Macroeconomic Dynamics, **22**, 2018, 1510–1534. Printed in the United States of America.
doi:10.1017/S1365100516000791

TESTING FOR WEAK SEPARABILITY USING STOCHASTIC SEMI-NONPARAMETRIC TESTS: AN EMPIRICAL STUDY ON US DATA

Ryan S. Mattson
West Texas A&M

Philippe de Peretti
Université Paris1 Panthéon-Sorbonne

In this paper, we use the weak separability criterion to check for the existence of six different monetary aggregates reported by the Center of Financial Stability (CFS). We implement an extended version of the semi-nonparametric tests introduced by Barnett and de Peretti on US monthly data from January 1967 to December 2012. The test, first, checks for the necessary existence conditions of an overall utility function and a monetary subutility function, and then tests for the separability of the latter. On different subsamples, our results suggest that only the DM1 aggregate meets the separability criterion. Implemented on macroeconomic data, we have tested a joint assumption about separability and the existence of a representative agent. Thus, the rejection of the null could also be due to the rejection of stringent Gorman's conditions. More advanced tests for weak separability are clearly required to confirm the results found in this paper.

Keywords: Weak Separability Criterion, Semi-Nonparametric Tests, CFS Data

1. INTRODUCTION

The goal of this paper is to use the weak separability criterion to check for the existence of six different monetary aggregates routinely computed by the Center of Financial Stability (CFS), namely DM1, DM2M, DM2 ALL, DM3, DM4-, and DM4. As emphasized by Barnett (1980), Deaton and Muellbauer (1980), and Blackorby et al. (1998), weak separability is a key condition for an aggregate to exist, ensuring that all substitution effects are internalized. It also returns important information about the structure of preferences, and natural clustering over goods and assets.[1] Empirical applications are found in Swofford and Whitney (1987,

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7-SSH/2007-2013) under grant agreement no. 320270 SYRTO. This work was achieved through the Laboratory of Excellence on Financial Regulation (Labex ReFi) supported by PRES heSam under the reference ANR10LABX0095. It benefited from French government support managed by the National Research Agency (ANR) within the project Investissements d'Avenir Paris Nouveaux Mondes (investments for the future Paris New Worlds) under the reference ANR11IDEX000602. Address correspondence to: Philippe de Peretti, Centre d'Economie de la Sorbonne (CES), Université Paris1 Panthéon-Sorbonne, 106-112 Boulevard de l'Hôpital, 75013 Paris, France; e-mail: philippe.de-peretti@univ-paris1.fr.

© 2017 Cambridge University Press 1365-1005/17

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

1988), Drake (1997), or Fisher and Fleissig (1997), among others. Generally, results support the separability of, at least, a narrow aggregate of money.

To avoid some drawbacks associated with separability tests in flexible parametric forms, i.e., the joint test about a parametric form and weak separability and the destruction of flexibility when imposing restrictions [see Blackorby et al. (1977) or Denny and Fuss (1977)], most studies have implemented nonparametric tests. These tests are based on Varian (1982, 1983), exploiting Afriat (1967, 1973) results within Samuelson's revealed preference framework. The Varian test for separability is sequential in that it tests different groups of prices and quantities for utility maximization by using the well-known Generalized Axiom of Revealed Preference (GARP). Nevertheless, as shown by Barnett and Choi (1989), the test is strongly biased toward rejection. Moreover, GARP is nonstochastic and the Varian empirical test is based on an only sufficient condition. Thus, results based on this approach are often inconclusive. The test is extended in several directions by Swofford and Whitney (1994), Fleissig and Whitney (2003, 2005), Elger and Jones (2008), Cherchye et al. (2015), and Barnett and de Peretti (2009), based upon de Peretti (2007). Varian (1985, 1990) and Hjertstrand and Swofford (2014) can be referenced for extensions of GARP in stochastic environments.

Concerning Fleissig and Whitney (2005) and Elger and Jones (2008), they introduce closely related frameworks that allow us to test for weak separability when observed quantities data are contaminated by measurement error. Whereas their approach is appealing from a theoretical point of view, empirical implementations are not that straightforward, since it requires the knowledge of the variance of the error terms, which is generally unknown. Moreover, as noticed in Jones and de Peretti (2005) and Elger and Jones (2008), the computed measurement is, by far, not an estimate of the "true" measurement error. In particular, the computed one is likely to be dramatically lower compared to the true one. As a result, their test may suffer from a huge distortion.

For these reasons, to test for weak separability, in this paper, we will use the Barnett and de Peretti (2009) necessary and sufficient semi-nonparametric procedure and, in particular, a slightly extended version. The Barnett and de Peretti (2009) test is a two-step procedure. First, the necessary existence conditions for both an overall utility function and a (monetary) subutility function are tested under the assumption that quantity data are contaminated with noise (Step 1). Next, if the data pass this step, the (monetary) subutility is tested for separability within the overall utility function (Step 2) using the Leontief–Sono theorem. The latter step exploits the theoretical independence between the marginal rate of substitution between two assets of the cluster and assets/goods outside the cluster. Compared to Barnett and de Peretti (2009), in Step 1, we use an adjustment procedure having the same structure, i.e., we solve, iteratively and sequentially, a quadratic program under sets of transitive constraints. The main difference is that we use a more complex set of constraints, allowing for a return of a smaller adjustment. In Step 2, we use a different test of independence designed for time series.

Two main conclusions are drawn regarding results and methodology. Considering the results, on five different samples, spanning January 1967–December

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

2012, and among the six aggregates tested, only the narrow DM1 meets the weak separability criterion on most periods, and no other aggregate is found to be separable. Rejection of weak separability is either caused by our inability to find monetary subutility functions rationalizing the data (Step 1), especially on periods covering June 1974–November 1982 and September 1991–February 2008, or caused by the rejection of the independence between the marginal rate of substitution between assets of the group and quantities/assets outside the group (Step 2). Regarding the methodology, one should keep in mind that the null is based on a composite assumption about maximization/separability and the existence of a representative agent. Hence, when rejecting, it could be the case that what we reject are in fact stringent Gorman's conditions for aggregating over agents [Gorman (1961)]. In other words, rejecting the null could indicate heterogeneity across agents, which is not taken into account by our test. This is a limitation of our work, which clearly calls for more research in this field. Our results should therefore be treated with caution, and our paper clearly calls for more advanced studies to confirm our main findings.

This paper is structured as follows. In Section 2, we introduce the extended Barnett and de Peretti (2009) test and discuss extensions. In Section 3, we present the data and implement the tests. Finally, in Section 4, we conclude and discuss our results, point out the limitation of our work, and suggest possible extensions.

2. SEMI-NONPARAMETRIC TESTS OF UTILITY MAXIMIZATION AND WEAK SEPARABILITY

In this section, we first introduce the weak separability concept, and then an extended version of the Barnett and de Peretti (2009) testing procedure. At the end of the section, differences with the original test are presented. For the purpose of clarity, all computational details concerning the adjustment procedure are given in the appendix.

Let $\mathbf{X}^*$ be a $(T \times k)$ matrix of real quantities, and let $\mathbf{x}_i^* = (x_{i1}^*, x_{i2}^*, \ldots, x_{ik}^*)'$ be the ith row of the matrix, $i \in (1, T)$. Define $\mathbf{P}$ as a $(T \times k)$ matrix of corresponding prices, wherein, similarly, $\mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{ik})'$ is the ith row of the matrix, $i \in (1, T)$. Now let the $(T \times a)$ $\mathbf{X}^{*(1)}$ matrix be a partition of $\mathbf{X}^*$, $a \in (1, k-1)$, with $\mathbf{x}_i^{*(1)} = (x_{i1}^*, x_{i2}^*, \ldots, x_{ia}^*)'$, and let the $[T \times (k-a)]$ $\mathbf{X}^{*(2)}$ matrix with $\mathbf{x}_i^{*(2)} = (x_{i(a+1)}^*, x_{i(a+2)}^*, \ldots, x_{ik}^*)'$ be the complementary one, with $\mathbf{X}^{*(1)} \cup \mathbf{X}^{*(2)} = \mathbf{X}^*$ and $\mathbf{X}^{*(1)} \cap \mathbf{X}^{*(2)} = \varnothing$. $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ are the corresponding prices with $\mathbf{p}_i^{(1)} = (p_{i1}, p_{i2}, \ldots, p_{ia})'$ and $\mathbf{p}_i^{(2)} = (p_{i(a+1)}, p_{i(a+2)}, \ldots, p_{ik})'$. Define weak separability over $\mathbf{X}^{(1)}$ as follows.

DEFINITION 1. *There is weak separability if (i) there exists a utility function (1) rationalizing the data and (ii) the latter admit a rewriting (2):*

$$U_i = U(\mathbf{x}_i^*),\ i = 1, 2, \ldots, T \tag{1}$$

$$U_i = V\left[\mathbf{x}_i^{*(2)}, f(\mathbf{x}_i^{*(1)})\right],\ i = 1, 2, \ldots, T, \tag{2}$$

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

where $U()$ is the overall twice differentiable utility function, $V()$ is a strictly increasing macrofunction, and $f()$ is the twice differentiable subutility function, also known as the microfunction.

To test for this particular preference structure, Barnett and de Peretti (2009) have introduced a necessary and sufficient semi-nonparametric test. Their procedure is a two-step one that amounts to (i) jointly testing for the existence of $U()$ and $f()$ (Step 1) and (ii) testing for the separability of $f()$ (Step 2).

2.1. Step 1: Jointly Testing for the Existence of $U()$ and $f()$

To test for compliance with utility maximization for $(\mathbf{x}_i^*, \mathbf{p}_i)_{i=1}^T$ and $(\mathbf{x}_i^{*(1)}, \mathbf{p}_i^{(1)})_{i=1}^T$, the GARP, introduced by Varian (1982), is used. Introducing GARP, first define the following three binary relations: $\mathbf{x}_i^*$ is said to be strictly directly revealed preferred to $\mathbf{x}_j^*$ if $\mathbf{p}_i \cdot \mathbf{x}_i^* > \mathbf{p}_i \cdot \mathbf{x}_j^*$, written as $\mathbf{x}_i^* P^0 \mathbf{x}_j^*$; $\mathbf{x}_i^*$ is said to be directly revealed preferred to $\mathbf{x}_j^*$ if $\mathbf{p}_i \cdot \mathbf{x}_i^* \geq \mathbf{p}_i \cdot \mathbf{x}_j^*$, written as $\mathbf{x}_i^* R^0 \mathbf{x}_j^*$; $\mathbf{x}_i^*$ is said to be revealed preferred to $\mathbf{x}_j^*$ if $\mathbf{x}_i^* R^0 \mathbf{x}_m^*, \mathbf{x}_m^* R^0 \mathbf{x}_k^*, \ldots, \mathbf{x}_p^* R^0 \mathbf{x}_j^*$, written as $\mathbf{x}_i^* R \mathbf{x}_j^*$, where R is the transitive closure of R^0. Let r_{ij} be an element of the transitive closure matrix (ith row, jth column). GARP is then defined as follows.

DEFINITION 2 (GARP). *For a couple of observations $(i, j) \in (1, T), i \neq j$, $\mathbf{x}_i^* R \mathbf{x}_j^* \Rightarrow \mathbf{p}_j \cdot \mathbf{x}_j^* \leq \mathbf{p}_j \cdot \mathbf{x}_i^*$.*

Using GARP, Varian (1982) proved the following theorem.

THEOREM 1 (Varian (1982), shortened). *For a set $(\mathbf{x}_i^*, \mathbf{p}_i)_{i=1}^T$, the two following. conditions are equivalent:*

(*i*) *There exists a locally nonsatiated utility function $U()$that rationalizes the data.*
(*ii*) *The data satisfy GARP, i.e., $\forall (i, j) \in (1, T), i \neq j : \mathbf{x}_i^* R \mathbf{x}_j^*$ implies **not** $\mathbf{p}_j \cdot \mathbf{x}_j^* > \mathbf{p}_j \cdot \mathbf{x}_i^*$.*

Barnett and de Peretti (2009) consider jointly testing for $U()$ and $f()$ under the additional assumption that quantity data are contaminated with an additive noise. Under the null, they assume that data are generated by a weakly separable utility but are not observable. Only $\mathbf{X}$ is observed, related to $\mathbf{X}^*$ by the additive relation

$$x_{ij} = x_{ij}^* + \varepsilon_{ij}, \tag{3}$$

where ε_{ij} is iid with zero mean and variance $\sigma_{\varepsilon_j}^2$, with distribution function $F_j(x)$. The distribution $F_j(x)$ is max and min-stable, $j = 1, \ldots, k$.
In (3), ε_{ij} stands for a measurement or a small optimization error. Then, given (3), the test for Step 1 is achieved by solving over z_{ij} the quadratic program

$$\text{obj} = \min_{z_{ij}} \sum_{i=1}^{T} \sum_{j=1}^{k} \left(x_{ij} - z_{ij}\right)^2 \tag{4}$$

subject to

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

$$\left| \; \forall (i,j) \in (1,T) : \mathbf{z}_i R \mathbf{z}_j \Rightarrow \mathbf{p}_j \cdot \mathbf{z}_j \leq \mathbf{p}_j \cdot \mathbf{z}_i, \right. \tag{C.1}$$

$$\left| \; \forall (i,j) \in (1,T) : \mathbf{z}_i^{(1)} R \mathbf{z}_j^{(1)} \Rightarrow \mathbf{p}_j^{(1)} \cdot \mathbf{z}_j^{(1)} \leq \mathbf{p}_j^{(1)} \cdot \mathbf{z}_i^{(1)}, \right. \tag{C.2}$$

where $\mathbf{z}_i^{(1)} = (z_{i1}, z_{i2}, \ldots, z_{ia})'$ and $\mathbf{p}_i^{(1)} = (p_{i1}, p_{i2}, \ldots, p_{ia})'$.

If in (4) obj $= 0$, then accept the null of utility maximization[2] for $(\mathbf{x}_i^*, \mathbf{p}_i)_{i=1}^T$ *and* $(\mathbf{x}_i^{*(1)}, \mathbf{p}_i^{(1)})_{i=1}^T$. If obj $\neq 0$, then data violate GARP when testing for $U()$ or $f()$ or for both. If violations of GARP appear, then (4) will return a matrix $\widehat{\mathbf{Z}}$ satisfying the two utility maximization programs such that the distance with the observed quantities is minimal. Define $\widehat{\mathbf{E}} = (\mathbf{X} - \widehat{\mathbf{Z}})$ as the matrix of theoretical adjustments.

To test for the null that the adjustments are nonsignificant, i.e., violations of GARP being caused by purely stochastic causes, use a test tracking excess adjustments in some goods. For this purpose, compare for each $j = 1, \ldots, k$ the theoretical extremal adjustments to some estimates of the quantiles of the "true" extremal measurement errors in the data, computed by simulations methods. Then, accept the null if for all $j = 1, \ldots, k$, the theoretical extremal adjustments are less than the corresponding quantiles.

To compute these quantiles, first define the state-space model:

$$x_{ij} = A\alpha_{ij} + \varepsilon_{ij}, \tag{5}$$

$$\alpha_{(i+1)j} = M\alpha_{(i+1)j} + Q_j, \tag{6}$$

where

$$A = \begin{bmatrix} 1 & 0 \end{bmatrix}, \; M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \; Q_j = \begin{bmatrix} \sigma_{\xi_j}^2 & 0 \\ 0 & \sigma_{\zeta_j}^2 \end{bmatrix},$$

$$\alpha_{ij} = \begin{bmatrix} x_{ij}^* \\ v_{ij} \end{bmatrix}, \alpha_{1j} \sim N(a_{1j}, P_{1j}).$$

The above model is the local linear trend model [see, e.g., Durbin and Koopman (2001)]. In the measurement equation (5), ε_{ij} is the measurement error for good j in period i with variance $\sigma_{\varepsilon_j}^2$ and x_{ij}^* is the unobserved quantity, treated as a state variable, in the transition equations (6).

Then, use the so-called simulation smoother [de Jong and Shephard (1995)]. In the latter, smoothed errors are generated conditional on the observed series $(x_{1j}, x_{2j}, \ldots, x_{Tj})$, which makes sense regarding the null assumption. Following Durbin and Koopman (2001), the recursions for the Kalman filter are given by the following equations:

$$\left| \begin{array}{ll} v_{ij} = x_{ij} - Aa_{ij} & F_i = AP_iA_i' + \widehat{\sigma}_{\varepsilon_j}^2 \\ K_i = MP_iA_i'F_i^{-1} & L_i = M - K_iA_i \\ a_{(i+1)j} = Ma_{ij} + K_iv_{ij} & P_{(i+1)j} = MP_iL_i' + \widehat{Q}_j, \end{array} \right. , i = 1, 2, \ldots, T \tag{7}$$

for some (diffuse) maximum likelihood estimates of $(\widehat{\sigma}_{\varepsilon_j}^2, \widehat{\sigma}_{\xi_j}^2, \widehat{\sigma}_{\zeta_j}^2)$.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

A sample of smoothed errors $(\widetilde{\varepsilon}_{ij})_{i=1}^{T}$ can then be computed by the following set of equations, solved recursively:

$$\left|\begin{array}{l} \widetilde{r}_{i-1} = A' F_i^{-1} v_{ij} - \widetilde{W}_i C_i^{-1} d_i + L_i' \widetilde{r}_i \\ \widetilde{N}_{i-1} = A' F_i^{-1} A + \widetilde{W}_i' C_i^{-1} \widetilde{W}_i + L_i' \widetilde{N}_i L_i \\ \widetilde{W}_i = \widehat{\sigma}_{\varepsilon_j}^2 (F_i^{-1} A - K_i' \widetilde{N}_i L_i) \\ d_i \sim N(0, C_i) \\ C_i = \widehat{\sigma}_{\varepsilon_j}^2 - \widehat{\sigma}_{\varepsilon_j}^2 \widetilde{D}_i \widehat{\sigma}_{\varepsilon_j}^2 \\ \widetilde{D}_i = F_i^{-1} + K_i' \widetilde{N}_i K_i \\ \widetilde{\varepsilon}_{ij} = d_i + \widehat{\sigma}_{\varepsilon_j}^2 (F_i^{-1} v_{ij} - K_i' \widetilde{r}_i). \end{array}\right. , \left|\begin{array}{l} i = T, \ldots, 1 \\ \widetilde{r}_T = 0 \\ \widetilde{N}_T = 0 \end{array}\right. \quad (8)$$

Run (8) for each good $j = 1, \ldots, k$ a large number of times, and store at each replication the extremal adjustments $\widetilde{\min}^j = \min(\widetilde{\varepsilon}_{1j}, \widetilde{\varepsilon}_{2j}, \ldots, \widetilde{\varepsilon}_{Tj})'$ and $\widetilde{\max}^j = \max(\widetilde{\varepsilon}_{1j}, \widetilde{\varepsilon}_{2j}, \ldots, \widetilde{\varepsilon}_{Tj})'$. Build the distributions for the two extremes and compute for each j the quantiles at 90%, 95%, and 99%, $Q_{90}^{\min^j}$, $Q_{95}^{\min^j}$, $Q_{99}^{\min^j}$ and $Q_{90}^{\max^j}$, $Q_{95}^{\max^j}$, $Q_{99}^{\max^j}$, respectively. Extremal theoretical adjustments, $\widehat{\min}^j = \min(\widehat{e}_{1j}, \widehat{e}_{2j}, \ldots, \widehat{e}_{Tj})$ and $\widehat{\max}^j = \max(\widehat{e}_{1j}, \widehat{e}_{2j}, \ldots, \widehat{e}_{Tj})$, and extremal measurement errors can now be compared at a given threshold. Given the set of assumptions:

$$\left\{\begin{array}{l} H_0 : (\mathbf{x}_i^*, \mathbf{p}_i)_{i=1}^T \text{ and } (\mathbf{x}_i^{*(1)}, \mathbf{p}_i^{(1)})_{i=1}^T \text{ are rationalized by utility functions,} \\ H_A : (\mathbf{x}_i^*, \mathbf{p}_i)_{i=1}^T \text{ or/and } (\mathbf{x}_i^{*(1)}, \mathbf{p}_i^{(1)})_{i=1}^T \text{ is (are) not rationalized} \\ \quad \text{by utility function(s).} \end{array}\right.$$

The decision rule at 5% is therefore

$$\left\{\begin{array}{l} \text{Accept } H_0 \text{ if: } \left\{\begin{array}{l} \text{obj} = 0, \text{ or,} \\ \widehat{\min}^j \geq Q_{95}^{\min^j} \text{ and } \widehat{\max}^j \leq Q_{95}^{\max^j}, \forall j \in (1, k) \end{array}\right. \\ \text{Reject } H_0 \text{ if } \exists j \in (1, k) : \widehat{\min}^j < Q_{95}^{\min^j} \text{and/or } \widehat{\max}^j > Q_{95}^{\max^j}. \end{array}\right.$$

We next turn to separability testing.

2.2. Step 2: Testing for Weak Separability

Following Deaton and Muellbauer (1980), weak separability implies that the marginal rate of substitution between any two goods of $\mathbf{X}^{(1)}$ (j, l) is independent of the goods (m) outside the group, ensuring that all substitutions effects are internalized, so an aggregate can exist:

$$\partial \left[\frac{\dfrac{\partial U(\mathbf{x}_i)}{\partial x_{ij}}}{\dfrac{\partial U(\mathbf{x}_i)}{\partial x_{il}}} \right] \Big/ \partial x_{im} = 0. \quad (9)$$

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

Equation (9), also known as the Leontief–Sono theorem, is a necessary and sufficient condition for weak separability. [See also Blackorby et al. (1998).] By the Konyus lemma, if the data satisfy GARP for the subutility, then

$$\frac{\mathbf{p}_i^{(1)'}}{\mathbf{p}_i^{(1)} \cdot \mathbf{x}_i^{(1)}} = \frac{\nabla f(\mathbf{x}_i^{(1)})}{\mathbf{x}_i^{(1)} \cdot \nabla f(\mathbf{x}_i^{(1)})}, i = 1, \ldots, T, \tag{10}$$

where $\nabla f(\mathbf{x}_i^{(1)})$ is the $(a \times 1)$ gradient vector evaluated at $\mathbf{x}_i^{(1)}$.

Dividing the jth row of the system (10) by the lth row, $j, l = 1, \ldots, a$, $j \neq l$, returns the well-known condition

$$\frac{p_{ij}^{(1)}}{p_{il}^{(1)}} = \frac{\dfrac{\partial f(\mathbf{x}_i^{(1)})}{\partial x_{ij}^{(1)}}}{\dfrac{\partial f(\mathbf{x}_i^{(1)})}{\partial x_{il}^{(1)}}}. \tag{11}$$

Testing for weak separability thus amounts to testing for the independence between uniquely defined price ratios of goods inside the group (in log-form) and (log) quantities outside the group. Note that (9) is a necessary and sufficient condition for $f()$ to be weakly separable. This is in contrast with the empirical implementation of Varian's test, in which to test for separability, solutions of the so-called Afriat inequalities are used. Since solutions are not unique, so when rejecting it is not clear if separability could be accepted for other sets of solutions. To perform this task, we use the multivariate independence test introduced by El Himdi and Roy (1997) for time series. This test amounts to checking for noncorrelation between the innovations of two multivariate series, computed as residuals of two independent Vector Auto Regressive (VAR) processes. Let $\mathbf{ps}_i = \log(p_{i1}^{(1)}/p_{i2}^{(1)}, p_{i1}^{(1)}/p_{i3}^{(1)}, \ldots, p_{i(a-1)}^{(1)}/p_{ia}^{(1)})'$ be the vector of unique price ratios of the separable group. Then, the two VAR representations are given by

$$\mathbf{b}_i^{(1)} = \mathbf{ps}_i' - \sum_{r=1}^{p_1} \Phi_r \mathbf{ps}_{i-r}', \tag{12}$$

$$\mathbf{b}_i^{(2)} = \log(\mathbf{x}_i^{(2)})' - \sum_{r=1}^{p_2} \Psi_r \log(\mathbf{x}_{i-r}^{(2)})'. \tag{13}$$

Define $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ as the correlation matrices for $(\mathbf{b}_i^{(1)})_{i=r+1}^T$ and $(\mathbf{b}_i^{(2)})_{i=r+1}^T$, and $\mathbf{R}_{12}$ as the cross-correlations. Then, the test statistic for noncorrelation is given by

$$Qh = T\,vec(\mathbf{R}_{12})'(\mathbf{R}_{11} \otimes \mathbf{R}_{22})^{-1} vec(\mathbf{R}_{12}). \tag{14}$$

Under the null, Qh is distributed as a chi-square with $d_1 d_2$ degrees of freedom, with d_1 and d_2 being, respectively, the dimensions of $\mathbf{ps}_i$ and $\mathbf{x}_i^{(2)}$. Note that the

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

test can deal with stationary and nonstationary series (possibly cointegrated) and is robust to additive outliers.

2.3. Differences with Barnett and de Peretti (2009)

There are three differences with regard to the original work of Barnett and de Peretti (2009). First, as shown in the appendix, program (4) is still sequentially and iteratively solved, but other sets of constraints are used. Specifically, in Barnett and de Peretti (2009), when adjusting quantities to rebuild a coherent preference chain, the budget for an observation i violating GARP was forced to remain unchanged, ensuring that z_i will not cause new violations with bundles located above in the preference chain. Here, we do not force the budget in i to remain unchanged but add an alternative set of constraints. On the one hand, the program is more complex, but, on the other hand, it returns smaller adjustments [see constraint (C.1.1) in (19)]. Second, we use a more general model (5) to extract measurement error, which is more relevant for trended data. Note that if $\sigma^2_{\zeta_j} = 0$, then we have the filtering process used in Barnett and de Peretti (2009). Finally, to test for separability, instead of using restriction tests, a noncorrelation test, which is relatively easier to implement and assumed to have more power, is used.

3. IMPLEMENTING THE TESTS

In this section, we first introduce our data. Next, since weak separability is a particular rewriting of the overall utility function, we first test the data for compliance with the overall maximization program, for a utility grouping goods, leisure, and money. Two tests are used, the standard nonstochastic GARP test, as given in Definition 2, and, if violations appear, we solve (4) with only the (C.1) constraint, i.e., ignoring the (C.2) constraint corresponding to the subutility maximization program. To test the significance of the theoretical adjustment, we compare extremal theoretical adjustments to the quantiles of the extremal measurement errors estimated by simulations methods, as explained in Section 2. Significance tests are given at the 5% threshold. Then, over each period for which the overall maximization program is not rejected, we implement the Barnett and de Peretti (2009) procedure to test for weak separability. First, we jointly test for an overall utility and a subutility function. Here, the adjustment procedure takes into account constraints (C.1) and (C.2), so data are adjusted to be compliant with both the subutility and the overall utility. We call this adjustment the overall adjustment. If the adjustment is nonsignificant, then we turn to the separability test.

3.1. Data Description

The test is implemented on monthly per capita data for consumption, leisure, and money in the United States from January 1967 to December 2012. Table 1 presents

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 1. Goods and assets entering the analysis

	Consumption goods
DUR	Durable Goods
NDUR	Nondurable Goods
SER	Services
	Leisure
LEIS	Leisure
	Monetary assets
CP	Commercial Paper
CURTC	Currency and Travelers Checks
DD	Demand Deposits at Commercial Institutions
IMMF	Institutional Money market Funds
LTD	Large Time Deposits
MMDA	Money Market Demand Accounts at Thrift and Commercial Institutions
OCD	Other Checkable Deposits at Thrift and Commercial Institutions
SAV	Savings Accounts at Thrift and Commercial Institutions
SAVM	Savings and Money Market Demand Accounts at Thrift and Commerical Institutions
STD	Small Time Deposits at Thrift and Commercial Institutions
REPO	Overnight and Term Repurchase Agreements
RMF	Retail Money Market Funds
TB	Short Term (3-Month) Treasury Bills

all goods and assets entering the econometric analysis. Data for consumption goods and services are taken from the Federal Reserve Bank of Saint Louis.

For durables, a stock is constructed from current expenditures. The associated price is a user cost or rental price. Following Diewert (1974) and Patterson (1991), the stock is computed as

$$S_i = I_i + (1-\delta)S_{i-1}, i = 2, \ldots, T, S_1 = I_1, \tag{15}$$

where S_i is the stock in period i, I_i in the expenditure in period i, and δ is a constant depreciation rate, set here to $\delta = 10\%$.

The rental price at time i is defined as

$$p_i = \frac{p_i^* \left[r_i^B + \delta(1+\Pi_i) - \Pi_i \right]}{\left(1 + r_i^B\right)}, \tag{16}$$

where p_i^* is the spot price for durable goods, δ is the depreciation rate, r_i^B is a benchmark rate, and Π_i is the period inflation rate for durable goods.

The measure used for leisure is nonmarket time. Following Swofford and Whitney (1987), nonmarket time is 98 hours less the average hours worked by nonsupervisory employees in the United States, multiplied by 4 to get monthly

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

observations. The price of leisure is a shadow cost as derived in Barnett (1979). The observed wage rate, in this case the average hourly earnings of nonsupervisory employees, does not take into account the effect of employment available in the economy on the price of leisure. For full employment, the shadow price of leisure would equal the wage rate; however, in times of higher unemployment, the price of leisure would not be the same as the wage rate. For some observed wage rate w_i and level of employment E_i, the shadow cost of leisure is defined as

$$\hat{w_i} = w_i E_i^{\gamma}, \tag{17}$$

where γ is a parameter, set here to $\gamma = 2.3$, according to Barnett (1979)

Monetary assets and user costs are taken from the data set provided on the Advances in Monetary and Financial Measures website, from the CFS. These data are collected from a variety of accessible resources described in detail in Barnett et al. (2013). Real quantities are the assets divided by a consumer price index, whereas nominal user cost is defined according to Barnett (1978):

$$\pi_{ij} = p_i^{\text{cpi}} \frac{R_i^B - R_{ij}}{R_i^B}, \tag{18}$$

where $R_i^B = 1 + r_i^B$, and r_i^B is the rate of a benchmark asset in period i, $R_{ij} = 1 + r_{ij}$, with r_{ij} being the asset j own rate in period i, and p_i^{cpi} is a consumer price index.

The benchmark rate is chosen as a short-term rate for bank loans to commercial and industrial customers. This rate approximates the highest possible return that a bank would provide, as the bank would not offer to pay more on interest than it loans out [Offenbacher and Shemesh (2011)]. Using this method avoids the use of an arbitrary addition of basis points to a highest rate of return chosen from the basket of monetary asset returns.[3]

Some components entered into the survey at later times, and others left the survey; for example, Money Market Demand Accounts (MMDA) became part of the survey in the early 1980s and were then placed into an overall measure of savings deposits and MMDA in the early 1990s. Also, paper currency and travelers checks are combined into one component "Currency and Travelers Checks" due to the consistent and precipitous drop in the use of travelers checks. The differing thrift and commercial bank series are also combined and an average of their rates of return is used to determine the nominal user cost. Savings, interest checking, MMDA, and small denomination time deposits for commercials and thrifts are therefore combined into series for savings, interest checking, MMDA, and small time deposits. For econometric analysis, the overall sample is then split into five periods, as presented in Table 2, based on the beginning and ending of the availability of these series of data. Some data series such as overnight repurchase agreements and commercial paper are spliced as necessary for major changes in data sources. For further details about the CFS data, see Barnett et al. (2013).

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 2. Goods and assets entering the analysis

	Periods				
	Period 1 January 1967–December 1969	Period 2 January 1970–May 1974	Period 3 June 1974–November 1982	Period 4 December 1982–August 1991	Period 5 September 1991–December 1992
	Consumption goods				
DUR	Yes	Yes	Yes	Yes	Yes
NDUR	Yes	Yes	Yes	Yes	Yes
SER	Yes	Yes	Yes	Yes	Yes
	Leisure				
LEIS	Yes	Yes	Yes	Yes	Yes
	Monetary assets				
CURTC	Yes	Yes	Yes	Yes	Yes
CP	Yes	Yes	Yes	Yes	Yes
DD	Yes	Yes	Yes	Yes	Yes
OCD	Yes	Yes	Yes	Yes	Yes
STD	Yes	Yes	Yes	Yes	Yes
LTD	Yes	Yes	Yes	Yes	Yes
TB	Yes	Yes	Yes	Yes	Yes
SAV	Yes	Yes	Yes	Yes	No
REPO	No	Yes	Yes	Yes	Yes
RMF	No	No	Yes	Yes	Yes
IMMF	No	No	Yes	Yes	Yes
MMDA	No	No	No	Yes	No
SAVM	No	No	No	No	Yes

3.2. Testing for Overall Utility Maximization

Table 3 presents the results of the analysis for the five subperiods. The upper part of Table 3 reports the number of GARP violations, whereas the lower part presents tests of the violations of GARP (if any). For the stochastic analysis, main entries are (i) the number of bundles altered so that the whole set complies with utility maximization, (ii) the overall adjustment (obj), i.e., the solution of the program (4) when only the (C.1) constraint is taken into account, (iii) the minimal and maximal adjustments [$\min(\text{obj}_i)$ and $\max(\text{obj}_i)$] observed and the corresponding bundle i, and (iv) the results of the significance test at 5%.

Focusing first on the deterministic GARP analysis, for periods 1 and 2, utility maximization is not rejected with a zero violation rate, but it is for periods 3, 4, and 5. Whereas it seems not severe for period 3 (2 violations), periods 4 and 5 exhibit more violations, respectively, 30 and 367. In order to extract some information about the distributions of violations, we use a graphical analysis for periods 4 and 5. Figure 1 presents four key graphs. The two upper graphs are scatter plots of all couples (i, j) such that $\mathbf{x}_i R \mathbf{x}_j$ and $\mathbf{x}_j P \mathbf{x}_i$. The two lower figures exhibit with how many observations an observation i violates GARP.

Two interesting features arise: (i) For both periods, violations of GARP are concentrated in the upper right quadrant $(i, j) \in (82, 105)$ (September 1989–August 1991) for period 4 and $(i, j) \in (161, 245)$ for period 5 (January 2005–December 2012). (ii) Concerning the two lower graphs, the problem with the violations of GARP for period 5 appears much more severe than for period 4, leading to doubt for utility maximization. For period 5, one possible explanation of the high violation rate could be related to the Global Financial Crisis and changes in the risk exposure, which are not accounted for in the user costs prices, see Barnett and Wu (2005). Our analysis is therefore challenged by the high violation rate during this crisis period.

Turning next to stochastic analysis, for period 3, only one bundle needs to be adjusted, with an overall (quadratic) adjustment of 0.1690 for bundle $i = 9$. As expected, at 5%, we fail to reject the null of utility maximization. A similar conclusion holds for period 4, where only seven bundles are altered for the data set to be compliant with utility maximization. For period 5, with a massive adjustment of 1128679.40, and 43 bundles altered, the null is strongly rejected.

Thus, as a conclusion, the overall utility maximization assumption holds for periods 1, 2, 3, and 4 but not for period 5. For the latter period, the large difference between the minimal and the maximal adjustments for some period i, as well as the standard error of obj_i (63870.55), indicates a high dispersion among the adjustments, suggesting that some violations may be caused by measurement errors and other not.

Since violations of GARP are linked by transitive relationships, finding subsamples for which GARP is not violated or where violations are nonsignificant is no easy task. Using a grid-search algorithm, we identify for period 5 two subsamples, (1–198) and (207–254) corresponding to September 1991–February 2008 and

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 3. Deterministic and stochastic tests for utility maximization

	Period 1 January 1967–December 1969	Period 2 January 1970–May 1974	Period 3 June 1974–November 1982	Period 4 December 1982–August 1991	Period 5 September 1991–December 1992
			Deterministic analysis		
# of GARP violations	0	0	2	30	367
			Stochastic analysis		
# of altered bundles	-	-	1	7	43
Adjustment, Obj			0.1690	2140.07	1128679.40
$\text{Min}(\text{Obj}_i)(i)$			0.1690 (9)	0.0005 (86)	0.0240 (234)
$\text{Max}(\text{Obj}_i)(i)$			0.1690 (9)	2136.51 (100)	319093.75 (210)
Significance[1] at 5%	-	-	No	No	Yes

[1]Test based on the empirical quantiles of the extremal smoothed errors, computed using the simulation smoother.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

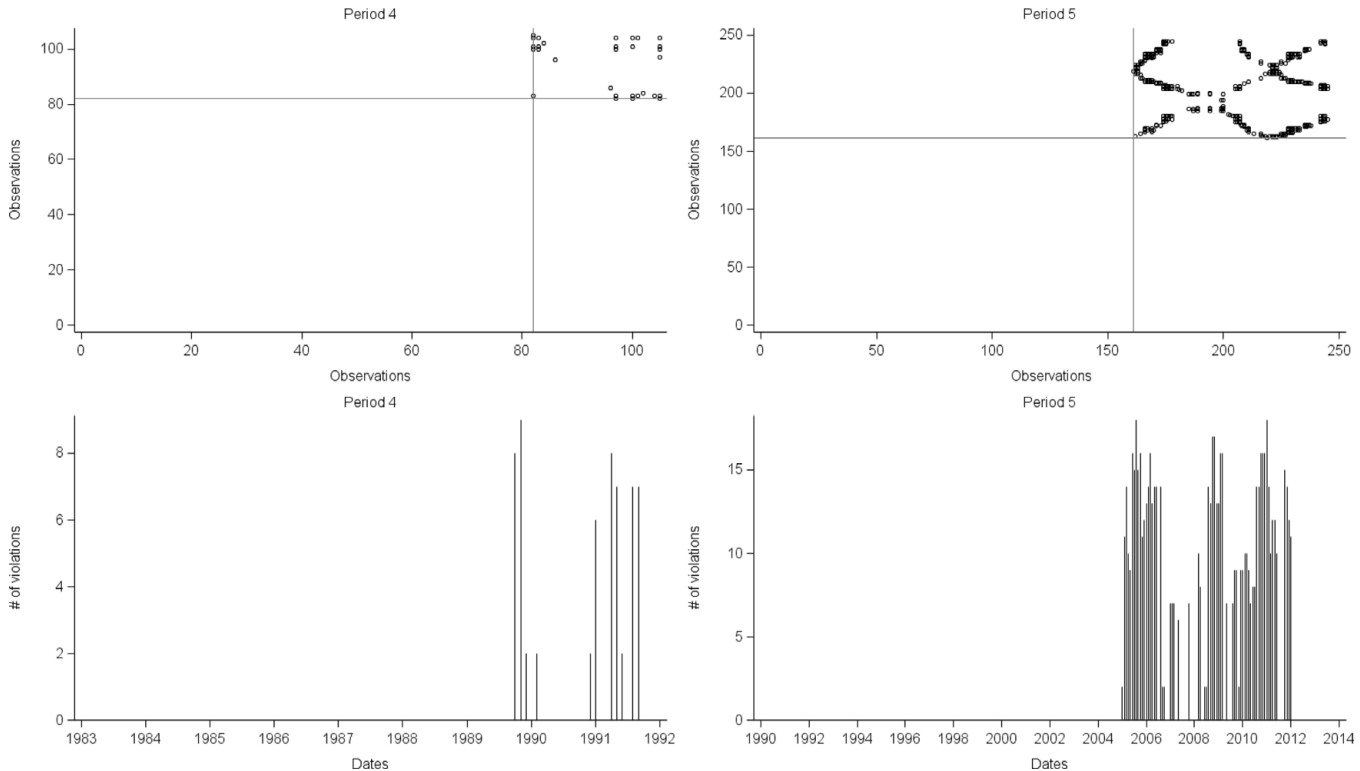


FIGURE 1. Distribution of violations. The two upper graphs return all couples (i, j) such that $\mathbf{x}_i R \mathbf{x}_j$ and $\mathbf{x}_j P \mathbf{x}_i$. The two lower graphs give with how many bundles a bundle $\mathbf{x}_i$ is violating GARP.

TABLE 4. Monetary aggregates components

	DM1	DM2M	DM2 ALL	DM3	DM4-	DM4
CURTC	X	X	X	X	X	X
DD	X	X	X	X	X	X
OCD	X	X	X	X	X	X
SAV		X	X	X	X	X
SAVM		X	X	X	X	X
MMDA		X	X	X	X	X
RMF		X	X	X	X	X
STD			X	X	X	X
IMMF			X	X	X	X
LTD				X	X	X
REPO				X	X	X
CP					X	X
TB						X

January 2009–December 2012, on which GARP is violated, but the violations are nonsignificant.[4]

Given these results, in the sequel, we implement weak separability tests on periods 1, 2, 3, and 4, but also on period $4'$, corresponding to December 1982–September 1989 (period with no violations of GARP), and periods $5'$, $5''$, and $5'''$, corresponding to, respectively, September 1991–February 2005 (no violation of GARP), September 1991–February 2008, and January 2009–December 2012 (nonsignificant violations).

3.3. Testing for Weak Separability

We next implement weak separability tests over six different aggregates, DM1, DM2M, DM2 ALL, DM3, DM4-, and DM4, as presented in Table 4, for the different subperiods defined above. Recall that our procedure is carried out in two steps. (i) The first step involves finding the minimal adjustment in order for the data to be compliant with both the overall and subutility maximization programs. At this stage, as shown in Barnett and de Peretti (2009) and in the appendix, program (4) is iteratively and sequentially solved. First, the data are adjusted to be compliant with the subutility $f()$. Next, the data are adjusted to be compliant with the overall utility $U()$ under the additional constraint that they must remain compliant with the subutility. (ii) In the second step, if the overall adjustment is nonsignificant, then we run the separability test. Hence, separability can be rejected if either the theoretical adjustment or the test statistic (14) is too large.

Tables 5–7 present the results for the different subsamples. For one subsample and one aggregate, main entries are as follows. (i) For the subutility, the number of violations of GARP, as well as the theoretical adjustment (if any), i.e., the value of the objective function, is included. Note that this adjustment is purely

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 5. Results of the extended Barnett and de Peretti (2009) procedure for weak separability, periods 1, 2, and 3

	DM1	DM2M	DM2 ALL	DM3	DM4-	DM4
Period 1: January 1967–December 1969						
Analysis for the subutility						
GARP	0	0	5	0	0	2
Adjustment	0	0	0.2631	0	0	0.0006
Analysis for the overall utility and the subutility						
GARP	0	0	0	0	0	0
Overall adjustment	0	0	0.2631	0	0	0.0006
Significance[1] (5%)	.	.	No	.	.	No
Separability tests						
Independence[2]	3.27 (0.952)	45.77 (0)	52.12 (0)	76.76 (0)	98.14 (0)	90.46 (0)
Separability	Yes	No	No	No	No	No
Period 2: January 1970–May 1974						
Analysis for the subutility						
GARP	2	2	47	33	22	29
Adjustment	0.0003	0.3309	56082.84	5155.21	6055.98	4193.57
Analysis for the overall utility and the subutility						
Significance of the Adjustment						
GARP	0	0	0	0	0	0
Overall adjustment	0.0003	0.3309	56082.84	5155.21	6055.98	4193.57
Significance[1] (5%)	No	No	Yes	No	No	No
Separability tests						
Independence[2]	8.10 (0.618)	100.36 (0)	.	152.18 (0)	152.74 (0)	145.43 (0)
Separability	Yes	No	No	No	No	No
Period 3: June 1974–November 1982						
Analysis for the subutility						
GARP	7035	937	N.R.	2131	2193	2554
Adjustment	No Sol.	4748201.70	N.R.	3369718.00	3549231.00	9887286.31
Analysis for the overall utility and the subutility						
GARP	.	6	N.R.	6	4	0
Overall adjustment	No Sol.	4698628.81	N.R.	3369718.20	3549232.80	9887286.31
Significance[1] (5%)	.	Yes	N.R.	Yes	Yes	Yes

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 5. Continued

	DM1	DM2M	DM2 ALL	DM3	DM4-	DM4
Separability tests						
Independence[2]	.	.	N.R.	.	.	.
Separability	No	No	N.R.	No	No	No

[1]Test based on the empirical quantiles of the extremal smoothed errors, computed using the simulation smoother.
[2]Test based on (14).
"N.R." stands for Not Relevant.
"No. Sol." stands for No Solution to the quadratic program (4).
p-values between parentheses.

indicative since it is not tested. Only the overall adjustment is tested. Nevertheless, compared to the overall adjustment, it gives some clues to understand why overall maximization could be rejected. (ii) For both the overall utility and the subutility, the same information plus a significance test is included. (iii) if the data pass the significance test, we test for separability and report the value of Qh as well the associated p-value. Thus, results of the separability tests are given only if there exist both an overall utility and a subutility rationalizing the data.

Table 5 presents results for periods 1, 2, and 3. To pick up an example, when testing for the existence of a subutility for DM1 (period 1), there are no GARP violations. A similar result holds when testing for both a subutility and an overall utility function. DM1 is also found to be separable using the noncorrelation test. Switching to period 2, when testing for separability, two violations appear with a very low adjustment of 0.0003. Adjusting the data, and testing for existence of both the subutility and the overall maximization programs, no additional violations are found, thus returning an overall adjustment of 0.0003. With a p-value of 0.618, data are still separable. This result is not found in period 3. With 7035 violations of GARP for the subutility, we were even not able to find a solution to the minimization problem for the subutility or for both utilities. Here, separability is strongly rejected.

More generally, over period 1 (January 1967–December 1969), when testing for GARP for the subutility, only two clusters present violations: DM2 ALL and DM4. Nevertheless, the overall required adjustment is very low, say, 0.2631 and 0.0006, and nonsignificant. When testing for separability, only the DM1 group is found separable, whereas for all other aggregates, separability is strongly rejected.

The analysis is more interesting for period 2 (January 1970–May 1974), since for all monetary clusters, violations appear. Except for DM2 ALL, the theoretical adjustments are nonsignificant. Again, DM1 is the only aggregate found to be separable. Note that when monetary data are adjusted, no violation occurs when testing for both the overall utility maximization program and the subutility program. Testing for a subutility function and an overall utility function for period 3 (June 1974–November 1982) produces a high number of violations for all clusters

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 6. Results of the extended Barnett and de Peretti (2009) procedure for weak separability, periods 4 and 4′

	DM1	DM2M	DM2 ALL	DM3	DM4-	DM4
Period 4: December 1982–August 1991						
Analysis for the subutility						
GARP	218	62	47	4	23	27
Adjustment	14006.75	2753.53	413.91	0.012	607.13	679.69
Analysis for the overall utility and the subutility						
GARP	31	32	30	30	30	30
Overall adjustment	14010.643	5238.98	3091.48	2540.63	3417.64	8342.94
Significance[1] (5%)	No	No	No	No	No	No
Separability tests						
Independence[2]	15.01 (0.306)	395.40 (0)	372.16 (0)	380.27 (0)	328.28 (0)	274.86 (0)
Separability	Yes	No	No	No	No	No
Period 4': December 1982–September 1989						
Analysis for the subutility						
GARP	84	44	45	2	10	18
Adjustment	5714.87	1724.66	413.90	0.0001	48.67	644.99
Analysis for the overall utility and the subutility						
GARP	0	2	0	0	0	0
Overall Adjustment	5714.87	1447.27	413.90	0.0001	48.67	644.99
Significance[1] (5%)	No	No	No	No	No	No
Separability tests						
Independence[2]	13.61 (0.401)	312.46 (0)	319.66 (0)	314.67 (0)	266.10 (0)	230.53 (0)
Separability	Yes	No	No	No	No	No

[1]Test based on the empirical quantiles of the extremal smoothed errors, computed using the simulation smoother.
[2]Test based on (14).
p-values between parentheses.

with huge and significant adjustments. As mentioned above, for DM1, we were not even able to solve (4). Compared to the results we obtained, when testing for the overall maximization (only two violations, Table 4), the rejection of weak separability seems to be entirely due to our inability to find a monetary subutility function rationalizing the data.

We now turn to periods 4 (Table 5) and 5 (Table 6), which revealed large numbers of violations. For the former period, we perform the tests on both the

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 7. Results of the extended Barnett and de Peretti (2009) procedure for weak separability, periods 5′, 5″, and 5‴

	DM1	DM2M	DM2 ALL	DM3	DM4-	DM4
Periods 5′: September 1991–February 2005						
Analysis for the subutility						
GARP	6	47	245	10	46	17
Adjustment	1.17	11550.95	113044.28	166.48	5060.95	194.62
Analysis for the overall utility and the subutility						
GARP	0	0	0	0	0	0
Overall adjustment	1.17	11550.95	113044.28	166.48	5060.95	194.62
Significance[1] (5%)	No	No	Yes	No	No	No
Separability tests						
Independence[2]	14.57 (0.265)	484.65 (0)	.	681.63 (0)	463.52 (0)	389.36 (0)
Separability	Yes	No	No	No	No	No
Period 5″: September 1991–February 2008						
Analysis for the subutility						
GARP	6	797	2174	709	770	714
Adjustment	1.17	274085.51	1629798.30	710457.29	944018.13	537723.51
Analysis for the overall utility and the subutility						
GARP	0	0	8	2	6	4
Overall adjustment	1.17	274085.51	1629798.30	710457.29	943954.36	538734.75
Significance[1] (5%)	No	Yes	Yes	Yes	Yes	Yes
Separability tests						
Independence[2]	22.84 (0.029)	.	.	.	.	.
Separability	No	No	No	No	No	No
Period 5‴: January 2009–December 2012						
Analysis for the subutility						
GARP	0	9	0	0	0	0
Adjustment	0	12.48	0	0	0	0
Analysis for the overall utility and the subutility						
GARP	19	19	19	7	19	19
Overall adjustment	1500.78	1509.78	1575.44	7.59	1497.29	1550.16
Significance[1] (5%)	No	No	No	No	No	No

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

TABLE 7. Continued

	DM1	DM2M	DM2 ALL	DM3	DM4-	DM4
			Separability tests			
Independence[2]	16.91 (0.152)	144.92 (0)	157.52 (0)	202.27 (0)	167.18 (0)	146.42 (0)
Separability	Yes	No	No	No	No	No

[1]Test based on the empirical quantiles of the extremal smoothed errors, computed using the simulation smoother.
[2]Test based on (14).
p-values between parentheses.

overall sample (December 1982–August 1991) and on the period when no violation appeared (December 1982–September 1989). Over the period December 1982–August 1991, the analysis produces GARP violations for both the monetary subutility functions and the overall utility function, as suggested in Table 3. Nevertheless, the adjustments are nonsignificant. Here again, only DM1 appears to be weakly separable. A similar conclusion is drawn when focusing on period 4′, except that less violations are found.

For the last period (Table 7) we perform the analysis over three subperiods, September 1991–February 2005, where no violations appear when testing for the overall maximization, September 1991–February 2008, and January 2009–December 2012, the latter two subperiods presenting no significant violations of GARP. During September 1991–February 2005, there exist both a subutility function and an overall utility function for all clusters, except for DM2ALL, but here again only DM1 is weakly separable. During September 1991–February 2008, the assumption of a subutility function rationalizing the data is strongly rejected for all clusters, except DM1, which is not separable at the 5% level. Finally, during January 2009–December 2012, still only DM1 presents empirical evidence of weak separability.

4. CONCLUSION AND DISCUSSION

There are several important conclusions to be drawn from this paper. Concerning our results, we have implemented an extended version of the Barnett and de Peretti (2009) test to check for weak separability of six different monetary aggregates, DM1, DM2M, DM2 ALL, DM3, DM4-, and DM4, all reported by the CFS. The test is a two-step procedure that first jointly tests for the existence of an overall utility function and a subutility function, and then tests for the separability of the latter within the overall utility function. In this framework, rejection of weak separability can be either due to the inability to find utility function(s) rationalizing the data or due to the nonseparability of the subutility function.

Among the six aggregates, our results suggest that the only aggregate found separable is DM1 and that no other aggregate meets the separability criterion. Expanding on the reasons why we rejected separability for the five other aggregates, it appears that on most subsamples, the null is rejected because of the

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

nonindependence between the marginal rate of substitution of the assets inside the group and quantities outside. On the other subsamples, i.e., June 1974–November 1982 and September 1991–February 2008, separability was rejected because we were unable to find well-defined monetary subutility functions rationalizing the data. Over the former period, we were even unable to find a utility function rationalizing the data for DM1.

It is crucial to stress the limitations of our approach and also what we did and did not in this paper. What we did was to test the separability of six existing aggregates, according to the Federal Reserve Bank of the United States' definitions. Failing to reject the null only for DM1 does not imply that other definitions of money, i.e., DM1 plus other assets, may not exist, since what we did not do was to test for all possible broad clusterings. This was beyond the scope of this paper and is left for future research. Second, and more importantly, we have tested for a composite hypothesis about weak separability and the existence of a representative agent, i.e., weak separability and Gorman's conditions for aggregating over agents. These conditions are extremely stringent, and it is likely that the rejection of weak separability may be due to the rejection of Gorman's conditions, i.e., due to heterogeneity across agents. Clearly, different groups of agents may value different clusters as money, causing the rejection of the null as we test for broader aggregates. To some extent, accepting DM1 as the only separable aggregate would support this assumption. In other words, rejection of other broader clusters is likely due to the heterogeneity of choices. This inability to conclude about the reasons why we reject separability clearly calls for more advanced works and testing procedures, taking into account heterogeneity across agents [see Barnett (2004)], to confirm the results found in this paper. An alternative would be implementing the tests on microeconomic data focusing on a class of agents according to the distribution of their income.

Finally, it is worth comparing our results with other studies that have found weak separability for a broader aggregate of money. Most have used quarterly or even annual data. In this case, jumps in budgets could be too large from a period to another, causing too few budgets hyperplanes intersections, thus dramatically decreasing the power of GARP. In this study, we use monthly data, ensuring a high number of intersections, which partly explains why we reach different conclusions. When dealing with this kind of analysis, it could then be useful to report results of the GARP analysis, together with a power index as suggested, for instance, by Andreoni et al. (2013).

NOTES

1. For justifications of money in the utility function, see Feenstra (1986), or Poterba and Rotemberg (1987).

2. To avoid any confusion, note that obj is by far not an estimate of the true objective function when data are measured with errors, as also emphasized by Elger and Jones (2008). obj is the minimal perturbation in order for the data to be compliant with both the utility and the subutility function. We will refer to it as the minimal theoretical perturbation.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

3. See also Anderson and Jones (2011).
4. Results not reported, available upon request.

REFERENCES

Afriat, Sydney (1967) The construction of a utility function from expenditure data. *International Economic Review* 8, 67–77.

Afriat, Sydney (1973) On a system of inequalities on demand analysis: An extension of the classical model. *International Economic Review* 14, 460–472.

Anderson, Richard A. and Barry E. Jones (2011) A comprehensive revision of the US Monetary Services (Divisia) Index. *Federal Reserve Bank of St. Louis Review* 93, 325–359.

Andreoni, James, Benjamin J. Gillen, and William T. Harbaugh (2013) The Power of Revealed Preference Tests: Ex-post Evaluation of Experimental Design. Working paper. University of California, San Diego.

Barnett, William A. (1978) The user cost of money. *Economic Letters* 1, 145–149. Reprinted in Barnett, William A. and Apostolos Serletis (2000) *The Theory of Monetary Aggregation*. Chapter 2, Amsterdam: North Holland.

Barnett, William A. (1979) The joint allocation of leisure and goods expenditure. *Econometrica* 47, 539–563.

Barnett, William A. (1980) Economic monetary aggregates: An application of index number and aggregation theory. *Journal of Econometrics* 14, 11–48. Reprinted in Barnett, William A. and Apostolos Serletis (2000) *The Theory of Monetary Aggregation*. Chapter 2, Amsterdam: North Holland.

Barnett, William A. (2004) Multilateral aggregation-theoretic monetary aggregation over heterogeneous countries. *Journal of Econometrics* 136, 457–482,

Barnett, William A. and Seungmook Choi (1989) A Monte Carlo study of tests of blockwise weak separability. *Journal of Business and Economic Statistics* 7, 363–377. Reprinted in William A. Barnett and Jane Binner (2004) *Functional Structure and Approximation in Econometrics*. Chapter 12, Amsterdam: North Holland.

Barnett, William A. and Philippe de Peretti (2009) Admissible clustering of aggregator components: A necessary and sufficient stochastic seminonparametric test for weak separability. *Macroeconomic Dynamics* 13, 317–334.

Barnett, William A. and Shu Wu (2005) On user costs of risky monetary assets. *Annals of Finance* 1(1), 35–50.

Barnett, William A., Jia Lu, Ryan S. Mattson, and Jeff van den Noort (2013) The new CFS divisia monetary aggregates: Design, construction, and data sources, manuscript. *Open Economies Review* 24, 101–124.

Blackorby, Charles, Daniel Primont, and Robert R. Russell (1977) On testing separability restrictions with flexible functional forms. *Journal of Econometrics* 5, 195–209.

Blackorby, Charles, Daniel Primont, and R. Robert Russell (1998) Separability: A survey. In Salvador Barbera, Peter Hammond, and Christian Seidl (eds.), *The Handbook of Utility Theory*, vol. 1, pp. 49–92. Kluwer: Dordrecht.

Cherchye, Laurens, Thomas Demuynck, Bram De Rock, and Per Hjerstrand (2015) Revealed preference tests for weak separability: An integer programming approach. *Journal of Econometrics* 186(1), 129–141.

Deaton, Angus and John Muellbauer (1980) *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

de Jong, Piet and Neil Shephard (1995) The simulation smoother for time series models. *Biometrika* 82, 339–350.

de Peretti, Philippe (2007) Testing the significance of the departures from weak separability. In William A. Barnett and Apostolos Serletis (eds.), *International Symposia in Economic Theory and Econometrics: Function Structure Inference*, pp. 3–22. Amsterdam: Elsevier.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

Denny, Michael and Melvyn Fuss (1977) The use of approximation analysis to test for separability and the existence of consistent aggregates. *American Economic Review* 67, 404–418.

Diewert, Walter E. (1974) Intertemporal consumer theory and the demand for durables. *Econometrica* 42, 497–516.

Drake, Leigh (1997) Nonparametric demand analysis of U.K. personal sector decisions on consumption, leisure, and monetary assets: A reappraisal. *The Review of Economics and Statistics* 79, 679–683.

Durbin, James and Siem Jan Koopman (2001) Time series analysis by state space methods. In *Oxford Statistical Science Series Book* vol. 38. Oxford: Oxford University Press.

El Himdi, Khalid and Roch Roy (1997) Tests for noncorrelations for two multivariate ARMA series. *Canadian Journal of Statistics* 25, 233–256.

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

Feenstra, Robert C. (1986) Functional equivalence between liquidity costs and the utility of money. *Journal of Monetary Economics* 22, 271–291.

Fisher, Douglas and Adrian R. Fleissig (1997) Monetary aggregation and the demand for assets. *Journal of Money, Credit and Banking* 29, 458–475.

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

Fleissig, Adrian R. and Gerald A. Whitney (2005) Testing the significance of violations of Afriat's inequalities. *Journal of Business and Economic Statistics* 23, 355–362.

Gorman, William M. (1961) On a class of preference fields. *Metroeconomica* 13, 53–56.

Hjertstrand, Per and James L. Swofford (2014) Are the choices of people stochastically rational? A stochastic test of the number of reveals preference violations. *Empirical Economics* 46, 1495–1519.

Jones, Barry E. and Philippe de Peretti (2005) A comparison of two methods for testing the utility maximization hypothesis when quantity data is measured with error. *Macroeconomic Dynamics* 9, 612–629.

Leontief, Wassily (1947) Introduction to a theory of the internal structure of functional relationships. *Econometrica* 15, 361–373.

Offenbacher, Edward Akiva and Shachar Shemesh (2011) Divisia Monetary Aggregates for Israel: Background Note and Metadata. Bank of Israel, Research Department: Monetary/Finance Division.

Patterson, Kerry D. (1991) A non-parametric analysis of personal sector decisions on consumption, liquid assets and leisure. *Economic Journal* 101, 1103–1116.

Poterba, James M. and Julio J. Rotemberg (1987) Money in the utility function: An empirical implementation. In William A. Barnett and Kenneth J. Singleton (eds.), *New Approaches to Monetary Economics*, pp. 219–240. New York: Cambridge University Press.

Swofford, James L. and Gerald A. Whitney (1987) Non-parametric tests of utility maximization and weak separability for consumption, leisure, and money. *Review of Economic Studies* 69, 458–464.

Swofford, James L. and Gerald A. Whitney (1988) A comparison of nonparametric tests of weak separability for annual and quarterly data on consumption, leisure, and money. *Journal of Business & Economic Statistics* 6, 241–246

Swofford, James L. and Gerald A. Whitney (1994) A revealed preference test for weakly separable utility maximization with incomplete adjustment. *Journal of Econometrics* 60, 235–49.

Varian, Hal (1982) The nonparametric approach to demand analysis. *Econometrica* 50, 945–973.

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

Varian, Hal (1985) Non-parametric analysis of optimizing behavior with measurement error. *Journal of Econometrics* 30, 445–458.

Varian, Hal (1990) Goodness-of-fit in optimizing models. *Journal of Econometrics* 46, 125–140.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

APPENDIX: COMPUTING THE MINIMAL THEORETICAL ADJUSTMENT

In this appendix, we present how the quadratic program (4) is solved. As in Barnett and de Peretti (2009), the program is sequentially and iteratively solved. Concerning the sequence, data are, first, adjusted to be compliant with the subutility and, next, to be compliant with both the overall utility and the subutility. Therefore, (4) is broken into two subprograms, the latter using all the information available in the former. Each program is iteratively solved. Introduce the two following two additional definitions.

DEFINITION 3 ($\mathbf{x}_i V R \mathbf{x}_j$). *Two observations* $\mathbf{x}_i$ *and* $\mathbf{x}_j$ *satisfy the binary relation* $\mathbf{x}_i V R \mathbf{x}_j$ *if there exists a sequence between* $\mathbf{x}_i$ *and* $\mathbf{x}_j$ *such that* $\mathbf{x}_i R \mathbf{x}_k$ *and* $\mathbf{x}_k P_i^0 \mathbf{x}_k$, $\mathbf{x}_k R \mathbf{x}_m$ *and* $\mathbf{x}_m P_k^0 \mathbf{x}, \ldots, \mathbf{x}_n R \mathbf{x}_j$ *and* $\mathbf{x}_j P_n^0 \mathbf{x}$.

DEFINITION 4 ($\mathbf{x}_i S R \mathbf{x}_j$). *Two observations* $\mathbf{x}_i$ *and* $\mathbf{x}_j$ *satisfy the binary relation* $\mathbf{x}_i S R \mathbf{x}_j$ *if* $S(i) = S(j)$, *where* $S(i) = (\sum_{j=1}^{T} r_{ij}) - 1$ *is a function returning the sum m of the* i*th row and the transitive closure matrix R, m returning to how many bundles* $\mathbf{x}_i$ *is revealed preferred to.*

As $\mathbf{x}_i V R \mathbf{x}_j$ implies $\mathbf{x}_i S R \mathbf{x}_j$, all the bundles violating GARP can be classified into n independent sets $B_l, l = 1, 2, \ldots, n$, each set corresponding to a particular rupture in the preference chain. For the sub-utility program, the following five-step procedure can therefore be used:

Step 1.1. Test $D^{(1)} = (\mathbf{x}_i^{(1)}, \mathbf{p}_i^{(1)})_{i=1}^{T}$ for compliance with GARP. If violations appear, go to Step 2.1; otherwise, stop the loop.
Step 2.1. Search for clusters of violations by classifying all bundles violating GARP into n independent sets B_l. In each set, the bundles violating GARP are related by $\mathbf{x}_i^{(1)} V R \mathbf{x}_j^{(1)}$ and $\mathbf{x}_i^{(1)} S R \mathbf{x}_j^{(1)}$. Search for the one, defined as B_1, such that all bundles included in are candidates to be at the same higher place in the preference chain. Go to Step 3.1.
Step 3.1. For all bundles $\mathbf{x}_i^{(1)} \in B_1$, solve the following quadratic program and store the objective function obj_i:

$$\text{obj}_i = \min_{z_{ij}^{(1)}} \sum_{j=1}^{a} \left(x_{ij}^{(1)} - z_{ij}^{(1)} \right)^2. \tag{19}$$

Subject to:

$$\left| \begin{array}{l} \mathbf{p}_i^{(1)} \cdot \mathbf{x}_n^{(1)} \geq \mathbf{p}_i^{(1)} \cdot \mathbf{z}_i^{(1)}, \\ \forall \mathbf{x}_n^{(1)} \text{ such that } \mathbf{x}_n^{(1)} R \mathbf{x}_i^{(1)} \text{ implies not } \mathbf{x}_i^{(1)} P^0 \mathbf{x}_n^{(1)}, \end{array} \right. \tag{C.1.1}$$

$$\left| \mathbf{p}_m^{(1)} \cdot \mathbf{x}_m^{(1)} \leq \mathbf{p}_m^{(1)} \cdot \mathbf{z}_i^{(1)}, \forall \mathbf{x}_m^{(1)} \in B_1, \; m \neq i. \right. \tag{C.1.2}$$

Then, go to Step 4.1.

Step 4.1. Select the bundle having the minimal objective function. Let $\widehat{\mathbf{z}}_i^{(1)}$ be the bundle solution.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press

Step 5.1. In $D^{(1)}$, replace $\mathbf{x}_i^{(1)}$ with $\widehat{\mathbf{z}}_i^{(1)}$. Go to Step 1.1.

The above five-step procedure returns a set $\widehat{D}^{(1)} = (\widehat{\mathbf{z}}_i^{(1)}, \mathbf{p}_i^{(1)})_{i=1}^T$ compliant with GARP and a coherent transitive closure matrix. The different sets of constraints clearly appear with regard to Barnett and de Peretti (2009). Whereas constraint (C.2.1) is equivalent but simplified, constraint (C.1.1) is quite different. Indeed, in the original program, the overall expenditure during period i was forced to remain unchanged $[\mathbf{p}_i^{(1)} \cdot \mathbf{z}_i^{(1)} = \mathbf{p}_i^{(1)} \cdot \mathbf{x}_i^{(1)}]$ in order for $\mathbf{z}_i^{(1)}$ to (i) remain at a given place in the preference chain, and (ii) cause no new violations with bundles located above in the preference chain, i.e., bundles revealed preferred to $\mathbf{x}_i^{(1)}$. Here, we use exactly the same idea but replace $\mathbf{p}_i^{(1)} \cdot \mathbf{z}_i^{(1)} = \mathbf{p}_i^{(1)} \cdot \mathbf{x}_i^{(1)}$ with $\mathbf{p}_i^{(1)} \cdot \mathbf{x}_n^{(1)} \geq \mathbf{p}_i^{(1)} \cdot \mathbf{z}_i^{(1)}, \forall \mathbf{x}_n^{(1)}$ for all $\mathbf{x}_n$ such that $\mathbf{x}_n^{(1)} R \mathbf{x}_i^{(1)}$ implies not $\mathbf{x}_i^{(1)} P^0 \mathbf{x}_n^{(1)}$. We thus allow the overall budget to change, providing for more flexibility in the program. We therefore expect smaller adjustments compared to Barnett and de Peretti (2009).

Focus now on the adjustment program for the subutility and the overall utility. Rename $\widehat{\mathbf{z}}_i^{(1)}$ as ϑ_i and define $\widehat{D} = [(\vartheta_i, \mathbf{x}_i^{(2)}),\ (\mathbf{p}_i^{(1)}, \mathbf{p}_i^{(2)})]_{i=1}^T$. Producing a set compliant with both the subutility and the overall utility maximization programs is then achieved by using the above five-step procedure, replacing $D^{(1)}$ with $\widehat{D}$, $\mathbf{x}_i^{(1)}$ with $(\vartheta_i, \mathbf{x}_i^{(2)})$, and (19) with (20).

$$\text{obj}_i = \min_{z_{ij}^{(1)}, z_{ij}^{(2)}} \left[\sum_{j=1}^{a} \left(x_{ij} - z_{ij}^{(1)} \right)^2 + \sum_{j=a+1}^{k} \left(x_{ij} - z_{i(j-a)}^{(2)} \right)^2 \right] \tag{20}$$

subject to

$$\left| \begin{array}{l} \mathbf{p}_i^{(1)} \cdot \vartheta_n + \mathbf{p}_i^{(2)} \cdot \mathbf{x}_n^{(2)} \geq \mathbf{p}_i^{(1)} \cdot \mathbf{z}_i^{(1)} + \mathbf{p}_i^{(2)} \cdot \mathbf{z}_i^{(2)}, \\ \forall (\vartheta_n, \mathbf{x}_n^{(2)}) \text{ such that} \\ (\vartheta_n, \mathbf{x}_n^{(2)}) R (\vartheta_i, \mathbf{x}_i^{(2)}) \text{ implies not } (\vartheta_i, \mathbf{x}_i^{(2)}) P^0 (\vartheta_n, \mathbf{x}_n^{(2)}), \end{array} \right. \tag{C.2.1}$$

$$\left| \begin{array}{l} \mathbf{p}_m^{(1)} \cdot \vartheta_m + \mathbf{p}_m^{(2)} \cdot \mathbf{x}_m^{(2)} \leq \mathbf{p}_m^{(1)} \cdot \mathbf{z}_i^{(1)} + \mathbf{p}_m^{(2)} \cdot \mathbf{z}_i^{(2)}, \\ \forall (\mathbf{p}_m^{(1)} \cdot \vartheta_m) \in B_1,\ m \neq i, \end{array} \right. \tag{C.2.2}$$

$$\left| \mathbf{p}_i^{(1)} \cdot \vartheta_o \geq \mathbf{p}_i^{(1)} \cdot \mathbf{z}_i^{(1)}, \forall \vartheta_j \text{ such that } \vartheta_o R \vartheta_i \right. \tag{C.2.3}$$

$$\left| \mathbf{p}_p^{(1)} \cdot \vartheta_p \leq \mathbf{p}_p^{(1)} \cdot \mathbf{z}_i^{(1)}, \forall \vartheta_p \text{ such that } \vartheta_i R \vartheta_p. \right. \tag{C.2.4}$$

The above program returns a set $\widehat{\widehat{D}} = [(\widehat{\mathbf{z}}_i^{(1)}, \widehat{\mathbf{z}}_i^{(2)}), (\mathbf{p}_i^{(1)}, \mathbf{p}_i^{(2)})]_{i=1}^T$ compliant with both the utility and the subutility maximization programs.

https://doi.org/10.1017/S1365100516000791 Published online by Cambridge University Press