c)* may easily become practically infeasible for large data sets. This follows because the inequalities (4) are non-linear in the term $\lambda_t(V_t - V_s)/\phi_t$, and any procedure based on this condition is therefore prone to solve a complex non-linear constrained optimization problem.

To remedy the computational complexity of implementing condition *(c)*, Cherchye, Demuynck, De Rock and Hjertstrand (2015) proposed a test procedure based on condition *(d)*. They show that (5)-(7) in condition *(d)* can be formulated as a mixed integer linear programming (MILP) problem. A MILP problem is very similar to a standard linear programming (LP) problem in the sense that it minimizes a linear objective function subject to a set of linear constraints. But in contrast to a LP problem, the variables in a MILP problem may also take integer or binary values.[12]

Cherchye et al.'s (2015) procedure is based on the result that the inequalities (5)-(7) in condition *(d)* are equivalent to that the following inequalities hold:[13]

$$V_s - V_t - \phi_t \boldsymbol{r}_t(\boldsymbol{y}_s - \boldsymbol{y}_t) \leq 0 \tag{8}$$

$$u_t - u_s - X_{t,s} \leq -\psi \tag{9}$$

$$\left(X_{t,s} - 1\right) \leq u_t - u_s \tag{10}$$

$$\phi_t \boldsymbol{z}_t(\boldsymbol{w}_t - \boldsymbol{w}_s) + (V_t - V_s) - X_{t,s}A_t \leq -\psi \tag{11}$$

$$\left(X_{t,s} - 1\right)A_s \leq \phi_s \boldsymbol{z}_s(\boldsymbol{w}_t - \boldsymbol{w}_s) + (V_t - V_s) \tag{12}$$

$$0 \leq V_t \leq 1 \tag{13}$$

$$0 \leq u_t \leq 1 - \psi \tag{14}$$

$$\psi \leq \phi_t \leq 1 \tag{15}$$

$$X_{t,s} \in \{0,1\} \tag{16}$$

The binary (0-1) variables $X_{t,s}$ captures the logical relation between the inequalities in (6) and (7), and equals one if and only if $u_t - u_s \geq 0$. ψ is a small positive number and A_t is a fixed number larger than $\boldsymbol{z}_t \boldsymbol{w}_t + 1$.

3. FITTING QUANTITY INDEXES

The purpose of this paper is to analyze how 'close' the aggregates come to solving the revealed preference conditions for weak separability. For this, we suggest a test procedure that combines Cherchye et al.'s (2015) MILP problem with Fleissig and Whitney's (2003) idea of perturbing the quantity index such that it satisfies the weak separability conditions. This procedure therefore inherits many of the computational advantages of Cherchye et al.'s (2015) problem.

Let Q_t be a period t quantity index or aggregate constructed from the price and quantity data $\{\boldsymbol{r}_t; \boldsymbol{y}_t\}$ for the separable group. Even if the utility function $U(\boldsymbol{y}, \boldsymbol{w})$ is weakly separable in the $\boldsymbol{y}$-goods such that $U(\boldsymbol{y}, \boldsymbol{w}) = u(V(\boldsymbol{y}), \boldsymbol{w})$, there is no guarantee that substituting Q_t in place of V_t in the inequalities (8)-(16) will preserve the revealed preference ordering. Following Fleissig and Whitney (2003) we propose to calculate the minimal perturbation of the quantity index Q_t such that the inequalities hold. Because the quantity index Q_t is normalized to some arbitrary constant in a given base period, it is inappropriate to use a model where the perturbation enters additively. Hence, we specify a proportional model which is invariant to the chosen base year:

$$\tilde{Q}_t = Q_t(1 + \varepsilon_t), \tag{17}$$

where $\tilde{Q}_t$ is the perturbed quantity index that satisfies weak separability, and $\varepsilon_t = \tilde{Q}_t / Q_t - 1$ is the required perturbation or error. In discussing the results we refer to $\tilde{Q}_t$ as the "corrected Afriat index".[14] We suggest calculating the minimum total squared percentage perturbation such that $\tilde{Q}_t$ satisfies the weak separability conditions by solving the following mixed integer quadratic programming (MIQP) problem:[15]

$$\min_{\{\tilde{Q}_t, \phi_t, u_t, X_{t,s}\}_{s,t \in \mathbb{T}}} \frac{100^2}{T} \sum_{t=1}^{T} \left(\frac{\tilde{Q}_t}{Q_t} - 1 \right)^2 \quad s.t.$$

$$\tilde{Q}_s - \tilde{Q}_t - \phi_t \boldsymbol{r}_t (\boldsymbol{y}_s - \boldsymbol{y}_t) \leq 0 \tag{18}$$

$$u_t - u_s - X_{t,s} \leq -\psi \tag{19}$$

$$\left(X_{t,s} - 1\right) \leq u_t - u_s \tag{20}$$

$$\phi_t \boldsymbol{z}_t(\boldsymbol{w}_t - \boldsymbol{w}_s) + \left(\tilde{Q}_t - \tilde{Q}_s\right) - X_{t,s} A_t \leq -\psi \tag{21}$$

$$\left(X_{t,s} - 1\right) A_s \leq \phi_s \boldsymbol{z}_s(\boldsymbol{w_t} - \boldsymbol{w}_s) + \left(\tilde{Q}_t - \tilde{Q}_s\right) \tag{22}$$

$$0 \leq \tilde{Q}_t \leq \eta_t \tag{23}$$

$$0 \leq u_t \leq 1 - \psi \tag{24}$$

$$\psi \leq \phi_t \leq 1 \tag{25}$$

$$X_{t,s} \in \{0,1\} \tag{26}$$

Let $\{\hat{Q}_t\}_{t \in \mathbb{T}}$ be the optimal solutions of the corrected Afriat indexes from this problem.[16] Define the optimal perturbations as $\hat{\varepsilon}_t = \hat{Q}_t / Q_t - 1$ for all $t \in \mathbb{T}$. Thus, $100 \times \hat{\varepsilon}_t$ is the minimal percentage amount of perturbation in period t that is required for the weak separability conditions (18)-(26) to hold.

We implement the MIQP problem to compute the required minimal perturbations with Q_t set to the Fisher, Sato-Vartia, Törnqvist, Walsh, Paasche, and Laspeyres indexes and simple summation. To assess how well Q_t "fit" the data we use several tests.

As a first test, because the model (17) is invariant to the chosen base year we can compare the amount of perturbation calculated from the different indexes. We can also compare the number of values of each aggregate that must be perturbed. In Table 1 we define three summary statistics used to compare the required perturbations of each aggregate.

[TABLE 1 HERE]

We use the percentage mean error to see if errors in aggregation are systematically biased in one direction or another. We also calculate the percentage root mean squared error since it avoids positive and negative errors canceling each other out and making the fit seem better than it is.[17]

Finally, we calculate the percentage of adjusted values to check if few or many values of each aggregate must be adjusted to satisfy the revealed preference conditions.

As a second test, we compare the direction of change between adjacent periods for the corrected Afriat index and the quantity index. That is, we check how often the sign of the growth rates $\hat{Q}_t/\hat{Q}_{t-1} - 1$ and $Q_t/Q_{t-1} - 1$ differ.

However, comparing growth rates between adjacent periods only allows for comparisons involving those particular consecutive time periods. With our third test we check the preference rankings between the corrected Afriat index and each aggregate to allow for comparisons involving non-adjacent periods. As such, we can interpret $\tilde{Q}_t$ and Q_t as representing utility levels in period t, and use that the period t bundle of goods and assets is directly revealed preferred to the period s bundle *if and only if* the utility level assigned to t is greater than or equal to that assigned to s. Correspondingly, we can define the $T \times T$ directly revealed preferred matrices $\hat{Q}M$ and QM where the $t, s -$ element, given by $\hat{Q}M_{ts}$ and QM_{ts}, equals 1 if $\hat{Q}_t \geq \hat{Q}_s$ and $Q_t \geq Q_s$, and 0 otherwise. An inconsistency between the corrected Afriat index and the quantity index occurs if $\hat{Q}M_{ts} \neq QM_{ts}$. In this case, the index indicates that the representative agent prefers the period t bundle over the period s bundle or vice versa, while the corrected Afriat index indicates the converse. In total, there can be $T \times (T - 1)$ inconsistences or unequal elements between $\hat{Q}M$ and QM of which $T \times (T - 1)/2$ are unique.

In the following section we discuss superlative index numbers and the particular index numbers used for the comparisons.

4. INDEX NUMBERS

Diewert (1976 and 1978) originated the concept of superlative index numbers. He defined a flexible functional form as a second order approximation to an arbitrary aggregator function, e.g.

a utility, cost, production or distance function. A superlative index number is exact when it can be derived from a particular flexible functional form. Among the index numbers that fit within Diewert's definition are the Fisher ideal and the Walsh indexes, while the Törnqvist index is on the open boundary of the function set in the definition.

Barnett and Choi (2008) pointed out that Diewert's definition only covers a strict subset of all superlative index numbers and suggested an alternative definition that spans all index numbers that are superlative. In addition, Barnett and Choi (2008) identified an alternative class of superlative indexes that they named the Theil-Sato indexes, which are in general terms log change index numbers with symmetric mean weights. The class of Theil-Sato index numbers spans, for example, the Törnqvist and Walsh indexes, and the Sato-Vartia index.

So in our comparisons we consider the Fisher ideal, the Sato-Vartia, the Törnqvist and the Walsh indexes. All these are consistent with Barnett and Choi's (2008) operational definition of a superlative index number. We include the Fisher and Sato-Vartia indexes because they are superior from the axiomatic approach (Diewert, 1993). We include the Törnqvist index because it is best from a stochastic view point and widely used in monetary economics. Moreover, we include the Walsh index because, as discussed in IMF (2005), it is best from the view point of a pure index number.

We also include the Paasche and Laspeyres indexes, which are exact for a first order approximation of some aggregator function. These indexes are commonly used by many government statistical agencies.

Barnett (1980) initiated a large literature criticizing the use of simple summation to construct monetary aggregates that Chrystal and MacDonald (1994) called the Barnett critique. Simple summation is only theoretically valid when the quantities are proportional to a base vector and, even more restrictively, it requires that the degree of proportionality equals one. Thus, simple summation implicitly assumes that all goods and assets in the aggregate are perfect substitutes and that the substitutability must be in a one to one ratio. Although this is very restrictive and highly unlikely, simple summation is still frequently used by, for example, central banks to construct monetary aggregates. Since some of our comparisons are from monetary data and monetary aggregation, we also include simple summation in our analysis.

Let a_{jt} be the period t expenditure share for good $j = 1, \dots, H$:

$$a_{jt} = \frac{r_{jt} y_{jt}}{\sum_{l=1}^{H} r_{lt} y_{lt}}.$$

The superlative indexes we consider are defined as follows:

- The Fisher ideal quantity index:

$$Q_t^F = Q_{t-1}^F \sqrt{\frac{\left(\sum_{j=1}^{H} r_{jt} y_{jt}\right)\left(\sum_{j=1}^{H} r_{jt-1} y_{jt}\right)}{\left(\sum_{j=1}^{H} r_{jt} y_{jt-1}\right)\left(\sum_{j=1}^{H} r_{jt-1} y_{jt-1}\right)}}.$$

- The Sato-Vartia quantity index:

$$Q_t^{SV} = Q_{t-1}^{SV} \prod_{j=1}^{H} \left(\frac{y_{jt}}{y_{jt-1}}\right)^{\left(\frac{a_{jt} - a_{jt-1}}{\log a_{jt} - \log a_{jt-1}}\right)}.$$

- The Törnqvist quantity index:

$$Q_t^T = Q_{t-1}^T \prod_{j=1}^{H} \left(\frac{y_{jt}}{y_{jt-1}}\right)^{\frac{1}{2}(a_{jt} + a_{jt-1})}.$$

- The Walsh quantity index:

$$Q_t^W = Q_{t-1}^W \prod_{j=1}^{H} \left(\frac{y_{jt}}{y_{jt-1}}\right)^{\sqrt{a_{jt} \times a_{jt-1}}}.$$

The non-superlative indexes we consider are defined as follows:

- The Paasche quantity index:

$$Q_t^P = Q_{t-1}^P \frac{\sum_{j=1}^H r_{jt} y_{jt}}{\sum_{j=1}^H r_{jt} y_{jt-1}}.$$

- The Laspeyres quantity index:

$$Q_t^L = Q_{t-1}^L \frac{\sum_{j=1}^H r_{jt-1} y_{jt}}{\sum_{j=1}^H r_{jt-1} y_{jt-1}}.$$

- The simple sum quantity aggregate:

$$Q_t^{ss} = \sum_{j=1}^H y_{jt}.$$

We next set forth the data used in the comparisons.

5. DATA

The data are the same as in Hjertstrand, Swofford and Whitney (2016). The data are annualized quarterly per person 16 years of age and older U.S. data and cover the period from the first quarter of 2000 to the third quarter of 2011. Expenditures are from the Bureau of Economic Analysis and civilian labor force over age 16 data is from the Current Population Survey. Hours and hourly wage rates were obtained from the Economic Report of the President.

Hjertstrand et al. (2016) examined four categories of consumption goods and leisure:

1. *SER*: real expenditures on services
2. *NDUR*: real expenditures on nondurables
3. *DUR*: real expenditures on durables
4. *LEIS*: hours of leisure.

The prices of services and nondurables are the respective implicit price deflators. The price of durables is a user cost. An annualized ten percent depreciation rate was applied each quarter to annualize expenditures on durables to make it compatible with annualized expenditures on services and nondurables. Leisure is calculated as 98 hours minus average hours worked per week during the quarter. Hours worked are subtracted from 98 hours to allow time each week for required non-market and non-leisure activities like sleeping and eating.

The monetary assets used were:

5. *CUR+DD*: currency plus demand deposits
6. *TC*: traveler's checks
7. *OCDCB* and *OCDTH*: other checkable deposits at commercial banks and thrifts.
8. *SDCB* and *SDTH*: savings deposits at commercial banks and thrifts.
9. *STDCB* and *STDTH*: small time deposits at commercial banks and thrifts.
10. *MMFR* and *MMFI*: retail and institutional money market mutual funds.
11. *TB*: treasury bills
12. *CP*: commercial paper.
13. *LTD*: large time deposits

These monetary assets and associated user costs were obtained from the Center for Financial Stability (CFS). Demand deposits, other checkable deposits and savings deposits were adjusted for retail sweeps.[18] The monetary goods were deflated by the implicit price deflator to obtain real per capita balances. The user costs for the monetary assets were multiplied by the implicit price deflator to yield nominal prices of a dollar of real balances.

Now that we have described the tests and the data set used, we present the results from analyzing how well the quantity indexes fit these data.

6. RESULTS

Using the data described in the previous section, Hjertstrand et al. (2016) applied the test procedure proposed by Cherchye et al. (2015) to test whether "money" is weakly separable from consumption and leisure and if major categories of consumption expenditures are weakly separable from monetary assets. This is important because weak separability between "money" and consumption and leisure is required for the existence of a monetary aggregate, and consequently, restricts the effect of money on real economic activity. Weak separability of consumption goods from leisure and monetary assets must hold in order for consumption to have a stable relationship with income.

Hjertstrand et al. (2016) found that the five utility structures reported in Table 2 are weakly separable.

[TABLE 2 HERE]

Thus, the goods and assets in the sub-utility function $V(\cdot)$ in each of the five structures are weakly separable from all other goods and assets. Structures 1-3 are different monetary aggregates. In particular, Structure 1 is the FED aggregate M1 and Structure 2 is the modern analog of what Friedman and Schwartz (1963) called money in the U.S. Structure 3 is equivalent to the FED old liquidity aggregate L. Structures 4 and 5 are narrow and broader real sector consumption aggregates, respectively.

For each of these five structures and for each quantity aggregate in Section 4, we generate solutions to the corrected Afriat index as described in Section 3. Figures 1-5 present graphs of the calculated perturbed errors for each index in every structure.

[FIGURE 1 HERE]

[FIGURE 2 HERE]

[FIGURE 3 HERE]

[FIGURE 4 HERE]

[FIGURE 5 HERE]

We see that the errors of all indexes are close to each other for the narrower monetary aggregates in Structures 1 and 2 (Figures 1 and 2). However, in the broader aggregate in Structure 3 (Figure 3), the errors from simple summation are considerably larger than for the other indexes. This pattern can also be seen between the narrow and broader consumption aggregates in Structures 4 and 5 (Figures 4 and 5), respectively. As in Barnett (1980), we find the four superlative indexes give aggregates that are almost indistinguishable from each other.

In Table 3, we present the calculated summary statistics of the required perturbation for each quantity index in every structure.

[TABLE 3 HERE]

The percentage mean error indicates whether there is a directional bias. The negative entries show that the indexes generally need to be adjusted downwards in order to satisfy the weak separability conditions. Moreover, consistent with Figures 1-5, we see that simple summation produces a larger systematic bias in Structures 3 and 5, where more goods are being aggregated.

The percentage root mean squared error is a measure of the perturbation required for the indexes to satisfy the weak separability conditions for which positive and negative errors do not cancel out. For Structures 1 and 2, where there are a low number of goods aggregated there is little difference between the aggregates, though the simple sum requires slightly less perturbation especially for Structure 2. For the broader aggregate in Structure 3, which contains 13 goods, the superiority of the other indexes, in relation to simple summation, shows up. The performance of the superlative indexes relative to the simple sum is particularly interesting as it provides support for the use of a superlative index to construct monetary aggregates. Similarly, for the real sector consumption aggregates in Structures 4 and 5, as the number of goods aggregated increases from two goods to four goods, the superiority of the other indexes, in relation to simple summation emerges with the widely used Laspeyres error slightly smaller than the superlative indexes. Finally, the statistic "% of adjusted values" in Table 3 shows that the indexes must be perturbed in almost every case in every time period to satisfy the revealed preference inequalities.[19]

We next compare the growth rates of the indexes with the growth rates of the corrected Afriat indexes in adjacent periods. Table 4 presents the percentage of times the growth rate of the index moves in the opposite direction of the growth rate of the corrected Afriat index.

[TABLE 4 HERE]

We find that the superlative indexes always move in the same direction as the corrected Afriat index, while the simple sum aggregate sometimes move in the wrong direction. This occurs in 2% of all comparisons in Structure 1 and in 6.5% of all comparisons in Structure 3. The results for the Laspeyres index are in line with that of the superlative indexes always getting the direction of change correct. However, the Paasche index moves in the wrong direction 2% of the times for Structure 3. These results show that while for some structures the superlative indexes may have to be perturbed more than the simple sum aggregate, the superlative indexes still perform better than the simple sum aggregates in terms of getting the direction of change correct.

In Table 5 we report the percentage of times the preference rankings of bundles between the index and the corrected Afriat index differs.

[TABLE 5 HERE]

As explained in Section 3, in contrast to comparing the growth rates as reported in Table 4, in Table 5 we are reporting results when each time period is compared to all other time periods. We find that the superlative indexes perform best and only rank the bundles incorrectly in at most 0.37% of the comparisons in Structure 5. Simple summation performs worse in every structure while the results of the Paasche and Laspeyres indexes are marginally worse than the superlative indexes.

Overall, we find that the superlative index numbers perform better than the non-superlative indexes when more goods and assets are being aggregated. For the consumption aggregates this may be because the components already have been aggregated up by government agencies using legitimate index numbers. Thus, services (*SUR*) and nondurables (*NDUR*) may already be legitimate aggregates and not much harm is done in this case by simply adding them together. Additionally, services and nondurables may be more substitutable with each other than durables (*DUR*) and leisure (*LEIS*), in which case there is fewer substitution effects for the simple sum to miss in the narrower aggregate in Structure 4 compared to the broader aggregate in Structure 5**.**

One reason for our results showing the superiority of the superlative indexes emerging as the number of goods increase is that smaller groupings may contain closer substitutes. The narrower monetary aggregates in Structures 1 and 2 contain assets including demand deposits and saving accounts that are likely highly substitutable, while the broader aggregate in Structure 3 contains less substitutable assets such as T-BILLS. In fact, the reason we view Structure 2 as a modern analogue of Friedman and Schwartz (1963) money is their argument that money for the U.S. in the mid 20th century should have included savings deposits due to the degree of substitutability between savings demand deposits and currency.[20] All this is consistent with our empirical results.

7. SUMMARY AND CONCLUSIONS

We have compared consumption and monetary aggregates based on four superlative indexes, two non-superlative indexes and the atheoretical simple sum that is only economically valid under very restrictive assumptions. We conduct several tests of whether these aggregates maintain the information and properties of weak separability found from revealed preference tests. Some of our main findings are that:

- In terms of the percentage root mean squared error, all indexes outperformed the simple sum as the number of goods aggregated increased.
- The superlative indexes always have the direction of change right while the Laspeyres and simple sum aggregate sometime have the direction of change wrong.
- In pairwise comparisons and comparisons over all time periods and structures, all indexes outperformed the simple sum. For all structures, no index outperformed the superlative indexes.

We view these results as supporting the use of superlative indexes. Thus, following Fisher (1922) and Barnett (1980) we provide more evidence to the large body of literature criticizing the use of simple summation.

NOTES

[1] The Törnqvist index is a discrete time approximation to the Divisia index.
[2] Vartia (1976) defines an index number formula to be consistent in aggregation if the value of the index calculated in two stages coincides with the value of the index as calculated in a single stage.
[3] See Balk (1995) for a detailed discussion.
[4] IMF (2005, p.451) discusses the relative merits of the Fisher, Törnqvist and Walsh indexes.
[5] Simple summation is an ad hoc or atheoretical aggregate unless all quantity vectors are proportional to some base quantity vector, with the degree of proportionality equal to one. Only in this very restrictive case is simple summation consistent with economic theory.
[6] Fisher (1922, p.29-30) writes in a famous passage "In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers". As one reader of this paper pointed out, Fisher was clearly referring to the Carli index – i.e., the simple average of ratios - and not specifically to the index form version of the simple sum, i.e. the Dutot index, which is a ratio of averages. However, Fisher most likely felt the same resentment towards the Dutot index. This view has been frequently echoed. In their consumer price manual, the ILO writes: "…the Dutot gives more weight to the price relatives for the products with high prices in period 0. It is nevertheless difficult to provide an economic rationale for this kind of weighting. Prices are not expenditures. If the products are homogeneous, very few quantities are likely to be purchased at high prices if the same products can be purchased at low prices. If the products are heterogeneous, the Dutot should not be used anyway, as the quantities are not commensurate and not additive." (ILO, 2004, p.15).
[7] The representative agent assumption is generally made to avoid restrictive assumptions on preferences such as homotheticity or quasi-homotheticity. However as Caves, Christensen and Diewert (1982) have shown even without homotheticity the Törnqvist index that provides a second order approximation to the Malmquist index is both superlative and tracks the distance function.

[8] The concept of Hicksian aggregation is analogous to Leontief aggregation in terms of prices. It states that if all price vectors are proportional to some base price vector then the long term bilateral price index for each time period equals the coefficient of proportionality (Hicks, 1946, p.312-313).
[9] A data set $\{\boldsymbol{f}_t; \boldsymbol{q}_t\}_{t\in\mathbb{T}}$ satisfies the Generalized Axiom of Revealed Preference (GARP) when $\boldsymbol{q}_t R \boldsymbol{q}_s$ implies $\boldsymbol{f}_s\boldsymbol{q}_s \leq \boldsymbol{f}_s\boldsymbol{q}_t$, where R is the revealed preference chain, $\boldsymbol{q}_t R \boldsymbol{q}_s \leftrightarrow \boldsymbol{f}_t\boldsymbol{q}_t \geq \boldsymbol{f}_t\boldsymbol{q}_{s_1}, \boldsymbol{f}_{s_1}\boldsymbol{q}_{s_1} \geq \boldsymbol{f}_{s_1}\boldsymbol{q}_{s_2}, \dots, \boldsymbol{f}_{s_{k-1}}\boldsymbol{q}_{s_{k-1}} \geq \boldsymbol{f}_{s_{k-1}}\boldsymbol{q}_{s_k}, \boldsymbol{f}_{s_k}\boldsymbol{q}_{s_k} \geq \boldsymbol{f}_{s_k}\boldsymbol{q}_s$ for the k additional observations $(s_1, \dots, s_k) \in \mathbb{T}$. Eq. (2) is obtained by defining the $(K+1)$-vectors $\boldsymbol{f}_t = ({\phi_t}^{-1}, \boldsymbol{z}_t)$ and $\boldsymbol{q}_t = (V_t, \boldsymbol{w}_t)$ for values ϕ_t and V_t that satisfies (2) and testing whether $\{\boldsymbol{f}_t; \boldsymbol{q}_t\}_{t\in\mathbb{T}}$ satisfies GARP.
[10] Varian (1983) suggested a different test-procedure based on condition *(b)*. See Hjertstrand (2009) for a discussion.
[11] Swofford and Whitney's (1994) procedure allows for incomplete adjustment of the expenditure and may therefore account for various forms of habit persistence such as adjustment costs and the formation of expectations. Fleissig and Whitney's (2008) and Elger and Jones' (2008) procedures are designed to handle measurement errors in the data and are implementations to test whether the 'true' data, i.e. without errors, satisfies weak separability.
[12] See Schrijver (1998) for a detailed discussion on linear and integer programming.
[13] See Theorem 4 in Cherchye et al. (2015) for a formal proof.
[14] We use the term "corrected Afriat index" because the revealed preference conditions for weak separability, from where the index is calculated, are essentially an extension of Afriat's Theorem (Afriat, 1967). The term "corrected" refers to that the index is perturbed or corrected to satisfy weak separability.
[15] We set $\psi = 10^{-6}, \eta_t = 1000$ and $A_t = \boldsymbol{z}_t\boldsymbol{w}_t + \eta_t + 1$ in our application.
[16] These numbers constitutes optimal solutions in the quadratic (L2) norm. Specifying the objective function in another norm may yield different solution values. As a robustness check, we also calculated the optimal perturbations in the L1-norm. This did not alter our conclusions.
[17] Notice that the objective function in the MIQP problem (18)-(26) minimizes the percentage root mean squared error.
[18] Jones, Dutkowsky and Elger (2005) found sweep adjustment to be important in identifying appropriate monetary aggregates. At one time the data could be adjusted for both retail and commercial sweeps. The FED, unfortunately in our view, has discontinued their data on sweeps. However, the CFS sweep adjusts on its own, using an econometric model.
[19] The only exception is the Sato-Vartia, the Walsh and the Paasche indexes which are not perturbed in one single time period in Structure 5.
[20] Friedman and Schwartz (1970) did point out the arbitrariness of simple summation and argued in favor of weight sum monetary aggregates.

REFERENCES

Afriat, Sydney N. (1967) The construction of utility functions from expenditure data. *International Economic Review* 8, 66-77.

Balk Bert M. (1995) Axiomatic price theory: a survey. *International Statistical Review* 63, 63-93.

Barnett, William A. (1980) Economic monetary aggregate: an application of index number and aggregation Theory. *Journal of Econometrics* 14, 11-48.

Barnett, William A. and Ki-Hong Choi (2008) Operational identification of the complete class of superlative index numbers: an application of Galois theory. *Journal of Mathematical Economics* 44, 603-612.

Caves, Douglas W., Laurits R. Christensen and W. Erwin Diewert (1982) The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50, 1393-1414

Cherchye, Laurens, Thomas Demuynck, Bram De Rock and Per Hjertstrand (2015) Revealed preference tests for weak separability: an integer programming approach. *Journal of Econometrics* 186, 129–141.

Chrystal, K. Alec and Ronald MacDonald (1994) Empirical evidence on the recent behavior and usefulness of simple-sum and weighted measures of the money stock. *Federal Reserve Bank of St. Louis Review* 76, 73-109.

Diewert, W. Erwin (1976) Exact and superlative index numbers. *Journal of Econometrics* 4, 115-145.

Diewert, W. Erwin (1978) Superlative index numbers and consistency in aggregation. *Econometrica*, 46, 883-900.

Diewert, W. Erwin (1993) Overview of volume 1. In W.E. Diewert and A.O. Nakamura (Eds.), *Essays in Index Number Theory*, pp. 1-31. Elsevier Science Publishers B.V.

Diewert, W. Erwin and Cedric Parkan (1985) Tests for the consistency of consumer data. *Journal of Econometrics* 30, 127–147.

Elger, Thomas and Barry E. Jones (2008) Can rejections of weak separability be attributed to random measurement error in the data? *Economic Letters* 99, 44-47.

Fleissig, Adrian R. and Gerald A. Whitney (2003) A new PC-based test for Varian's weak separability conditions. *Journal of Business and Economic Statistics* 21, 133-144.

Fleissig, Adrian R. and Gerald A. Whitney (2008) A nonparametric test of weak separability and consumer preferences. *Journal of Econometrics* 147, 275–281.

Fisher, Irving (1922) *The Making of Index Numbers*. Boston, MA: Houghton-Mifflin.

Friedman, Milton and Anna J. Schwartz (1963) *A Monetary History of the United States, 1867-1960*. Study by the National Bureau of Economic Research. Princeton: Princeton University Press.

Friedman, Milton and Anna J. Schwartz (1970). *Monetary Statistics of the United States: Estimates, Sources, Methods*. Study by the National Bureau of Economic Research. New York: Columbia University Press.

Hicks, John R. (1946) *Value and Capital*, 2nd edition. Oxford: Claredon Press.

Hjertstrand, Per (2009) A Monte Carlo study of the necessary and sufficient conditions for weak separability. *Advances in Econometrics* 24, 151-182.

Hjertstrand, Per, James L. Swofford and Gerald Whitney (2016) Mixed integer programming revealed preference tests of utility maximization and weak separability of consumption, leisure and money. *Journal of Money Credit and Banking* 58, 1547-1561.

ILO (2004) *Consumer Price Index Manual: Theory and Practice*. International Labour Organization, Geneva. Downloadable at: www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/presentation/wcms_331153.pdf

IMF (2005) *Producer Price Index Manual: Theory and Practice*. International Monetary Fund, Washington, D.C. Downloadable at: https://www.imf.org/external/pubs/ft/ppi/2010/manual/ppi.pdf

Jones, Barry E., Donald H. Dutkowsky and Thomas Elger (2005) Sweep programs and optimal monetary aggregation. *Journal of Banking & Finance* 29, 483-508.

Leontief, Wassily V. (1936) Composite commodities and the problem of index numbers. *Econometrica* 4, 35-59.

Samuelson, Paul A. (1983) *Foundations of Economic Analysis*, Enlarged Edition. Cambridge, MA: Harvard University Press.

Schrijver, Alexander (1998) *Theory of Linear and Integer Programming*. Chichester, England: John Wiley & Sons.

Swofford, James L. and Gerald A. Whitney (1994) A revealed preference test for weakly separable utility maximization with incomplete adjustment. *Journal of Econometrics* 60, 235-249.

Varian, Hal R. (1983) Nonparametric tests of consumer behavior. *Review of Economic Studies* 50, 99-110.

Vartia, Yrjö (1976) Ideal log-change index numbers. *Scandinavian Journal of Statistics* 3, 121–126.

Table 1: Summary statistics on the size and number of perturbations

Summary statistic	Formula
% mean error:	$100 \times \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t$
% root mean squared error:	$100 \times \sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2}$
% of adjusted values:	$100 \times \frac{1}{T}\sum_{t=1}^{T} I_{\hat{\varepsilon}_t \neq 0}$

Note: $I_{\varepsilon \neq 0}$ denotes the indicator function: $I = 1$ if $\varepsilon \neq 0$, and zero otherwise.

Table 2: Utility structures from Hjertstrand et al. (2016)

1. $u(\mathrm{DUR}, \mathrm{NDUR}, \mathrm{SER}, \mathrm{LEIS}, \mathbf{V}(\mathbf{CUR} + \mathbf{DD}, \mathbf{TC}, \mathbf{OCD} - \mathbf{CB}, \mathbf{OCD} - \mathbf{TH}), \mathrm{SD} - \mathrm{CB}, \mathrm{SD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{R}, \mathrm{STD} - \mathrm{CB}, \mathrm{STD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{I}, \mathrm{T} - \mathrm{BILLS}, \mathrm{CP}, \mathrm{LTD})$

2. $u(\mathrm{DUR}, \mathrm{NDUR}, \mathrm{SER}, \mathrm{LEIS}, \mathbf{V}(\mathbf{CUR} + \mathbf{DD}, \mathbf{TC}, \mathbf{OCD} - \mathbf{CB}, \mathbf{OCD} - \mathbf{TH}, \mathbf{SD} - \mathbf{CB}, \mathbf{SD} - \mathbf{TH}), \mathrm{MMMF} - \mathrm{R}, \mathrm{STD} - \mathrm{CB}, \mathrm{STD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{I}, \mathrm{T} - \mathrm{BILLS}, \mathrm{CP}, \mathrm{LTD})$

3. $u\big(\mathrm{DUR}, \mathrm{NDUR}, \mathrm{SER}, \mathrm{LEIS}, \mathbf{V}(\mathbf{CUR} + \mathbf{DD}, \mathbf{TC}, \mathbf{OCD} - \mathbf{CB}, \mathbf{OCD} - \mathbf{TH}, \mathbf{SD} - \mathbf{CB}, \mathbf{SD} - \mathbf{TH}, \mathbf{MMMF} - \mathbf{R}, \mathbf{STD} - \mathbf{CB}, \mathbf{STD} - \mathbf{TH}, \mathbf{MMMF} - \mathbf{I}, \mathbf{T} - \mathbf{BILLS}, \mathbf{CP}, \mathbf{LTD})\big)$

4. $u(\mathbf{V}(\mathbf{NDUR}, \mathbf{SER}), \mathrm{DUR}, \mathrm{LEIS}, \mathrm{CUR} + \mathrm{DD}, \mathrm{TC}, \mathrm{OCD} - \mathrm{CB}, \mathrm{OCD} - \mathrm{TH}, \mathrm{SD} - \mathrm{CB}, \mathrm{SD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{R}, \mathrm{STD} - \mathrm{CB}, \mathrm{STD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{I}, \mathrm{T} - \mathrm{BILLS}, \mathrm{CP}, \mathrm{LTD})$

5. $u(\mathbf{V}(\mathbf{DUR}, \mathbf{NDUR}, \mathbf{SER}, \mathbf{LEIS}), \mathrm{CUR} + \mathrm{DD}, \mathrm{TC}, \mathrm{OCD} - \mathrm{CB}, \mathrm{OCD} - \mathrm{TH}, \mathrm{SD} - \mathrm{CB}, \mathrm{SD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{R}, \mathrm{STD} - \mathrm{CB}, \mathrm{STD} - \mathrm{TH}, \mathrm{MMMF} - \mathrm{I}, \mathrm{T} - \mathrm{BILLS}, \mathrm{CP}, \mathrm{LTD})$

Note: The sample period is 2000.1-2011.3. Source: Table 1 in Hjertstrand et al. (2016).

Table 3: Calculated summary statistics of fit

	Fisher	Sato-Vartia	Törnqvist	Walsh	Paasche	Laspeyres	Simple-Sum
Structure 1: M1							
% mean error	-0.0697	-0.0697	-0.0697	-0.0697	-0.0697	-0.0697	-0.0650
% root mean squared error	2.6400	2.6398	2.6396	2.6303	2.6401	2.6402	2.5489
% of adjusted values	100	100	100	100	100	100	100
Structure 2: Friedman and Schwartz Money							
% mean error	-0.0255	-0.0255	-0.0255	-0.0255	-0.0261	-0.0250	-0.0175
% root mean squared error	1.5977	1.5971	1.5959	1.5977	1.6158	1.5798	1.3214
% of adjusted values	100	100	100	100	100	100	100
Structure 3: Broad Money							
% mean error	-0.2885	-0.2872	-0.2929	-0.2885	-0.3108	-0.2672	-0.4833
% root mean squared error	5.3713	5.3591	5.4122	5.3712	5.5752	5.1690	6.9518
% of adjusted values	100	100	100	100	100	100	100
Structure 4: Narrow Consumption Aggregate							
% mean error	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003
% root mean squared error	0.1732	0.1732	0.1732	0.1732	0.1733	0.1731	0.1739
% of adjusted values	100	100	100	100	100	100	97.8723
Structure 5: Broader Real Sector Aggregate							
% mean error	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002	-0.0027
% root mean squared error	0.1501	0.1500	0.1500	0.1501	0.1501	0.1500	0.5190
% of adjusted values	100	97.8723	100	97.8723	97.8723	100	100

Note: Structures 1-5 contain 3, 6, 13, 2 and 4 components, respectively.

Table 4: Percentage of sign inconsistencies between corrected Afriat index growth rates and the growth rates of the various aggregates

	Fisher	Sato-Vartia	Törnqvist	Walsh	Paasche	Laspeyres	Simple sum
Structure 1	0	0	0	0	0	0	2.1739
Structure 2	0	0	0	0	0	0	0
Structure 3	0	0	0	0	2.1739	0	6.5217
Structure 4	0	0	0	0	0	0	0
Structure 5	0	0	0	0	0	0	0

Note: There are in total 46 comparisons in the sample period 2000.1-2011.3. Structures 1-3 are monetary aggregates, M1 with 3 components, Friedman and Schwartz money with 6 components, and broad money with 13 components, respectively. Structures 4-5 are real sector aggregates with 2 and 4 components, respectively.

Table 5: Percentage of inconsistencies between the rankings from the aggregates and the rankings from the corrected Afriat index

	Fisher	Sato-Vartia	Törnqvist	Walsh	Paasche	Laspeyres	Simple sum
Structure 1	0	0	0	0	0	0	0.1850
Structure 2	0.1850	0.1850	0.1850	0.1850	0.1850	0.2775	0.3700
Structure 3	0.0925	0.0925	0.0925	0	0.2775	0.0925	2.5902
Structure 4	0	0	0	0	0	0.0925	0.0925
Structure 5	0.3700	0.3700	0.3700	0.3700	0.3700	0.3700	0.8326

Note: There are in total 1081 unique comparisons in the sample period 2000.1-2011.3 Structures 1-3 are monetary aggregates, M1 with 3 components, Friedman and Schwartz money with 6 components, and broad money with 13 components, respectively. Structures 4-5 are real sector aggregates with 2 and 4 components, respectively.

Figure 1: Structure 1 - M1

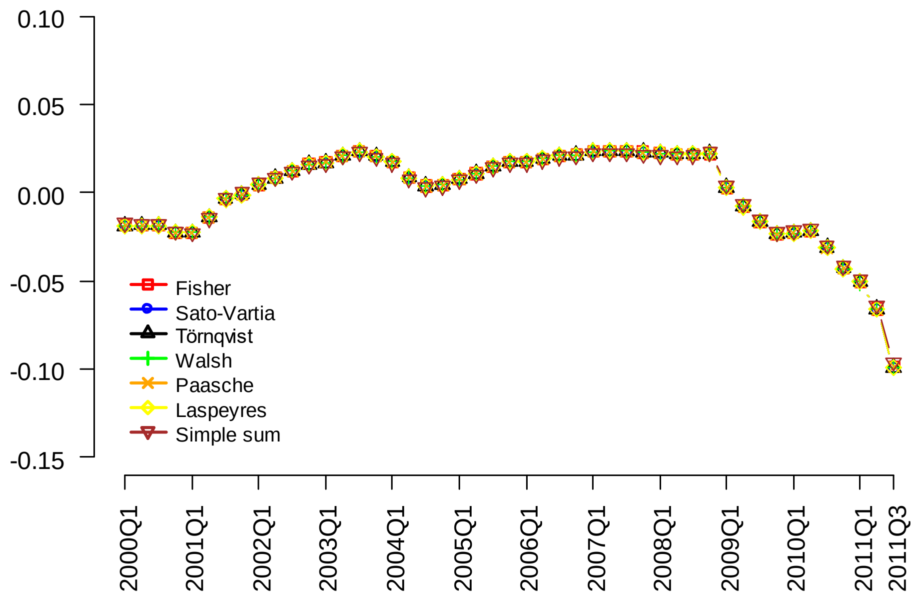

Figure 2: Structure 2 - Friedman and Schwartz Money

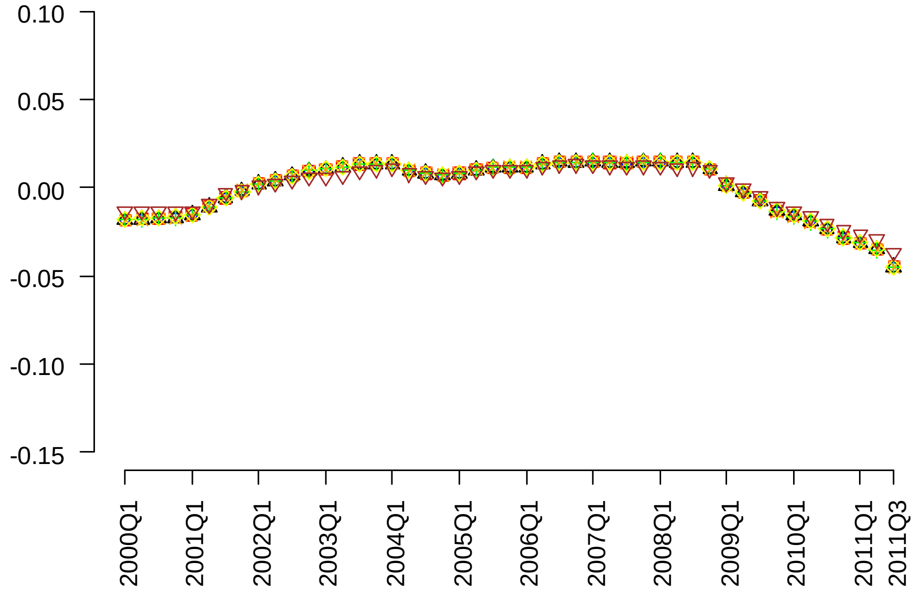

Figure 3: Structure 3 - Broad Money

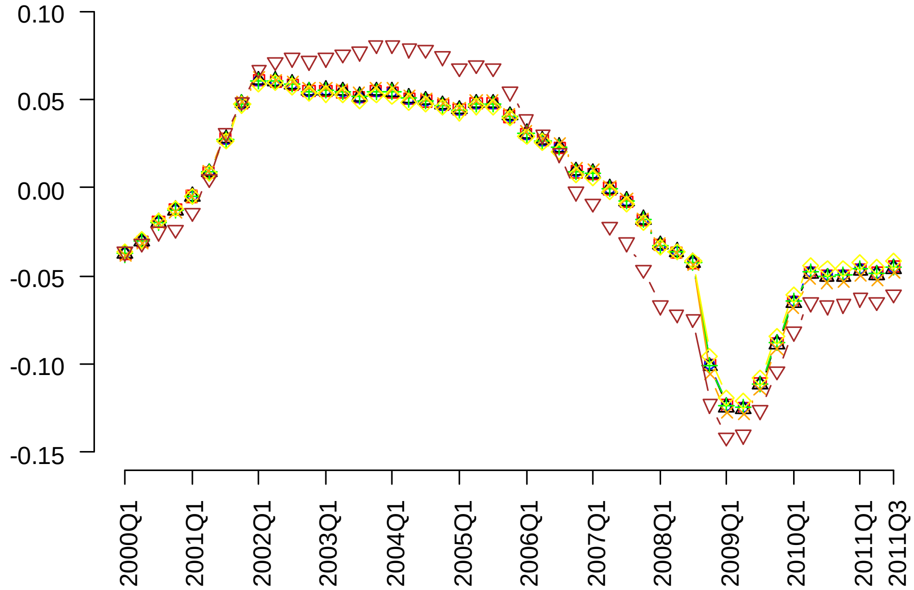

Figure 4: Structure 4 - Narrow Consumption Aggregate

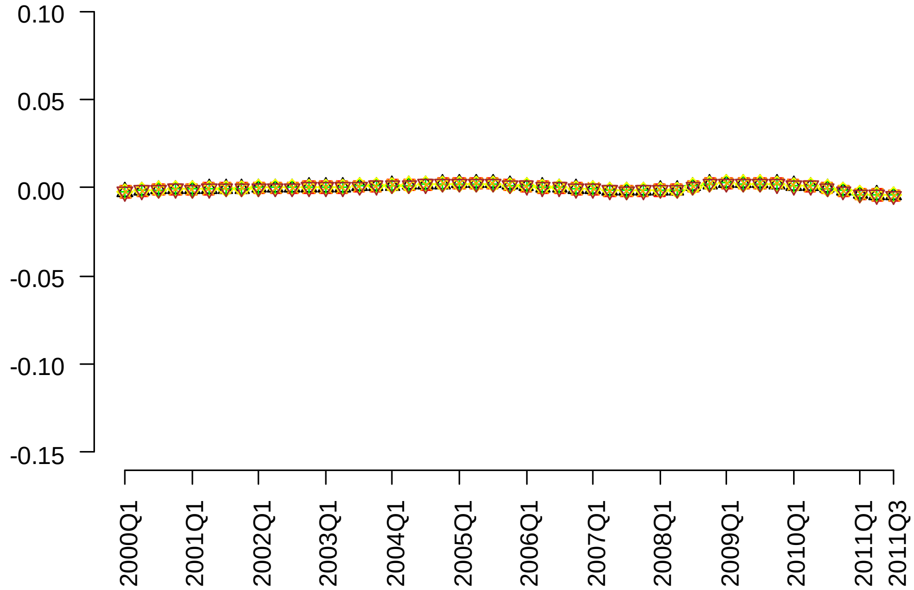

Figure 5: Structure 5 - Broader Real Sector Aggregate