Macroeconomic Dynamics, **4**, 2000, 547–572. Printed in the United States of America.

SEPARABILITY, AGGREGATION, AND EULER EQUATION ESTIMATION

ADRIAN R. FLEISSIG
California State University at Fullerton

A. RONALD GALLANT
University of North Carolina at Chapel Hill

JOHN J. SEATER
North Carolina State University

We derive a seminonparametric utility function containing the constant relative risk aversion (CRRA) function as a special case, and we estimate the associated Euler equations with U.S. consumption data. There is strong evidence that the CRRA function is misspecified. The correctly specified function includes lagged effects of durable goods and perhaps nondurable goods, is bounded as required by Arrow's Utility Boundedness Theorem, and has a positive rate of time preference. Constraining sample periods and separability structure to be consistent with the generalized axiom of revealed preference affects estimation results substantially. Using Divisia aggregates instead of the NIPA aggregates also affects results.

Keywords: Euler Equations, Separability, Aggregation

1. INTRODUCTION

Euler equation estimation, which has become a mainstay of macroeconometrics, attempts to find the utility function of the representative agent. Whether such an exercise is useful is the subject of some disagreement, with a number of economists skeptical of the representative-agent framework for well-known reasons. We do not join that debate here. Our purpose is to examine some conditions that must be satisfied for Euler equation estimation to be valid once one has accepted the representative-agent framework. Three such conditions occupy our attention: The data must be consistent with the existence of a well-behaved aggregate utility function, the broad categories of goods (such as nondurables, services, and durables) used in estimation must be consistent with the utility function's

We thank William Barnett, Alastair Hall, Douglas Fisher, Philip Meguire, Douglas Pearce, Kerry Smith, and two referees for helpful comments. Address correspondence to: Adrian R. Fleissig, California State University, Fullerton Department of Economics, Box 6848, Fullerton, CA 92834-6848, USA; e-mail: afleissig@fullerton.edu.

 1365-1005/00 $9.50

separability structure, and the aggregation method used to construct categories of goods also must be consistent with that separability structure. A substantial microeconomic literature discusses conditions for the existence of a utility function, its separability structure, and methods for aggregating data. However, that literature generally has been ignored in the Euler equation studies applied to aggregate data, which typically assume (implicitly) that the relevant conditions are satisfied and proceed straight to estimation. In this paper we apply the microeconomic literature to Euler equation estimation. We find that assumptions often made in Euler equation studies using aggregate U.S. data apparently are invalid, and correcting the problems has a substantial effect on estimation results.

A necessary condition for using the representative-agent approach is that there must exist a nonsatiated, continuous, concave, monotonic utility function that is consistent with the data set under examination. Such a utility function exists if and only if the data satisfy the Generalized Axiom of Revealed Preference (GARP); see Varian (1982). Using tests developed by Varian (1983), Fleissig et al. (2000) find that the U.S. National Income and Product Account (NIPA) consumption data generally are *not* consistent with GARP over the full periods of data availability. Estimating Euler equations over those sample periods therefore may be invalid. Over subsample periods for which the data are GARP-consistent, two important results emerge: First, the aggregate utility function is separable in nondurables and services but not in durables; second, durables can be aggregated up to two or three categories of expenditures but not all the way up to one single overall quantity. The first of these results implies that one cannot estimate Euler equations from a utility function (or, strictly speaking, a subutility function) that includes only durables; if durables are to be included, then nondurables and services also must be included. The second result implies that if durables are included, then one cannot use simply an overall aggregate of all durables but instead must enter several subaggregates separately. Also, the method of aggregating the data should be consistent with the utility function's separability structure.

Most published Euler equation studies using aggregate data have ignored the foregoing issues and have used sample periods, separability structures, and/or aggregation methods that apparently are inconsistent with the data. The estimation results of those studies therefore are of uncertain validity. We explore the sensitivity of the results by restricting attention to GARP-consistent sample periods, separability structures, and aggregation methods. We find that doing so affects estimation substantially. We compare results obtained for a standard constant relative risk aversion (CRRA) function with those obtained for a seminonparametric (SNP) function that nests that CRRA.

With no restrictions imposed, the CRRA function appears to be adequate, but as one restricts the sample period, separability structure, and aggregation method to those consistent with GARP, the CRRA becomes misspecified. In contrast, the SNP function is not misspecified. The SNP function is essentially free from specification error (up to a certain structure, discussed later) and allows for nonseparability between expenditure on durable and nondurable goods. The overidentifying

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

restrictions are not rejected, the rate of time preference is significantly positive, and the utility function is bounded, as required by Arrow's (1971) Utility Boundedness Theorem. Lagged durables and possibly lagged nondurables are significant, with coefficients consistent with habit persistence. Estimates of the index of relative risk aversion differ considerably across specifications, data frequency, and separability assumptions. This parameter is central to much applied analysis, such as evaluation of certain kinds of public policies and assessments of some effects of economic development. We find that its magnitude can vary by as much as 30 times across the model specifications we study. We also find that monthly data give results much different from those for annual or quarterly data, whose results are consistent with each other, suggesting a possible problem with the use of monthly data.

The rest of the paper is organized as follows: Section 2 presents the model to be estimated and derives the orthogonality conditions. Section 3 discusses the data and the issues of GARP consistency, separability, and choice of aggregation method. Section 4 presents the estimation results. Section 5 concludes the paper.

2. MODEL TO BE ESTIMATED

2.1. Utility Maximization and Service Flow of Goods

The consumer maximizes the expected value of an intertemporal multigood utility function

$$\text{Max}\, \mathcal{E}_t \sum_{s=0}^{\infty} \beta^s u^*\left(\boldsymbol{c}^*_{t+s}, \boldsymbol{d}^*_{t+s}\right), \qquad 0 < \beta < 1, \tag{1}$$

where β is the discount factor, $\boldsymbol{c}^*$ is a vector of service flows derived from expenditure on the nondurables ($\boldsymbol{c}$), $\boldsymbol{d}^*$ is a vector of service flows from the stock of durable goods ($\boldsymbol{k}$), $\mathcal{E}_t(\cdot)$ is the conditional expectation $\mathcal{E}_t(\cdot \mid \mathcal{F}_t)$, and $\mathcal{F}_t$ is the agent's information set at time t. We allow nondurables to generate service flows, both because many goods classified as nondurable have at least some durability [Darby (1974) Dunn and Singleton (1986)] and because utility may exhibit habit-persistence. Equation (1) is fairly general and includes as special cases many formulations from the literature.[1] We follow the consumption Euler equation literature in ignoring leisure and treating income as exogenous to the household's consumption choice problem.[2]

The consumer maximizes the expected utility function (1) subject to the period-by-period budget constraint:

$$A_{t+1} = (1 + r_t)A_t + y_t - (c_t + p_{d,t} d_t),$$

where A_t is assets, y_t is income, r_t is the interest rate, c_t is expenditure on nondurable and services, d_t is expenditure on durable goods, and $p_{d,t}$ is the price of durables in terms of nondurables. The consumer also must satisfy the terminal condition of exactly exhausting the lifetime sources. Using the period-by-period constraint gives the expected lifetime constraint[3]:

$$\mathcal{E}\sum_{s=0}^{\infty}\left[y_{t+s}\prod_{i=0}^{S}(1+r_{t+i})^{-1}\right]=\mathcal{E}\sum_{s=0}^{\infty}\left[(c_{t+s}+p_{d,t+s}d_{t+s})\prod_{i=0}^{S}(1+r_{t+i})^{-1}\right].$$

The possible sources of uncertainty are future income y_{t+s}, future durables prices $p_{d,t+s}$, and future interest rates r_{t+s}. We assume at time t, when the consumer formulates a consumption plan, that the agent knows the values of y, p, and r for period t but not for any future period $t+s$.

Data do not exist for consumption service flows from durable goods, and so, we must transform expenditures into service flows. We assume the service flow from the ith nondurable good, c^*_{it}, depends on contemporaneous and J lagged expenditures on c_{it}:

$$c^*_{it}=f^{(i)}(c_{it},c_{it-1},\ldots,c_{it-J}),\qquad 0\leq J<\infty$$

or, in vector notation,

$$\boldsymbol{c}^*_t=\boldsymbol{f}(\boldsymbol{c}_t,\boldsymbol{c}_{t-1},\ldots,\boldsymbol{c}_{t-J}),\qquad 0\leq J<\infty. \tag{2}$$

If $J=0$, then c_{it} generates service flows only in the period of purchase and is literally nondurable. Each durable good evolves according to the usual accumulation equation

$$k_{it}-k_{it-1}=d_{it}-\mu_i k_{it-1}, \tag{3}$$

where μ_i is the rate of depreciation, assumed here to be known a prior.[4] Equation (3) has the solution

$$k_{it}=[1-(1-\mu)L]^{-1}d_{it}. \tag{4}$$

We assume that the period t service flow from the ith durable good d^*_{it} is proportional to the average stock of the good during the period

$$\begin{aligned}d^*_{it}&=\nu_i\tfrac{1}{2}(k_{it-1}+k_{it}),\qquad 0<\tilde{\nu}_i<1\\&=\nu_i(k_{it-1}+k_{it})\end{aligned} \tag{5}$$

in which the only unknown variable on the right-hand side is ν.

Collecting all stocks into vectors gives the vector equivalent of (5):

$$\boldsymbol{d}^*_t=\boldsymbol{\nu}(\boldsymbol{k}_t+\boldsymbol{k}_{t-1}). \tag{6}$$

We then can rewrite the utility function as

$$u^*\left(\boldsymbol{c}^*_t,\boldsymbol{d}^*_t\right)=u(\boldsymbol{c}_t,\ldots,\boldsymbol{c}_{t-J};\boldsymbol{k}_t+\boldsymbol{k}_{t-1}) \tag{7}$$

by substituting (2) and (6) into (1). We can simplify the discussion without loss of generality by considering a composite nondurable c_t (a scalar) and a composite durable good d_t (also a scalar). Then, equation (7) simplifies to

$$u^*\left(c^*_t,d^*_t\right)=u(c_t,\ldots,c_{t-J};k_t+k_{t-1}). \tag{8}$$

2.2. Seminonparametric Utility Function

We derive a seminonparametric generalization of the CRRA utility function by assuming that conditions for exact aggregation of individual utility functions are satisfied.[5] For example, suppose individual utility is degree $\delta\gamma$ homogeneous in c and degree $\theta\gamma$ in k; that is, u^* is homogeneous of degree $(\delta+\theta)\gamma$. Then, extracting c_t and k_t from (8) results in

$$\begin{aligned} u^*\left(c_t^*, d_t^*\right) &= \left(c_t^\delta k_t^\theta\right)^\gamma u^*\left(1, \boldsymbol{C}_t^\#, \boldsymbol{K}_t^\#\right) \\ &= \left(c_t^\delta k_t^\theta\right)^\gamma w\left(\boldsymbol{C}_t^\#, \boldsymbol{K}_t^\#\right), \end{aligned}$$

which is a product of two functions, the first being homogeneous of degree $(\delta+\theta)\gamma$ and the other homogeneous of degree zero, where $C_t^\# = (c_{t-1}/c_t, c_{t-2}/c_t, \ldots, c_{t-J}/c_t)$ and $K_t^\# = (1 + k_{t-1}/k_t)$. Taking an affine translation gives

$$\tilde{u}^*\left(c_t^*, d_t^*\right) = a\left(c_t^\delta k_t^\theta\right)^\gamma w\left(\boldsymbol{C}_t^\#, \boldsymbol{K}_t^\#\right) + b$$

and setting $a = 1/\gamma$ and $b = -1/\gamma$ gives the utility function[6]

$$\tilde{u}^*\left(c_t^*, d_t^*\right) = \frac{\left(c_t^\delta k_t^\theta\right)^\gamma w\left(\boldsymbol{C}_t^\#, \boldsymbol{K}_t^\#\right) - 1}{\gamma}. \tag{9}$$

The resulting utility function thus is homothetic (an affine translation of a homogeneous function) and so is consistent with the Gorman (1953) conditions for exact aggregation of individual preferences. Exact aggregation is sufficient but not necessary for constructing an aggregate utility function.

The function $w(\boldsymbol{C}^\#, \boldsymbol{K}^\#)$ is unknown but can be approximated with arbitrary accuracy by the multivariate polynomial

$$w(\boldsymbol{C}^\#, \boldsymbol{K}^\#) = a_0 + \gamma^2 \sum_{\lambda=1}^{L} A^\lambda(\boldsymbol{C}^\#, \boldsymbol{K}^\#), \tag{10}$$

where $A^\lambda(\boldsymbol{C}^\#, K^\#)$ is itself a polynomial of all possible terms of order λ in the elements of $(\boldsymbol{C}^\#, \boldsymbol{K}^\#)$ and the accuracy of the approximation improves as L increases.[7] Under the mild restriction that the ratios c_{t-h}/c_t and k_{t-h}/k_t (equivalently, the growth rates of c and k) are bounded above and below, the multivariate polynomial is dense in a Sobolev norm. A Sobolev norm requires the function to approximate *globally* the true unknown function *and* its derivatives.[8] Approximating derivatives well is important because the Euler equations are obtained by differentiating the utility function.[9] We determine a value for L in (10) empirically by standard upward testing, adding polynomial terms until they are no longer statistically significant.

We can test specific forms of the utility function by restricting the polynomial. For example, the utility function is the standard (multigood) CRRA in c and k only if $w(\boldsymbol{C}^\#, \boldsymbol{K}^\#) = 1$, which requires either that all polynomial coefficients be zero or that $\gamma = 0$. In the latter case, the utility function is logarithmic, which is a special case of the CRRA. We therefore can test the CRRA through the two null hypotheses that the polynomial coefficients all are zero and that $\gamma = 0$. Failure to reject either null implies the CRRA.

2.3. Orthogonality Conditions

We derive orthogonality conditions from two equilibrium conditions of the model. The control variables are current expenditures on nondurables and durables, c_t and d_t. Associated with them are two first-order conditions that can be written

$$\mathcal{E}_t[MU_c(t) - MU_c(t+1)r_{t+1}] = 0 \tag{11}$$

and

$$MU_d(t) - p_{dt} MU_c(t) = 0, \tag{12}$$

where r_t equals one plus the interest rate and p_{dt} is the real cost of durable purchases. Equation (11) is the *intertemporal* first-order condition, equating the disutility of foregoing current consumption c_t to the expected utility gained from future consumption c_{t+1}. Equation (12) is the *intratemporal* first-order condition, equating the marginal rate of substitution of a unit of nondurable consumption for a unit of durable expenditure at time t to the relative price of durables and nondurables. Equation (11) yields one set of testable orthogonality conditions. In contrast, equation (12) is not useful directly for deriving testable restrictions. Because the marginal utility of durables in (12) involves expectations infinitely far into the future, we follow Dunn and Singleton (1986) and use the related condition

$$MU_{d^*}(t) = p_{d^*t} MU_c(t), \tag{13}$$

where p_{d^*t} is the real user cost of durables, $p_{dt} - (1-\mu)(1+r_t)^{-1} E_t p_{dt+1}$. This condition compresses all expectations about the future into p_{d^*t}, which involves only an expectation one period into the future.

The orthogonality conditions for our general SNP utility function given by (9) and (10) are cumbersome, and so, to give an example, we report here the conditions when (10) is quadratic and nondurables provide service flows for, at most, two periods (i.e., $J = 1$). In applications, more lags may be required. Under this example, the SNP utility function is

$$u^*\left(c_t^*, d_t^*\right) = \frac{\left(c_t^\delta k_t^\theta\right)^\gamma \left\{1 + \gamma^2 \left[a_1 \frac{c_{t-1}}{c_t} + a_2 \frac{k_{t-1}}{k_t} + a_3 \left(\frac{c_{t-1}}{c_t}\right)^2 + a_4 \left(\frac{c_{t-1}}{c_t}\right)\left(\frac{k_{t-1}}{k_t}\right) + a_5 \left(\frac{k_{t-1}}{k_t}\right)^2\right]\right\} - 1}{\gamma} \tag{14}$$

or

$$u^*(c^*, d^*) = \frac{[\text{CRRAF}_t][\text{Poly}_t] - 1}{\gamma},$$

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

where CRRAF is the constant relative risk aversion factor $(c^{\delta}k^{\theta})^{\gamma}$ and Poly is the quadratic polynomial $w(\boldsymbol{C}^{\#}, \boldsymbol{K}^{\#})$. Using (14), we have

$$MU_c(t+s) = \mathcal{E}_t\left[\frac{\delta\beta^s}{c_{t+s}}(\text{CRRAF}_{t+s})(\text{Poly}_{t+s}) + \sum_{j=0}^{1}\beta^{s+j}\frac{(\text{CRRAF}_{t+s+j})}{\gamma}\left(\frac{\partial\,\text{Poly}_{t+S+j}}{\partial c_{t+s}}\right)\right], \quad \textbf{(15)}$$

where

$$\frac{\partial\,\text{Poly}_{t+s+j}}{\partial c_{t+s}} = (-1)^{j+1}\frac{\gamma^2}{c_{t+s}}\left[a_1\left(\frac{c_{t+s+j-1}}{c_{t+s+j}}\right) + 2a_3\left(\frac{c_{t+s+j-1}}{c_{t+s+j}}\right)^2 + a_4\left(\frac{c_{t+s+j-1}}{c_{t+s+j}}\right)\left(\frac{k_{t+s+j-1}}{k_{t+s+j}}\right)\right]$$

for $j = 0$ or 1. Substituting (15) into (11) gives the intertemporal orthogonality condition used in estimation.

The SNP utility function (14) does not contain the service flow variable d^*, which has been replaced by k_t and k_{t-1} using (5). We obtain from (14)

$$MU_{d^*}(t) = \frac{\theta\beta}{k_t}(\text{CRRAF}_t)(\text{Poly}_t) + \sum_{j=0}^{1}\beta^j\frac{(\text{CRRAF}_{t+j})}{\gamma}\left(\frac{\partial\,\text{Poly}_{t+j}}{\partial k_t}\right), \quad \textbf{(16)}$$

where

$$\frac{\partial\,\text{Poly}_{t+j}}{\partial k_t} = (-1)^{j+1}\frac{\gamma^2}{k_t}\left[a_2\left(\frac{k_{t+j-1}}{k_{t+j}}\right) + a_4\left(\frac{c_{t+j-1}}{c_{t+j}}\right)\left(\frac{k_{t+j-1}}{k_{t+j}}\right) + 2a_5\left(\frac{k_{t+j-1}}{k_{t+j}}\right)^2\right]$$

for $j = 0$ or 1. Substituting (16) and (15) with $s = 0$ into (13) gives the intratemporal orthogonality conditions used in estimation.

3. DATA AND SEPARABILITY

Having derived the seminonparametric generalization of the CRRA, we now need aggregate data to estimate the Euler equations. To keep estimation tractable, we must reduce the thousands of goods consumed to a small number of aggregates. We thus must choose a sample period to analyze, a set of aggregates to construct, and an aggregation method for constructing them.

The typical approach is to use the longest sample period for which data are available, the groups of goods reported by NIPA (such as nondurables, services, and durables), and the aggregation method (Laspeyres quantity aggregates) that NIPA uses to construct those groups. This approach simply assumes, without any formal testing, the existence of a well-behaved aggregate utility function with an appropriate separability structure. An alternative approach is to examine the properties of the data with formal tests. Such tests can be either parametric or nonparametric. Parametric tests [e.g., Eichenbaum and Hansen (1990)] are attractive in that they are stochastic and have asociated with them the usual well-defined statistics. They have the limitation that one must write down explicit functional forms to test; failure to include the true form may lead to meaningless results. Varian's (1982, 1983) nonparametric tests have the opposite characteristics. They are attractive in that they are totally independent of the underlying form of the utility function, but they have a shortcoming in that they are nonstochastic and so have no associated statistics of significance. None of the three possible specification approaches (no testing, parametric testing, nonparametric testing) is perfect. We base the analysis in this study on Varian's nonparametric tests and use the results of Fleissig et al. (2000). We turn now to a brief discussion of the methodology underlying those constructs.

3.1. Nonparametric Tests

Varian's (1982, 1983) nonparametric approach tests the existence of a utility function that "rationalizes" a set of data, that is, that yields the observed consumption path as the optimal path among all feasible paths. The nonparametric tests do not depend on the form of the utility function. Varian has provided a computer program, NONPAR, for conducting the tests.

A limitation of NONPAR, already mentioned, is that the tests are deterministic rather than statistical. According to the tests, if even a single violation of revealed preference is found, GARP is rejected. Is one rejection out of the hundreds or thousands of inequalities imposed by GARP really a rejection? This question cannot be answered in the usual way of checking against a significance level. No distribution theory has yet been developed for the NONPAR tests, and so, there are no confidence intervals or levels of significance for test results. The NONPAR program tests the sufficient condition but not the necessary condition, making interpretation of rejections of GARP difficult. In particular, measurement error could lead to spurious conclusions concerning GARP consistency. However, if conclusions regarding GARP consistency arise solely from measurement error, which is random, we would not expect to see them vary in a systematic manner. We therefore report below the results of Euler equation estimation conducted over three increasingly GARP-consistent samples and check if the results change in a systematic manner across the various samples. If they do, it is less likely that measurement error is the reason for the GARP-inconsistency of the less consistent sample periods. In fact, the tests show just such systematic behavior, suggesting that measurement error is not the driving force behind our results.

The derivation of GARP assumes certainty. A rejection of GARP might arise solely because agents' state of uncertainty differed in two periods, leading them to choose different consumption bundles even if they face the same relative prices in the two periods. In such a case, our tests would indicate a failure of GARP even though the true aggregate utility function was unchanged and was consistent with the data in both periods. If this sort of problem is serious, then our GARP tests will be largely spurious and should have no systematic relation to the results we obtain from Euler equation estimation. In particular, we should see no systematic change in estimation results as we restrict the estimation to sample periods that are increasingly GARP-consistent. That we do find this kind of systematic change suggests again that the GARP tests are valid.

The tests have some limitations, but the alternative of not testing at all seems considerably more limited.

3.2. GARP-Consistency of the NIPA Data

Fleissig et al. (2000) apply NONPAR to the NIPA data on consumption. The GARP tests require prices for all goods. The relevant price for the stock of durables is the user cost, which involves expectations about future prices of the goods in question. Fleissig et al. examined perfect foresight (assumes the current expected value equals the realized future value of the random variable), static expectations (assumes the current expected value equals the current realized value, which is optimal if the random variable is a driftless random walk), and ARIMA models of expectations. They found that ARIMA models gave negative user costs and yielded results essentially like those obtained from perfect foresight. We therefore confine attention to perfect foresight and static expectations in the empirical work that follows.

The revealed preference tests of Fleissig et al. are for monthly (1959:1–1990:12), quarterly (1959:1–1990:4), and annual (1929–1990) data. All of the data fail GARP over the full periods of data availability, except for static expectations for annual data, but pass over various subsamples. For annual and quarterly data, there are long GARP-consistent subsamples under both expectations schemes. Under static expectations, the quarterly (1960:1–1990:4) data, and annual (1929–1990) data are consistent with GARP.[10] Two subsamples are GARP-consistent under perfect foresight: quarterly (1960:1–1980:4 and 1981:4–1990:4) and annual (1935–1981 and 1982–1990) data. For monthly data, there is no GARP-consistent subsample of any econometrically usable length under static expectations, but there is a 20-year GARP-consistent subsample of 1970:5–1990:12 under perfect foresight.

3.3. Separability of the Utility Function

Weak separability of the utility function is a necessary condition for aggregation. Goods within a group can be substituted for each other with no regard for the allocation of goods excluded in that particular group. This kind of substitutability is justified only if the utility function is weakly separable in the appropriate groups of goods; see Gorman (1959).

Varian (1983) provides tests of separability, and Fleissig et al. (2000) have applied them to the NIPA consumption data. For annual data, the utility function is separable in (nondurables, services), (motor vehicles), (other autos), and (remaining durables), and for quarterly data the function is separable in (nondurables, services), (motor vehicles), and (furniture, remaining durables). Thus the utility function has four arguments with annual data and three with quarterly data. Also, durables are never separable from nondurables and services, as some of the existing literature has assumed.

For monthly data, there is no separable utility function, which means no aggregation at all is permissible. This result raises questions about the validity of results based on monthly data reported in the literature, which generally uses aggregates of goods that omit durables. Fleissig et al. feel that the monthly data are simply inadequate for econometric use, and their conclusion is buttressed by Wilcox's (1992) closely related criticisms of the data. Despite these difficulties, we use monthly data to compare the results of our methods with those of the literature and also with our own results from the annual and quarterly data.

3.4. Aggregation Method

The Commerce Department uses Laspeyres quantity indices to construct the NIPA data,[11] but several better aggregation methods exist. One alternative is Divisia aggregation, which is superlative or exact for a flexible form [Diewert (1976, 1978)], allows for less than perfect substitution, and is derived from the first-order conditions of consumer optimization; see Barnett et al. (1992). Moreover, the Divisia index provides a second-order approximation to an arbitrary unknown aggregator function, whereas the Laspeyres quantity index can only give a first-order approximation. To see what difference the aggregation method makes, we compare results from NIPA aggregates and Divisia aggregation.

4. ESTIMATION RESULTS

We turn now to the estimation. We seek answers to three groups of questions: (1) Are the estimation results affected by imposing GARP-consistent sample periods and separability structures, (2) does the SNP utility function offer significant improvement over the standard CRRA function, and (3) what are the final estimation results and their implications?

We follow Dunn and Singleton (1986) in scaling the orthogonality conditions to correct for nonstationarity induced by economic growth. The scaling is done by dividing (15) and (16) by $c_t^{\delta\gamma-1}k_t^{(1-\delta)\gamma}$, making the disturbances in (11) and (13) functions of the ratios c_{t+j}/c_t, d_{t+j}/d_t, and $c_t/(p_{dt}d_t)$. The disturbances then will be stationary if purchases of nondurables and durables grow according to geometric trends. Estimation is by iterated GMM using TSP International. The convergence criterion is set at 0.0001. To correct for any autocorrelation in the errors, we use Gallant's (1987) estimator for the weighting matrix. The data are from NIPA.

As mentioned earlier, we add terms to the polynomial (10) until they no longer are statistically significant. We started with the quadratic (14), but none of the quadratic terms ever was significant, so we report here only the fits for the linear model. The final utility functions for annual and quarterly data are, respectively,

$$u^*\left(c_t^*, d1_t^*, d2_t^*, d3_t^*\right) = \frac{\left(c_t^{\delta} k1_t^{\theta_1} k2_t^{\theta_2} k3_t^{\theta_3}\right)^{\gamma} \left\{1 + \gamma^2 \left[a_1 \frac{c_{t-1}}{c_t} + a_2 \frac{k1_{t-1}}{k1_t} + a_3 \frac{k2_{t-1}}{k2_t} + a_4 \frac{k3_{t-1}}{k3_t}\right]\right\} - 1}{\gamma} \quad \textbf{(17)}$$

and

$$u^*\left(c_t^*, d1_t^*, d2_t^*\right) = \frac{\left(c_t^{\delta} k1_t^{\theta_1} k2_t^{\theta_2}\right)^{\gamma} \left\{1 + \gamma^2 \left[a_1 \frac{c_{t-1}}{c_t} + a_2 \frac{k1_{t-1}}{k1_t} + a_3 \frac{k2_{t-1}}{k2_t}\right]\right\} - 1}{\gamma}. \quad \textbf{(18)}$$

We discuss the monthly equation later.

4.1. Annual Data

We impose the separability structure found by Fleissig et al. and use as arguments of the utility function (nondurables and services), (motor vehicles), (other autos), and (remaining durables). The latter three together constitute the set of durable goods, but the separability tests indicate that they must not be aggregated into one overall durable good. Later we explore the consequences of using such an overall aggregate. We estimate Euler equations derived from (17) over both GARP-consistent and total sample periods and also with both NIPA and Divisia aggregation to compare results. Estimates of the CRRA parameter γ and the diagnostic test statistics differ substantially across sample periods and expectations methods, but the other parameter estimates and their standard errors are about the same. Therefore, for comparing fits, Table 4 presents only the estimates for γ and the test statistics; estimates of the other parameters will be discussed presently.

Panel A of Tables 1 and 2 presents the results for the annual data. The joint test for the SNP polynomial terms always strongly rejects the null hypothesis of insignificance (Wald1, Table 1). Thus the CRRA specification often used in the literature [e.g., Hansen and Singleton (1983), Dunn and Singleton (1986)] seems misspecified for the annual data. We discuss the consequences of using the CRRA presently.

The concavity tests (Table 2) report percentage of years for which the Hessian was not negative semidefinite. Under static expectations, the full sample is GARP-consistent, so only one set of concavity results is reported. The concavity results suggest that the SNP function is inconsistent with the data under static expectations. Under NIPA aggregation, the SNP function fails concavity 50% of the time, and under Divisia aggregation, it fails 100% of the time. Under perfect foresight, the full sample is not GARP-consistent, and so, Table 2 reports three sets of

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

results: one for the full sample (the one used in most of the literature), one for a possibly GARP-consistent sample, and one for the longest definitely GARP-consistent sample. Imposition of GARP-consistency has dramatic effects. Over the full sample, neither the CRRA nor the SNP ever fails the concavity restriction. As the sample is made more GARP-consistent, the SNP function begins violating the concavity restriction with NIPA data. Over the partially GARP-consistent and full sample, it violates 38% and 100% of the time, respectively. In contrast, the SNP function never violates concavity when aggregation is by Divisia. We have strong prior reasons to prefer Divisia aggregation, and so, it seems that the Divisia results are the relevant ones and that the SNP function passes the concavity tests. Apparently, both GARP-consistency and the aggregation method make a difference to estimation results.

The only SNP function that produces useful estimates for the annual data is that for the 1935–1981 sample under perfect foresight and Divisia aggregation. Table 3 reports the full estimation results for the Euler equations derived from (17) in that case. Wald2 is a joint test of the significance of the durables polynomial coefficients a_2, a_3, and a_4, and Wald3 is a t-test of the significance of the nondurables and services coefficient a_1. The null hypothesis is rejected, indicating significance of coefficients. Their negative signs are consistent with habit persistence. The parameter γ is significantly different from zero, and so, consumer preferences are not logarithmically separable over the annual data, which contradicts the results reported by Dunn and Singleton (1986) for monthly data. The discount factor β is significantly below one, implying a significantly positive rate of time preference, a result frequently not found in the literature.

The standard CRRA is not rejected by the overidentifying restrictions or the concavity tests (except for NIPA aggregation under static expectations). Furthermore, the estimates of the parameters δ, θ_1, θ_2, and θ_3 are much the same as with the SNP function. However, the CRRA specification is rejected, for the Wald1 test strongly supports significance of the SNP linear polynomial terms. The main difference between the CRRA and SNP functions is the very different estimates of γ. For the CRRA, γ always is positive except under NIPA aggregation with static expectations, when it is insignificantly negative. In contrast, γ always is significantly negative for the SNP utility function.

This difference in the sign of γ has an important economic implication. Under quite mild restrictions on preferences and probabilistic beliefs, any preference ordering can be represented by a utility function and conversely. Arrow (1971) calls this conclusion the Expected Utility Theorem. Arrow proves the remarkable Utility Boundedness Theorem, which states that any utility function satisfying the conditions of the Expected Utility Theorem must be bounded above and below. Boundedness of both CRRA and SNP utility depends on the sign of γ. When γ is positive, both functions are unbounded above as consumption goes to infinity but bounded below (by $-1/\gamma$) as consumption goes to zero, whereas both are bounded above (again by $-1/\gamma$) but unbounded below when γ is negative. Unboundedness below as c goes to zero is not important, because both utility functions can be made bounded below by a minor translation that has no impact on any other important properties of the utility function or on estimation.[12] Our estimates of γ then imply that the CRRA is unbounded above and inconsistent with Arrow's Theorem, whereas the SNP function (17) is bounded above and consistent with the theorem.

There are two ways to look at the differing results for static expectations and perfect foresight. An SNP function can globally approximate the "true" utility function arbitrarily well; given the insignificance of higher-order terms, our linear polynomial SNP therefore should be very accurate. If one is willing to maintain the hypothesis of maximization of concave utility by a representative agent, then the concavity failures under static expectations indicate the inadequacy of that expectations scheme. On the other hand, if one is not willing to take a stand on the validity of either expectations scheme (because neither really is likely to be correct), then the differing concavity results leave open the question of whether maximization of a concave utility function by a representative agent is a useful model of the aggregate data.

Finally, the NONPAR tests assume time-separable preferences. A failure of GARP might therefore indicate that preferences are not time-separable rather than that no rationalizing utility function exists at all. Our results, taken as a whole, suggest that is not what is going on. If non-time-separability were the reason for the GARP failures, then the nonseparable SNP function should take care of the problem and should show no systematic differences in estimation results over increasingly GARP-consistent sample periods. As we have seen, however, such differences do occur, strikingly, implying that GARP-inconsistency really does indicate nonexistence of a rationalizing utility function over the sample periods in question.

4.2. Quarterly Data

With quarterly data, Fleissig et al. find the utility function is separable in (nondurables, services), (motor vehicles), and (furniture, remaining durables). The results are essentially the same as for the annual data. For GARP-consistent samples and Divisia aggregation, we have no violations of concavity at all, irrespective of expectations scheme. Thus, there is no conflict in the concavity results for the two expectations schemes, in contrast to the annual data. Also, for GARP-consistent samples and Divisia aggregation, Wald1 implies significance of the SNP linear polynomial terms. Once again, we see important effects of imposing GARP-consistency and appropriate aggregation: For perfect foresight expectations, Wald1 fails to reject the null of insignificance for the SNP polynomial under NIPA aggregation, irrespective of sample period, and for Divisia aggregation over the GARP-inconsistent sample period. As with the annual data, the overidentifying restrictions never are rejected. Also, the estimates of γ always are significantly positive for the CRRA and significantly negative for the SNP (except for Divisia perfect foresight over the GARP-inconsistent sample period, when γ is insignificantly negative).

The Divisia, GARP-consistent estimates are reported in Table 3 for both static expectations and perfect foresight. There is no substantive difference between the two sets of results, and so, it does not matter much that we have no way to choose between them.

In summary, the results for quarterly data are essentially the same as those for annual data. The only substantive difference between them is that the a_1 coefficient on nondurables and services is insignificant for the quarterly data, in contrast to the result for the annual data. This result is a bit odd in that it implies that consumers display habit persistence in durables but not in nondurables and services. Why such a dichotomy should prevail is unclear.

4.3. Monthly Data

Much of the Euler equation literature has used monthly data [e.g., Hansen and Singleton (1982, 1983), Dunn and Singleton (1986), Eichenbaum et al. (1988), Gallant and Tauchen (1989), and Eichenbaum and Hansen (1990)]. Monthly data are appealing because it seems likely that they are closer to the decision period of the representative agent than are either quarterly or annual data. However, Fleissig et al.'s (in press) results suggest that monthly data are mostly inconsistent with GARP and may be inappropriate for Euler equation estimation. Also, no groupings of the monthly data satisfied the GARP separability tests, which means that utility function estimation ought to use as independent consumption variables all 13 categories of consumer expenditures reported by Citibase. No previous Euler equation studies have taken account of these characteristics of the data. All have used sample periods inconsistent with GARP, and all have used aggregates of consumption inconsistent with the separability results (typically total expenditure on nondurables and services as one variable and in some cases total expenditures on durables as another).

Under perfect foresight, we use a moderately long sample of 1970:5–1990:12 for the monthly data and also report for comparison purposes results for the sample 1960:1–1990:12, which is the longest GARP-consistent sample for the quarterly data. We use the same groupings as for the quarterly data: (nondurables, services), (motor vehicles), and (furniture, remaining durables) because entering the 13 categories of consumption expenditure would make the problem intractable.

The monthly results are reported in Tables 1C and 2C. Concavity is violated in every case except for the SNP utility function under NIPA aggregation with static expectations over a GARP-inconsistent sample period. The overidentifying restrictions always are strongly rejected. The model clearly is unacceptable.

If we had only the monthly results, we would be unable to say much about why the model fails. It could be that there are problems with the data, or it could be that the model is misspecified. However, based on the results of Fleissig et al. (in press) and Wilcox (1992), it seems most likely that the monthly results stem from problems with the data. The annual and quarterly results are consistent with each other and differ sharply from the monthly results.

4.4. Separability and Excessive Aggregation

Through all the foregoing tests, we impose the separability structure suggested by Fleissig et al.'s (2000) GARP tests. Several articles in the literature have used a single total measure of durables, and so, we now turn to a brief examination of how sensitive the estimation is to overaggregating durables into a single quantity.

Table 4 reports the results. Point estimates of γ usually are substantially larger than in Table 1, sometimes over 30 times as large; they also sometimes have the opposite sign from the corresponding estimates in Table 4. These differences often are statistically significant. For the quarterly data, there are some rejections of the overidentifying restrictions, whereas previously there were none. Overaggregation therefore does lead to some deterioration of the estimation.

The differences between the magnitudes of γ reported in Tables 1 and 4 are economically significant. In steady state, $\gamma - 1$ is both the index of relative risk aversion and the elasticity of marginal utility with respect to total expenditure. Both interpretations make γ useful in applied economic analysis. For example, interpreted as the index of risk aversion, γ is used in evaluating public policies that affect the probability of uncertain events [Freeman (1993)]. Interpreted as the elasticity of marginal utility of total expenditure, γ is used to assess the effect of economic development on patterns of consumption [Lluch et al. (1977)]. Changing the magnitude of γ by a factor of two or three would have a substantial impact on any such evaluation, but the estimated magnitude of γ changes by a factor of 30 or even more simply because of overaggregation of durable goods. Clearly, failure to restrict the estimation appropriately has an effect on results that is economically important.

5. CONCLUSION

In this study, we derive an SNP utility function to explore several issues in Euler equation estimation and to examine the adequacy of the often-used CRRA utility function. We find that estimation results are sensitive to restricting sample periods to be consistent with GARP, to the separability structure imposed on the utility function, and to the aggregation method used to construct the U.S. consumption data. We also show that the SNP function can reverse conclusions from the CRRA often found to be inconsistent with economic theory. Our results suggest that the CRRA function is severely misspecified. The main results are summarized below:

(1) According to microeconomic theory, a nonsatiated, continuous, concave, monotonic aggregate utility function does not exist unless the data satisfy GARP. Evidence presented by Fleissig et al. (in press) suggests that the NIPA data on consumption generally do not satisfy GARP over the full periods of data availability. That evidence presumes a time-separable utility function and can only be treated as a guideline for our estimation of the non-time-separable SNP function. However, using GARP tests is preferable to the usual approach of doing no testing at all and is not obviously inferior to the alternative approach of using parametric tests. Our results suggest that restricting sample periods to those consistent with GARP has important effects on estimation results.

(2) Tests of the utility function's separability structure suggest that it has as arguments two or three separate categories of durables. Using a single aggregate of total durable expenditure leads to a deterioration of results, although the effects are not as severe as using GARP-inconsistent sample periods.
(3) Aggregation theory suggests that the aggregate Laspeyres quantity indexes used by NIPA are deficient. No other Euler equation study except for that by Hayashi (1982) has used anything but NIPA aggregation. We find, as did Hayashi, that using Divisia aggregates instead of NIPA aggregates affects the results.
(4) Our results are consistent with those of Fleissig et al.'s (1997) conclusion that monthly NIPA data are of little value in empirical work. The annual and quarterly estimation results are well behaved and mutually consistent; the monthly results are not well behaved and are inconsistent with those for annual and quarterly data.
(5) When we confine attention to appropriate sample periods, separability structures, and aggregation methods, and to annual and quarterly data, we find that the CRRA utility function is rejected in favor of an SNP form that nests the CRRA. The preferred SNP function includes a linear polynomial in lagged service flows from nondurables and durables. The nondurables term is marginally significant in the annual data but not in the quarterly data, suggesting that measured nondurables may contain durable elements but probably do not. The durables term in the SNP function always is significant.
(6) The SNP function is bounded, whereas the CRRA is not. Boundedness is required for a utility function to be consistent with certain aspects of behavior toward risk [Arrow (1971)].
(7) The point estimates of the rate of time preference are always positive and usually significantly so, results often not obtained in the Euler equation literature.

Our overall conclusion is that imposing the restrictions implied by revealed preference theory has significant effects on estimation, leading to results often different from those reported in the existing literature. This conclusion suggests a need for further research on two fronts. First, revealed-preference theory needs to be extended. The existing theory presumes time separability, but our evidence, as well as that in previous empirical literature, supports non-time-separability of preferences. Nonetheless, restricting estimation of our non-time-separable utility function to sample periods and separability structures consistent with the current, theoretically limited version of GARP substantially affects estimation results. The nonparametric test results concerning GARP inconsistency of the data and the utility function's separability structure across goods thus seem to reflect something more than the failure of those tests to account for non-time-separability. It therefore would be most useful to extend revealed-preference theory to encompass non-time-separable preferences. Second, further empirical research should examine the sensitivity of our results to alternative instrument sets and the inclusion of leisure in the utility function.

NOTES

1. For example, Hansen and Singleton (1982) and Gallant and Tauchen (1989) are special cases that omit durable goods, and Eichenbaum and Hansen (1990) and Dunn and Singleton (1986) only consider two composite goods: (1) services plus nondurables and (2) durables.

2. See Abowd and Card (1987), for example. We omit leisure because of the difficulty of measuring the exact nature of the employment decision and the opportunity cost of time [e.g., Swofford and Whitney (1987); Gordon (1973)]. Including leisure would be a useful extension.

3. The terminal condition together with the period-by-period budget constraint implies the expected lifetime budget constraint, but the converse implication does not hold. We do not use the terminal condition in deriving the Euler equations that we test later, so we do not dwell here on these details. See Blanchard and Fischer (1989, p. 286) for further discussion.

4. In the empirical work that follows, we use the BEA estimates for μ_i.

5. See Gallant and Nychka (1987) and Gallant and Tauchen (1989) for details on seminonparametric utility functions and their estimation.

6. Setting $b = -1/\gamma$ is necessary when $u(\boldsymbol{C}, \boldsymbol{K}) = 1$ and γ tends to zero.

7. For example, if $j = 1$, then $A^1(\boldsymbol{C}^\#, \boldsymbol{K}^\#) = a_{1,1}(c_{t-1}/c_t) + a_{1,2}[1 + (k_{t-1}/k_t)]$, $A^2(\boldsymbol{C}^\#, \boldsymbol{K}^\#) = a_{2,1}(c_{t-1}/c_t)^2 + a_{2,2}(c_{t-1}/c_t)[1 + (k_{t-1}/k_t)] + a_{2,3}[1 + (k_{t-1}/k_t)]^2$ and so on. If $j = 2$, then $A^1(\boldsymbol{C}^\#, \boldsymbol{K}^\#) = a_{1,1}(c_{t-1}/c_t) + a_{1,2}[1 + (c_{t-2}/c_t)] + a_{1,3}[1 + (k_{t-1}/k_t)]$, $A^2(\boldsymbol{C}^\#, \boldsymbol{K}^\#) = a_{2,1}(c_{t-1}/c_t)^2 + a_{2,2}(c_{t-1}/c_t)^2 + a_{2,3}[1 + (k_{t-1}/k_t)]^2 + a_{2,4}(c_{t-1}/c_t)(c_{t-2}/c_t) + a_{2,5}(c_{t-1}/c_t)[1 + (k_{t-1}/k_t)] + a_{2,6}(c_{t-2}/c_t)[1 + (k_{t-1}/k_t)]$, and so on.

8. This requirement goes beyond Diewert's (1974) definition for a flexible form, i.e., a local second-order approximation to the unknown utility function at some fixed point x^* and no requirements about the degree of approximation for the derivatives of the function. Examples of Diewert-flexible functional forms are the translog and AIDS models. See Gallant (1981), Gallant and Nychka (1987), and Gallant and Tauchen (1989) for discussions of Sobolev norms.

9. Despite its considerable generality, the SNP function has some limitations, as does any specific utility function. An important one, given some of the recent literature, is that the SNP function maintains a link between the degree of intertemporal substitution and the degree of risk aversion. Extending our work to consider other kinds of utility functions without that link would be useful.

10. It is interesting that the violations of GARP occur in the early 1930s and in 1981. Several often-mentioned episodes in economic history, in particular the October 1979 change in Fed operating procedure, produce no GARP violations and therefore cause no special problems in estimating Euler equations.

11. Recently, the BEA has also begun to publish national accounts data using a chained Fisher ideal index. Since the Fisher ideal index and the Divisia index are both superlative indices, our results, which use the Divisia index, should be similar to those based on the Fisher ideal index.

12. Write the CRRAF in both functions as $(c^\delta k^\theta + 1)^\gamma$. For the magnitudes of c and k in the data, adding 1 has an imperceptible effect on estimation, but both utility functions now go to zero as either c or k goes to zero.

REFERENCES

Abowd, J.M. & D. Card (1987) Intertemporal labor supply and long-term employment contracts. *American Economic Review* 77, 50–68.

Arrow, K.J. (1971) Essays in the Theory of Risk-Bearing. Chicago: Markham.

Barnett, W.A., D. Fisher & A. Serletis (1992) Consumer theory and the demand for money. *Journal of Economic Literature* 4, 2086–2119.

Blanchard, O.J. & S. Fischer (1989) *Lectures in Macroeconomics*. Cambridge, MA: MIT Press.

Darby, M.R. (1974) The permanent income theory of consumption—a restatement. *Quarterly Journal of Economics* 88, 228–250.

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

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

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

Dunn, K.B. & K.J. Singleton (1986) Modelling the term structure of interest rates under nonseparable utility and durability of goods. *Journal of Financial Economics* 17, 27–55.

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

Eichenbaum, M.S. & L.P. Hansen (1990) Estimating models with intertemporal substitution using aggregate time series data. *Journal of Business and Economic Statistics* 8, 53–69.

Eichenbaum, M.S., L.P. Hansen & K.J. Singleton (1988) A time series analysis of representative agent models of consumption and leisure choice under uncertainty. *Quarterly Journal of Economics* 103, 51–78.

Fleissig, A.R., A.R. Hall & J.J. Seater (2000) GARP, separability, and the representative agent. *Macroeconomic Dynamics* 4, 324–342.

Freeman, A.M. (1993) *The Measurement of Environmental and Resource Values*. Washington, DC: Resources for the Future.

Gallant, A.R. (1981) On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form. *Journal of Econometrics* 15, 211–245.

Gallant, A.R. (1987) *Nonlinear Statistical Models*. New York: John Wiley.

Gallant, A.R. & D.W. Nychka (1987) Semi-nonparametric maximum likelihood estimation. *Econometrica* 55, 363–390.

Gallant, A.R. & G. Tauchen (1989) Semi-nonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing implications. *Econometrica* 57, 1091–1120.

Gordon, R.J. (1973) The welfare cost of higher unemployment. *Brookings Papers on Economic Activity* Spring, 133–195.

Gorman, W.M. (1953) Community preference fields. *Econometrica* 21, 63–80.

Gorman, W.M. (1959) Separable utility and aggregation. *Econometrica* 27, 469–481.

Hansen, L.P. & K.J. Singleton (1982) Generalized instrumental variables estimation of non-linear rational expectations models. *Econometrica* 50, 1269–1286.

Hansen, L.P. & K.J. Singleton (1983) Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy* 91, 249–265.

Hayashi, F. (1982) The permanent income hypothesis: Estimation and testing by instrumental variables. *Journal of Political Economy* 90, 895–916.

Lluch, C., A.A. Powell & R.A. Williams (1977) *Patterns in Household Demand and Saving*. New York: Oxford University Press.

Swofford, J.L. & G.A. Whitney (1987) Non-parametric tests of utility maximization and weak separability for consumption, leisure, and money. *Review of Economics and Statistics* 69, 458–464.

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

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

Wilcox, D.W. (1992) The construction of U.S. consumption data: Some facts and their implications for empirical work. *American Economic Review* 82, 922–941.

Table 1. Comparative regression results for CRRA and linear polynomial SNP utility functions[a]

A. Annual data[b]

	CRRA		Linear polynomial SNP				
Time period	γ	OIR	γ	OIR	Wald1	Wald2	Wald3
			NIPA static expectations				
1931–1990	−0.007	23.1	−0.420	22.6	102.690	35.240	−0.862
	(0.015)	[51.4]	(0.055)	[47.0]	{0.000}	{0.000}	{0.389}
			Divisia static expectations				
1931–1990	0.127	24.2	−0.376	23.7	70.763	15.289	1.811
	(0.019)	[51.4]	(0.100)	[47.0]	{0.000}	{0.002}	{0.070}
			NIPA perfect foresight				
1931–1990	0.028	22.2	−0.092	22.4	47.880	31.943	−2.858
	(0.023)	[51.4]	(0.030)	[47.0]	{0.000}	{0.000}	{0.004}
1935–1990	0.071	22.0	−0.150	20.9	33.822	33.035	3.932
	(0.013)	[51.4]	(0.024)	[47.0]	{0.000}	{0.000}	{0.000}
1935–1981	0.058	20.8	−0.163	20.3	47.880	31.943	1.120
	(0.011)	[51.4]	(0.023)	[47.0]	{0.000}	{0.000}	{0.263}
			Divisia perfect foresight				
1931–1990	0.083	23.5	−0.118	23.2	10.124	8.786	−3.142
	(0.025)	[51.4]	(0.023)	[47.0]	{0.038}	{0.033}	{0.002}
1935–1990	0.133	22.7	−0.210	16.7	22.751	20.550	−2.891
	(0.015)	[51.4]	(0.052)	[47.0]	{0.000}	{0.000}	{0.002}
1935–1981	0.130	20.0	−0.155	20.0	5.613	8.066	−2.369
	(0.015)	[51.4]	(0.034)	[47.0]	{0.018}	{0.045}	{0.018}

[a] γ = utility function parameter [see equation (14)], the standard error is in parentheses.
Wald1 = test of all linear polynomial terms; the p-value is in braces.
Wald2 = test of durable-goods polynomial terms; the p-value is in braces.
Wald3 = test of nondurable-goods polynomial term; the p-value is in braces.
OIR = χ_2 test of overidentifying restrictions, with 1% critical value in brackets.
[b] The instruments set is {constant, c_1^*, c_2^*, c_3^*, c_4^*, p_1^*, p_2^*, p_3^*, p_4^*}, where $x_i^* = (x_{it} - x_{it-1})/x_{it-1}$ for $i = 1, \ldots, 4$ and where
c_{1t} = aggregate of nondurables and services,
c_{2t} = aggregate of motor vehicles,
c_{3t} = aggregate of other autos,
c_{4t} = aggregate of other durables,
p_{it} $i = 1, \ldots, 4$ are the corresponding price aggregates.
There are four equations and nine instruments. CRRA has five parameters; SNP has nine parameters. Therefore, CRRA has $4 \times 9 - 5 = 31$ overidentifying restrictions, and SNP has $4 \times 9 - 9 = 27$ overidentifying restrictions.

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

TABLE 1. (Continued.)

B. Quarterly data[c]

Time period	CRRA γ	CRRA OIR	Linear polynomial SNP γ	Linear polynomial SNP OIR	Wald1	Wald2	Wald3
			NIPA static expectations				
1960:01–1990:04	0.935	12.4	−1.064	11.9	9.449	7.800	0.851
	(0.106)	[33.4]	(0.202)	[29.1]	{0.024}	{0.020}	{0.395}
			Divisia static expectations				
1960:01–1990:04	0.895	12.2	−0.617	11.3	23.885	6.782	0.821
	(0.080)	[33.4]	(0.109)	[29.1]	{0.000}	{0.034}	{0.412}
			NIPA perfect foresight				
1960:01–1990:04	1.476	11.6	−0.291	11.8	2.553	1.779	−1.280
	(0.098)	[33.4]	(0.119)	[29.1]	{0.466}	{0.411}	{0.201}
1960:01–1980:04	1.656	9.2	0.124	8.9	5.419	4.330	2.178
	(0.046)	[33.4]	(0.025)	[29.1]	{0.144}	{0.115}	{0.029}
			Divisia perfect foresight				
1960:01–1990:04	1.418	12.5	−0.013	12.2	3.786	2.431	−0.159
	(0.122)	[33.4]	(0.079)	[29.1]	{0.286}	{0.296}	{0.868}
1960:01–1980:04	1.406	9.1	−0.398	8.9	16.552	11.361	−1.238
	(0.056)	[33.4]	(0.127)	[29.1]	{0.001}	{0.003}	{0.216}

[c] For quarterly data, the instrument set is {constant, $c_1^*, c_2^*, c_3^*, p_1^*, p_2^*, p_3^*$} where $x_i^* = (x_{it} - x_{it-1})/x_{it-1}$ for $i = 1, \ldots, 4$ and where
c_{1t} = aggregate of nondurables and services,
c_{2t} = aggregate of motor vehicles,
c_{3t} = aggregate of other durables,
p_{it} $i = 1, 2, 3$ are the corresponding price aggregates.
There are three equations and seven instruments. CRRA has four parameters; SNP has seven parameters. Therefore, CRRA has $3 \times 7 - 4 = 17$ overidentifying restrictions, and SNP has $3 \times 7 - 7 = 14$ overidentifying restrictions.

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

Table 1. (Continued.)

C. Monthly data[d]

	CRRA		Linear polynomial SNP				
Time period	γ	OIR	γ	OIR	Wald1	Wald2	Wald3
			NIPA static expectations				
1960:1–1990:12	3.570	88.8	22.528	85.3	8.438	8.334	1.530
	(0.715)	[33.4]	(3.931)	[29.1]	{0.038}	{0.016}	{0.126}
			Divisia static expectations				
1960:1–1990:12	7.566	81.9	21.115	56.8	14.396	14.39	2.873
	(0.819)	[33.4]	(2.800)	[29.1]	{0.002}	{0.001}	{0.090}
			NIPA perfect foresight				
1960:1–1990:12	2.869	84.6	18.642	77.1	8.802	8.703	1.908
	(0.002)	[33.4]	(3.163)	[29.1]	{0.032}	{0.012}	{0.167}
1970:5–1990:12	0.580	72.1	11.448	70.7	3.571	3.568	0.091
	(0.520)	[33.4]	(3.065)	[29.1]	{0.311}	{0.168}	{0.763}
			Divisia perfect foresight				
1960:1–1990:12	7.594	75.1	−2.660	70.8	1.554	1.461	0.517
	(0.875)	[33.4]	(1.404)	[29.1]	{0.670}	{0.482}	{0.678}
1970:5–1990:12	3.909	72.1	12.175	56.1	3.928	3.923	−0.141
	(0.767)	[33.4]	(3.107)	[29.1]	{0.269}	{0.269}	{0.888}

[d] The instrument set, number of parameters, number of equations, and therefore the number of overidentifying restrictions are the same as for quarterly data.

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

Table 2. Violations of concavity[a]

A. Annual data

	1931–1990		1935–1990		1935–1981	
Method[b]	CRRA	SNP	CRRA	SNP	CRRA	SNP
SSE	100	50	—	—	—	—
DSE	0	100	—	—	—	—
SPF	0	0	0	38	0	100
DPF	0	0	0	0	0	0

B. Quarterly data

	1960:1–1990:4		1960:1–1980:4	
Method[b]	CRRA	SNP	CRRA	SNP
SSE	0	0	—	—
DSE	0	0	—	—
SPF	0	0	0	100
DPF	0	0	0	0

C. Monthly data

	1960:1–1990:12		1970:5–1980:4	
Method[b]	CRRA	SNP	CRRA	SNP
SSE	100	2	—	—
DSE	100	84	—	—
SPF	100	22	48	26
DPF	100	100	100	45

[a] Numbers are percentages of observation periods for which the Hessian matrix was not negative semidefinite.
[b] DPF = Divisia perfect foresight, DST = Divisia static expectation, SPF = NIPA perfect foresight, and SST = NIPA static expectations.

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

TABLE 3. Preferred models: Full estimation results

A. Annual data

Divisia Perfect Foresight 1935–1981[a]:

$$u\left(c_t^*, d1_t^*, d2_t^*, d3_t^*\right) = \frac{1}{\gamma}\left(c_t^{\delta} k1_t^{\theta_1} k2_t^{\theta_2} k3_t^{\theta_3}\right)^{\gamma}\left\{1+\gamma^2\left[a_1\frac{c_{t-1}}{c_t}+a_2\frac{k1_{t-1}}{k1_t}+a_3\frac{k2_{t-1}}{k2_t}+a_4\frac{k3_{t-1}}{k3_t}\right]\right\}$$

Parameter	Estimate	Standard error
β	0.977	0.009
γ	−0.155	0.034
δ	0.154	0.008
θ_1	0.388	0.019
θ_2	0.244	0.009
a_1	−2.011	0.849
a_2	−10.939	4.531
a_3	−2.978	1.053
a_4	−9.240	3.735

B. Quarterly data

Divisia Static Expectations 1960:1–1990:4[b]:

$$u\left(c_t^*, d1_t^*, d2_t^*\right) = \frac{1}{\gamma}\left(c1_t^{\delta} k1_t^{\theta_1} k2_t^{\theta_2}\right)^{\gamma}\left\{1+\gamma^2\left[a_1\frac{c_{t-1}}{c_t}+a_2\frac{k1_{t-1}}{k1_t}+a_3\frac{k2_{t-1}}{k2_t}\right]\right\}$$

Parameter	Estimate	Standard error
β	0.992	0.002
γ	−0.617	0.109
δ	0.254	0.009
θ_1	0.404	0.005
a_1	0.279	0.340
a_2	−1.351	0.522
a_3	−1.132	0.445

[a] $\theta_3 = 1 - \delta - \theta_1 - \theta_2$ because the parameters are constrained to result in degree-one homogeneity; c = aggregate of nondurables and services; $k1$ = motor vehicles; $k2$ = other autos; and $k3$ = other durables.
[b] $\theta_2 = 1 - \delta - \theta_1$ because the parameters are constrained to result in degree-one homogeneity; c = aggregate of nondurables and services; $k1$ = motor vehicles; and $k2$ = other durables.

TABLE 3. (Continued.)

Divisia Perfect Foresight 1960:1–1980:4[c]:

$$u\left(c_t^*, d1_t^*, d2_t^*\right) = \frac{1}{\gamma}\left(c_t^{\delta} k1_t^{\theta_1} k2_t^{\theta_2}\right)^{\gamma}\left\{1 + \gamma^2\left[a_1 \frac{c_{t-1}}{c_t} + a_2 \frac{k1_{t-1}}{k1_t} + a_3 \frac{k2_{t-1}}{k2_t}\right]\right\}$$

Parameter	Estimate	Standard error
β	0.988	0.002
γ	−0.398	0.127
δ	0.356	0.011
θ_1	0.351	0.005
a_1	−2.059	1.663
a_2	−2.266	1.199
a_3	0.371	1.765

[c] $\theta_2 = 1 - \delta - \theta_1$ because the parameters are constrained to result in degree-one homogeneity; c = aggregate of nondurables and services; $k1$ = motor vehicles; and $k2$ = other durables.

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

TABLE 4. Effects of excessive aggregation

A. Annual data		
	Linear polynomial SNP	
Time period	γ	OIR
	NIPA static expectations	
1935–1990	−0.911	10.995
	(0.224)	[15.086]
	Divisia static expectations	
1935–1990	−0.282	7.899
	(0.009)	[15.086]
	NIPA perfect foresight	
1931–1990	1.228	10.727
	(0.261)	[15.086]
1935–1990	−2.158	3.844
	(0.020)	[15.086]
1935–1981	2.773	8.181
	(1.353)	[15.086]
	Divisia perfect foresight	
1931–1990	−3.641	8.901
	(1.207)	[15.086]
1935–1990	−3.591	7.790
	(1.093)	[15.086]
1935–1981	−3.549	7.336
	(1.231)	[15.086]
B. Quarterly data		
	NIPA static expectations	
1960:1–1990:4	−1.014	19.462
	(0.358)	[15.086]
	Divisia static expectations	
1960:1–1990:4	−0.639	17.883
	(0.184)	[15.086]
	NIPA perfect foresight	
1960:1–1990:4	1.194	11.560
	(0.310)	[15.086]
1960:1–1980:4	1.288	11.351
	(0.298)	[15.086]
	Divisia perfect foresight	
1960:1–1990:4	1.147	11.488
	(0.345)	[15.086]
1960:1–1980:4	1.210	10.098
	(0.292)	[15.086]

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press

TABLE 4. (Continued.)

C. Monthly data

	Linear polynomial SNP	
Time period	γ	OIR
	NIPA static expectations	
1960:1–1990:4	−0.962	90.488
	(0.636)	[18.475]
	Divisia static expectations	
1960:1–1990:4	27.683	31.927
	(4.541)	[18.475]
	NIPA perfect foresight	
1960:1–1990:4	−1.591	74.817
	(0.550)	[18.475]
	Divisia perfect foresight	
1960:1–1990:4	−3.709	59.676
	(0.602)	[18.475]

https://doi.org/10.1017/S1365100500017077 Published online by Cambridge University Press