Divisia Monetary Services Indexes: Theory and Practice

Barry E. Jones[1,2]

"No one speaks of 'temperature' index…it is enough to say 'temperature'. Whenever it is known what is being referred to the reference is simply to that thing…Similarly, with the price of a commodity we do not have an 'index' of the price but just the price…The 'index' language comes in economics from complex compulsions to give numbers even when the meaning is not and perhaps cannot be known."

S. N. Afriat, The Price Index,
Cambridge University Press,
Cambridge, 1977, p. 17.

I. Introduction

The financial system of an advanced economy provides an array of monetary assets, which vary considerably in their ability to facilitate transactions, term to maturity, and rates of return. In other words, monetary assets differ in terms of the kinds of "monetary services" they provide and in terms of their functions as "stores of value". These differences lead, in turn, to two natural questions. How do we decide which assets to treat as money? And, how do we construct an index that measures the flow of monetary services provided by the assets that we decide to treat as money?

Friedman and Schwartz (1970, pp. 151-152) constructed a set of monetary aggregates by simple summation of the quantities of four different combinations of monetary assets. They offered the following criticism of this commonly used aggregation procedure:

> The restriction of our attention to these four combinations seems a less serious limitation to us than our acceptance of the common procedure of taking the quantity of money as equal to the aggregate value of the assets it is decided to treat as money. This procedure is a very special case of the more general approach [which] consists of regarding each asset as a joint product having different degrees of "moneyness," and defining the quantity of money as the weighted sum of the aggregate value of all assets, the weights for individual assets varying from zero to unity with a weight of unity assigned to that asset or assets regarded as having the largest quantity of "moneyness" per dollar of aggregate value. The procedure we have followed implies that all weights are either zero or unity.

Simple sum monetary aggregates are now regularly published by many central banks. An alternative aggregation-theoretic approach to constructing monetary aggregates was initially developed in a series of influential papers: Barnett (1978, 1980, 1982), Donovan (1978), and Barnett and Spindt (1982); See Barnett (1987), Barnett, Fisher, and Serletis (1992), Belongia (1995), and Anderson, Jones, and Nesmith (1997a) for surveys. Today, the Bank of England and the Federal Reserve Bank of St. Louis both publish Divisia monetary aggregates based on the aggregation-theoretic approach; See Fisher, Hudson, and Pradhan (1993) and Hancock (2005) for the UK and Anderson, Jones, and Nesmith (1997b) and Anderson and Buol (2005) for the US.[3]

The aggregation-theoretic approach treats monetary assets as durable goods that provide a flow of services to a representative agent over time. Barnett (1978) and Donovan (1978) showed that the user cost price of a monetary asset is the interest foregone as a result of holding the monetary asset rather than an alternative (benchmark) asset that earns a higher interest rate, but which does not provide any monetary services. Thus, a monetary asset that has a high return relative to the benchmark rate will have a low user cost. Building upon Diewert (1976), Barnett (1980) constructed superlative indexes for a set of monetary assets based upon their quantities and user costs. The most commonly used superlative index number formulae are Fisher's ideal index and the Törnqvist index. In the literature, superlative monetary indexes are usually referred to as either "Divisia indexes" or as "monetary services indexes". Barnett and Spindt (1982) and Farr and Johnson (1985) constructed Divisia indexes for the Federal Reserve Board at the same levels of aggregation as the conventional simple sum monetary aggregates. Thornton and Yue (1992), Anderson, Jones, and Nesmith (1997b), and Anderson and Buol (2005) constructed Divisia indexes for the Federal Reserve Bank of St. Louis.

There are two steps in constructing a monetary aggregate. The first step is to choose the components of the aggregate and the second step is to choose a method for constructing the aggregate. Barnett (1982) suggested a procedure that provides a method for determining the components of the monetary aggregate, which is consistent with aggregation theory. The procedure is based upon testing groups of monetary assets for weak separability. Weak separability implies that the marginal rates of substitution between pairs of assets included in the aggregate are not affected by the quantities of variables that are not included in the aggregate.

Belongia (1996), Anderson, Jones, and Nesmith (1997a), Lucas (2000), Schunk (2001), Stracca (2001, 2004), Duca and VanHoose (2004), and Drake and Mills (2005) provide recent discussions on the merits of the Divisia index relative to the conventional simple sum index. The simple sum index implicitly assumes that all monetary assets are regarded as being perfect substitutes by the owners of those assets. Belongia (1996, p. 1067) states "[a]s this condition is strongly rejected by empirical evidence, the simple sum aggregates cannot internalize pure substitution effects and, as such, they are prone to spurious shifts that would suggest a change in the utility derived from money holdings when, in fact, no such change has occurred." In contrast, the Divisia index approach, developed by Barnett (1978, 1980, 1982), builds upon an explicit microeconomic model of a utility maximizing representative consumer. In the model, the representative consumer chooses quantities of consumer goods and services as well as monetary assets so as to maximize a lifetime utility function subject to a sequence of inter-temporal budget constraints. The assumption that a set of current period monetary assets are weakly separable from all other decision variables in the utility function implies the existence of a monetary aggregate, which can be tracked up to a second-order approximation error using a superlative index.

The Federal Reserve Bank of St. Louis bases its Divisia monetary services indexes on the superlative Törnqvist index. The level of Törnqvist index does not have a simple relationship to the levels of the monetary asset quantities. Instead, the growth rate of the Törnqvist index is a weighted average of the growth rates of the component quantities. The weight on the growth rate of each component is its average expenditure share over the two periods, which is calculated from the user costs of the various monetary assets. Consequently, if a particular monetary asset has a relatively high average expenditure share in a particular period, then its growth rate will be relatively highly weighted in the growth rate of the index. The idea behind this is that the marginal rates of substitution between the monetary assets will be equal to the corresponding ratios of user costs, assuming utility maximization. Over time, as the expenditure shares change, the weights on the growth rates of the components will change as well. Lucas

(2000, pp. 270-1) emphasized that this "…avoids the awkward necessity of classifying financial assets as either entirely money or not at all, and lets the data do most of the work in deciding how monetary aggregates should be revised over time as interest rates change and new instruments are introduced."

In this paper, I use a state of the art non-parametric revealed preference test to identify weakly separable groups of monetary assets for the United States over the period 1993-2003. The test is based upon computing indexes that satisfy the Afriat inequalities using an algorithm recently proposed by Fleissig and Whitney (2003). The tests are applied to the M1, M2, and M3 collections of assets as well as to two zero-maturity measures (M2M and MZM). I also test the components of M1 plus total savings deposits, which I refer to as M1+. I find that M2M, MZM, and M3 are all admissible, but M1, M1+, and M2 violate the necessary conditions for weak separability. The test results also indicate that the superlative Törnqvist index needs only small adjustments to serve as a measure of the utility derived from the weakly separable monetary assets within the Afriat inequalities.

I also compare and contrast the Divisia indexes with the conventional simple sum monetary aggregates over the period 1970-2004. I show that the most economically significant differences between the Divisia and simple sum indexes occurred during the critical period between 1978 and 1983. Finally, I investigate the effects of alternative empirical proxies for the benchmark rate, which is used to construct user costs for the monetary assets. I compare published Divisia indexes that are based on an envelope of monetary asset returns and a long-term bond rate with alternative Divisia indexes that are based on an envelope that includes only monetary asset returns. I find that the benchmark rate does seem to have some important effects on the index. In particular, the alternative benchmark rate lead to reduced correlation between the growth rates of the Divisia indexes and the growth rates of the simple sum indexes.

The remainder of the paper is organized as follows: Sections II and III survey the literatures on monetary aggregation theory and weak separability testing. Section IV presents empirical results using the weak separability tests and Section V presents empirical comparisons between the Divisia and simple sum indexes. Section VI concludes.

II. Monetary Aggregation Theory and Superlative Index Numbers

In this section, I explain the main ideas from monetary aggregation theory within an inter-temporal decision model for a representative consumer. In particular, I derive the user costs of the monetary assets within the model and I develop the importance of weak separability. Finally, I build upon these results to motivate the use superlative index numbers such as the Törnqvist index.

A. An Inter-temporal Decision Model

The basic theoretical framework is a perfect-certainty inter-temporal decision model. In the model, an infinitely-lived agent maximizes lifetime utility, $\mathcal{U}(c_t, \mathbf{m}_t, c_{t+1}, \mathbf{m}_{t+1}, ...)$, subject to a sequence of inter-temporal budget constraints. The consumer's utility function takes the additively (strong) time separable form:

$$\mathcal{U} = \sum_{s=t}^{\infty} \beta^{s-t} u(c_s, \mathbf{m}_s), \tag{1}$$

where c_s denotes real consumption and $\mathbf{m}_s = (m_{1s}, m_{2s}, ..., m_{Ns})$ denotes the real stocks of N monetary assets in period s. The real monetary asset stocks are the nominal stocks divided by p_s, which is the price of consumption. The monetary asset stocks provide a flow of monetary services to the agent, which is modeled by including the monetary assets in the agent's utility function. The instantaneous utility function, u, is assumed to be monotonically increasing and strictly concave.

In real terms, the inter-temporal budget constraint for period s is as follows:

$$x_s + \sum_i \frac{(1+R_{i,s-1}) m_{i,s-1} p_{s-1}}{p_s} + \frac{(1+R_{s-1}) a_{s-1} p_{s-1}}{p_s} = \sum_i m_{i,s} + a_s + c_s . \tag{2}$$

In the constraints, a_s denotes the real stock of a risk-free benchmark asset that does not provide any monetary services, and x_s denotes exogenous income in real terms. R_{is} is the nominal interest rate on monetary asset i and R_s is the nominal interest rate on the benchmark asset. Interest is assumed to be paid at the end of each period. In some expressions, it will convenient to adopt the following notation: $1+r_s \equiv (1+R_s)p_s / p_{s+1}$.

I assume, for simplicity, that all asset stocks are constrained to be non-negative: *i.e.* $m_{is} \geq 0$ and $a_s \geq 0$ for all $i = 1,...,n$ and for all *s*. The inter-temporal decision problem is to choose a sequence of consumption and asset holdings $c_s, \mathbf{m}_s, a_s, s = t, t+1,...$ so as to maximize lifetime utility subject to these inter-temporal constraints and given initial real asset holdings: $\mathbf{m}_{t-1}, a_{t-1}$. The agent is assumed to solve the problem in each period.

B. User Costs Under Perfect Certainty

Let $c_s^*, \mathbf{m}_s^*, a_s^*$ denote the optimal values for consumption, monetary assets, and the benchmark asset in period *s*. Let $\pi_{it}^* \equiv (R_t - R_{it})/(1+R_t)$ denote the *real user cost* of the i^{th} monetary asset and $\pi_{it} \equiv p_t \pi_{it}^*$ denote the *nominal user cost* of the i^{th} asset. Using this notation, the Euler equations that characterize the optimal solution are as follows:

$$\frac{\frac{\partial u}{\partial c}(c_t^*, \mathbf{m}_t^*)}{\frac{\partial u}{\partial c}(c_{t+1}^*, \mathbf{m}_{t+1}^*)} = \beta \frac{p_t (1+R_t)}{p_{t+1}} = \beta(1+r_t) \tag{3}$$

$$\frac{\frac{\partial u}{\partial m_i}(c_t^*, \mathbf{m}_t^*)}{\frac{\partial u}{\partial c}(c_{t+1}^*, \mathbf{m}_{t+1}^*)} = \beta \frac{p_t (R_t - R_{it})}{p_{t+1}} = \beta(1+r_t)\pi_{it}^* \text{ for } i = 1,...,N, \tag{4}$$

(3) is the standard Euler equation for consumption. It means that the marginal rate of substitution between current and future consumption is equated to the gross real interest rate. (3) and (4) imply that the marginal rate of substitution between the i^{th} monetary asset and consumption in period *t* is equated to the i^{th} asset's real user cost in period *t*:

$$\frac{\frac{\partial u}{\partial m_i}(c_t^*, \mathbf{m}_t^*)}{\frac{\partial u}{\partial c}(c_t^*, \mathbf{m}_t^*)} = \pi_{it}^* \text{ for } i = 1,...,N. \tag{5}$$

Similarly, the marginal rate of substitution between the i^{th} and j^{th} monetary assets in period t is equated to the ratio of the asset's user costs in period t (in either nominal or real terms). The user costs are, therefore, acting like the prices (or opportunity costs) of the monetary assets in the optimality conditions, which is consistent with how they are defined. The user costs represent the difference between the rate of return being earned on the monetary asset and the rate of return that could have been earned on the benchmark asset (the interest rate differential is discounted to present value, because interest is assumed to be paid at the end of the period in the model); See Donovan (1978) and Barnett (1978).[4]

C. User Costs Under Risk

The standard perfect-certainty user costs are not appropriate if the asset returns are stochastic, which would be the case if stock and bond mutual funds were treated as monetary assets. To account for such assets, I next consider an extension of the model, in which the agent maximizes expected lifetime utility and the nominal returns on at least some monetary assets are stochastic. In a stochastic extension of the model, Barnett (1995) shows that the Euler equations are as follows:

$$\frac{\partial u}{\partial c}(c_t^*, \mathbf{m}_t^*) = E_t\left[\beta \frac{p_t(1+R_t)}{p_{t+1}} \frac{\partial u}{\partial c}(c_{t+1}^*, \mathbf{m}_{t+1}^*)\right] \tag{6}$$

$$\frac{\partial u}{\partial m_i}(c_t^*, \mathbf{m}_t^*) = E_t\left[\beta \frac{p_t(R_t - R_{it})}{p_{t+1}} \frac{\partial u}{\partial c}(c_{t+1}^*, \mathbf{m}_{t+1}^*)\right], \tag{7}$$

where E_t denotes the conditional expectations operator.

We can obtain some intuition by assuming that the current period prices and asset returns are all known with perfect certainty, even though future prices and asset returns are stochastic. Under this assumption, p_t , R_t , and R_{it} can be pulled out of the conditional expectations in (6) and (7). This, in turn, implies that (5) is correct even under risk. Thus, although future prices and returns are stochastic, the perfect certainty user costs are valid as long as the current period prices and returns are known with perfect certainty; See Barnett (1995). This is a highly unrealistic assumption, however, since interest is assumed to be paid at the end of each period.

Barnett, Jensen, and Liu (1997) and Barnett and Wu (2005) provide more general results under risk neutrality and risk aversion. Under risk neutrality, the real user costs are defined as follows: $\frac{E_t[R_t]-E_t[R_{it}]}{1+E_t[R_t]}$, *i.e.* where the asset returns are replaced by their conditional expectations. In the more general risk aversion case, the real user cost of a monetary asset is equal to the risk-neutral user cost plus a risk-adjustment term that depends on the covariance between the asset's rate of return and the marginal utility of consumption. Barnett, Jensen, and Liu (1997) proposed a specific risk-adjustment derived within a CCAPM model. In that model, the risk adjustments are based upon the covariance between the asset return and the growth rate of consumption. They found, however, that the risk adjustments are small for reasonable values of the risk-aversion coefficient due to the low contemporaneous covariance between the asset returns and the growth rate of consumption. Barnett and Wu (2005) extend the results in Barnett, Jensen, and Liu (1997) to allow for inter-temporal non-separability, which can generate far more substantial risk adjustments, even with a reasonable estimate of the risk-aversion coefficient.

D. Weak Separability

The optimal solutions for current-period consumption, c_t^*, and real monetary assets, $\mathbf{m}_t^*$, from the inter-temporal decision problem also solve a conditional current-period decision problem. The current-period decision is to maximize the utility function, $u(c,\mathbf{m})$, subject to the budget constraint: $p_t c + \boldsymbol{\pi}_t \cdot \mathbf{m} = Y_t$, where $Y_t = p_t c_t^* + \boldsymbol{\pi}_t \cdot \mathbf{m}_t^*$ is the optimal expenditure on consumption and monetary assets from the inter-temporal decision problem and $\boldsymbol{\pi}_t$ is the vector of user costs; See Barnett (1987) for further discussion. This result is an implication of additive (strong) time separability of the lifetime utility function, $\mathscr{U}$, but it can also be obtained under somewhat weaker assumptions. The conditional current-period decision can be used to develop the importance of weak separability.[5]

Weak separability means that there exists a well-defined utility function for the monetary assets (or a subset of them) that separates them from all other decision variables in the instantaneous utility function. Formally, the instantaneous utility function, u, is said to be weakly separable in $\mathbf{m}$ if there exists a macro function, U, and a monotonic and strictly concave category sub-utility function, V, such that

$$u(c,\mathbf{m}) = U(c, V(\mathbf{m})) \,. \tag{8}$$

Weak separability implies that the marginal rates of substitution between pairs of current-period monetary assets are independent of current-period consumption:

$$\frac{\frac{\partial U}{\partial m_i}(c_t^*, \mathbf{m}_t^*)}{\frac{\partial U}{\partial m_j}(c_t^*, \mathbf{m}_t^*)} = \frac{\frac{\partial U}{\partial V}\left(c_t^*, V(\mathbf{m}_t^*)\right)\frac{\partial V}{\partial m_i}(\mathbf{m}_t^*)}{\frac{\partial U}{\partial V}\left(c_t^*, V(\mathbf{m}_t^*)\right)\frac{\partial V}{\partial m_j}(\mathbf{m}_t^*)} = \frac{\frac{\partial V}{\partial m_i}(\mathbf{m}_t^*)}{\frac{\partial V}{\partial m_j}(\mathbf{m}_t^*)} = \frac{\pi_{it}^*}{\pi_{jt}^*} = \frac{\pi_{it}}{\pi_{jt}} \text{ for } i = 1, \ldots, N \tag{9}$$

Thus, under weak separability, the optimal solution for current-period monetary assets, $\mathbf{m}_t^*$, also solves the current period decision problem of maximizing the sub-utility function, $V(\mathbf{m})$, subject to the budget constraint $\boldsymbol{\pi}_t \cdot \mathbf{m} = \boldsymbol{\pi}_t \cdot \mathbf{m}_t^*$. Therefore, the system of demand equations for the set of weakly separable monetary assets are functions of only the nominal user costs of those assets and total expenditure on the set.[6]

E. Aggregation

In this paper, I will focus on aggregation-theoretic results under homogeneous weak separability, although weaker theoretical results are available in the non-homothetic case; See Barnett (1987) for detailed discussion. If the sub-utility function, V, is linearly homogeneous, then u is said to be homogeneously weakly separable in $\mathbf{m}$. Under this assumption, $Q_t = V(\mathbf{m}_t^*)$ is a monetary quantity aggregate. The unit expenditure function, $\Pi_t = e(\boldsymbol{\pi}_t, 1)$, is the corresponding dual price aggregate, where $e(\boldsymbol{\pi}, \tilde{v}) = \min_{\mathbf{h}} \{\boldsymbol{\pi} \cdot \mathbf{h} : V(\mathbf{h}) = \tilde{v}\}$. The unit expenditure function is linearly homogeneous, which implies that if all user costs increase by the same percentage then the price aggregate does as well. Linear homogeneity of the sub-utility function implies an analogous property for the quantity aggregate. Linear homogeneity also implies that the price and quantity aggregates satisfy the strong factor reversal test:

$$\Pi_t Q_t = e(\boldsymbol{\pi}_t, 1) V(\mathbf{m}_t^*) = e\left(\boldsymbol{\pi}_t, V(\mathbf{m}_t^*)\right) = \boldsymbol{\pi}_t \cdot \mathbf{m}_t^*, \tag{10}$$

which means that the product of the quantity and price aggregates equals total expenditure on monetary services.

Finally, homogeneous weak separability implies that the inter-temporal decision problem can be solved in two stages: In the first stage, the agent determines the quantities of current and future consumption, future quantities of the monetary assets, and total expenditure on monetary assets in the current period. In the second stage, the consumer determines the quantities of the monetary assets in the current period. The agent behaves, therefore, as if Q_t and Π_t are the quantity and price of an elementary good.[7]

F. Admissibility

Barnett (1982) proposed an aggregation-theoretic procedure for selecting the components of a monetary aggregate, which builds upon the theoretical results from the previous section; See also Barnett, Fisher, and Serletis (1992). Barnett, Hinich, and Yue (1991) discuss a related selection procedure under risk neutrality. The procedure involves testing the components of the proposed monetary aggregate for three conditions:

C1 (Existence): The components must be weakly separable;
C2 (Consistency): The components must be homogeneously weakly separable;
C3: The components must include currency and must only include monetary assets.

He referred any group of assets that satisfied C2 and C3 as admissible as a monetary aggregate. He used the term admissible as an indicator for a group that only satisfied C1. C1 and C2 are derived directly from the theoretical results developed in the previous section. C3, however, merits some additional discussion. It says that all admissible monetary aggregates must be "nested about 'hard core money'", which must include currency at a minimum; See Barnett (1982, p. 697). Barnett argued that C3 could be tightened or loosened depending on the scope of the research involved. The condition could be modified for the US, for example, to state that all of the assets in M1 should be included in any admissible aggregate.

It should be noted that the admissibility criteria proposed by Barnett (1982) does not necessarily provide a criteria for uniquely deciding what assets should be considered to be money. The reason is that a weakly separable group of monetary assets could be a subset of a broader group that is also weakly separable, thereby forming nested admissible groups. There is no aggregation-theoretic rationale to consider the narrower grouping to be money and not the broader one or *vice versa*; See Fisher, Hudson, and Pradhan (1993, p. 13) for further discussion. If multiple admissible groupings exist, then some alternative criteria must be used to determine if a particular grouping should be preferred over the others. For example, Barnett (1982) argued that the broadest admissible grouping might be preferable for certain purposes. If, however, a particular asset grouping is not weakly separable, then there does not exist a monetary aggregate for those assets and an index should not be constructed aver them. Moreover, there is no reason to suspect *a priori* that weak separability is more plausible for broader asset groupings than for narrower asset groupings, and simply constructing an index over the broadest set of assets that could possibly be considered money is suspect.[8]

Once one or more admissible groupings of assets have been identified, a method needs to be chosen to construct a quantity index over the components of a particular admissible group. The index needs to be able to track the true quantity aggregate, $Q_t = V(\mathbf{m}_t^*)$, in discrete time using the available data on the quantities and user costs. The conventional simple sum index just adds up the component quantities for the included assets, which will only be correct if those assets are perfectly substitutable in identical ratios (*i.e.* if the sub-utility function is linear with equal coefficients on each asset). In other words, all monetary assets must have the same "degree of moneyness" (Friedman and Schwartz, 1970, p. 151-2), an assumption that is not justifiable on either conceptual or empirical grounds.

Swofford and Whitney (1991, p. 760) argued that by choosing the components of a monetary aggregate through testing for weak separability and constructing the monetary aggregate using a superlative index, consistency with economic theory is maintained in both steps. In the next section, I describe several superlative index numbers that are able to track the unknown quantity aggregate without making any assumptions about the degree of substitutability between the included assets.

G. Superlative Index Numbers and the Simple Sum Index

As above, I will focus on results under the assumption that V is linearly homogeneous, although additional results are available in the literature for the non-homothetic case.

In the context of monetary aggregation, a quantity index, $f_Q(\mathbf{m}_t, \mathbf{m}_s, \boldsymbol{\pi}_t, \boldsymbol{\pi}_s)$, is a function of the quantities and user costs in two periods that approximates Q_t / Q_s. A quantity index is *superlative* if (assuming utility maximization) it is exactly correct for a linearly homogeneous direct utility function that can provide a second-order differential approximation to an arbitrary (C^2) linearly homogeneous utility function; See Diewert (1976a, 1981). We can also define dual price indexes, $f_\Pi(\mathbf{m}_t, \mathbf{m}_s, \boldsymbol{\pi}_t, \boldsymbol{\pi}_s)$, through the weak factor reversal test:

$$f_Q(\mathbf{m}_t, \mathbf{m}_s, \boldsymbol{\pi}_t, \boldsymbol{\pi}_s) f_\Pi(\mathbf{m}_t, \mathbf{m}_s, \boldsymbol{\pi}_t, \boldsymbol{\pi}_s) = \frac{\boldsymbol{\pi}_t \cdot \mathbf{m}_t}{\boldsymbol{\pi}_s \cdot \mathbf{m}_s}. \tag{11}$$

The most commonly used superlative indexes are Fisher's ideal index and the Törnqvist index. Fisher's ideal quantity index is defined as follows:

$$f_Q^F(\mathbf{m}_t, \mathbf{m}_s, \boldsymbol{\pi}_t, \boldsymbol{\pi}_s) = \sqrt{\frac{\boldsymbol{\pi}_t \cdot \mathbf{m}_t}{\boldsymbol{\pi}_t \cdot \mathbf{m}_s} \frac{\boldsymbol{\pi}_s \cdot \mathbf{m}_t}{\boldsymbol{\pi}_s \cdot \mathbf{m}_s}}, \tag{12}$$

which is the geometric mean of the Paasche and Laspeyres quantity indexes. It is exact for the (flexible) homogeneous quadratic direct utility function, $V(\mathbf{m}) = \sqrt{\sum\sum A_{ij} m_i m_j}$.

The Törnqvist quantity index is defined as follows:

$$f_Q^T(\mathbf{m}_t, \mathbf{m}_s, \boldsymbol{\pi}_t, \boldsymbol{\pi}_s) = \prod_i \left(m_{it} / m_{is} \right)^{(w_{it} + w_{is})/2}, \tag{13}$$

where $w_{it} = \pi_{it} m_{it} / \boldsymbol{\pi}_t \cdot \mathbf{m}_t$ is the expenditure share of asset *i* for period *t*. This index is exact for the (flexible) linearly homogeneous Translog direct utility function.[9]

In this article, I construct superlative quantity indexes using the Törnqvist formula. The standard method is to construct chained indexes, *i.e.* where the index is computed using quantity and user costs data from adjacent periods. The notation, Q_t^T , will be used to denote the chained quantity index defined by $Q_t^T / Q_{t-1}^T = f_Q^T(\mathbf{m}_t, \mathbf{m}_{t-1}, \boldsymbol{\pi}_t, \boldsymbol{\pi}_{t-1})$. The log-change of the Törnqvist index is easy to interpret, since

$$\ln(Q_t^T) - \ln(Q_{t-1}^T) = \sum_i \frac{w_{it} + w_{i,t-1}}{2} \left(\ln(m_{it}) - \ln(m_{i,t-1}) \right). \tag{14}$$

Thus, the log-change (growth rate) of the index is a weighted sum of the log-changes (growth rates) of each component, where the weights are the average expenditure shares between the two adjacent periods. The growth rates of components that have relatively large expenditure shares will be weighted relatively heavily in the growth rate of the quantity index and, similarly, the growth rates of components with relatively small expenditures will not receive high weights in the growth rate of the index; See Barnett (1983) for further discussion.

Finally, the simple sum index, *S*, is defined as the sum of the component quantities: $S_t = \sum_i m_{it}$. The growth rate of the simple sum index is given by the following expression:

$$\frac{S_t - S_{t-1}}{S_{t-1}} = \sum_i \frac{m_{i,t-1}}{S_{t-1}} \frac{m_{it} - m_{i,t-1}}{m_{i,t-1}} . \tag{15}$$

Thus, the growth rate of the simple sum index is also a weighted sum of the growth rates of the component quantities. Instead of weighting by the expenditure shares, however, the simple sum index weights the growth rates of each asset by the ratio of the quantity of that asset to the level of the simple sum index.

III. Non-Parametric Weak Separability Tests

In this section, I will discuss non-parametric revealed preference tests, which provide a practical method for testing groups of monetary assets for admissibility. The revealed preference approach is designed to test a dataset consisting of a finite number of observations on the quantities and prices/user costs for a set of goods, services, and assets for consistency with utility maximization and weak separability; See Samuelson (1948), Houthakker (1950), Richter (1966), Afriat (1977), Varian (1982, 1983, 1985), and Swofford and Whitney (1994).

A. Tests of Utility Maximization

Let $\mathbf{x}_i = (x_{1i},..., x_{Ni})$ and $\mathbf{p}_i = (p_{1i},..., p_{Ni})$ denote observed quantity and price data for *N* goods, where $i = 1,..., I$ denotes a specific observation in the dataset.[10] A utility function, *u*, is said to rationalize the dataset if $u(\mathbf{x}_i) \geq u(\mathbf{x})$ for all **x** such that $\mathbf{p}_i \cdot \mathbf{x}_i \geq \mathbf{p}_i \cdot \mathbf{x}$. Varian (1982, 1983) provided necessary and sufficient conditions for the existence of a utility function that rationalizes the dataset. I will begin with a heuristic argument to motivate the main result; See Varian (1982, p. 970) and Swofford and Whitney (1994, p. 237).

Suppose that u is a non-satiated, differentiable, concave, monotonic utility function that rationalizes the dataset. A property of differentiable concave functions is

$$u(\mathbf{x}) \le u(\mathbf{x}_i) + \nabla u(\mathbf{x}_i) \cdot (\mathbf{x} - \mathbf{x}_i) \text{ for all } i, j = 1, ..., I, \tag{16}$$

where ∇u denotes the gradient of the utility function. The first-order necessary conditions for utility maximization are $\nabla u(\mathbf{x}_i) = \lambda_i \mathbf{p}_i$ for all $i = 1, ..., I$, where λ_i is the Lagrange multiplier for the budget constraint. We can substitute these conditions into (17) to obtain the well-known Afriat inequalities:

$$U_j - U_i \le \lambda_i (\mathbf{p}_i \cdot \mathbf{x}_j - \mathbf{p}_i \cdot \mathbf{x}_i) \text{ for all } i, j, \tag{17}$$

where $U_i = u(\mathbf{x}_i)$ and $U_j = u(\mathbf{x}_j)$.

Varian (1982) proved the equivalence of the following four conditions:

i) There exists a non-satiated utility function that rationalizes the dataset;

ii) There exists a non-satiated, continuous, concave, monotonic utility function that rationalizes the dataset;

iii) The existence of strictly positive Afriat indexes, U_i and λ_i, $i = 1, ..., I$ satisfying (17); and

iv) The dataset satisfies the Generalized Axiom of Revealed Preference (GARP).

The utility function $U(\mathbf{x}) = \min_{k=1,...,I} \left\{ U_k + \lambda_k (\mathbf{p}_k \cdot \mathbf{x} - \mathbf{p}_k \cdot \mathbf{x}_k) \right\}$ rationalizes the observed dataset. Thus, U_i can be interpreted as the utility level provided by $\mathbf{x}_i$, since $U(\mathbf{x}_i) = U_i$ for all $i = 1, ..., I$. Similarly, the index, λ_i, measures the marginal utility of expenditure at $\mathbf{x}_i$; See Varian (1983, pp. 100-101).

A dataset can easily be tested for GARP using standard revealed preference relations. If $\mathbf{p}_i \cdot \mathbf{x}_i > \mathbf{p}_i \cdot \mathbf{x}_j$, then $\mathbf{x}_i$ is strictly directly revealed preferred to $\mathbf{x}_j$ ($\mathbf{x}_i P^0 \mathbf{x}_j$). Similarly, if $\mathbf{p}_i \cdot \mathbf{x}_i \ge \mathbf{p}_i \cdot \mathbf{x}_j$, then $\mathbf{x}_i$ is directly revealed preferred to $\mathbf{x}_j$ ($\mathbf{x}_i R^0 \mathbf{x}_j$). $\mathbf{x}_i$ is revealed preferred to $\mathbf{x}_j$ ($\mathbf{x}_t R \mathbf{x}_s$) if there are *K* observations, $\mathbf{x}_{k_1}, \ldots, \mathbf{x}_{k_K}$, such that $\mathbf{x}_i R^0 \mathbf{x}_{k_1} R^0 \mathbf{x}_{k_2} \ldots R^0 \mathbf{x}_{k_K} R^0 \mathbf{x}_j$. A dataset violates GARP, and therefore cannot be rationalized by a utility function, if there are any two observations $\mathbf{x}_i$ and $\mathbf{x}_j$ such that $\mathbf{x}_i R \mathbf{x}_j$ and $\mathbf{x}_j P^0 \mathbf{x}_i$. Intuitively, if revealed preferences are intransitive then the dataset is inconsistent with utility maximization.

In practice, revealed preference tests are often implemented empirically using time series data. In this case, $\mathbf{x}_i$ and $\mathbf{p}_i$ would represent observed quantities and prices for *N* goods, services, and assets in *I* different time periods.[11] The use of time series data in revealed preference tests requires the implicit assumption of strong time separability of preferences; See Swofford and Whitney (1987).

B. Revealed Preferences and Index Numbers

Index numbers and the revealed preference tests are both data driven tools that are used to make comparisons between two bundles using only observed data on quantities and prices/user costs. Among the superlative index numbers, Fisher's ideal index has an apparently unique relationship to the revealed preference axioms; See Diewert (1976a, 1976b, 1981). Suppose that $\mathbf{x}_i P^0 \mathbf{x}_j$, *i.e.* $\mathbf{p}_i \cdot \mathbf{x}_i > \mathbf{p}_i \cdot \mathbf{x}_j$. If $\mathbf{x}_i$ and $\mathbf{x}_j$ do not violate GARP, then it must be the case that $\mathbf{p}_j \cdot \mathbf{x}_i > \mathbf{p}_j \cdot \mathbf{x}_j$, since otherwise $\mathbf{x}_j R^0 \mathbf{x}_i$ causing a GARP violation. Thus, if the two bundles do not violate GARP, then

$$f_Q^F(\mathbf{x}_i, \mathbf{x}_j, \mathbf{p}_i, \mathbf{p}_j) = \sqrt{\frac{\mathbf{p}_i \cdot \mathbf{x}_i}{\mathbf{p}_i \cdot \mathbf{x}_j} \frac{\mathbf{p}_j \cdot \mathbf{x}_i}{\mathbf{p}_j \cdot \mathbf{x}_j}} > 1 .$$

Fisher's ideal index indicates that the quantity aggregate increases if $\mathbf{x}_j$ is replaced by $\mathbf{x}_i$, which is consistent with $\mathbf{x}_i$ being strictly directly revealed preferred to $\mathbf{x}_j$.[12]

This result shows that Fisher's ideal index makes bilateral comparisons that are consistent with the direct revealed preference relations provided that the data satisfy the appropriate axioms of revealed preference theory. This seems to be a unique property amongst the superlative class of index numbers; See Diewert (1981).

C. Necessary and Sufficient Conditions for Weak Separability

The key theorem for non-parametric tests of weak separability was provided by Varian (1983); See also Swofford and Whitney (1994).[13] We partition the dataset into two groups of goods. Let $\mathbf{y}_i = (y_{1i},..., y_{Mi})$ denote the observed quantities for a group of M goods with corresponding prices, $\mathbf{p}_i^y = (p_{1i}^y,..., p_{Mi}^y)$. Let $\mathbf{z}_i = (z_{1i},..., z_{N-M,i})$ denote the observed quantities of all other goods with prices, $\mathbf{p}_i^z = (p_{1i}^z,..., p_{N-M,i}^z)$. Finally, let $\mathbf{x}_i$ and $\mathbf{p}_i$ continue to denote the quantities and prices for the complete set of N goods.

Varian (1983) shows that there exists a concave, monotonic, continuous and weakly separable utility function, $u(\mathbf{x}) = U(\mathbf{z}, V(\mathbf{y}))$, that rationalizes the data if and only if there exist strictly positive indexes, U_i, V_i, λ_i, and μ_i, $i = 1,..., I$ satisfying the following inequalities:

$$V_j - V_i \leq \mu_i (\mathbf{p}_i^y \cdot \mathbf{y}_j - \mathbf{p}_i^y \cdot \mathbf{y}_i) \text{ for all } i, j, \tag{18}$$

$$U_j - U_i \leq \lambda_i \left((\mathbf{p}_i^z \cdot \mathbf{z}_j - \mathbf{p}_i^z \cdot \mathbf{z}_i) + \frac{V_j - V_i}{\mu_i} \right) \text{ for all } i, j \tag{19}$$

These conditions are necessary and sufficient for weak separability.

The first set of conditions (18) implies that the $\mathbf{y}, \mathbf{p}^y$ data satisfy GARP. The sub-utility function, $V(\mathbf{y}) = \min_{k=1,...,I} \left\{ V_k + \mu_k (\mathbf{p}_k^y \cdot \mathbf{y} - \mathbf{p}_k^y \cdot \mathbf{y}_k) \right\}$, rationalizes the $\mathbf{y}, \mathbf{p}^y$ dataset. Therefore, V_i can be interpreted as the sub-utility level provided by $\mathbf{y}_i$. The two sets of conditions together imply that the $\mathbf{x}, \mathbf{p}$ data (containing all N goods) satisfies GARP. Thus, the two necessary conditions for weak separability are that the $\mathbf{y}, \mathbf{p}^y$ and $\mathbf{x}, \mathbf{p}$ datasets both satisfy GARP.

The interpretation of (19) is a bit more subtle. Suppose we replace the quantities of the weakly separable goods, $\mathbf{y}_i$, with V_i for each observation and we replace the prices of those goods, $\mathbf{p}_i^y$, with $1/\mu_i$, within the $\mathbf{x},\mathbf{p}$ dataset. Let $\tilde{\mathbf{x}}_i = (z_{1i},...,z_{Mi},V_i)$ and $\tilde{\mathbf{p}}_i = (p_{1i}^z,...,p_{Mi}^z,1/\mu_i)$ denote the resulting quantities and prices. (19) is equivalent to the condition that the $\tilde{\mathbf{x}},\tilde{\mathbf{p}}$ dataset satisfies GARP. In the remainder of this paper, I will refer to the indexes V and μ collectively as Afriat indexes and I will refer to V as a sub-utility index.

D. Non-Parametric Tests of Weak Separability

The necessary and sufficient conditions for weak separability can be tested through a three-step sequence of three GARP tests:

Step 1 - Test the $\mathbf{x},\mathbf{p}$ dataset for GARP and reject weak separability if it is violated.

Step 2 - Test the $\mathbf{y},\mathbf{p}^y$ dataset for GARP and reject weak separability if it is violated.

If the $\mathbf{y},\mathbf{p}^y$ and $\mathbf{x},\mathbf{p}$ datasets are both consistent with GARP, then use a numerical algorithm to compute Afriat indexes satisfying (18).

Step 3 - Test the $\tilde{\mathbf{x}},\tilde{\mathbf{p}}$ dataset for GARP and reject weak separability if it is violated.

Varian (1982) provided a numerical algorithm to construct the Afriat indexes, and a weak separability test based on that algorithm was provided in the widely-used computer program NONPAR. In spite of its widespread use in empirical studies, it is well-known that the NONPAR test is extremely biased toward rejecting weak separability; See, for examples, Fisher (1989), Barnett and Choi (1989), Swofford and Whitney (1994), and Fleissig and Whitney (2003). The bias could be reduced by using an alternative algorithm to compute the Afriat indexes, since these indexes are not uniquely determined by the data.

Fleissig and Whitney (2003) proposed an alternative linear programming (LP) algorithm to construct the Afriat indexes. Their algorithm computes Afriat indexes that minimize

$$F = \sum_{i=1}^{I} \left\{ \left| V_i - Q_i^T \right| + \left| \mu_i - \frac{Q_i^T}{\mathbf{y}_i \cdot \mathbf{p}_i^y} \right| \right\}$$

subject to the constraints in (18), where Q_i^T is a Törnqvist quantity index constructed for the $\mathbf{y}, \mathbf{p}^y$ dataset.[14] The basic idea behind the algorithm is that the Törnqvist index should approximately satisfy the Afriat inequalities as a sub-utility index, since it is a superlative index number. The proportional root mean square error formula

$$PRMSE(V) = \sqrt{\frac{1}{I} \sum_{i=1}^{I} \left(\frac{V_i - Q_i^T}{Q_i^T} \right)^2} \tag{20}$$

provides a natural measure of the differences between the sub-utility index and the Törnqvist index.

Fleissig and Whitney's algorithm appears to significantly reduce the bias towards rejection; See Fleissig and Whitney (2003) and Jones, Dutkowsky, and Elger (2005). Swofford and Whitney (1994) proposed a joint test of the necessary and sufficient conditions for weak separability, which can completely eliminate the bias toward rejection in non-parametric weak separability tests. Their test involves computing strictly positive indexes, $U_i, V_i, \lambda_i, \mu_i$ and ϕ_i, $i = 1, ..., I$ that minimize

$$G = \sum_{i=1}^{I} (\mu_i \phi_i - \lambda_i)^2$$

subject to following conditions:

$$V_j - V_i \le \mu_i (\mathbf{p}_i^y \cdot \mathbf{y}_j - \mathbf{p}_i^y \cdot \mathbf{y}_i) \text{ for all } i, j = 1, ..., I, \tag{21}$$

$$U_j - U_i \le \lambda_i (\mathbf{p}_i^z \cdot \mathbf{z}_j - \mathbf{p}_i^z \cdot \mathbf{z}_i) + \phi_i (V_j - V_i) \text{ for all } i, j = 1, ..., I. \tag{22}$$

If a feasible solution to the constraints can be found such that G is minimized to zero, then $\phi_i = \lambda_i / \mu_i$ and (21) and (22) are equivalent to (18) and (19). In that case, the necessary and sufficient conditions for weak separability are satisfied.[15] The Swofford and Whitney test is unbiased, since the necessary and sufficient conditions for weak separability are tested jointly.

E. Parametric Tests of Weak Separability

I focus on non-parametric revealed preference tests of weak separability in this paper, but weak separability can also be tested parametrically. Parametric weak separability tests can be formulated through restrictions on the estimated parameters of a flexible functional form. Barnett and Choi (1989) evaluate parametric tests based on several standard flexible functional forms including the Translog and Rotterdam models. They emphasize that flexible functional forms lose their local flexibility property if weak separability is imposed globally. Consequently, the parametric tests are often formulated as tests of approximate separability (*i.e.* exact separability at a point and approximate separability elsewhere).[16]

In a Monte Carlo study, Barnett and Choi (1989) found that "[r]elative to the separability testing criteria, all models usually performed poorly… the problem is not resolved by conducting the test at a point, rather than globally, or by the use of Varian's non-parametric test." Barnett and Choi's negative results regarding the non-parametric test were based on the NONPAR test. The unbiased joint test of those conditions proposed by Swofford and Whitney (1994) may perform better in such experiments, although more study is needed; See also Fleissig and Whitney (2003).

Swofford and Whitney (1994) discuss the advantages and disadvantages of the non-parametric approach relative to the parametric approach. The main advantage of the non-parametric approach is that it does not require any assumptions about the form of the utility function; See also Varian (1982, 1983, 1990), Swofford and Whitney (1987), and Gross (1995). In contrast, the parametric tests are actually "joint tests" of weak separability and the choice of functional form.

The main disadvantage of the standard non-parametric approach is that it is non-stochastic and cannot account for measurement errors. In contrast, parametric weak separability tests explicitly incorporate stochastic elements through the error term in the demand system. Varian (1985) and Epstein and Yatchew (1985) proposed extensions to the standard revealed preference tests of cost minimization (WACM) and utility maximization (GARP), which can be used to determine whether or not violations of optimizing behavior could be attributed to random measurement errors. The approach is based on computing adjusted quantity data needed to make the observed data compliant with the required axiom (*e.g.* GARP or WACM); See Varian (1985), Jones, Dutkowsky and Elger (2005), and Jones and de Peretti (2005) for some empirical applications. A weakness of the approach is that the tester needs to have some knowledge about the second moment of the distribution of random measurement errors in the data, which can be compared with a bound or "critical value" computed from the magnitude of the adjustments made to the observed data.[17]

The more advanced forms of the non-parametric tests suggested by Varian (1985) and Swofford and Whitney (1994) are based on solving complex non-linear programming problems. The complexity of these problems increases with the square of the sample size, *I*, since the number of non-linear inequality constraints is a function of $I(I-1)$. Thus, the most advanced non-parametric revealed preference techniques are, at present, limited to datasets with moderate sample sizes.

IV. Admissible Asset Groupings for the United States

In this section, I illustrate the use of non-parametric revealed preference tests using data for the United States for the period 1993-2003. Specifically, I test various conventional groupings of monetary assets for admissibility using Fleissig and Whitney's weak separability test and discuss some of the implications of the test results.

Revealed preference tests have been widely applied in monetary economics; See, for examples using US data, Swofford and Whitney (1987, 1994), Belongia and Chalfant

(1989), Fisher (1989), Fisher and Fleissig (1997), and Jones, Dutkowsky, and Elger (2005). Patterson (1991), Belongia and Chrystal (1991), and Drake (1997) apply non-parametric tests to UK monetary data and Spencer (1997) tests European monetary assets for GARP using data for the principle members of the Exchange Rate Mechanism (ERM). The empirical application in this paper builds most directly upon Jones, Dutkowsky, and Elger (2005).

A. Data Description

The current set of monetary assets comprising the M3 monetary aggregate for the US are described in Table 1. The components differ slightly in earlier periods due to the introduction of new monetary assets and other changes in the data. Table 1 breaks the assets up into those that comprise the conventional M1, M2M, M2, and MZM, monetary aggregates as they are currently defined. M2M (M2 minus small time deposits) and MZM (M2M plus institution only money market mutual funds) are usually referred to as zero-maturity monetary aggregates. Zero-maturity monetary aggregates have been advocated by various authors as an alternative to M2; See Duca and Van Hoose (2004) for a recent survey of the issues.

The quantities and user costs for the monetary assets in M3 are provided by the Federal Reserve Bank of St. Louis in the monetary services index (MSI) database; See Anderson, Jones, and Nesmith (1997b) and Anderson and Buol (2005). The benchmark rate of return used to calculate the user costs is the envelope of the own rates of the monetary assets and the BAA bond rate. The data used in this study are the revised data described in Anderson and Buol (2005), except that I assume that demand deposits (DD) are non-interest bearing in the user cost calculations. Anderson and Buol (2005) provide separate user costs for BDMMMF and IOMMMF among other improvements in the data.

Jones, Dutkowsky, and Elger (2005) tested groups of monetary assets for weak separability from consumer goods and services, leisure, and other monetary assets using quarterly data for the sample period 1993-2001. In this paper, I test the M1, M2M, M2, MZM, and M3 asset groups for weak separability over the extended sample period 1993Q1-2003Q4 using Fleissig and Whitney's (2003) test. In addition, I also test an asset group denoted by M1+, which is comprised of the components of M1 and the two savings deposits components (SAVMMDAC and SAVMMDAT).

The dataset includes quarterly data for the 13 monetary assets in M3 and three categories of consumption that make up personal consumption expenditures (PCE): PCE Durables (*DUR*), PCE Non-Durables (*NDUR*), and PCE Services (*SER*). The quantities of the monetary assets are quarterly-averaged nominal asset stocks from the MSI database converted to real per-capita terms using the PCE deflator and the quarterly-averaged US civilian population.[18] The nominal user costs of the monetary assets are quarterly-averaged monthly real user costs from the MSI database multiplied by the PCE deflator. The quantities of the consumption variables are real per-capita expenditures and the corresponding prices of the consumption variables are implicit deflators. All consumption data is from FRED II.

B. Test Results

The weak separability test is carried out in three steps, as described in Section III. The first step of the test is to check the $\mathbf{x},\mathbf{p}$ dataset for GARP, which is satisfied. The second step of the test is to check the $\mathbf{y},\mathbf{p}^{y}$ datasets for GARP on a case by case basis. This condition is satisfied for M2M, MZM, and M3, but is violated for M1, M1+, and M2. There are 4 violations for M1, 9 violations for M1+, and 7 violations for M2. Therefore, these three groups violate the necessary conditions for weak separability. For the remaining groups, I computed Afriat indexes using Fleissig and Whitney's LP algorithm. The third step of the test is to check the $\tilde{\mathbf{x}},\tilde{\mathbf{p}}$ datasets for GARP, again on a case by case basis. The third step indicates that M2M, MZM, and M3 are all weakly separable.

C. The Törnqvist Index and the Sub-Utility Index Compared

As discussed above, the index, V_i, can be interpreted as the sub-utility levels at each observation. Jones, Dutkowsky, and Elger (2005) found that the differences between the sub-utility index and the Törnqvist quantity index were generally small as measured by *PRMSE*(*V*); See equation (20). The test results in this paper, using the revised and extended dataset, corroborate those findings. In particular, *PRMSE*(*V*) was 0.0133 for M2M, 0.0089 for MZM, and 0.0057 for M3.

Figures 1 and 2 compare the two indexes in levels for M2M and M3 respectively. The figures show that the Törnqvist quantity index needs only small adjustments to satisfy the Afriat inequalities as a sub-utility index, especially for M3. The adjustments are larger for M2M, but the largest adjustments are confined to 1995. The corresponding graph for MZM (not shown) has adjustments along the lines of those for M3. Several interpretations of these findings may be possible. If we regard the flow of utility provided by the monetary assets as a measure of "monetary services", then this result indicates that the Törnqvist quantity index is tracking this monetary service flow with a high degree of accuracy. Indeed, this result may provide some additional justification for the use of the term "monetary services indexes" that is often associated with the Törnqvist quantity index; See Farr and Johnson (1985) and Anderson, Jones, and Nesmith (1997b).

On the other hand, the finding may be somewhat surprising, because the theoretical ability of superlative quantity indexes to track the utility function depends on the property of linear homogeneity, which is usually rejected empirically. In the non-homothetic case, the tracking ability of superlative quantity indexes are usually defined relative to the distance function rather than the utility function; See, for example, Diewert (1981, pp. 211-213).

D. Analysis

The weak separability tests find that M2M, MZM, and M3 are all admissible as indicators over the period 1993-2003, but M1, M1+, and M2 are not. Given that M3 is a very broad collection of assets, the results seem to suggest focusing on one of the two zero-maturity monetary aggregates.

The weak separability test results in this paper are for the recent period 1993-2003. There are many studies that test for admissibility in preceding periods. One of the most comprehensive studies was Fisher and Fleissig (1997). They tested the M1, M2, M3, and L asset groups for weak separability over various sub-periods covering 1960-1993. The only group that was weakly separable in all of the sub-periods was a group that included all of the components of M1, except for business demand deposits. Similarly, among the conventional asset groups, M1 received the most support in their study. Swofford and Whitney (1987) studied the period 1970-1985. They found that a relatively narrow group of monetary assets including currency, household demand deposits, other checkable deposits and savings deposits was weakly separable, which is similar to the M1+ measure tested in this paper. They also found that expanding the collection of assets beyond savings deposits (for example, to include small-denomination time deposits) resulted in rejecting weak separability. Thus, on balance, these earlier studies tended to support relatively narrow monetary aggregates.

Moving beyond the test results, zero-maturity aggregates are probably preferable to M1 in my sample period, because measures of transactions deposits and M1 have been distorted by retail sweep programs; See Dutkowsky and Cynamon (2003), Jones, Dutkowsky, and Elger (2005), and Dutkowsky, Cynamon, and Jones (2005). Anderson and Rasche (2001) provide a detailed discussion of the growth of retail sweeping in the US. In a retail sweep program, banks reclassify balances in their customers' transactions deposits to a money market deposit account (MMDA). This reduces the bank's required reserves, but does not affect the customer's perceived transactions account balance since retail sweeping is essentially invisible to the customer.

Retail sweep programs were first introduced in January of 1994, but began to spread rapidly beginning in May of 1995. According to Anderson (2003, p. 4), "[t]oday, households (and firms) perceive themselves to own approximately *twice* as many transactions deposits in depository institutions as those same depository institutions report to the Federal Reserve (and the Federal Reserve includes in M1)…". Dutkowsky and Cynamon (2003) find that retail sweeps help explain systematic over-prediction in post-sample simulations of M1 demand.

V. Divisia Indexes for the United States

In this section, I compare and contrast Divisia and simple sum indexes for the United States over the period 1970-2004.[19] My analysis will have three main goals: First, to highlight the most economically significant differences between the indexes; Second, to use expenditure shares to interpret the differences between the indexes; Third, to determine the robustness of the Divisia index to the empirical proxy of the benchmark interest rate used to calculate the user costs of the monetary assets. The empirical analysis draws mainly upon Barnett (1984), Barnett, Fisher, and Serletis (1992), and Stracca (2001, 2005).[20]

A. Comparison of Divisia and Simple Sum Indexes

I begin my analysis by comparing the Divisia and simple sum indexes for M1, M2, M2M, and M3. Figures 3 through 6 plot the monthly levels of the nominal simple sum and Divisia indexes for M1, M2M, M2, and M3 respectively over 1970:1 - 2004:12. The Divisia indexes are from the Federal Reserve Bank of St. Louis (Anderson and Buol, 2005) and the simple sum indexes are from FRED II. All indexes have been rescaled to equal 100 in 1970:1.

It is apparent from the Figure 3 that Divisia and simple sum M1 are very similar and that the levels of both indexes declined significantly beginning in the mid-1990s reflecting the spread of retail sweeping. Based upon these two observations, I will not directly consider the M1 indexes further in this paper. Figures 5 and 6 show that the simple sum index has had a higher average growth rate than the Divisia index for M2 and M3 over the sample period and the spread between the indexes widens after 1978. Barnett, Fisher, and Serletis (1992) provide similar graphs for M1, M2, M3, and L using data covering 1970-1987 and draw similar conclusions. Figure 4 shows that M2M is somewhat of an intermediate case. Divisia and simple sum M2M are very similar until approximately 1980, but differ to some extent afterwards.

The most economically significant differences between the growth rates of the Divisia and simple sum indexes show up during the critical monetary policy episode from 1978 to 1982. In order to highlight these differences, Figures 7 and 8 plot the annualized monthly growth rates of the indexes for M2 and M3 respectively over 1974:1 – 1987:12.[21] The growth rates of the two indexes are fairly similar from 1974 to 1977, especially for M2. The growth rates differ significantly, however, towards the end of 1978 and continuing through early 1982. In particular, the figures show that the growth rates of Divisia M2 and M3 were negative at multiple points during the period from late 1978 to mid 1981, whereas the growth rates of the simple sum indexes were not. It is also apparent from the figures that the growth rates of Divisia M2 and M3 are very similar over this period. At the time, Barnett (1984) argued that monetary policy as measured by Divisia M2 and M3 was, consequently, both much tighter and "much more volatile than was suggested by the simple sum aggregates".[22] These conclusions are clearly upheld using much more recent vintages of the data.

Statistical correlations between the indexes tell a similar story as the figures. The correlation between the growth rates of Divisia and simple sum M2 were 0.86 over the period 1970-1987, but were 0.76 over the shorter period 1979-1987. Similarly, the correlation between the growth rates of Divisia and simple sum M3 were 0.69 over the period 1970-1987, but were 0.49 over the shorter period 1979-1987. In contrast, the full sample correlations (1970-2004) are 0.9 for M2 and 0.85 for M3. The simple sum and Divisia indexes for M1 and M2M are highly correlated over the full sample and over all of the sub-samples.

B. Explaining the Differences

The Divisia indexes are based on the superlative Törnqvist quantity index formula. As described above, the growth rate of the Törnqvist index is a weighted sum of the growth rates of the component quantities, where the weights are the average expenditure shares between the two periods; See Barnett (1983) for further discussion. The levels of the simple sum indexes are unweighted sums of the levels of the asset quantities. As discussed above, the growth rate of the simple sum index is a weighted sum of the growth rates of the component quantities, where the weight on the growth rate of a particular asset is the ratio of the quantity of that asset to the level of the simple sum index; See equations (14) and (15). I will refer to these weights as expenditure weights and simple sum weights respectively. Mike Belongia (in this volume) provides further empirical results along that compliment the analysis presented here.

Mathematically, the expenditure weight on the growth rate of the i^{th} monetary asset in period t is $(w_{it} + w_{i,t-1})/2$, where $w_{it} = \pi_{it} m_{it} / \boldsymbol{\pi}_t \cdot \mathbf{m}_t$ is the expenditure share of the i^{th} asset in period t. The user costs are defined as $\pi_{it} = (R_t - R_{it})/(1 + R_t)$ as described above. The simple sum weight on the growth rate of the i^{th} monetary asset in period t is $m_{i,t-1} / \mathbf{1} \cdot \mathbf{m}_{t-1}$, where $\mathbf{1}$ is a vector having 1's in each element.

Stracca (2001) analyzed a synthetic Divisia index for the Euro area by comparing the expenditure weights for the growth rate of the Törnqvist index with the corresponding simple sum weights. In particular, he graphed these weights for the components of Euro area M1 relative within indexes for M3. He found that the expenditure weights always exceeded the simple sum weights from 1980 to 2000, which he interpreted this as meaning that "…the Divisia monetary aggregate [for M3] is 'more liquid' than simple sum M3 and 'less liquid' than simple sum M1." In other words, the growth rates of the liquid assets in M1 are (collectively) weighted more heavily in the growth rate of the Divisia index than they are in the corresponding simple sum index, both of which contain relatively illiquid (Non-M1) monetary assets.

Figure 9 plots analogous expenditure and simple sum weights for an aggregate of the components of M1 within the M2 indexes for the US for 1970-1987. Figure 10 plots the weights for 1988-2004. I assume that DD is non-interest bearing in these calculations to simplify interpretation, but this assumption does not qualitatively affect the results or subsequent analysis. The weights within the M3 indexes (not shown) are very similar to the weights within M2. The figures show that the expenditure weights for the components of M1 exceed the simple sum weights for both periods, similar to Stracca's results for the Euro area. In Figure 9, the weights tend to decline for both simple sum and Divisia throughout most of the 1970s, but the gap remains relatively stable throughout the decade. The gap declines somewhat, however, in the early to mid-1980s.

In Figure 10, the gap between the weights is less stable. The weights on the M1 components tend to decline for both indexes beginning in 1994, which reflects (at least in part) the spread of retail sweeping, as discussed above. Figure 11 plots the benchmark rate (R_t in the user cost formula) used to compute the Divisia index against the Federal Funds rate for 1988-2004. A comparison of Figures 10 and 11 reveals an interesting relationship in the most recent sample period. When the Federal Funds rate is relatively close to the benchmark rate, the spread between the expenditure weights and the simple sum weights on the M1 components widens. Similarly, when the Federal Funds rate is relatively far below the benchmark rate (*e.g.* in 2002 and 2003), the spread between the expenditure and simple sum weights on the M1 components narrows considerably.[23]

This relationship can be interpreted by considering the relationship between the own rates on the various monetary assets and the Federal Funds rate. If the own rates on all of the monetary assets were zero, then their user costs would all be the same and the expenditure and simple weights would be approximately equal. In terms of the components of M1, the own rates on CC and DD are assumed to be zero in my calculations and the own rates on OCDCTOT and OCDTTOT are low relative to the own rates on the other monetary assets. In contrast, the own rates on many of the Non-M1 components of M2 (*e.g.* BDMMMF) are fairly highly correlated with the Federal Funds rate. So, when the Federal Funds rate is low, the own rates on many of the Non-M1 components of M2 will also be fairly low and, consequently, the expenditure and simple sum weights will be relatively close together. When the Federal Funds rate rises, the own rates on many of the Non-M1 components of M2 also rise, but CC and DD are assumed to be non-interest bearing. This, in turn, implies that the spread between the expenditure weight on the M1 components will increase relative to the simple sum weight on M1, as reflected in Figure 10.[24]

C. The Benchmark Rate

The empirical proxy for the benchmark rate used to construct the user costs of the monetary assets has been discussed in several recent papers; See Stracca (2001), Barnett (2003), Hancock (2005), and Drake and Mills (2005). Hancock (2005, p. 40) explains that "[t]he optimal benchmark asset should provide at least as good a store of value as the components of the money supply, but have no use for transactions." The benchmark rate of return should also be for a risk-free asset. In practice, an empirical proxy must be chosen for the benchmark rate, since no such asset exists in reality. The Federal Reserve Bank of St. Louis (Anderson, Jones, and Nesmith, 1997b) uses an envelope of the own rates of the monetary assets and the BAA bond rate as the benchmark rate. Barnett (2003) and Stracca (2001, 2004) criticize the use of long-term bond rates in proxies of the benchmark rate. Stracca (2001, p. 16), for example, states "…the 10 year maturity is too long and not representative of agents' 'normal investment horizon'. At shorter horizons bond yields are, of course, not risk-free." The problem of choosing an empirical proxy for the benchmark rate is essentially the same problem as choosing the alternative rate of return in a conventional aggregate money demand function, an issue that has been the focus of numerous empirical studies.

In order to assess the empirical importance of this issue, I constructed Divisia indexes using an alternative proxy for the benchmark rate. Specifically, the benchmark rate used in my alternative Divisia indexes is the upper envelope of the interest rates on only the monetary assets in M3 (*i.e.* the maximum own rate on the assets in M3, but not including the BAA rate).[25] I computed statistical correlations between the growth rates of the alternative Divisia indexes and the growth rates of the published Divisia and simple sum indexes for M2 and M3 over 1970-2004. The correlation between the growth rates of the alternative and published Divisia indexes is 0.87 for M2 and 0.78 for M3, whereas the correlation between the growth rates of the alternative Divisia indexes and the growth rates of the simple sum indexes is 0.67 for M2 and 0.50 for M3. As discussed above, the correlation between the growth rates of the published Divisia and simple sum indexes is 0.90 for M2 and 0.85 for M3. Thus, the change in benchmark does appear to have an impact on the properties of the Divisia index. In particular, the correlation between the growth rates of the Divisia and simple sum indexes is substantially lower for the alternative Divisia indexes. These results suggest that future research might be profitably directed toward improving empirical proxies of the benchmark rates and toward understanding how the benchmark rate affects the information content of the Divisia index as an indicator variable.

VI. Conclusions

Divisia monetary services indexes now provide an accepted alternative to the simple sum monetary aggregates published by many central banks. Divisia indexes have been widely studied in academia and are currently published by the Federal Reserve Bank of St. Louis for the US and the Bank of England for the UK.

In this paper, I have surveyed the theoretical literature on monetary aggregation theory with a focus on two key issues: How do we choose the components of a monetary aggregate so as to be consistent with aggregation theory? And, how do we construct an index that measures the flow of monetary services derived from those monetary components? I have argued that by identifying admissible groups of monetary assets through testing for weak separability and by using a superlative index number (such as the Törnqvist) to measure the flow of monetary services over admissible asset groups, we can resolve both issues in a way that is consistent with economic theory. In contrast, the simple sum index is based on the implicit assumption that all of the monetary assets are perfect substitutes for one another.

Turing from theory to practice, I illustrate the use of non-parametric weak separability tests on data for the United States for the period 1993-2003. I find that two zero-maturity asset groups (M2M and MZM) and M3 are admissible, but that M1 and M2 are not. I also find (using a method proposed by Fleissig and Whitney, 2003) that the superlative Törnqvist quantity index needs only small adjustments to measure the utility derived from the admissible monetary asset groups. Next, I consider the differences between the Törnqvist (Divisia) and simple sum indexes for the United States over the period 1970-2004. The most economically significant differences between the Divisia and simple sum indexes are found to have occurred during the critical period from 1978 to 1983. Specifically, monetary policy would have been viewed as being much tighter and also more volatile if Divisia monetary aggregates had been used to gauge the effects of policy actions instead of conventional simple sum monetary aggregates over that period, key differences that were noted by contemporary observers.

The 1990s have seen the focus on monetary aggregates in the US largely shift from narrow aggregates (such as M1) towards broader aggregates (such as MZM or M2), although much controversy remains about the best broad measure to use in empirical studies; See Duca and VanHoose (2004) for a through discussion of these issues. In addition, the "missing M2" episode of the early 1990s has led to increasing interest in the degree of substitutability between small-denomination time deposits and bond and stock mutual funds. These issues, however, point to continued or increased interest in Divisia measures, because broader measures of money contain more assets that are imperfect substitutes for the medium of exchange assets in M1, which are themselves imperfect substitutes for one another. Moreover, the potential inclusion of capital uncertain assets in a broad monetary aggregate implies greater attention to appropriately measuring the opportunity costs of monetary assets, which has been a traditional focus in monetary aggregation theory. As stated succinctly by Lucas (2000, p.270-1), "I share the widely held opinion that M1 is too narrow an aggregate for this period [the 1990s], and I think that the Divisia approach offers much the best prospects for resolving the difficulty.”

References

Afriat, S. (1977) *The Price Index*, Cambridge: Cambridge University Press.

Anderson, R. (2003) "Retail Deposit Sweep Programs: Issues for Measurement, Modeling, and Analysis",Federal Reserve Bank of St. Louis Working Paper 2003-026A.

Anderson, R., B. Jones, and T. Nesmith (1997a) "Monetary Aggregation Theory and Statistical Index Numbers", Federal Reserve Bank of St. Louis *Review* 79, Jan/Feb, 31-51.

Anderson, R., B. Jones, and T. Nesmith (1997b) "Building New Monetary Services Indexes: Concepts, Data, and Methods", Federal Reserve Bank of St. Louis *Review* 79, Jan/Feb, 53-82.

Anderson, R., and J. Buol (2005) "Revisions to User Costs for the Federal Reserve Bank of St. Louis Monetary Services Indexes." Federal Reserve Bank of St. Louis Review, forthcoming.

Anderson, R. and R. Rasche (2001) "Retail Sweep Programs and Bank Reserves: 1994-1999", Federal Reserve Bank of St. Louis *Review* 83, Jan/Feb, 51-72.

Barnett, W. (1978) "The User Cost of Money", *Economics Letters*, 1, 145-149.

Barnett, W. (1980) "Economic Monetary Aggregates: An Application of Index Numbers and Aggregation Theory", *Journal of Econometrics*, 14, 11-14.

Barnett, W. (1982) "The Optimal Level of Monetary Aggregation", *Journal of Money, Credit, and Banking*, 14, 687 - 710.

Barnett, W. (1983) "Understanding the New Divisia Monetary Aggregates", *Review of Public Data Use* 11, 349-355.

Barnett, W. (1984) "Recent Monetary Policy and the Divisia Monetary Aggregates", *American Statistician* 38, 165-172.

Barnett, W. (1987) "The Microeconomic Theory of Monetary Aggregation", in *New Approaches to Monetary Economics*, W. Barnett and K. Singleton (eds.) Cambridge: Cambridge University Press, 115-168.

Barnett, W. (1995) "Exact Aggregation Under Risk", in *Social Choice, Welfare, and Ethics*, W. Barnett, H. Moulin, M. Salles, and N. Schofield (eds.) Cambridge: Cambridge University Press, 252-274.

Barnett, W. (2003), "Aggregation Theoretic Monetary Aggregation over the Euro Area When Countries are Heterogeneous", *ECB Working Paper* # 260.

Barnett, W. and P. Spindt (1982) "Divisia Monetary Aggregates: Compilation, Data, and Historical Behavior", *Staff Study* # 116, Federal Reserve Board.

Barnett, W., E. Offenbacher and P. Spindt (1984) "The New Divisia Monetary Aggrgeates", *Journal of Political Economy* 92, 1049-1085.

Barnett, W. and S. Choi, (1989), "A Monte Carlo Study of Tests of Blockwise Weak Separability", *Journal of Business and Economic Statistics*, 7, 363-377.

Barnett, W., M. Hinich, and P. Yue (1991) "Monitoring Monetary Aggregates Under Risk Aversion", in *Monetary Policy on the 75th Anniversary of the Federal Reserve System*, M. Belongia (ed.) Boston: Kluwer Academic, 189-222.

Barnett, W., D. Fisher, and A. Serletis, "Consumer Theory and the Demand for Money", *Journal of Economic Literature* 30, December 1992, 2086-2119.

Barnett, W., Y. Liu and M. Jensen (1997) "CAPM Risk Adjustment for Exact Aggregation over Financial Assets", *Macroeconomic Dynamics* 1, 485-512.

Barnett, W. and A. Serletis (2000), *The Theory of Monetary Aggregation*, Amsterdam: North Holland.

Barnett, W. and J. Binner (2004) *Functional Structure and Approximation in Econometrics*, Amsterdam: North Holland.

Barnett, W. and S. Wu (2005) "On User Costs of Risky Monetary Assets", *Annals of Finance* 1, January, 35-50.

Belongia, M. (1995), "Weighted Monetary Aggregates: A Historical Survey", *Journal of International and Comparative Economics* 4, 87-114.

Belongia, M. (1996), "Measurement Matters: Recent Results from Monetary Economics Reexamined", *Journal of Political Economy*, 104, 1065-1083.

Belongia, M. and J. Chalfant (1989) "The Changing Empirical Definition of Money: Some Estimates from a Model of the Demand for Money Substitutes", *Journal of Political Economy* 97, April 1989, 387-97.

Belongia, M. and A. Chrystal (1991), "An Admissible Monetary Aggregate for the United Kingdom", *Review of Economics and Statistics*, 93, 497-503.

Belongia, M. and J. Binner (2000) *Divisia Monetary Aggregates: Theory and Practice*, London: Macmillan Press Ltd.
Binner J. and S. Wattam (2003) "A new composite leading indicator of inflation for the UK; a Kalman filter approach", *Global Business and Economics Review,* **5**, 242-264.
Binner J., R. Bissoondeeal, T.Elger, A. Gazely and A. Mullineux (2005) "A comparison of linear forecasting models and neural networks; An application to Euroinflation and EuroDivisia", *Applied Economics* 37, 665-680.
Blackorby, C., D. Primont, and R. Russell (1978) *Duality, Separability, and Functional Structure*, Amsterdam: North-Holland.
Blundell, R. and J. Robin (2000) "Latent Separability: Grouping Goods without Weak Separability", *Econometrica* 68, 53-84.
Crawford, I. (2004) "Necessary and Sufficient Conditions for Latent Separability", *IFS CEMMAC Working Paper* CWO02-04.
de Peretti, P. (2005) "Testing the Significance of the Departures from Utility Maximization", *Macroeconomic Dynamics*, forthcoming.
Deaton, A. and J. Muellbauer (1980) *Economics and Consumer Behavior*, Cambridge University Press.
Diewert, W. (1974) "Intertemporal Consumer Theory and the Demand for Durables", *Econometrica* 42, 497-516.
Diewert, W. (1976a) "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 115-145.
Diewert, W. (1976b) "Harberger's Welfare Indicator and Revealed Preference Theory", *The American Economic Review* 66, March, 143-152.
Diewert, W. (1981) "The Economic Theory of Index Numbers: A Survey", in *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone* A. Deaton (ed.), Cambridge: Cambridge University Press, 163-208.
Donovan, D. (1978) "Modeling the Demand for Liquid Assets: An Application to Canada", *IMF Staff Papers* 25, 676-704.
Drake, L. (1992), "The Substitutability of Financial Assets in the UK and the Implications for Monetary Aggregation", *Manchester School of Economics and Social Studies* 60, 221-248.
Drake, L. (1997), "Nonparametric Demand Analysis of UK Personal Sector Decisions on Consumption, Leisure and Monetary Assets: A Reappraisal"*, Review of Economic Statistics*, 70, 679-684.
Drake, L., A. Mullineux, and J. Agung (1997) "One Divisia Money for Europe?" *Applied Economics* 29, 775-786.
Drake L., and T. Mills (2005) "A New Empirically weighted Monetary Aggregate for the United States." *Economic Inquiry* 43, 138-157.
Duca, J. and D. VanHoose (2004) "Recent Developments in Understanding the Demand for Money" *Journal of Economics and Business* 56, July/Aug, 247-272.
Dutkowsky, D. and B. Cynamon (2003) "Sweep Programs: The Fall of M1 and the Rebirth of Medium of Exchange", *Journal of Money, Credit, and Banking* 35, 263-79.
Dutkowsky, D., B. Cynamon, and B. Jones (2005) "US Narrow Money for the Twenty-First Century", *Economic Inquiry* forthcoming.
Epstein, L. and A. Yatchew (1985) "Nonparametric Hypothesis Testing and Applications to Demand Analysis", *Journal of Econometrics* 30, 149-169.
Farr, H., and D. Johnson. (1985) "Revisions in the Monetary Services (Divisia) Indexes of the Monetary Aggregates", Staff Study #147, Federal Reserve Board.
Fase, M. and C. Winder (1996) "Divisia Aggregates and the Demand for Money in Core EMU", DNB Staff Reports #5.
Fisher, D. (1989) *Money Demand and Monetary Policy*, Ann Arbor: The University of Michigan Press.
Fisher, D., and A. Serletis (1989) "Velocity and the Growth of Money in the United States, 1970-1985", *Journal of Macroeconomics* 11, summer, 323-332.
Fisher, D. and A. Fleissig (1997), "Monetary Aggregation and the Demand for Assets", *Journal of Money, Credit, and Banking*, 29, 458-475.
Fisher, P., S. Hudson and M. Pradhan (1993), "Divisia Measures of Money: an Appraisal of Theory and Practice", *Bank of England Working Paper*, 9.
Fleissig, A. and G. Whitney (2003), "A New PC-Based Test for Varian's Weak Separability Conditions", *Journal of Business and Economic Statistics* 21, 133-144.
Fleissig, A. and G. Whitney (2005), "Testing for the Significance of Violations of Afriat's Inequalities",

Journal of Business and Economic Statistics 23, 355-362.
Ford, J. and J. Morris (1996) “The Money Multiplier, Simple Sum, Divisia and Innovation-Divisia Monetary Aggregates: Cointegration Tests for the UK”, *Applied Economics* 28, June, 705-714.
Friedman, M. and A. Schwartz (1970) *Monetary Statistics of the United States: Estimates, Sources, Methods, and Data*, New York: Columbia University Press (for the NBER)
Green, H. (1964) *Aggregation in Economic Analysis: An Introductory Survey*, Princeton: Princeton University Press.
Gross, J. (1995) “Testing Data for Consistency with Revealed Preference”, *The Review of Economics and Statistics* 77, 701-710.
Hancock, M. (2005) “Divisia Money”, *Bank of England Quarterly Bulletin* Spring, 39-46.
Houthakker, H. (1950) “Revealed Preference Theory and the Utility Function”, *Economica* XVII, May, 159-174.
Jones, B. and P. de Peretti (2005) “A Comparison of Two Methods for Testing the Utility Maximization Hypothesis when Quantity Data are Measured with Error”, *Macroeconomic Dynamics*, forthcoming.
Jones, B., D. Dukowsky, and T. Elger (2005) “Sweep Programs and Optimal Monetary Aggregation”, *Journal of Banking and Finance* 29, 483-508.
Jones, B., T. Elger, D. Edgerton and D. Dutkowsky (2005), “Towards a Unified Approach to Testing for Weak Separability”, *Economics Bulletin* 3, 1-7.
Lucas, R. (2000), “Inflation and Welfare”, *Econometrica*, 68, 247-274.
Patterson, K. (1991) "A Non-Parametric Analysis of Personal Sector Decisions on Consumption, Liquid Assets and Leisure", *Economic Journal*, 101, 1103-1116.
Pollak, R., and T. Wales (1992) *Demand System Specification andEstimation*, Oxford: Oxford University Press.
Richter, M. (1966) “Invariance Axioms and Economic Indexes”, *Econometrica* 34, 739-755.
Samuelson, P. (1948) “Consumption Theory in Terms of Revealed Preference”, *Economica* XV, Nov, 243-253.
Schunk, D. (2001) “The Relative Forecasting Performance of the Divisia and Simple Sum Monetary Aggregates”, *Journal of Money, Credit, and Banking* 33, May, 272-283.
Serletis, A. (1987) “Monetary Asset Separability Tests”, in *New Approaches to Monetary Economics*, W. Barnett and K. Singleton (eds.) Cambridge: Cambridge University Press,.
Serletis, A. (1988) “The Empiricl Relationship between Money, Prices, and Income Revisited”, *Journal of Business and Economic Statistics* 6, July, 351-358.
Serletis, A, and M. King (1993) “The Role of Money in Canada”, *Journal of Macroeconomics* 15, Winter, 91-107.
Spencer, P. (1997) “Monetary Integration and Currency Substitution in the EMS: The Case of a European Monetary Aggregate”, *European Economic Review* 41, 1403-1419.
Stracca, L. (2001), “Does Liquidity Matter? Properties of a Synthetic Divisia Monetary Aggregate in the Euro Area”, *ECB Working Paper* # 79.
Stracca, L. (2004), “Does Liquidity Matter? Properties of a Synthetic Divisia Monetary Aggregate in the Euro Area”, *Oxford Bulletin of Economics and Statistics*, 66, 309-331.
Swofford, J. and G. Whitney (1987) “Nonparametric Tests of Utility Maximization and Weak Separability for Consumption, Leisure, and Money”, *Review of Economics and Statistics* 69, 458-464.
Swofford, J. and G. Whitney (1991) “The Composition and Construction of Monetary Aggregates”, *Economic Inquiry* 29, 752-61.
Swofford, J. and G. Whitney (1994) “A Revealed Preference Test for Weakly Separable Utility Maximization with Incomplete Adjustment”, *Journal of Econometrics* 60 (1-2), Jan/Feb, 235-49.
Thornton, D. and P. Yue (1992) “An Extended Series of Divisia Monetary Aggregates”, Federal Reserve Bank of St. Louis *Review* 74, 35-52.
Varian, H. (1982), "The Nonparametric Approach to Demand Analysis", *Econometrica*, 50, 945-974.
Varian, H. (1983), "Non-parametric Tests of Consumer Behaviour", *Review of Economic Studies*, 50, 99-110.
Varian, H. (1985), "Non-parametric Analysis of Optimising Behaviour with Measurement Error", *Journal of Econometrics*, 30, 445-458.
Varian, H. (1990), "Goodness-of-Fit in Optimising Models", *Journal of Econometrics*, 46, 125-40.

Wesche, K. (1991) "The Demand for Divisia Money in a Core Monetary Union", Federal Reserve Bank of St. Louis *Review*, Sept/Oct, 51-60.

Table 1: Components of Monetary Aggregates

M1 = Currency and Travelers Checks (CC)
+ Demand Deposits (DD)
+ Other Checkable Deposits at Commercial Banks (OCDCTOT)
+ Other Checkable Deposits at Thrift Institutions (OCDTTOT)
M2M = M1
+ Savings Deposits (including MMDA) at Commercial Banks (SAVMMDAC)
+ Savings Deposits (including MMDA) at Thrift Institutions (SAVMMDAT)
+ Money Market Mutual Funds (BDMMMF)
MZM = M2M
+ Institution Only Money Market Mutual Funds (IOMMMF)
M2 = M2M
+ Small Denomination Time Deposits at Commercial Banks (SDTDC)
+ Small Denomination Time Deposits at Thrift Institutions (SDTDT)
M3 = M2
+ Institution Only Money Market Mutual Funds (IOMMMF)
+ Large Denomination Time Deposits (LDTD)
+ Total Repurchase Agreements (TOTRP)
+ Total Eurodollar Deposits (TOTED)

Table 2: Weak Separability Test Results

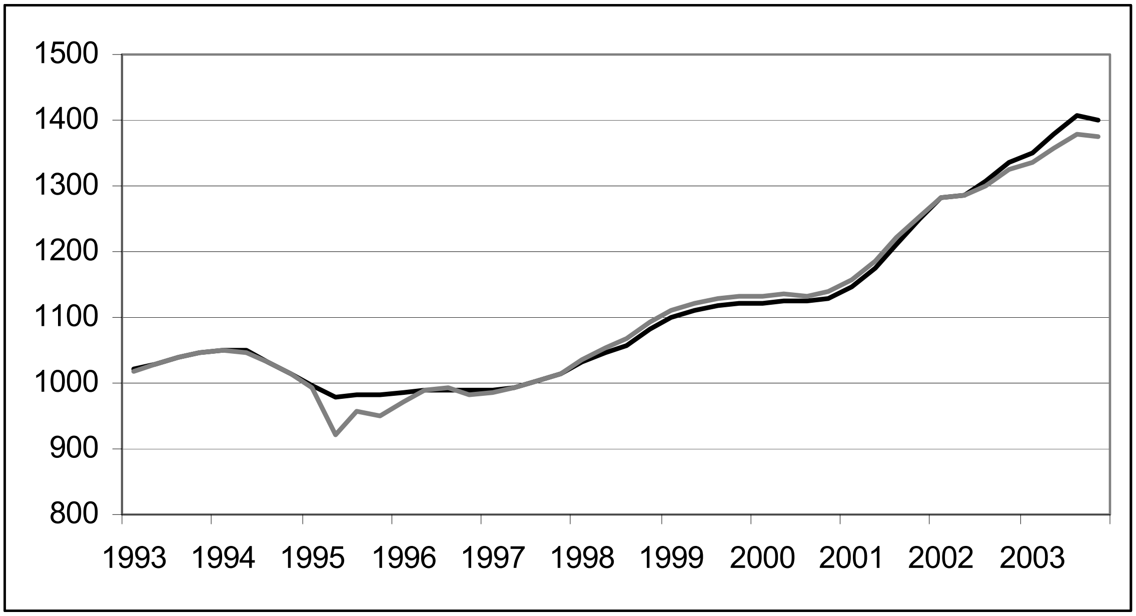
	Necessary Condition	Weak Separability
M1	N	------
M1+	N	------
M2M	Y	Y
MZM	Y	Y
M2	N	------
M3	Y	Y

Notes: Y for the necessary conditions indicates that the $\mathbf{y}, \mathbf{p}^{y}$ dataset is consistent with GARP, whereas an N indicates that it is not.

Y for weak separability indicates that the $\tilde{\mathbf{x}}, \tilde{\mathbf{p}}$ dataset is consistent with GARP (using the Fleissig and Whitney LP algorithm), whereas a number indicates the number of GARP violations. If the necessary condition is violated, then this condition cannot be tested.

Figure 1: Törnqvist Index versus Sub-Utility Index

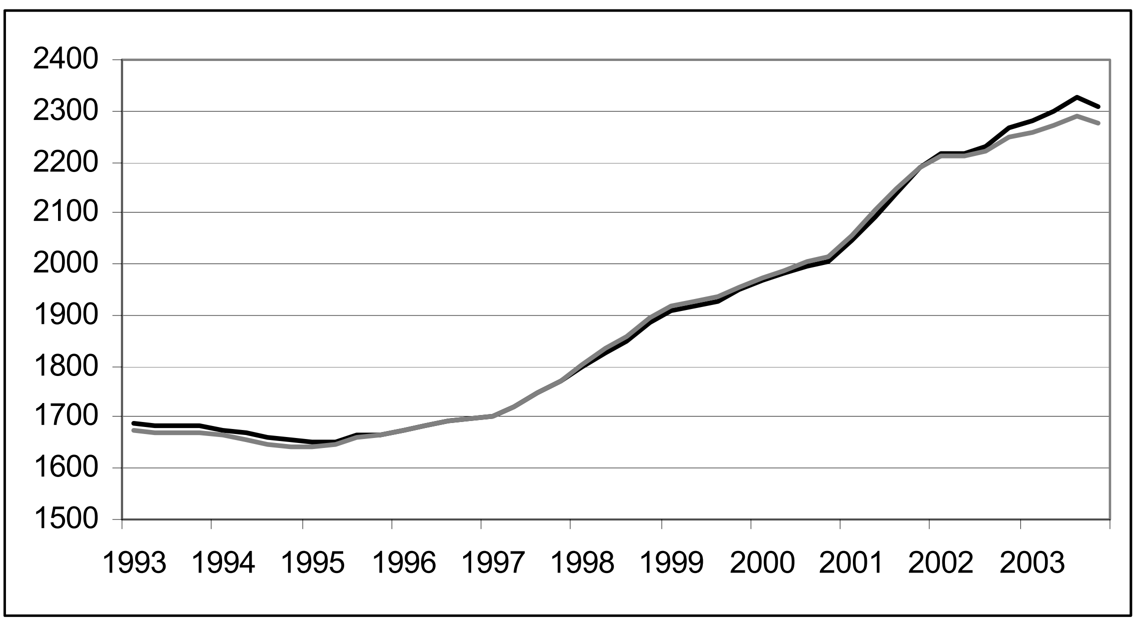
M2M Level of Aggregation



Notes: The black line is Törnqvist index computed from real per-capita asset stocks and nominal user costs. The grey line is the sub-utility index satisfying the Afriat inequalities.

Figure 2: Törnqvist Index versus Sub-Utility Index

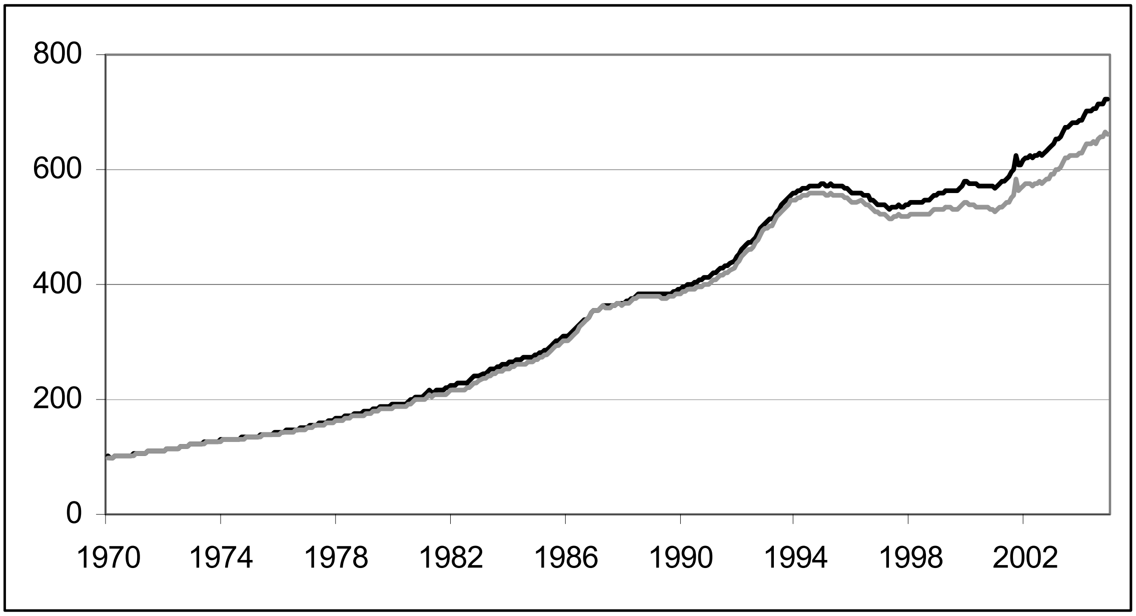
M3 Level of Aggregation



Notes: The black line is Törnqvist index computed from real per-capita asset stocks and nominal user costs. The grey line is the sub-utility index satisfying the Afriat inequalities.

Figure 3: Divisia versus Simple Sum Index

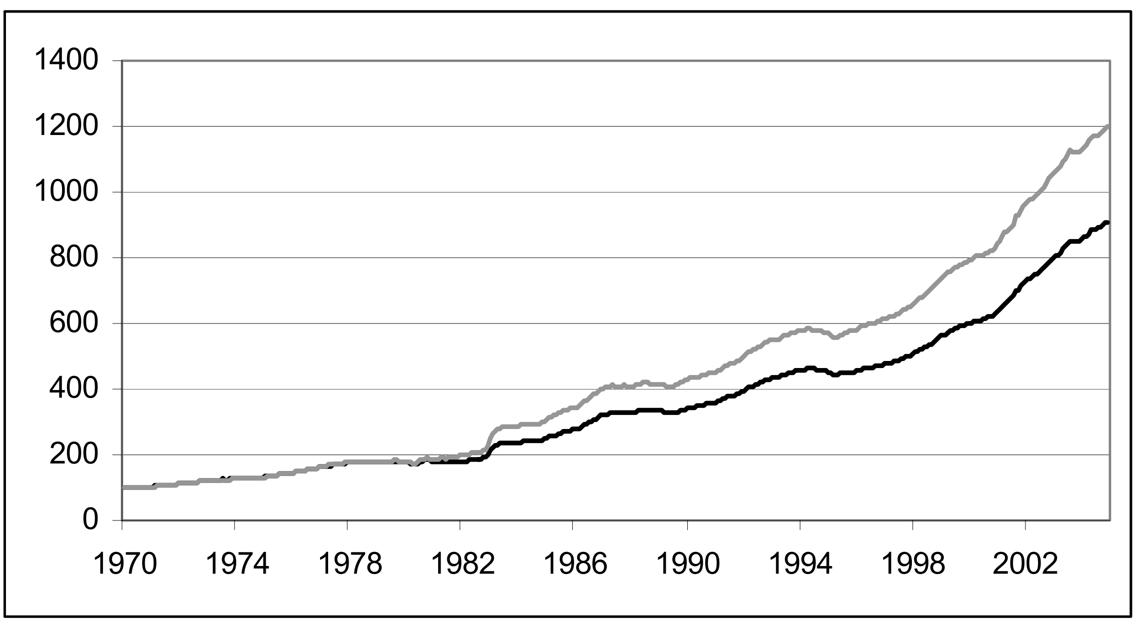
M1 Level of Aggregation



Notes: The black line is the nominal Divisia index and the grey line is the nominal simple sum index.

Figure 4: Divisia versus Simple Sum Index

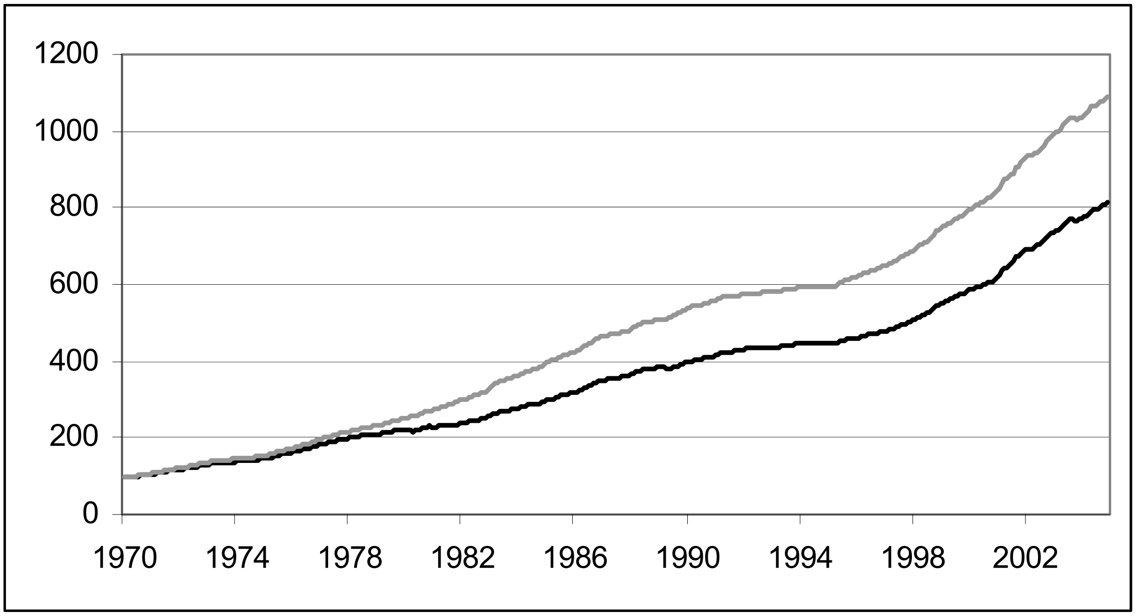
M2M Level of Aggregation



Notes: The black line is the nominal Divisia index and the grey line is the nominal simple sum index.

Figure 5: Divisia versus Simple Sum Index

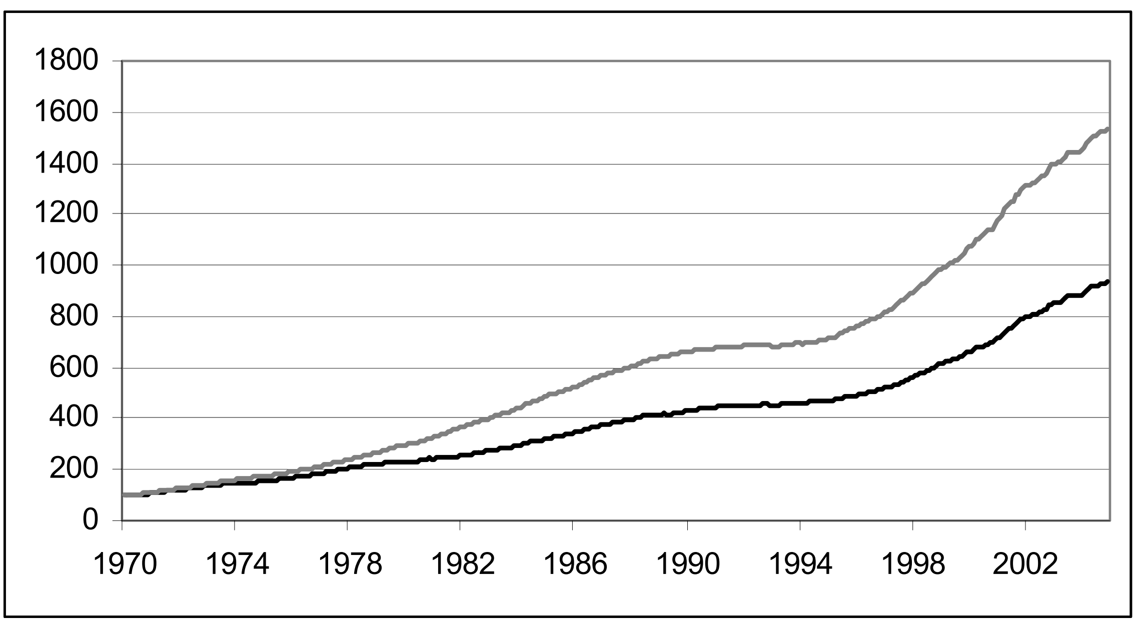
M2 Level of Aggregation



Notes: The black line is the nominal Divisia index and the grey line is the nominal simple sum index.

Figure 6: Divisia versus Simple Sum Index

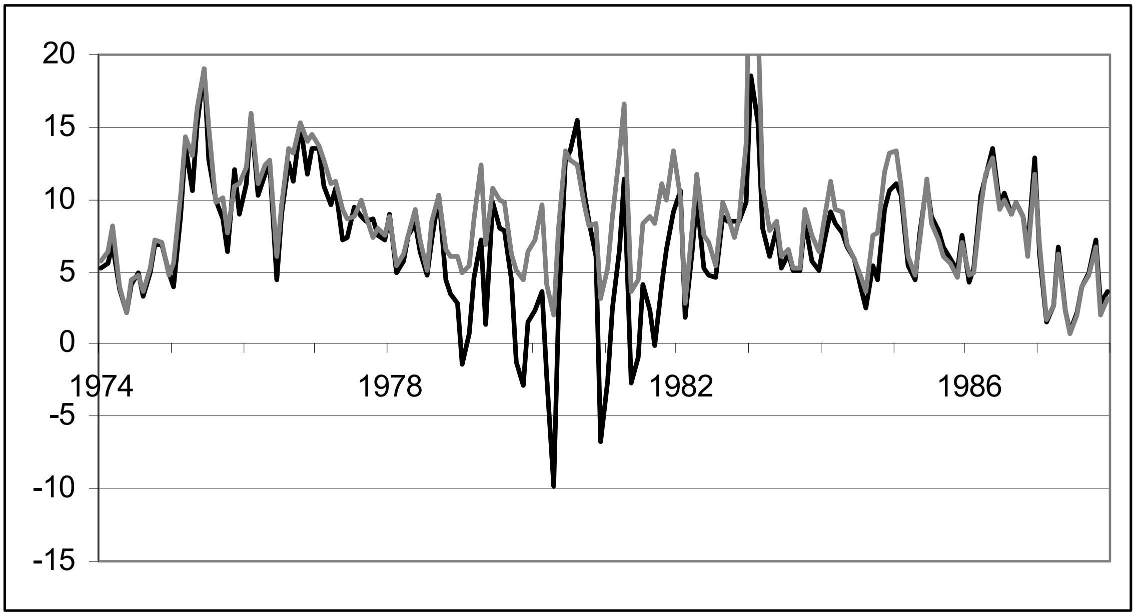
M3 Level of Aggregation



Notes: The black line is the nominal Divisia index and the grey line is the nominal simple sum index.

Figure 7: Growth Rates of Divisia and Simple Sum Indexes

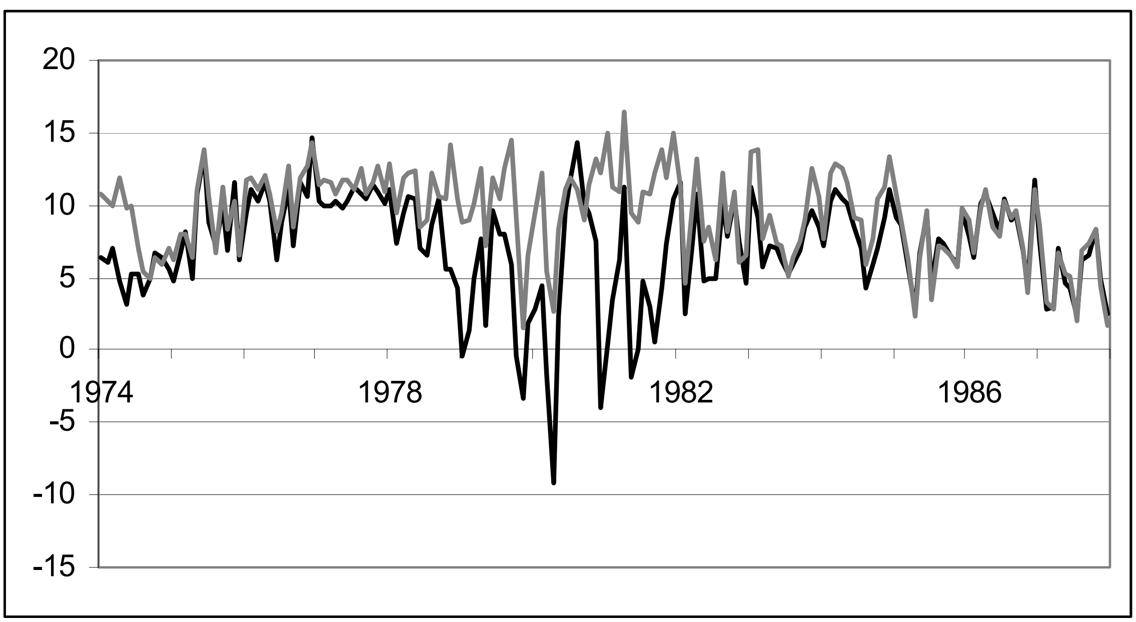
M2 Level of Aggregation



Notes: The black line is the log first difference of the Divisia index and the grey line is the log first difference of the simple sum index both multiplied by 1200. The growth rate of simple sum M2 exceeds 20% in the first two months of 1983, but the graph is cropped to preserve its comparability with Figure 8.

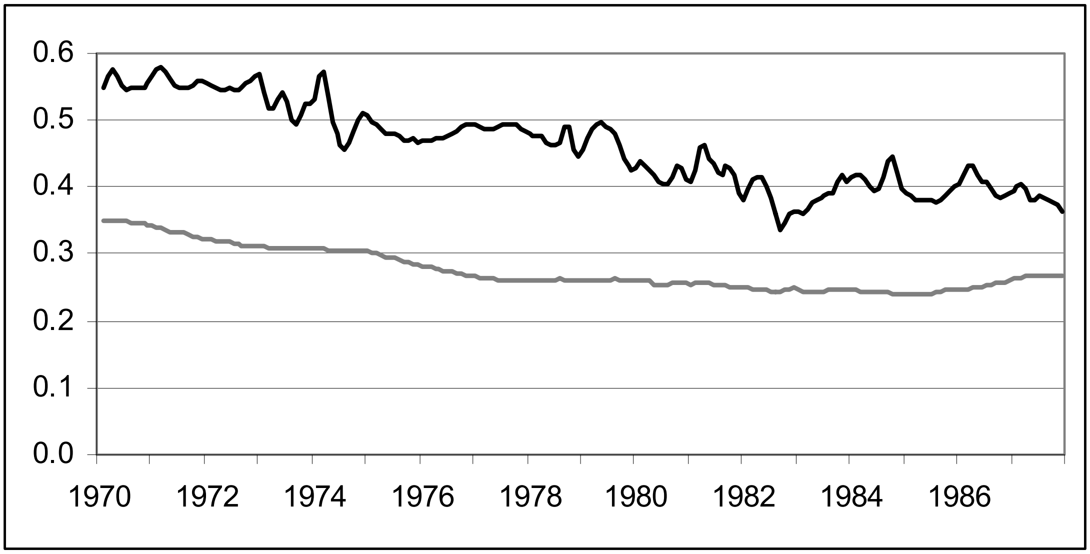
Figure 8: Growth Rates of Divisia and Simple Sum Indexes

M3 Level of Aggregation



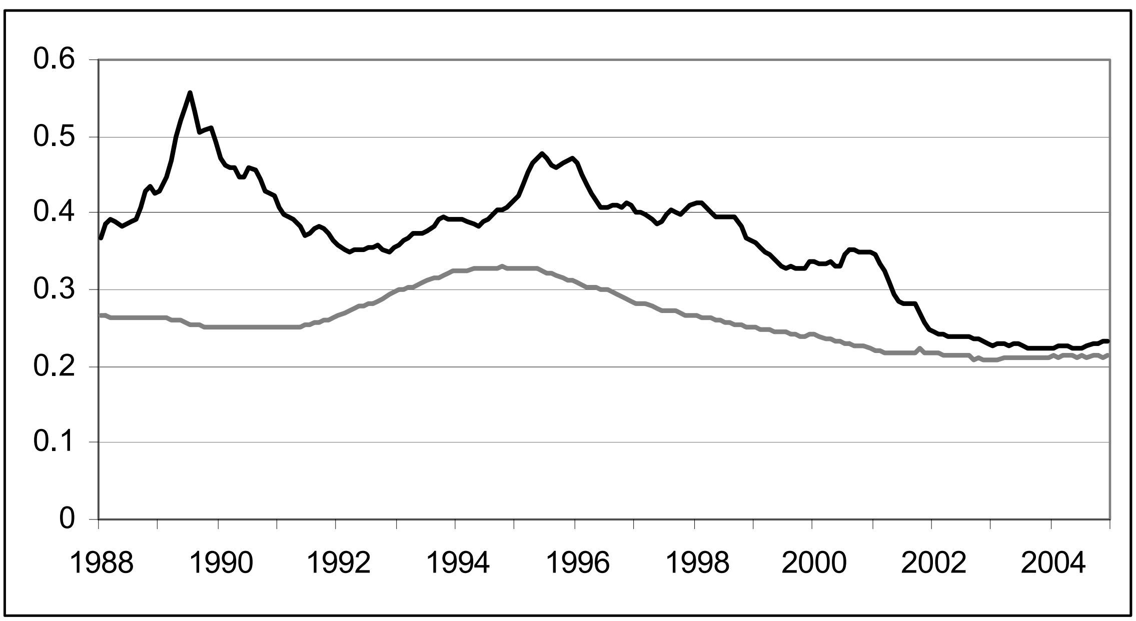
Notes: The black line is the log first difference of the Divisia index and the grey line is the log first difference of the simple sum index both multiplied by 1200.

Figure 9: Weights of M1 Monetary Assets in the Growth Rates of Divisia and Simple Sum M2



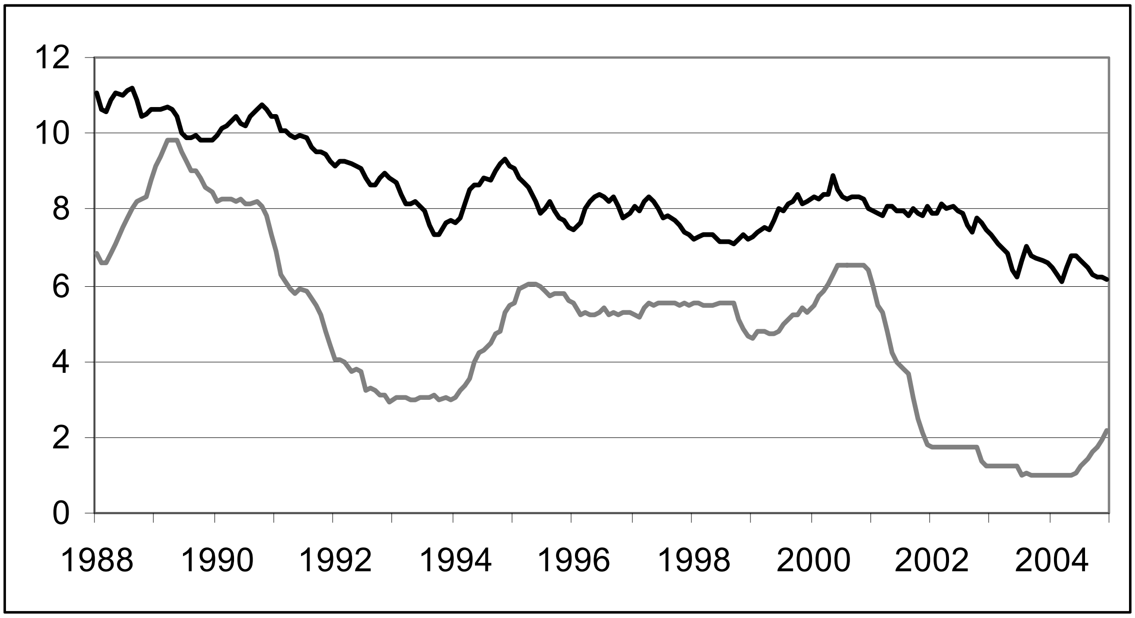
Notes: The black line is total expenditure on monetary assets in M1 divided by total expenditure on monetary assets in M2. The grey line is the ratio of simple sum M1 to simple sum M2. DD is assumed to be non-interest bearing in these calculations.

Figure 10: Weights of M1 Monetary Assets in the Growth Rates of Divisia and Simple Sum M2



Notes: The black line is total expenditure on monetary assets in M1 divided by total expenditure on monetary assets in M2. The grey line is the ratio of simple sum M1 to simple sum M2. DD is assumed to be non-interest bearing in these calculations.

Figure 11: Benchmark Rate versus Federal Funds Rate



Notes: The black line is the benchmark rate used to compute the Divisia indexes and the grey line is the Federal Funds rate.

Endnotes

[1] Contact Information: Department of Economics, State University of New York at Binghamton, P.O. Box 6000, Binghamton NY, 13902-6000. Email: bjones@binghamton.edu
[2] I would like to thank William Barnett, Per Gunnar Berglund, Jane Binner, Thomas Elger, Chris Hanes, and Utz-Peter Reich for helpful comments or discussions. I also thank Richard Anderson, Sharon Van Stratton, and Marisa de Natalia for help with data. I acknowledge financial support for this project from the Jan Wallander and Tom Hedelius Foundation (J03/19). All errors are my responsibility.
[3] Stracca (2001, 2004) constructed a "synthetic Divisia index" for the Euro area; See also Fase and Winder (1996), Wesche (1997), and Drake, Mullineaux and Agung (1997). Barnett (2003) has recently developed a heterogeneous agent theory, which could be used to construct Euro area Divisia indexes. See also Belongia and Binner (2000) for Divisia indexes for the core EMU as well as for a number of individual countries including: Australia, Canada Germany, Japan, and Korea.
[4] See Barnett (1987, 1995), Barnett, Jensen, and Liu (1997), and Barnett and Wu (2005) for generalizations.
[5] See Green (1964), Blackorby, Primont, and Russel (1978), Deaton and Muelbauer (1980), Pollack and Wales (1992), Barnett and Serletis (2000), and Barnett and Binner (2004).
[6] If a sub-group of the monetary assets are weakly separable from consumption and all other monetary assets in u, then an equivalent statement could be made for the weakly separable sub-group of monetary assets.
[7] For many of these purposes, quasi-homotheticity is sufficient. In general, however, the theoretical results are considerably weaker if the sub-utility function is not homothetic. In particular, two-stage budgeting is not generally valid. In the non-homothetic case, the theoretical quantity and price aggregates are defined via the dual distance and expenditure functions in the non-homothetic case: $Q_t = d(\mathbf{m}_t^*, \tilde{v})$ and $\Pi_t = e(\boldsymbol{\pi}_t, \tilde{v})$, where $d(\mathbf{m}, \tilde{v}) = \min_{\boldsymbol{\pi}} \{\boldsymbol{\pi} \cdot \mathbf{m} : e(\boldsymbol{\pi}, \tilde{v}) = 1\}$. These aggregates are linearly homogeneous in quantities and user costs respectively, but they do not satisfy factor reversal in general; See Deaton and Muelbauer (1980). The quantity and price aggregates both depend on the reference utility level, $\tilde{v}$, in the non-homothetic case; See Barnett (1987) for detailed analysis.
[8] On the contrary, many of the empirical studies that have tested US data for weak separability have supported relatively narrow monetary aggregates; See, for examples, Swofford and Whitney (1987) and Fisher and Fleissig (1997).
[9] If the utility function for the monetary assets is not homothetic, then Q_t / Q_s depends on the reference utility level, $\tilde{v}$. The Törnqvist quantity index can, nevertheless, be shown to be exactly correct (assuming utility maximization) if the distance function is Translog and the reference utility level is taken to be $\tilde{v} = \sqrt{\tilde{v}_t \tilde{v}_s}$, where $\tilde{v}_t = V(\mathbf{m}_t^*)$ and $\tilde{v}_s = V(\mathbf{m}_s^*)$; See Diewert (1976a, 1981). Similar results can also be obtained for other superlative quantity indexes as well as for corresponding superlative price indexes; See Diewert (1976a, 1976b, 1981).
[10] The group of goods may contain goods, services, monetary assets, and leisure, but I use the terms "goods" and "prices" for clarity of exposition.
[11] The revealed preference tests can also be implemented on cross-sectional data for consumers: *i.e.* where each observation represents the choice of a different consumer. Gross (1995) studies the impact of taste heterogeneity on the revealed preference tests and proposes a related goodness-of-fit measure when there are violations.
[12] It can similarly be shown that if $\mathbf{x}_i R^0 \mathbf{x}_j$, then $f_Q^F(\mathbf{x}_i, \mathbf{x}_j, \mathbf{p}_i, \mathbf{p}_j) > 1$ provided that the two bundles do not violate the weak axiom of revealed preference (WARP).
[13] Crawford (2004) derives conditions for latent separability; See Blundell and Robin (2000).
[14] See Fleissig and Whitney (2003) or Jones, Dutkowsky, and Elger (2005) for details about writing the objective function and constraints in the standard linear programming form.
[15] The objective function for the test is somewhat arbitrary; See Jones, Elger, Edgerton, and Dutkowsky (2005) for an alternative objective function. If a feasible solution can be found for the constraints, but G is not able to be minimized to zero, then (19) and (22) are not equivalent. Swofford and Whitney (1994)

interpreted such an outcome in terms of weak separability with incomplete or partial adjustment of expenditure on the weakly separable goods.

[16] Barnett and Choi (1989) discuss the appropriate parameter restrictions for a set of prices to be separable in the Translog and Generalized Leontief indirect utility functions. In their discussion of the Translog model, they point out that separability of a set of prices in the indirect utility function is not equivalent to separability of the corresponding quantities in the direct utility function unless the direct utility function is homothetic. See Serletis (1987) for discussion of quasi-homotheticity in this context.

[17] See Fleissig and Whitney (2005) and de Peretti (2005) for some alternative approaches.

[18] I use the quarterly average of a smoothed population series provided by Marisa de Natalia (BLS), which is available from 1990:1-2003:12.

[19] In this section, I use the term "Divisia index" to mean the superlative Törnqvist quantity index.

[20] See also Barnett, Offenbacher, and Spindt (1984), Serletis (1988), Fisher and Serletis (1989), Belongia and Chrystal (1991), Barnett, Fisher, and Serletis (1992), Drake (1992), Serletis and King (1993), Belongia (1996), Ford and Morris (1996), Schunk (2001), Stracca (2001, 2004), Binner and Wattam (2003), Binner, Bissoondeeal, Elger, Gazely, and Mullineux (2005), Drake and Mills (2005), and Belongia and Binner (2005).

[21] I do not show the corresponding figures for the M2M indexes, because they experienced very high growth rates in early 1983 due to the introduction of Super NOW accounts and MMDA.

[22] He was using data for the period from January 1979 to May 1983.

[23] Over this sample period, the benchmark rate is always equal to the BAA bond rate.

[24] This analysis only applies to the spread between the expenditure and simple sum weights not to the values of the weights themselves.

[25] For consistency, I also assume that DD are non-interest bearing, since the implicit rate of return on demand deposits used in the published Divisia indexes is based on a long term bond rate (5 year government bonds). I also calculated Divisia indexes under this assumption using the standard benchmark rate to determine whether or not it was empirically significant. I found that the growth rates of the resulting indexes were very highly correlated with the growth rates of the published indexes.