Macroeconomic Dynamics, **21**, 2017, 1119–1140. Printed in the United States of America.
doi:10.1017/S1365100515000085

TESTING CONSUMPTION OPTIMALITY USING AGGREGATE DATA

FÁBIO AUGUSTO REIS GOMES
University of São Paulo—Ribeirão Preto

JOÃO VICTOR ISSLER
Graduate School of Economics—EPGE,
Getulio Vargas Foundation

This paper tests the optimality of consumption decisions at the aggregate level, taking into account popular deviations from the canonical constant-relative-risk-aversion (CRRA) utility function model—rule of thumb and habit. First, we provide extensive empirical evidence of the inappropriateness of linearization and testing strategies using Euler equations for consumption—a drawback for standard rule-of-thumb tests. Second, we propose a novel approach to testing for consumption optimality in this context: nonlinear estimation coupled with return aggregation, where rule-of-thumb behavior and habit are special cases of an all-encompassing model. We estimated 48 Euler equations using GMM. At the 5% level, we only rejected optimality twice out of 48 times. Moreover, out of 24 regressions, we found the rule-of-thumb parameter to be statistically significant only twice. Hence, lack of optimality in consumption decisions represent the exception, not the rule. Finally, we found the habit parameter to be statistically significant on four occasions out of 24.

Keywords: Consumption Optimality, Intertemporal Substitution, Risk Aversion, Aggregate Return, Habit, Rule-of-Thumb Behavior, Representative-Consumer Model.

1. INTRODUCTION

For the U.S. economy, there has been a large early literature using time-series data rejecting optimizing behavior in consumption, which generated some relevant

We are especially grateful to two anonymous referees, Caio Almeida, Marco Bonomo, Luis Braido, Carlos E. Costa, Russell Davidson, Pedro C. Ferreira, Karolina Goraus, Fredj Jawadi (Editor), Anwar Khyat, and Naércio A. Menezes Filho for their comments and suggestions on earlier versions of this paper. We also benefited from comments given by the participants of ISCEF conferences in Paris, 2014, where this paper was presented. The usual disclaimer applies. Fabio Augusto Reis Gomes and João Victor Issler gratefully acknowledge support given by CNPq-Brazil. Issler also acknowledges the support given by CAPES, Pronex, FAPERJ, and INCT. We gratefully acknowledge research assistance given by Rafael Burjack, Marcia Waleria Machado, and Marcia Marcos. Address correspondence to: João Victor Issler, Graduate School of Economics—EPGE, Getulio Vargas Foundation, Praia de Botafogo 190 s. 1100, Rio de Janeiro, RJ 22250-900, Brazil; e-mail: jissler@fgv.br.

 1365-1005/16

puzzles; see Campbell and Deaton (1989), Flavin (1981), Hall (1978), Hansen and Singleton (1982, 1983, 1984), Mark (1985), Mehra and Prescott (1985), and Weil (1989). Most of these studies employed the constant-relative-risk-aversion (CRRA) utility function with exponential discounting of future utility in defining welfare. These rejections have led to two different strands of the consumption literature. The first investigated whether changing preferences could accommodate optimizing behavior; see Abel (1990) and Campbell and Cochrane (1999) for research on habit, and Epstein and Zin (1989, 1991) for research on nonexpected utility. The second strand introduced explicit forms of nonoptimizing behavior for consumption decisions. In that regard, the most influential study is that of Campbell and Mankiw (1989, 1990), who extended the basic optimizing model incorporating what they have labeled *rule-of-thumb* behavior: there are two types of consumers, of which the first type consumes according to optimizing behavior but the second consumes only his/her current income.[1] In this setup, changes in aggregate consumption respond to expected changes in aggregate income, and the response is a function of the importance of rule-of-thumb consumers.

In this context, rejecting optimizing behavior using aggregate data (time series) is an important setback in macroeconomics, where an optimizing representative-consumer framework with a CRRA utility function is commonly assumed. Moreover, this rejection has far-reaching implications: It raises the issue of whether or not we can postulate optimizing behavior in macroeconomics—if one cannot defend optimizing behavior at the aggregate level, one can question whether it is applicable at all.

This paper makes two original contributions to the literature on consumption optimality at the aggregate level. Our setup tests for optimality, having as a basis the standard CRRA framework for the representative consumer, where the generalized method of moments (GMM) is used in estimation and testing. We employ an encompassing model that simultaneously allows for the existence of rule-of-thumb behavior and habit in preferences, which are tested as exclusion restrictions. Our limited focus on these two departures stems from historical reasons: most of the literature in macroeconomics has employed the CRRA utility function, whereas the main deviations from optimality are the presence of rule-of-thumb consumers in Campbell and Mankiw (1989, 1990), and the extension in Weber (2002), dealing jointly with rule of thumb and habit in preferences.[2] Next, we detail our main contributions.

First, at least since Carroll (2001), it has been well known that ignoring higher-order terms in log-linearization of Euler equations yields inconsistent estimates of their respective parameters, invalidating hypothesis testing. This happens because past observed values do not constitute valid instruments, but these are exactly the instruments the previous literature has used. As shown in the following, this critique applies directly to linear or log-linear rule-of-thumb tests. In a direct test of the omission of higher-order terms in log-linearized models, we confirm empirically their misspecification. Results of auxiliary misspecification tests concur with the latter.

Second, we circumvent the problem of lack of instruments in log-linearized regressions using a nonlinear setup for estimation and testing. Our approach has two main ingredients. The first is to exploit the nonlinearity of the Euler equation of the optimizing agent, where, under rule of thumb, her/his consumption is a linear combination of consumption and income [Weber (2002)].[3] The second is to aggregate returns in the Euler equation for the optimizing agent. This is possible because the latter allows linear aggregation of gross returns, although it is a nonlinear function of consumption and preference parameters.

Aggregating returns has several benefits: (i) From a theoretical point of view, we know that only pervasive variation of returns affects intertemporal substitution in consumption.[4] Because of the law of large numbers, aggregation preserves the pervasive variation of returns, throwing away idiosyncratic variation. This allows a representative-consumer interpretation of utility parameters, where aggregate consumption is matched with aggregate returns—not with a handful of individual returns. (ii) Estimating Euler equations for several assets requires knowledge of —what assets are used to transfer wealth across time in every period; see Attanasio et al. (2002) and Vissing-Jørgensen (2002).[5] Although this may be a problem for panel-data studies, participation at the aggregate level is readily available from financial markets, wealth surveys, and national accounts. (iii) Standard GMM estimation employing a large number of returns is usually infeasible because the number of time periods is small vis-à-vis the number of assets. Return aggregation preserves the pervasive portion of return variation, allowing feasible estimation. On the other hand, if one focuses on a subset of returns in empirical tests, as it is commonly done in the literature, asset-return information is thrown away—which is suboptimal from an econometric point of view.

Our empirical implementation for testing optimality in consumption decisions, rule-of-thumb behavior, and habit in preferences requires the use of an aggregate return measure for the economy as a whole—what Mulligan (2002) labeled the *return to aggregate capital*. Here, we employ proxies of the return to aggregate capital at two different frequencies: the annual measures computed by Mulligan (2002) and Mulligan and Threinen (2010), and the quarterly measures computed by Mulligan and Threinen (2010). When these measures are used in estimation and testing, we provide unequivocal evidence that the log-linearization of the representative-consumer Euler equation is problematic. Beyond linearization, and within the CRRA model context, we provide strong evidence against rule-of-thumb behavior for U.S. consumers and against habits in consumer preferences. Our results are in sharp contrast to those in Campbell and Mankiw regarding rule of thumb and to those in Weber regarding habit. Indeed, we show that we can appropriately represent preferences for the U.S. representative consumer using a CRRA utility function with an annual discount rate of 0.95 and a relative-risk-aversion coefficient roughly between 1 and 2, depending on whether we employ consumption of nondurables or consumption of nondurables and services in estimation. Regarding the main objective of this paper—testing consumption optimality at the aggregate level—we found very little evidence for rejection of optimality in consumption decisions when overidentifying-restriction tests are employed, although, on occasion, there were rejections.

Once proper models and econometric techniques are applied to aggregate consumption, income, and aggregate return data, there is no reason to challenge optimizing behavior in consumption, as was the case with previous rule-of-thumb tests. We also show that augmented models for preferences such as consumption with habit formation are unnecessary to characterize intertemporal substitution having the canonical CRRA model as a starting point. Of course, that does not mean that a broader model could not also fit the data.[6] All in all, our evidence reduces the fear that optimizing behavior is the exception, not the rule.

The paper proceeds as follows. Section 2 presents the consumption models, the linear and the nonlinear consumer Euler equations, and the asset returns aggregation. Section 3 presents the econometric methodology. Section 4 presents the empirical results, and Section 5 concludes.

2. CONSUMPTION MODELS

2.1. The Standard Approach in Macroeconomics: CRRA Utility with Aggregate Data

The standard approach in macroeconomics consists of a single-good economy of identical consumers, whose utility functions are of the CRRA type:

$$u\left(C_t\right)=\frac{C_t^{1-\phi}-1}{1-\phi}, \qquad (1)$$

where C_t is consumption in period t and ϕ is the constant relative risk-aversion coefficient. Subject to a budget constraint and transversality conditions, consumers choose consumption and asset holdings to maximize the lifetime utility, given by $\mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(C_t\right)$, where $\beta \in (0, 1)$ is the intertemporal discount factor, and the mathematical expectation operator $\mathbf{E}_t(\cdot)$ is formed conditional on information available to the consumer up to period t. The representative agent can transfer wealth from one period to the next by buying individual assets, indexed by i, $i = 1, 2, \ldots, N$, whose returns are defined as $R_{i,t}=\frac{P_{i,t}+D_{i,t}}{P_{i,t-1}}$,where $P_{i,t}$ and $D_{i,t}$ are respectively their price and dividend. In this setup, the well-known nonlinear Euler equation is given by

$$\mathbf{E}_{t-1}\left\{\beta\left(\frac{C_t}{C_{t-1}}\right)^{-\phi} R_{i,t}\right\}=1, \qquad i=1,2,\ldots,N; \qquad (2)$$

see Hansen and Singleton (1982, 1983, 1984).

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

2.2. Testing Consumption Rule of Thumb in a Linear Framework

If one assumes a CRRA utility function, arriving at the Euler equation (2), further assuming joint conditional lognormality and homoskedasticity of $\left(\frac{C_t}{C_{t-1}}, R_{1,t}, R_{2,t}, \cdots, R_{N,t}\right)'$, the usual time series log-linear representation of consumption growth rate is obtained:

$$\Delta \ln C_t = \alpha + \frac{1}{\phi} r_{i,t} + \mu_{i,t}, \tag{3}$$

where $r_{i,t} \equiv \ln\left(R_{i,t}\right)$, $\alpha \equiv \frac{\left(\ln \beta + \frac{1}{2}\sigma^2\right)}{\phi}$, and $\sigma^2 = \text{VAR}\left[\Delta \ln C_t - \frac{1}{\phi} r_{i,t}\right]$. The error term $\mu_{i,t}$ is unpredictable, because it is an innovation regarding the optimizing agent's information set. The coefficient of the rate of return, $1/\phi$, is the elasticity of intertemporal substitution (EIS), which is the reciprocal of the CRRA coefficient, ϕ.

The conditions under which equation (3) is derived are very stringent: $\mu_{i,t} \left|\Omega_{t-1} \sim \mathcal{N}\left(0, \sigma^2_{\mu_i}\right)\right.$ for all i, with Ω_{t-1} representing the information set of the optimizing agent. The fact that $\mu_{i,t}$ is conditionally Gaussian and uncorrelated with elements of the conditioning set Ω_{t-1} implies that $\mu_{i,t}$ and $\mu_{i,t-s}$, $s > 0$ are independent. Moreover, $\mu_{i,t}$ must be independent of any function of the variables in Ω_{t-1}. In principle, residual-based tests of normality, conditional homoskedasticity, and serial correlation can be used to ensure that these restrictions apply to $\mu_{i,t}$. These will be used subsequently as auxiliary tests in log-linearized Euler equations.

Campbell and Mankiw (1989, 1990) proposed rule-of-thumb behavior for consumers at the aggregate level. There are two types of consumers: Type 1 consumes according to optimizing behavior. Type 2, on the other hand, is restricted to consuming her/his current income $\left(y_{2,t}\right)$. Income of the nonoptimizing agent holds a fixed proportion to aggregate income, $\lambda = \frac{y_{2,t}}{y_t}$, leading to $C_{2,t} = y_{2,t} = \lambda y_t$, where $C_{2,t}$ is agent's 2 consumption. For the optimizing agent, the literature considers two benchmark cases. The first imposes Hall's (1978) quadratic utility setup, where consumption of the optimizing agent follows a martingale process, i.e., $\mathbf{E}_t\left(\Delta C_{1,t+1}\right) = 0$, where $C_{1,t}$ is type 1 consumption. A broader benchmark case imposes CRRA utility for the optimizing agent, leading to a log-linear equation for testing $H_0 : \lambda = 0$, with aggregate consumption growth as the regressand:

$$\Delta \ln\left(C_t\right) = \lambda \Delta \ln\left(y_t\right) + (1-\lambda)\left(\alpha + \frac{1}{\phi} r_{i,t}\right) + (1-\lambda)\,\mu_{i,t}. \tag{4}$$

2.3. A Critique of Current Rule-of-Thumb Tests

The first modern study to focus on approximations to the Euler equation of consumption decisions was Carroll (2001). He states that

> In principle, the theoretical problems with Euler equation estimation stem from approximation error. The standard procedure has been to estimate a log-linearized, or first-order approximated, version of the Euler equation. This paper shows, however, that the higher order terms are endogenous with respect to the first-order terms (and also with respect to omitted variables), rendering consistent estimation of the log-linearized Euler equation impossible. Unfortunately, the second-order approximation fares only slightly better.

Along these lines, Araujo and Issler (2011) generalized this result, showing that estimation of approximations that omit higher-order terms does not have standard valid instruments, which consist of lagged values of observables. Their setup exploits the generalized Taylor expansion (not an approximation) of the exponential function around x, with increment h, showing that it does not depend on x. Applied to the asset-pricing equation (2), their results allow writing its log-linearized version as follows:

$$\Delta \ln(C_t) = \frac{\ln(\beta)}{\phi} + \frac{1}{\phi} r_{i,t} + \frac{\mathbf{E}_{t-1}(z_{i,t})}{\phi} + \mu_{i,t}, \qquad i = 1, 2, \ldots, N, \tag{5}$$

where $\frac{\mathbf{E}_{t-1}(z_{i,t})}{\phi}$ captures the effect of the higher-order terms of the general Taylor expansion, with $z_{i,t}$ being the higher-order term in each equation. In general, $\mathbf{E}_{t-1}(z_{i,t})$ will be a function of the variables in the conditioning set used by the econometrician to compute $\mathbf{E}_{t-1}(\cdot)$. Therefore, omission of $\frac{\mathbf{E}_{t-1}(z_{i,t})}{\phi}$ (or of parts of it) in estimating (5) will generate an omitted-variable bias. This will turn out to be a major problem for versions of (5). It must be stressed that the only reason this term is present in (5) is that we use a log-linear approximation of (2). Thus, we can circumvent the problem if we do not try to log-linearize it.

2.4. Nonlinear Euler Equation and Return Aggregation

Using a nonlinear instrumental variable estimator—e.g., a generalized method-of-moments (GMM) estimator—we can estimate the following system and test hypotheses of interest:

$$\mathbf{E}_{t-1}\left\{M_t R_{i,t}\right\} = \mathbf{E}_{t-1}\left\{\beta\left(\frac{C_t}{C_{t-1}}\right)^{-\phi} R_{i,t}\right\} = 1, \qquad i = 1, 2, \ldots, N, \tag{6}$$

where $\{M_t\}$ represents the process for the stochastic discount factor—taken here to be the CRRA model, recalling that the system is valid for all N assets in the economy.

Efficient estimation of preference and other parameters in system (6) requires estimating the whole system instead of just a portion of it. This happens for the same reason that single-equation OLS estimation is less efficient than SURE

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

estimation in the context of a system of linear regressions. However, system estimation may pose a problem in this context, because, in practice, the number of traded assets (N) in a real economy is large relative to the number of time-series observations (T). For that reason, most of the literature has opted to limit the size of N, e.g., $N = 2$: a risky and a "riskless" asset or, at most, a handful of assets or portfolios with limited asset coverage. Of course, this solution is suboptimal in terms of econometric efficiency.

In an interesting paper, Mulligan (2002) shows that an alternative to estimating the system as a whole is cross-sectional aggregation, where we do not throw away useful information contained in $R_{i,t}$, $i = 1, 2, \ldots, N$, but rather aggregate returns across i to isolate the common component of asset returns; see also the alternative approach in Araujo and Issler (2011). Support for cross-sectional aggregation in this context is based on the idea that idiosyncratic risk, uncorrelated with M_t, must be irrelevant to intertemporal substitution, and cross-sectional aggregation naturally eliminates it. If N is sufficiently large, return aggregation will deliver the common component of returns associated with intertemporal substitution, allowing matching aggregate consumption with an aggregate return.

In what follows we present a stylized version of Mulligan's approach. Consider the sequence of deterministic weights $\{\omega_i\}_{i=1}^{N}$, such that $|\omega_i| < \infty$ uniformly on N, with $\sum_{i=1}^{N} \omega_i = 1$ or $\lim_{N\to\infty} \sum_{i=1}^{N} \omega_i = 1$, depending on whether we allow the existence of an infinite number of assets. Cross-sectional aggregation of (5) implies that

$$\Delta \ln (C_t) = \frac{1}{\phi} \ln (\beta) + \frac{1}{\phi} r_t + \frac{1}{\phi} \mathbf{E}_{t-1} (z_t) + \mu_t, \tag{7}$$

where $r_t = \sum_{i=1}^{N} \omega_i r_{i,t}$ is the logarithm of the *return to the geometric mean of aggregate capital*, $z_t = \sum_{i=1}^{N} \omega_i z_{i,t}$, and $\mu_t = \sum_{i=1}^{N} \omega_i \mu_{i,t}$. Notice that we can specialize $\omega_i = 1/N$ to use equal weights in aggregation. Despite aggregating returns, Mulligan omits the term $\frac{1}{\phi}\mathbf{E}_{t-1}(z_t)$ in estimating (7), potentially leading to omitted-variable bias.

From an econometric point of view, the cross-sectional aggregation leading to (7) is very similar to the theoretical approach of Driscoll and Kraay (1998). They use orthogonality conditions in the form $\mathbf{E}\left(h\left(\theta, w_{i,t}\right)\right) = 0$, $i = 1, 2, \ldots, N$. If N is large relative to T, GMM estimation is not feasible, because we cannot estimate consistently the long-run variance–covariance matrix of the sample moments. Despite this, because for all i the orthogonality conditions hold, we can form a cross-sectional average $\widetilde{h}(\theta, w_t) = \frac{1}{N}\sum_{i=1}^{N} h\left(\theta, w_{i,t}\right)$ and estimate θ by GMM from $\mathbf{E}\left(\widetilde{h}(\theta, w_t)\right) = 0$. Under a set of standard assumptions, Driscoll and Kraay prove consistency and asymptotic normality for the GMM estimates of θ. That happens whether N is fixed or $N \to \infty$.

It is important to stress that, although the approach in Mulligan may be inappropriate because of the use of a linear consumption model, it is a clever way

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

of preserving information on all returns that would otherwise be lost if N were large relative to T. Using aggregate returns, we show next how to construct an encompassing consumption model that allows simultaneously for the existence of rule-of-thumb behavior and habit in preferences. These two departures from the standard CRRA model can then be tested as exclusion restrictions for the parameters of the encompassing model.

As in Weber (2002), the key issue is to note that the optimizing agent—type 1—obeys the Euler equation. Because we want to allow for habit and rule of thumb, we start with preferences with habit for the optimizing agent:

$$u\left(C_{1,t}, C_{1,t-1}\right) = \frac{\left(C_{1,t} - \gamma C_{1,t-1}\right)^{1-\phi} - 1}{1-\phi}, \qquad \phi \neq 1. \tag{8}$$

The optimizing agent Euler equation is

$$E_{t-1}\left\{\begin{array}{c}\left(C_{1,t-1} - \gamma C_{1,t-2}\right)^{-\phi} - \beta\left(C_{1,t} - \gamma C_{1,t-1}\right)^{-\phi}\left[\gamma + R_{i,t}\right] \\ +\gamma\beta\left(C_{1,t+1} - \gamma C_{1,t}\right)^{-\phi} R_{i,t}\end{array}\right\} = 0, \quad \text{for all } i. \tag{9}$$

Recall that aggregate consumption must be the sum of the consumption of the two types. Thus, $C_{1,t} = C_t - \lambda y_t$. Substituting the latter into (9),

$$E_{t-1}\left\{\begin{array}{c}\left[\left(C_{t-1} - \lambda y_{t-1}\right) - \gamma\left(C_{t-2} - \lambda y_{t-2}\right)\right]^{-\phi} \\ -\beta\left[\left(C_t - \lambda y_t\right) - \gamma\left(C_{t-1} - \lambda y_{t-1}\right)\right]^{-\phi}\left[\gamma + R_{i,t}\right] \\ +\gamma\beta\left[\left(C_{t+1} - \lambda y_{t+1}\right) - \gamma\left(C_t - \lambda y_t\right)\right]^{-\phi} R_{i,t}\end{array}\right\} = 0, \quad \text{for all } i. \tag{10}$$

Notice that (10) is a general model that encompasses rule of thumb and habit, under CRRA. It has three special cases: habit alone ($\lambda = 0$), rule of thumb alone ($\gamma = 0$), and neither habit nor rule of thumb ($\lambda = \gamma = 0$), which is the case of CRRA utility. Equation (10) depends only on observables, although $C - \lambda y$ is highly persistent, which can be dealt with by transformations using ratios: C_t/C_{t-1}, C_t/y_t, etc. Although (10) is nonlinear in consumption, it allows (linear) aggregation and averaging of $R_{i,t}$ across i, leading to GMM estimation as long as instruments are not indexed by i. Start with

$$\mathbf{E}_{t-1}\left\{\beta\frac{u'\left(C_t - \lambda y_t, C_{t-1} - \lambda y_{t-1}\right)}{u'\left(C_{t-1} - \lambda y_{t-1}, C_{t-2} - \lambda y_{t-2}\right)} R_{i,t}\right\} = 1, \quad \text{for all } i. \tag{11}$$

Center, postmultiply by instruments X_{t-1}, and use the law of iterated expectations:

$$\mathbf{E}\left\{\left[\beta\frac{u'\left(C_t - \lambda y_t, C_{t-1} - \lambda y_{t-1}\right)}{u'\left(C_{t-1} - \lambda y_{t-1}, C_{t-2} - \lambda y_{t-2}\right)} R_{i,t} - 1\right] \otimes X_{t-1}\right\} = 0, \quad \text{for all } i. \tag{12}$$

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

From Driscoll and Kraay, cross-sectionally aggregate (12), using weights w_i, $0 \leq w_i \leq 1$, $\sum_{i=1}^{N} w_i = 1$, with $R_t = \sum_{i=1}^{N} w_i R_{i,t}$. Denote the terms in brackets in (12) by $h\left(\theta, w_{i,t}\right)$. Then their aggregate version is

$$\begin{aligned}\widetilde{h}(\theta) &= \sum_{i=1}^{N} w_i h\left(\theta, w_{i,t}\right) \\ &= \left[\beta \frac{u'\left(C_t - \lambda y_t, C_{t-1} - \lambda y_{t-1}\right)}{u'\left(C_{t-1} - \lambda y_{t-1}, C_{t-2} - \lambda y_{t-2}\right)} R_t - 1\right] \otimes X_{t-1},\end{aligned}$$

where it becomes clear that we can estimate $\theta = (\beta, \phi, \gamma, \lambda)'$ by GMM using

$$\mathbf{E}\left\{\widetilde{h}(\theta)\right\} = \mathbf{E}\left\{\left[\beta \frac{u'\left(C_t - \lambda y_t, C_{t-1} - \lambda y_{t-1}\right)}{u'\left(C_{t-1} - \lambda y_{t-1}, C_{t-2} - \lambda y_{t-2}\right)} R_t - 1\right] \otimes X_{t-1}\right\} = 0. \quad \textbf{(13)}$$

The Euler equation behind the moment restrictions (13) is interpretable and can be viewed as that of the optimizing agent who holds a portfolio $R_t = \sum_{i=1}^{N} \omega_i R_{i,t}$ in every period.[7] A natural way to construct weights $\left(\omega_i \text{ or } \omega_{i,t}\right)$ is to look at *participation* of different assets in the portfolio of aggregate wealth in every period. This is motivated by the fact that Euler equations of the form in (10) only hold as an equality in t if asset i is being used to transfer wealth from $t-1$ to t. This is a crucial issue in testing for optimality, Because the Euler equation must hold under the null hypothesis.

Participation is discussed by Attanasio et al. (2002) and Vissing-Jørgensen (2002) in a panel-data context. There, the main problem is that we do not possess the information on specific assets used to smooth out consumption across time for every individual. However, for the representative consumer, one has information on the composition of aggregate wealth. Mulligan referred to the composite return R_t in the following terms: "the interest rate in aggregate theory is not the promised yield on a Treasury Bill or Bond, but should be measured as the expected return on a representative piece of capital." In our view, this is the return that should be used to recover interpretable preference parameters for the representative consumer. For that reason, optimality tests here will be conducted using the encompassing model (13) in the form of a J-test (Sargan test) as discussed in Hansen (1982).

3. ECONOMETRIC METHODOLOGY

3.1. Data

The critical series used in this study is the aggregate real interest rate represented by R_t in (13), which is used to uncover (or identify) the structural preference parameters of the representative consumer. R_t here is measured in different forms. The first measure is the *capital rental rate after income and property taxes* in the United States, as computed by Mulligan (2002), and its updated version measured

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

as the annual and quarterly estimates of the *net marginal product of capital* in the United States, as computed by Mulligan and Threinen (2010).[8] These two measures are identical, in a context where aggregate capital exists and its marginal product is net of depreciation.[9] The first proxy is available at annual frequency from 1947 to 1997, whereas the second is available from 1930 to 2009 at annual frequency. We use only the annual postwar data from 1950 onward and the quarterly data from 1950:1 onward.

The rest of the data used here were extracted from the U.S. National Income and Product Account (NIPA) and from the U.S. Census Bureau. From NIPA, we extracted annual data for real disposable personal income, nominal consumption of nondurables and its price index, and nominal services consumption and its price index. We used two measures of consumption in this paper, following almost all of the consumption literature: real consumption of nondurables and real consumption of nondurables and services. Unfortunately, there is no deflator for nondurables plus services. Thus, we aggregated nondurables and services using Irving Fisher's ideal price index – an equally weighted geometric average of the Laspeyres and Paasche price indices. Intuitively, by employing Fisher's method, we allow rebalancing the weights of the parts on the sum of the components. Simply summing up the deflated parts implies keeping these weights fixed throughout the whole postwar sample, which is obviously inappropriate. To obtain per capita series, we used population data from the U.S. Census Bureau.

3.2. Estimating and Testing Log-Linear Models

Our approach to testing the appropriateness of log-linear models has a direct test for omitted higher-order terms and an auxiliary approach that employs diagnostic tests verifying whether or not the stringent conditions under which an exact log-linear model holds are fulfilled.

The direct test employed here is a modified Ramsey's RESET test, designed to work within an instrumental-variable (IV) context. The standard version of Ramsey's RESET test is based on low-order polynomials in the predicted value of the dependent variable, e.g., $\hat{y}^2$, $\hat{y}^3$, and $\hat{y}^4$. The significance of any of these terms indicates misspecification. Under simultaneity, Pagan and Hall (1983) and Pesaran and Taylor (1999) extended the application of RESET tests. Here, we perform their RESET tests considering the significance of three groups of higher-order terms: $\hat{y}^2$; $\hat{y}^2$ and $\hat{y}^3$; $\hat{y}^2$, $\hat{y}^3$, and $\hat{y}^4$.

The auxiliary testing procedure is designed to apply the following diagnostic tests: homoskedasticity, no serial correlation, and normality tests, all designed to work within an IV context. The homoskedasticity tests applied here are Pagan and Hall's (1983) test, the White/Koenker $T \cdot R^2$ test, and the Breusch-Pagan/Godfrey/Cook-Weisberg test. No serial correlation of the error term is investigated by means of the Cumby and Huizinga (1992) test. Finally, we employ Shapiro–Wilk, Jarque–Bera, and Shapiro–Francia normality tests.

3.3. Estimating and Testing Nonlinear Models under Rule of Thumb and Habit

Because consumption and income are known to have roots of the autoregressive polynomial equal (or nearly equal) to unity, we transform the Euler equations to achieve stationarity. In the context of rule-of-thumb tests, Weber (2002) discusses this issue at some length, dividing Euler equations by specific powers of y_t to generate nonintegrated terms. These powers depend on preferences. Here, we opted for a slightly different route. For the preferences in (1) or (8), dividing Euler equations by C_{t-1} generates terms that have the following formats: gross growth rates of consumption or income, income-to-consumption or consumption-to-income ratios, or products of these.

We list the four Euler equations in untransformed and transformed formats, highlighting the transformations performed in order to achieve stationarity.

1. Optimizing agent with external habit and rule of thumb. The untransformed model is given in equation (10). Collecting terms and multiplying (10) by $\frac{1}{C_{t-1}}$ gives the following transformed model:

$$\mathbf{E}_{t-1}\left\{\begin{array}{l} \beta\left[R_t+\gamma\right]\left[\frac{C_t}{C_{t-1}}-\lambda\frac{y_t}{C_{t-1}}-\gamma\left(1-\lambda\frac{y_{t-1}}{C_{t-1}}\right)\right]^{-\phi} \\ -R_t\beta^2\gamma\left[\frac{C_{t+1}}{C_{t-1}}-\lambda\frac{y_{t+1}}{C_{t-1}}-\gamma\left(\frac{C_t}{C_{t-1}}-\lambda\frac{y_t}{C_{t-1}}\right)\right]^{-\phi} \\ -\left[1-\lambda\frac{y_{t-1}}{C_{t-1}}-\gamma\left(\left(\frac{C_{t-1}}{C_{t-2}}\right)^{-1}-\lambda\left(\frac{C_{t-1}}{y_{t-2}}\right)^{-1}\right)\right]^{-\phi} \end{array}\right\}=0. \quad \textbf{(14)}$$

2. Optimizing agent with external habit and no rule of thumb. The untransformed model is given in equation (10), with $\lambda = 0$. Collecting terms and multiplying the resulting equation by $\frac{1}{C_{t-1}}$ gives the following transformed model:

$$\mathbf{E}_{t-1}\left\{\begin{array}{c} \beta\left[R_t+\gamma\right]\left(\frac{C_t}{C_{t-1}}-\gamma\right)^{-\phi}-R_t\beta^2\gamma\left(\frac{C_{t+1}}{C_{t-1}}-\gamma\frac{C_t}{C_{t-1}}\right)^{-\phi} \\ -\left(1-\gamma\frac{C_{t-2}}{C_{t-1}}\right)^{-\phi} \end{array}\right\}=0. \quad \textbf{(15)}$$

3. Optimizing agent with CRRA utility, rule of thumb, and no habit. The untransformed model is given in equation (10), with $\gamma = 0$. Dividing the numerator and denominator of $\frac{C_t-\lambda y_t}{C_{t-1}-\lambda y_{t-1}}$ by C_{t-1} gives

$$\mathbf{E}_{t-1}\left\{\beta\left(\frac{\frac{C_t}{C_{t-1}}-\lambda\frac{y_t}{C_{t-1}}}{1-\lambda\frac{y_{t-1}}{C_{t-1}}}\right)^{-\phi}R_t-1\right\}=0. \quad \textbf{(16)}$$

4. Optimizing agent with CRRA utility—no rule of thumb and no habit. The model is already in stationary form:

$$\mathbf{E}_{t-1}\left[\beta\left(\frac{C_t}{C_{t-1}}\right)^{-\phi}R_t-1\right]=0. \quad \textbf{(17)}$$

The transformed models [i.e., equations (14)–(17)] are estimated by the GMM using the continuous updating method of Hansen et al. (1996), which has shown superior properties vis-à-vis alternative methods in empirical simulations. As instruments, we employ only lags of observables in each equation being considered.[10]

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

4. EMPIRICAL RESULTS

First, we present the results for log-linear models. Table 1 reports the estimation of log-linear models of the form in (3) for consumption of nondurables and consumption of nondurables and services. First, notice that the model is not rejected by the J-test on any occasion at the 5% level. However, in additional misspecification tests (direct test of omission of higher-order terms and auxiliary tests) it is rejected on every occasion by at least one test, with three exceptions—annual frequency, 1950–1997, for consumption of nondurables with lags 2 and 3, and annual frequency, 1950–1997, for consumption of nondurables and services, with lag 2 or with lags 2 and 3. Overall, most rejections occur in normality and serial-correlation tests, followed by rejections in RESET tests.

In Table 2 we test the log-linear model for rule of thumb under the same two alternative measures of consumption. The direct RESET tests for omitted higher-order terms rejected the null of their exclusions in 8 out of 12 cases. When we also consider the results of auxiliary tests, with one exception—consumption of nondurables and services, annual frequency from 1950 to 1997, lags 2 and 3 as instruments—for every regression run, there is a rejection on at least one of the specification tests discussed earlier. Given the poor performance of the log-linearized model so far, our next step is to focus on nonlinear estimation results.

Table 3 presents a GMM estimation of the encompassing model allowing for habit and rule of thumb—equation (14). We first look at annual data collected by Mulligan (2002) and Mulligan and Threinen (2010). Regardless of whether one uses consumption of nondurables or of nondurables and services, there are no rejections using Hansen's (1982) J-test of overidentifying restrictions. Moreover, on no occasion did we reject either $\gamma = 0$ or $\lambda = 0$ using robust t-ratios. Evidence with quarterly data is not so overwhelming: we still find no rejections of optimality using J-tests. However, when we employ nondurables and services, there is evidence that the habit parameter is statistically significant at the 5% level, with $\widehat{\gamma} \simeq 0.95$ on two occasions.[11] Changing the measure of consumption to nondurables takes us back to the same results with annual data. We still find no rejection on J-tests, and neither γ nor λ is statistically significant anywhere. Taking the whole evidence into account points toward simplifying the encompassing model in both dimensions, one at a time [equation (15) or equation (16)].

We consider next restricting the encompassing model with $\lambda = 0$, resulting in a pure habit model—equation (15). The results are displayed in Table 4. For annual data, we find no rejections for overidentifying-restriction tests (optimality), as well as no rejection of $\gamma = 0$ with robust t-ratios at 5%. For quarterly data, we rejected $\gamma = 0$ on two occasions when we employed nondurables and service.s[12] On one of them, we also rejected optimality using the J-test. Still, when we used consumption of nondurables alone, we neither found γ statistically significant nor rejected optimality when the J-test was employed.

Table 5 presents a GMM estimation of rule-of-thumb models for the constrained agent. J-tests never reject the overidentifying restrictions implied by the model.

TABLE 1. Instrumental-variable estimation for consumption and capital aggregate return $\Delta \ln C_t = \alpha + \frac{1}{\phi} r_t + \text{error}_t$

	Consumption of nondurables						Consumption of nondurables and services					
Aggregate return	Mulligan (2002)		MT (2011)		MT (2011)		Mulligan (2002)		MT (2011)		MT (2011)	
Frequency	Annual		Annual		Quarterly		Annual		Annual		Quarterly	
Period	1950–1997		1950–2009		1950q1–2009q4		1950–1997		1950–2009		1950q1–2009q4	
Instruments	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3
r_t	0.749	0.779*	0.425	0.473*	0.479*	0.495*	0.658	0.646	0.370	0.399	0.515*	0.511*
(s.e.)	(0.423)	(0.382)	(0.239)	(0.221)	(0.231)	(0.231)	(0.357)	(0.320)	(0.221)	(0.196)	(0.170)	(0.169)
Diagnostic tests												
J test	0.977	1.379	0.656	4.076	1.327	4.478	0.367	0.566	0.010	1.099	4.303*	6.158
(p-value)	(0.323)	(0.711)	(0.418)	(0.253)	(0.249)	(0.214)	(0.545)	(0.904)	(0.921)	(0.777)	(0.038)	(0.104)
	Null hypothesis is rejected?											
RESET	Yes*	No	No	No	No	Yes*	No	No	Yes**	No	Yes**	Yes**
Homoskedasticity	No	No	No	No	Yes*	Yes**	No	No	No	No	No	No
Serial correlation	No	No	Yes*	Yes*	No	No	No	No	Yes*	Yes*	Yes*	Yes*
Normality	No	No	Yes*	Yes*	Yes**	Yes**	No	No	No	No	Yes**	Yes**

Note: MT (2010) refers to Mulligan and Threinen (2010). Regression estimated by two-stage least squares using Newey and West's (1987) procedure for robust S.E. The instrument lists are composed of lags of the observables in the equation being estimated. RESET linearity tests used here are described in Pagan and Hall (1983) and Pesaran and Taylor (1999). Error serial correlation is investigated by means of the test in Cumby and Huizinga (1992). The null of Homoskedasticity is investigated by tests in Pagan and Hall (1983), the White-Koenker test, and Breusch-Pagan/Godfrey/Cook-Weisberg test. Finally, we employ Shapiro-Wilk, Jarque-Bera, and Shapiro-Francia Normality tests.
** and * Significant at 1% and 5% levels, respectively.

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

TABLE 2. Instrumental-variable estimation for consumption and capital aggregate return $\Delta \ln(C_t) = \lambda \Delta \ln(y_t) + (1-\lambda)(\alpha + \frac{1}{\phi} r_t) +$ error$_t$

	Consumption of nondurables						Consumption of nondurables and services					
Aggregate return	Mulligan (2002)		MT (2011)		MT (2011)		Mulligan (2002)		MT (2011)		MT (2011)	
Frequency	Annual		Annual		Quarterly		Annual		Annual		Quarterly	
Period	1950–1997		1950–2009		1950q1–2009q4		1950–1997		1950–2009		1950q1–2009q4	
Instruments	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3
r_t	−0.201	−0.040	0.053	0.082	0.759	0.242	0.283	0.246	0.181	0.203	0.803	0.330*
(s.e.)	(0.518)	(0.519)	(0.183)	(0.185)	(0.565)	(0.257)	(0.412)	(0.404)	(0.172)	(0.183)	(0.708)	(0.163)
$\Delta \ln Y_t$	0.943	0.548	0.732	0.584	−0.654	0.600	0.339	0.248	0.378	0.318	−0.719	0.440*
(s.e.)	(0.642)	(0.406)	(0.377)	(0.302)	(0.948)	(0.395)	(0.297)	(0.222)	(0.272)	(0.231)	(1.353)	(0.216)
Diagnostic tests												
J test	0.005	2.720	0.077	1.790	0.189	8.579	0.032	3.473	0.089	4.114	1.230	4.837
(p-value)	(0.941)	(0.606)	(0.781)	(0.774)	(0.664)	(0.073)	(0.857)	(0.482)	(0.766)	(0.391)	(0.267)	(0.304)
	Null hypothesis is rejected?						Null Hypothesis is rejected?					
RESET	Yes*	Yes*	No	Yes**	No	Yes**	Yes*	Yes*	No	No	Yes*	Yes**
Homoskedasticity	Yes**	Yes*	Yes*	Yes*	Yes**	Yes**	No	No	No	No	Yes**	Yes**
Serial correlation	No	No	No	No	No	Yes*	No	No	Yes**	No	No	No
Normality	Yes*	No	No	No	Yes**	Yes**	No	No	No	No	Yes**	Yes**

Note: See Note to Table 1.

https://doi.org/10.1017/S1365100515000085 Published online by Cambridge University Press

TABLE 3. GMM estimation for consumption, aggregate capital return, and income

$$\mathbf{E}_{t-1}\left\{\begin{array}{l}\beta\left[R_t+\gamma\right]\left[\frac{C_t}{C_{t-1}}-\lambda\frac{y_t}{C_{t-1}}-\gamma\left(1-\lambda\frac{y_{t-1}}{C_{t-1}}\right)\right]^{-\phi}\\ -R_t\beta^2\gamma\left[\frac{C_{t+1}}{C_{t-1}}-\lambda\frac{y_{t+1}}{C_{t-1}}-\gamma\left(\frac{C_t}{C_{t-1}}-\lambda\frac{y_t}{C_{t-1}}\right)\right]^{-\phi}\\ -\left[1-\lambda\frac{y_{t-1}}{C_{t-1}}-\gamma\left(\left(\frac{C_{t-1}}{C_{t-2}}\right)^{-1}-\lambda\left(\frac{C_{t-1}}{y_{t-2}}\right)^{-1}\right)\right]^{-\phi}\end{array}\right\}=0$$

	Consumption of nondurables						Consumption of nondurables and services					
Aggregate return	Mulligan (2002)		MT (2011)		MT (2011)		Mulligan (2002)		MT (2011)		MT (2011)	
Frequency	Annual		Annual		Quarterly		Annual		Annual		Quarterly	
Period	1950–1997		1950–2009		1950q1–2009q4		1950–1997		1950–2009		1950q1–2009q4	
Instruments	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3
β	0.944**	0.952**	0.946**	0.939**	0.988**	0.992**	0.989**	0.971**	0.964**	0.954**	0.812**	0.776
(s.e.)	(0.073)	(0.019)	(0.018)	(0.007)	(0.007)	(0.009)	(0.044)	(0.012)	(0.044)	(0.028)	(0.278)	(0.707)
ϕ	1.135	1.207	2.289	1.910**	2.781	3.408	2.376	1.377*	3.217	2.795*	7.096**	15.169
(s.e.)	(3.376)	(1.547)	(1.152)	(0.440)	(1.559)	(1.972)	(2.189)	(0.587)	(2.449)	(1.339)	(2.635)	(10.866)
λ	0.138	0.076	−0.564	−0.432	−0.971	−2.025	0.224	0.075	0.150	0.385	0.018	0.008
(s.e.)	(0.325)	(0.094)	(1.279)	(0.324)	(1.285)	(3.198)	(0.161)	(0.129)	(0.674)	(0.714)	(0.093)	(0.087)
γ	0.460	0.566	−1.059	−1.084	−0.622	−0.702	0.262	0.002	−0.223	−0.658	0.966**	0.949**
(s.e.)	(0.487)	(0.563)	(4.028)	(1.900)	(0.733)	(0.875)	(0.420)	(0.276)	(0.931)	(4.317)	(0.029)	(0.082)
J-test p-value	0.277	0.667	0.151	0.473	0.565	0.906	0.585	0.861	0.985	0.912	0.714	0.900

Note: MT (2010) refers to Mulligan and Threinen (2010). Models are estimated by the continuously updating GMM method of Hansen, Heaton, and Yaron (1996). The instrument lists is composed of lags of the observables in the equation being estimated.
** and * Significant at 1% and 5% levels, respectively.

TABLE 4. GMM estimation for consumption and aggregate capital return

$$\mathbf{E}_{t-1}\left\{\begin{array}{c}\beta\left[R_t+\gamma\right]\left(\frac{C_t}{C_{t-1}}-\gamma\right)^{-\phi}\\ -R_t\beta^2\gamma\left(\frac{C_{t+1}}{C_{t-1}}-\gamma\frac{C_t}{C_{t-1}}\right)^{-\phi}-\left(1-\gamma\frac{C_{t-2}}{C_{t-1}}\right)^{-\phi}\end{array}\right\}=0$$

	Consumption of nondurables						Consumption of nondurables and services					
Aggregate return	Mulligan (2002)		MT (2011)		MT (2011)		Mulligan (2002)		MT (2011)		MT (2011)	
Frequency	Annual		Annual		Quarterly		Annual		Annual		Quarterly	
Period	1950–1997		1950–2009		1950q1–2009q4		1950–1997		1950–2009		1950q1–2009q4	
Instruments	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3
β	0.959**	0.958**	0.943**	0.937**	0.982**	0.977**	0.876**	0.938**	1.287	1.041*	0.988	0.895*
(s.e.)	(0.006)	(0.006)	(0.023)	(0.018)	(0.004)	(0.011)	(0.153)	(0.098)	(0.754)	(0.294)	(1.813)	(0.414)
ϕ	1.258*	1.121*	2.976	2.528	2.084*	0.569	1.437	1.452	2.932*	2.804	2.053**	2.008**
(s.e.)	(0.523)	(0.436)	(1.817)	(1.313)	(0.927)	(2.919)	(0.851)	(0.727)	(1.362)	(1.380)	(0.187)	(0.052)
γ	−1.038	0.099	−1.047	0.053	−0.176	0.870	−0.261	−0.075	−0.707	−0.126	0.778**	0.900**
(s.e.)	(12.533)	(0.261)	(2.717)	(0.215)	(0.472)	(0.839)	(0.598)	(0.240)	(3.083)	(0.361)	(0.240)	(0.025)
J-test p-value	0.520	0.750	0.690	0.721	0.130	0.603	0.968	0.943	0.842	0.974	0.000**	0.478

Note: See Note to Table 3.

TABLE 5. GMM estimation for consumption, aggregate capital return, and income

$$\mathbf{E}_{t-1}\left\{\beta\left(\frac{\frac{C_t}{C_{t-1}}-\lambda\frac{y_t}{C_{t-1}}}{1-\lambda\frac{y_{t-1}}{C_{t-1}}}\right)^{-\phi}R_t-1\right\}=0$$

	Consumption of nondurables						Consumption of nondurables and services					
Aggregate return	Mulligan (2002)		MT (2011)		MT (2011)		Mulligan (2002)		MT (2011)		MT (2011)	
Frequency	Annual		Annual		Quarterly		Annual		Annual		Quarterly	
Period	1950–1997		1950–2009		1950q1–2009q4		1950–1997		1950–2009		1950q1–2009q4	
Instruments	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3
β	0.976**	0.981**	0.937**	0.881**	0.971**	0.988**	0.961**	0.966**	0.951**	0.921**	0.990**	0.989**
(s.e.)	(0.019)	(0.017)	(0.018)	(0.011)	(0.002)	(0.006)	(0.033)	(0.006)	(0.027)	(0.005)	(0.005)	(0.005)
ϕ	2.573	2.866*	2.326*	1.314**	1.016	3.590**	0.781	1.128**	2.527	1.052**	2.978**	2.847**
(s.e.)	(1.413)	(1.250)	(1.099)	(0.382)	(0.843)	(1.319)	(1.628)	(0.275)	(1.348)	(0.249)	(0.906)	(0.933)
λ	0.003	−0.028	−0.059	0.199**	0.183**	0.019	1.082	−0.036	0.043	−0.283	−0.175	0.237
(s.e.)	(0.051)	(0.054)	(0.099)	(0.000)	(0.012)	(0.069)	(0.605)	(0.095)	(0.261)	(0.226)	(0.273)	(0.146)
J-test p-value	0.354	0.483	0.586	0.663	0.258	0.249	0.507	0.815	0.903	0.343	0.153	0.254

Note: See Note to Table 3.

Moreover, in all but two cases, we did not find the rule-of-thumb parameter λ to be statistically significant at the 5% level. A significant rule-of-thumb parameter λ occurred for nondurable consumption with annual data (1950–2009), instrument lags 2 and 3, and quarterly data with instruments lagged twice. It is worth noting that, in these two instances, estimated values of λ are close to 0.2, well below the 0.5 values found by Campbell and Mankiw. Still, in Table 5, the relative-risk-aversion coefficient ϕ and the discount factor β are significant almost everywhere with plausible values: between 1 and 2 for the former and around 0.95 (annually) for the latter. So our next step is to examine the CRRA case.

Finally, Table 6 presents a GMM estimation of the basic CRRA model. J-tests only reject the restrictions implied by overidentifying restrictions once: lag 2 instruments with quarterly frequency. Still, the relative-risk-aversion coefficient ϕ and the discount factor β are significant almost everywhere with plausible values: ϕ is not statistically different from 1 or 2, depending on the consumption measure used, and β is statistically equal to 0.95 (annually) mostly everywhere.

All in all, we estimated 48 Euler equations using the GMM, with encouraging results vis-à-vis the optimality of consumption decisions—the title of this paper. If we take the level of significance to be 5%, we only rejected optimality twice out of 48 times. Regarding the issue of whether we can still rely on the canonical CRRA model, our opinion is that the evidence here supports its use with a few caveats: after all, out of 24 regressions testing the significance of habit or rule of thumb, we found the rule-of-thumb parameter λ to be statistically significant at the 5% level only twice, and the habit parameter γ to be statistically significant on four occasions. So, the overall evidence supports optimality under CRRA utility whenever an aggregate return is used.[13]

5. CONCLUSIONS

This paper makes the following contributions to the literature on consumption optimality. First, following up on the critique in Carroll (2001), we show empirically that the omission of higher-order terms in the log-linear approximation of Euler equations yields inconsistent estimates of the structural parameters when lagged observables are used as instruments. This critique extends to standard rule-of-thumb tests using a log-linearized model. Second, we show that the nonlinear estimation of a system of N asset-pricing equations can be done efficiently even if the number of asset returns (N) is high vis-a-vis the number of time-series observations (T), where system estimation is infeasible. We argue that efficiency can be restored by aggregating returns into a single measure that fully captures intertemporal substitution. Indeed, there is no reason that return aggregation cannot be performed in the nonlinear setting of the asset-pricing equation, because the latter allows linear aggregation of individual returns. Third, aggregation of the nonlinear Euler equation forms the basis of a novel optimality test and tests of deviations from the canonical CRRA model of consumption in the presence of rule-of-thumb and habit behavior.

TABLE 6. GMM estimation for consumption and aggregate capital return

$$\mathbf{E}_{t-1}\left[\beta\left(\frac{C_t}{C_{t-1}}\right)^{-\phi}R_t-1\right]=0$$

	Consumption of nondurables						Consumption of nondurables and services					
Aggregate return	Mulligan (2002)		MT (2011)		MT (2011)		Mulligan (2002)		MT (2011)		MT (2011)	
Frequency	Annual		Annual		Quarterly		Annual		Annual		Quarterly	
Period	1950–1997		1950–2009		1950q1–2009q4		1950–1997		1950–2009		1950q1–2009q4	
Instruments	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3	Lag 2	Lags 2, 3
β	0.959**	0.957**	0.932**	0.920**	0.981**	0.982**	0.973**	0.974**	0.953**	0.941**	0.984**	0.986**
(s.e.)	(0.007)	(0.005)	(0.015)	(0.005)	(0.003)	(0.003)	(0.014)	(0.013)	(0.031)	(0.014)	(0.003)	(0.004)
ϕ	1.225*	1.069**	2.199*	1.362**	2.003*	2.123**	1.471*	1.533*	2.634	1.997**	1.972**	2.279**
(s.e.)	(0.542)	(0.333)	(1.066)	(0.358)	(0.791)	(0.715)	(0.722)	(0.665)	(1.546)	(0.656)	(0.546)	(0.630)
J-test p-value	0.328	0.735	0.425	0.332	0.234	0.224	0.555	0.891	0.928	0.761	0.035*	0.120

Note: See Note in Table 3.

One of our main empirical results was to be able to back out plausible and precise preference-parameter estimates for the representative consumer, where the corresponding Euler-equation restrictions were not rejected by overidentifying-restriction tests. All in all, our estimates show that we can describe reasonably well the U.S. representative consumer with an annual discount rate of 0.95 and a relative-risk-aversion coefficient roughly between 1 and 2, depending on whether we employ consumption of nondurables or consumption of nondurables and services in estimation.

NOTES

1. Campbell and Mankiw conclude that about 50% of total income belongs to rule-of-thumb consumers.

2. Note that, by itself, habit does not constitute a deviation from optimality.

3. For the habit specification, it is also a function of a linear combination of lagged consumption and income.

4. Mulligan (2002) summarizes well why aggregating returns is a good strategy:

> If we were interested, say, in the willingness of consumers to substitute food for other goods, then we should look at the correlation between food expenditure and a food price index. This correlation would have little relation with the correlation between food expenditure and the price of carrots, unless there were a perfect correlation between the price of carrots and the price of all other foods. There may be theories of food demand implying that the price of carrots is always in the same proportion to other food prices, but if in fact there were something moving the price of carrots apart from the prices of other foods, then a price index for all foods is needed. By analogy, my paper compares consumption growth with the return on a large portfolio of capital assets, rather than the return on a particular asset.

5. Gross and Souleles (2002) and Vissing-Jørgensen (2002) have split households into groups. High estimates were obtained of the intertemporal elasticities of substitution in consumption—around 0.8—when stock return was used for stockholders or credit card interest rate for credit card debtors.

6. We leave this for future research.

7. For simplicity of notation we use weights that do not depend on time t, ω_i. However, there is no problem with having $\omega_{i,t-1}$, as long as these weights are *measurable*.

8. We thank Casey Mulligan for providing data on both papers to us.

9. To show that the marginal return to aggregate capital is identical to $R_t = \sum_{i=1}^{N} \omega_i R_{i,t}$, up to addition of a constant term, notice that, if we give weights according to participation of the value of total capital in the previous period, $\omega_i = \frac{V_{i,t-1}}{\sum_{i=1}^{N} V_{i,t-1}}$, recalling that the marginal return of each asset i is $R_{i,t} - 1 = \frac{Y_{i,t}}{V_{i,t-1}}$, where $Y_{i,t}$ is the current income accrued from asset i, including dividends and capital gains, and that $\sum_{i=1}^{N} \omega_i = 1$, then

$$R_t - 1 = \sum_{i=1}^{N} \omega_i \left(R_{i,t} - 1\right) = \sum_{i=1}^{N} \frac{V_{i,t-1}}{\sum_{i=1}^{N} V_{i,t-1}} \frac{Y_{i,t}}{V_{i,t-1}} = \frac{\sum_{i=1}^{N} Y_{i,t}}{\sum_{i=1}^{N} V_{i,t-1}},$$

the last term being exactly the measure constructed by Mulligan (2002) and Mulligan and Threinen (2010)—total accrued aggregate income divided by the last period's total capital value. Thus, for our estimated regressions, we added one to the marginal return to aggregate capital, expressed as a real number (not in percentage terms).

10. Because of time aggregation problems, lags start at order two, as recommended by Hall (1988).

11. It is interesting no note that, in these two rejections of $\gamma = 0$, the estimates of ϕ jumped from the interval [1, 3.5] to 7.1 and 15.2, respectively. Empirically, the continuously updating estimator of Hansen et al. (1996) displays fat tails, which may be the case here on these occasions.

12. These are the exact same data and instrument set that had significant γ's in Table 3.

13. This raises the question of whether we would reach a similar conclusion if we had a large sample of time periods and returns with feasible GMM estimation, i.e., if we performed system estimation with a large number of returns. We leave this to future research.

REFERENCES

Abel, A. (1990) Asset prices under habit formation and catching up with the Joneses. *American Economic Review Papers and Proceedings* 80, 38–42.

Araujo, F. and J.V. Issler (2011) A Stochastic Discount Factor Approach to Asset Pricing Using Panel Data Asymptotics. Working paper, Graduate School of Economics, Getulio Vargas Foundation, ensaios econômicos da EPGE 717.

Attanasio, O.P., J. Banks, and S. Tanner (2002) Asset holding and consumption volatility. *Journal of Political Economy* 110(4), 771–792.

Campbell, J.Y. and J. Cochrane (1999) Force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107(2), 205–251.

Campbell, J.Y. and A. Deaton (1989) Why is consumption so smooth? *Review of Economic Studies* 56, 357–374.

Campbell, J.Y. and N.G. Mankiw (1989) Consumption, income and interest rates: Reinterpreting the time series evidence. In O.J. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual,* pp. 185–214. Cambridge, MA: MIT Press.

Campbell, J.Y. and N.G. Mankiw (1990) Permanent income, current income, and consumption. *Journal of Business and Economic Statistics* 8, 265–280.

Carroll, C.D. (2001) Death to the log-linearized consumption euler equation! (And very poor health to the second-order approximation). *The B.E. Journal of Macroeconomics* 1(1), 1–38.

Cumby, R.E. and J. Huizinga (1992) Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions. *Econometrica*60(1), 185–195.

Driscoll, J.C. and A.C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80(4), 549–560.

Epstein, L.G. and S.E. Zin (1989) Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57, 937–968.

Epstein, L.G. and S.E. Zin (1991) Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* 99, 263–286.

Flavin, M.A. (1981) The adjustments of consumption to changing expectations about future income. *Journal of Political Economy* 89, 974–1009.

Hall, R.E. (1978) Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy* 86, 971–987.

Hall, R.E. (1988) Intertemporal substitution in consumption. *Journal of Political Economy* 96, 339–357.

Hansen, L.P., J. Heaton and A. Yaron (1996). Finite-sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics* 14(3), 262–280.

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

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

Hansen, L.P. and K.J. Singleton (1984) Erratum of the article " Generalized instrumental variables estimation of nonlinear rational expectations models." *Econometrica* 52(1), 267–268.

Mehra, R. and E. Prescott (1985) The equity premium: A puzzle. *Journal of Monetary Economics* 15, 145–161.

Mulligan, C. (2002) Capital, Interest, and Aggregate Intertemporal Substitution. NBER working paper 9373.

Mulligan, C. and L. Threinen (2010) The Marginal Products of Residential and Non-residential Capital through 2009. NBER working paper 15897.

Pagan, A.R. and D. Hall (1983) Diagnostic tests as residual analysis. *Econometric Reviews* 2(2), 159–218.

Pesaran, M.H. and L.W. Taylor (1999) Diagnostics for IV regressions. *Oxford Bulletin of Economics and Statistics* 61(2), 255–281.

Vissing-Jørgensen, A. (2002) Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy* 110, 825–853.

Weber, C.E. (2002) Intertemporal non-separability and "rule of thumb" consumption. *Journal of Monetary Economics* 49, 293–308.

Weil, P. (1989) The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24, 401–421.