Sensitivity, Moment Conditions, and the Risk-free Rate in Yogo (2006)

Nicola Borri
Giuseppe Ragusa

[1]*LUISS University, Rome, nborri@luiss.it*
[2]*LUISS University, Rome, gragusa@luiss.it*

ABSTRACT

In this paper we show that results presented in the seminal paper by Yogo, *A Consumption Based Explanation of Expected Stock Returns*, cannot be replicated. We find different estimates for the parameters and we obtain values of over-identified statistics that being much larger than those in the original paper indicate rejection of the durable consumption asset pricing model. By careful inspection of Yogo's replication files, we were able to track down the inconsistency to a coding bug. The rejection of the durable model is exemplified by its inability to simultaneously explain the risk-free rate and excess stock returns.

Keywords: equity premium, nonlinear GMM estimation, durable model

JEL Codes: G12, C58

Yogo (2006), *A Consumption Based Explanation of Expected Stock Returns*, presents evidence of a highly pro-cyclical stock of durable relative to non-durables. This evidence leads to a pricing model augmented to include the dynamics of durable consumption. In this model, households enjoy utility from the consumption of the non-durable good, and from the flow services of durable goods. When the elasticity of substitution between consumption of the durable and non-durable good is higher than the elasticity of intertemporal substitution, marginal utility is counter-cyclical. Yogo tests this model and finds that it is not rejected by the data. On the contrary, data reject traditional alternatives like the time-separable utility model (Eichenbaum and Hansen, 1987) and the non-separable expected utility model (Epstein and Zin, 1991), that are nested in the durable model via restrictions on the parameters.

We show that the non-linear durable model is rejected by exactly the same data used by Yogo. According to our own GMM estimates, the utility weight attached to consumption of the durable good is much smaller than in the original paper. The J-statistics for the unconditional moments should all be multiplied at least by a factor of 5, while the J-statistics for the conditional moments is equal to approximately 68, against a value of 42 in Yogo. The failure of the durable model has an old explanation: it cannot simultaneously match the low risk-free rate and the large equity risk-premium.

The remainder of this brief paper is organized as follows. In section 1 we briefly layout the durable model. In section 2 we present our results of the non-linear estimates in Yogo (2006). In section 3 we conclude.

1 The model

Yogo (2006) lays out a consumption-based asset pricing model where the representative household must choose how much to spend on consumption of non-durables (C) and on a durable consumption good (E). The main assumptions of the model are that the intratemporal utility has constant elasticity of substitution (CES)

$$u(C, D) = \left[(1-\alpha)C^{1-1/\rho} + \alpha D^{1-1/\rho} \right]^{1/(1-1/\rho)}, \qquad (1)$$

where $\alpha \in (0,1)$ and $\rho \geq 0$ is the elasticity of substitution (ES) between the two goods; and that the household's intertemporal utility is specified

by the following recursive function

$$U_t = \left\{ (1-\beta)u(C_t, D_t)^{1-1/\sigma} + \beta(E_t[U_{t+1}^{1-\gamma}])^{1/\kappa} \right\}^{1/(1-1/\sigma)}, \qquad (2)$$

where $\kappa = (1-\gamma)/(1-1/\sigma)$; $\beta \in (0,1)$ is the standard time-discount factor; $\sigma \geq 0$ is the elasticity of intertemporal substitution (EIS); and γ determines relative risk-aversion.

We refer to the original paper for a detailed description of the model. We note here that this model nests several standard consumption-based asset pricing models. First, when the ES = EIS (i.e., $\sigma = \rho$), we have the additively separable model of Epstein and Zin (1991). Second, when the EIS is the inverse of risk-aversion (i.e., $\sigma = 1/\gamma$) we have the non-separable expected utility model of Dunn and Singleton (1986). Third, when $\sigma = 1/\gamma = \rho$ we have the standard power-utility additively separable expected utility model. Denote with $R_{W,t+1}$ the return on wealth, Yogo shows that the intertemporal stochastic discount factor (SDF) is

$$M_{t+1} = \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-1/\sigma} \left(\frac{v(D_{t+1}/C_{t+1})}{v(D_t/C_t)} \right)^{-1/\rho-1/\sigma} R_{W,t+1}^{1-1/\kappa} \right]^{\kappa}, \qquad (3)$$

where

$$v\left(\frac{D}{C} \right) = \left[1 - \alpha + \alpha \left(\frac{D}{C} \right)^{1-1/\rho} \right]^{1/(1-1/\rho)}. \qquad (4)$$

2   Replication of the estimation and testing of the model

The Euler equations and intertemporal condition imply the following moment restrictions

$$
\begin{align}
0 &= E[(M_{t+1}R_{0,t+1} - 1)z_t] \qquad &(5) \\
0 &= E[M_{t+1}(R_{i,t+1} - R_{0,t+1})z_t], \quad i = 1,\ldots,N \qquad &(6) \\
0 &= E\left[\left(1 - \frac{u_{Dt}}{P_t u_{Ct}} - (1-\delta)M_{t+1}\frac{P_{t+1}}{P_t} \right) z_t \right], \qquad &(7)
\end{align}
$$

where $R_{0,t}$ is the 3-month T-bill rate and $R_{i,t}$, $i = 1,\ldots,N$ are the returns on N test assets. The test assets are 25 Fama and French (1993) portfolios sorted by size and book-to-market equity, 24 portfolios sorted by book-to-market equity within industry, 25 portfolios sorted by market and HML

betas. The instrument z_t is a $I \times 1$ vector of variables belonging to the information set of the representative household at time t. Using these moment conditions, the parameters of the durable consumption model can be estimated by the Generalized Method of Moments (GMM). Importantly, the validity of the model can be assessed by the over-identified restrictions test of Hansen (1982).

There are two versions of the model that depend on the set of instruments employed. The so called cross-sectional version uses a constant instrument; in the conditional version the vector of instruments comprises lagged variables. In particular, Yogo uses an instrument vector containing second lags of non-durable and durable consumption growth, dividend-price ratio, size spread, value spread, yield spread, and a constant.

2.1   Yogo's results

Yogo reports the results of the estimation of the non-linear model's unconditional moments for different sets of test assets and the conditional moments for the three Fama-French factors. For convenience, we report Yogo's estimates in Panel B of table 1. For all moments considered, the over-identified test statistic is below the relevant 5% critical value with the exception of the conditional version.

These empirical results seem to give empirical substance to the claim that the durable consumption model can successfully price both the risk-free rate and the discrepancy between the safe and risky rates. However, we have come to realize that these results present some inconsistencies which undermine the empirical relevance of the durable consumption model.

2.2   Replication

We have extensively investigated the replication files, written in GAUSS, that Yogo very transparently makes available on his website, and concluded that when $\{g_t(\theta)\}$ is evaluated, for certain combinations of the parameters, the SDF is not defined. The routine that Yogo used to calculate the moment function gives an evaluation error, but the GAUSS' optimization routine does not stop at the error; rather, it returns, as first step estimate, the value of θ at the step immediately before the optimization algorithm failed because of the evaluation error of $g_t(\theta)$[1]. When an estimate of the inverse

---

[1] The optimization routines used in Yogo's code are from the Constrained Optimization package from Aptech. The return code of the optimization routine is 3, which corresponds

of V is calculated using the parameter values returned by the GUASS' routine, the matrix $\hat{V}^{-1}$ assigns basically zero weights to the first and last moment condition. The first moment condition restricts the expected discounted (gross) risk-free rate to be equal to 1. This is why cannot reject the durable model, even if it does not price correctly the safe asset, as it is clear from figure 1 that plots the simulated SDF, defined by equation (3), using actual consumption data for the sample 1951:1-2001:4 and the parameter estimates for the unconditional moments tested on "all portfolios" taken from Yogo (2006) (we set $\beta = 0.939$, $\sigma = 0.023$, $\gamma = 205.905$, $\rho = 0.700$, and $\alpha = 0.802$).

For the durable model to successfully price the risk-free asset, the SDF should be equal to the inverse of the gross risk-free rate. A consistent estimator of $E[M_{t+1}]$ is given by the sample average of $\{M_{t+1}\}$ which, for the given parameters, is approximately equal to 0.24. But this means that Yogo (2006)'s durable consumption model implies a quarterly net risk-free rate of approximately 316%[2]. Note that this result is not specific to the unconditional model applied to "all portfolios", but rather holds also for the other test assets. This should not come as a surprise given the small estimates for the intertemporal elasticity parameter σ in Yogo (2006). Consider as back-of-the-envelope calculation a model with a single consumption good and Epstein and Zin (1991)'s utility. In this case, the log risk-free rate is (Campbell, 2003)

$$r_f = -\log \beta + \frac{g}{\sigma} + \text{2nd order terms,}$$

where β is close to 1 and g is the mean consumption growth. A small σ implies a large risk-free rate.

---

to "function calculation failed." In the appendix we explain in detail the coding bug, which occurs when the optimization routine evaluates the SDF for combinations of the parameters for which it is not defined. For example, when the test assets are 25 Fama and French (1993) portfolios sorted by size and book-to-market equity, the algorithm fails when evaluating the following combination of parameters: $\sigma = 0$; $\gamma = 189.26$; $\rho = 0.404$; $\alpha = 0.798$; $\beta = 0.919$.

[2]We are not the first to note that the durable model fails to explain the risk-free rate. For example, Lustig and Verdelhan (2007) use the durable model to explain the cross-section of foreign currency excess returns for portfolios sorted on the basis of the interest rate differential with respect to the US. They set all the parameters of the model according to Yogo's estimate but for γ, which is chosen as to minimize the squared pricing errors on the currency portfolios. As the resulting γ is large, Lustig and Verdelhan conclude: "our [durable] model cannot match the risk-free rate."

Figure 1: Stochastic discount factor from Yogo (2006).
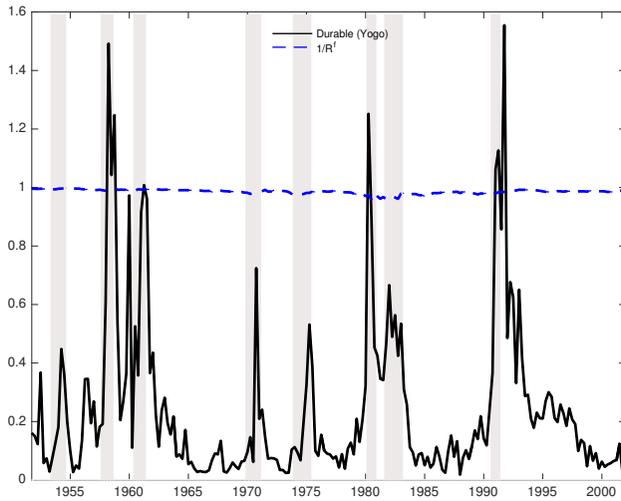
Description: The figure plots the time-series of the stochastic discount factor M_{t+1} of the durable model (black solid line) and the inverse of the realized gross risk-free rate (dashed blue line). M_{t+1} is constructed using estimates in Yogo (2006) for the nonlinear estimation of the unconditional moments on "all portfolios" (see table 1 panel B). Data are from Yogo (2006) at quarterly frequency for the period 1951:1-2001:4.

Interpretation: For the durable model to successfully price the risk-free asset, M_{t+1} should be equal to the inverse of the gross risk-free rate.

To address the discrepancy between Yogo's results and the performances of his SDF to correctly price the risk-free rate we re-estimate the durable consumption model by GMM. Given the empirical counterpart of equations (5)-(7), estimates of the parameter vector are obtained by the efficient GMM. In order to implement the estimation procedure several choices have to be made. In particular, the efficient GMM estimator requires a consistent estimator of θ to construct a consistent estimator of the inverse of the long run variance of $\{g_t(\theta)\}$. A typical approach is to obtain such preliminary estimators optimizing the GMM criterion function with a weighting function set to the identity matrix. The matrix V can be estimated using a semi-parametric estimator. As in Yogo (2006) we use a Vector Auto-regressive HAC estimator of V (Haan and Levin, 1998) where the lags of the VAR specification for $\{g_t(\tilde{\theta}_T)\}$ are chosen to minimize the AIC information criterion. The parameter space is restricted to be a compact subset of $\mathbb{R}^5$[3].

The estimation results, reported in panel A of table 1, are markedly different from the ones in the original paper.

In fact, based on the J-test, the non-durable model is always rejected, but for the case where 25 portfolios sorted by market and HML betas are used. The estimate for α, the utility weight attached to the consumption of the durable good, is much smaller than in Yogo (2006) (approximately, 0.1 vs. 0.8). Furthermore, the magnitude of the estimated parameters is very different and not robust to the choice of different test assets.

3 Conclusions

While a commonly held view is that GMM is computationally straightforward, its implementation is rather not trivial. Besides the coding bug, we have found that the GMM objective function of the nonlinear durable model has several local modes and that starting values have a great influence on the local mode found by the optimization algorithm. A good GMM implementation should always try to account for multi-modal objective function, using, for instance, global solvers[4]. Given the two step nature of

---

[3]In particular, we use the same restrictions in Yogo (2006)'s replication files. Therefore, the parameters are restricted as follows: $0.00001 \leq \sigma \leq 2; 0.00001 \leq \gamma \leq 300; 0.00001 \leq \rho \leq 2; 0.00001 \leq \alpha \leq 1; 0.01 \leq \beta \leq 1$.

[4]Our results in Table 1 are obtained using a combination of a genetic and grid search algorithm. First, we use RGENOUD (see (Mebane Jr, Sekhon, *et al.*, 2011) for a detailed

Table 1: Estimation of the Preference Parameters through the Euler Equations.

| Parameter | Unconditional | | | | Conditional |
	Fama-French	Industry & BE/ME	Beta-sorted	All portfolios	FF3
			Panel A: Borri-Ragusa		
σ	1.97	0.86	0.82	1.09	2.00
	(0.49)	(0.09)	(0.11)	(0.06)	(0.05)
γ	1.94	0.72	0.71	1.17	1.87
	(0.24)	(0.21)	(0.22)	(0.11)	(0.02)
ρ	1.86	1.71	1.10	1.12	0.75
	(2.72)	(2.17)	(1.25)	(0.42)	(0.00)
α	0.11	0.12	0.13	0.13	0.14
	(0.03)	(0.03)	(0.04)	(0.01)	(0.00)
β	0.99	0.99	0.99	0.99	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
J−test	61.67	43.97	27.56	103.31	68.39
	(0.00)	(0.00)	(0.19)	(0.01)	(0.00)
			Panel B: Yogo (2006)		
σ	0.024	0.023	0.024	0.023	0.023
	(0.009)	(0.007)	(0.009)	(0.002)	(0.005)
γ	191.438	199.496	185.671	205.905	174.455
	(49.868)	(44.280)	(43.924)	(11.785)	(23.340)
ρ	0.520	0.554	0.870	0.700	0.554
	(0.544)	(0.604)	(1.955)	(0.247)	(0.026)
α	0.827	0.821	0.786	0.802	0.816
	(0.089)	(0.091)	(0.156)	(0.027)	(0.006)
β	0.900	0.935	0.926	0.939	0.884
	(0.055)	(0.054)	(0.570)	(0.018)	(0.030)
J−test	12.050	9.583	1.866	5.065	42.500
	(0.956)	(0.984)	(1.000)	(1.000)	(0.065)

Description: This table reports the estimates of the preference parameters through the Euler equations of the nonlinear model. Panel A reports our estimates of the durable model; panel B reports estimates from Yogo (2006) (table II, page 552). Columns 2 to 4 report estimates obtained through the unconditional moment restrictions. From left to right, the test assets are 25 Fama–French portfolios sorted by size and book-to-market equity, 24 portfolios sorted by book-to-market equity within industry, 25 portfolios sorted by market and HML betas, and all 74 portfolios. The last column reports preference parameters estimated through the conditional moment restrictions. The test assets in this case are the Fama-French three factors: i.e., the market portfolio, SMB portfolio, and HML portfolio. The instruments are second lags of nondurable and durable consumption growth, dividend-price ratio, size spread, value spread, yield spread, and a constant. All estimates include the Euler equation for the three-month T-bill and the intratemporal FOC as additional moment restrictions. Estimation is by two-step GMM. HAC standard errors are in parentheses. Data are from the replication files available on Yogo's website.

Interpretation: Our own estimates of the nonlinear model (panel A) are markedly different from those in panel B from Yogo (2006) and not robust to the different test assets used. The J-test always rejects the non-durable model, but for the case where 25 portfolios sorted by market and HML betas are used.

the efficient GMM estimator, parameter estimates that are local modes of the GMM objective translate to different weighting matrices and to estimates that depend on which (first-stage) local mode is selected (which in turns depends on the initial value).

Even when one is very careful with the optimization step, another issue is the sensitivity of estimates and relative test statistics to the weighting matrix. In theory, under the null hypothesis of correct specification, that is $E[g_t(\theta_0)] = 0$, the selection of the initial weighting matrix is a second order problem as it will only affect the sampling distribution of the estimator. Correct specification (together with a host of regularity conditions) is enough to guarantee that the GMM estimator converges in probability to the the true value θ_0. One could dismiss efficiency consideration and consider a pre-specified weighting matrix. A pre-specified weighting matrix chooses which moments (or linear combination of moments) GMM will value in the minimization of the objective function. Cochrane (2005) argues that having this freedom is a useful feature since it allows downplaying the importance of moments involving either illiquid assets or assets which suffer from measurement errors. However, under correct specification, the difference between a GMM estimate with pre-specified weighting matrix and an efficient one should be (relatively) small. In the case of the durable model, these differences are instead quite large. What can explain these differences? Our conclusion is that the moment conditions (5) and (6)-(7) are not compatible, that is, the parameters that set the sample counterpart to (5) to zero tend to set the sample counterpart of (6) and (7) to large values. The results in Yogo (2006) corresponds to pre-specified weighting matrices that downweight the importance of (5) with respect to moment

---

discussion of this genetic algorithm) to randomly sampling 20,000 sets of parameters in the domain and evaluate the GMM objective functions over these points; second, the 10,000 set of points with the lowest values for the objective functions are retained; third, these points are "genetically mutated" to obtain further 10,000 points; fourth, on the 20,000 resulting sets of parameters, the objective function is evaluated again and a local solver is started at the set of parameters associated with the lowest GMM objective function. These steps are repeated 100 times. Our estimates are the combinations of parameters associated with the lowest overall GMM objective function. With our strategy, it is much more likely to find a global optima rather than stop at a local mode. Note that if we start the optimization routine from the same initial values in Yogo ($\sigma = 0.0018090568$, $\gamma = 189.12650$, $\rho = 0.78954919$, $\alpha = 0.90711062$, $\beta = 0.92$), the local solvers do not converge. On the other hand, Yogo's own GAUSS code returns our own estimates when we use them as starting values of the optimization routine. Details on the solution algorithm and replication files are available on our web pages.

conditions (6) and (7). These considerations aside, using a weighting matrix that effectively annihilate a key moment condition—the one pinning down the risk-free rate—cannot be considered consistent with the goal of the paper which was exactly to reconcile the dynamics of risk-free and excess asset return. While it is still possible that the observed differences in the estimates and test statistics reflect sampling variation, it is more likely that they are due to misspecification of the model. Note that if the model is misspecified, both the original parameter estimates and the one we report are meaningless from an economic point of view. However, the fact that the weighting matrix in Yogo is not consistent for the variance of the moment conditions implies that the J-statistic is not correct, i.e., does not have the asymptotic chi-square calibration. On the other hand, the J-statistic we report, and that leads to rejection of the model, has the correct asymptotic distribution under the null (provided that the model does not suffer from weak identification).

We point out that many papers have raised concerns about the GMM objective function being flat on large part of the parameter space in linearized factor models (Gospodinov *et al.*, 2014; Peñaranda and Sentana, 2015; Burnside, 2015). Some of these studies have analyzed the sensitivity of Yogo's results in the linearized version of the non-durable model. We are instead the first to address the non-linear version of the model. Although inference robust to weak identification in the nonlinear GMM case could be addressed using the results in Stock and Wright (2000), we believe that misspecification concerns are first order with respect to those of identification. A way to address weak identification would be to use existent empirical evidence, economic theory, or estimates from the linearized version of the model, to restrict the parameter space, for example constraining risk aversion to be large, in order to explain the equity premium, or the weight on durable consumption to be large enough. However, in this latter cases, the pricing errors—already significant—will be even larger.

Appendix: Additional details on the bug

In this appendix we briefly describe the coding error behind the wrong results reported in Yogo (2006). Yogo's code, available on his website, is written in GAUSS and uses the co, or constrained optimization library by Aptech. The relevant file for the GMM estimation of the non-linear durable model is nlest_dur.prg, which can be found in the folder Durable/programs/. To understand the bug, it is convenient to start from the following code snippet (line 122-26 of the original code):

```
/* GMM: first step */
W = eye(Nmom);
{param,fval,grad,retcode} = co(&nlobj_dur,param);
```

The first line initializes the initial weighting matrix to be diagonal (Nmom is the number of moment conditions). The second calls the optimization routine. The inputs are &nlobj_dur, which is the function that calculates the objective function and its gradient, and param, which contains the initial values of the parameters from which the solver starts finding a solution. The output values from a call to co are param,fval,grad,retcode, which are, respectively, the minimizer vector, the value of the function at the minimum, the value of the gradient, and the return code of the algorithm. The bug consists in the fact that, in all the first stage calls to co, the algorithm fails to converge. In fact, we found that the value of retcode is 3, which means that the GMM objective value failed to evaluate to a finite number for the proposal vector. Unfortunately, and differently from standard practice, the GAUSS solver does not report this as an error, but rather silently returns: retcode, which is the only way to confirm that there was a value with the objective function.

For example, take the case of the 25 Fama–French portfolios sorted by size and book-to-market equity. The starting vector, param, is equal to

```
>> param;
      0.002
    189.126
      0.789
      0.907
      0.920
```

The first stage values returned by `co` for `param`, `fval`, `grad` and `retcode` are

```
>> param;         >> fval;         >> grad;         >> retcode;
      0.024           259.421           144.252             3
    189.126                               0.443
      0.521                             583.398
      0.827                            3550.877
      0.873                             -34.023
```

Therefore, the solver did not get even close to a minimum, as it can be gauged by the magnitude of the gradient vector. In this case, the problem depends on the fact that, during the descent toward the minimum, the solver asked to evaluate the objective function to a point for which the moment restrictions are not defined and, consequently, `nlobj_dur` returned a `NaN` and the solver stopped. In this case, the value of the vector at which the objective function failed to evaluate is

```
>> param;
      0.000
    189.126
      0.404
      0.798
      0.919
```

The first element of the vector corresponds to the parameter σ. The problem is that the SDF is not defined for $\sigma = 0$. It is important to note that the GAUSS' solver does not return the parameter values at which the evaluation failed, but rather the last valid value of `param`. Therefore, the first stage estimate of Yogo are arbitrary because they depend on the path of the algorithm. For instance, changing the specifics of the `co` algorithm (line 14 of the original code) still results in a `retcode = 3`, but the returned parameter vector is different because the descent followed a different path.[5]

The second step of the GMM procedure starts by calculating an estimate of the long-run variance of the moment function, taking its inverse and then re-running the solver (line 127 and then 153-57)

---

[5]The algorithm used by Yogo is a Newton-Raphson method. Using a Broyden, Fletcher, Goldfarb, Shanno (BFGS) method results in similar problems, but with different returned parameter vector.

```
/* GMM: second step */
{rT,drT,Omega} = nlgmm_dur(param,lag);
W = invpd(Omega);
{param,fval,grad,retcode} = co(&nlobj_dur,param);
```

The first line calls the objective function and returns the moment function
(rT), the gradient of the moment function (drT), and an estimate of the
long-run variance (Omega). This variance is then inverted to obtain the
new weighting function to use in the GMM second step. An obvious by-
product of the failure of the solver in the first step, is that the resulting
weighting matrix is obtained by evaluating the moments at points that are
not minimizer of the GMM objective function. This would not necessarily
be a problem if it were not for the actual weighting matrix that results from
it. The next display shows the diagonal entries of W

```
>> diag(W);
      2.737
  32648.556
  58984.415
 136462.175
 206152.806
 134516.550
  43815.318
 102733.113
 141527.181
 144095.337
 141909.684
 110723.342
 121896.587
 193005.681
 175962.742
  69677.133
 140115.512
 128871.379
 169170.345
 112639.369
  55495.691
  43931.831
 100673.023
  76191.738
  61633.372
  35365.836
      4.708
```

These are the weights that each moment restriction receives in the GMM
objective function. The weight of the first and last moment conditions are

low relative to the weights of the other moment conditions (by a factor of 10^4). These two moment restrictions correspond to those in equation (5) and (7), respectively. Effectively, the second stage GMM is down-weighting the moment condition that "prices" the risk-free rate and the moment condition implied by the intratemporal FOC, to the point that their effect becomes negligible on the resulting estimates of the parameters. It is then no surprise that, as shown in figure 1, the SDF and the inverse of the gross risk-free rate diverge when the SDF is calculated at the Yogo's estimated parameters: these parameters do not even try to correctly price the risk-free rate. Although in this appendix we highlighted the problem for the case in which the test assets are the 25 Fama–French portfolios sorted by size and book-to-market equity, the same issues apply, almost unchanged, to the other test assets considered.

References

Burnside, C. 2015. "Identification and inference in linear stochastic discount factor models with excess returns". *Journal of Financial Econometrics*: nbv018.

Campbell, J. Y. 2003. "Consumption-based asset pricing". In: *Handbook of the Economics of Finance*. Ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Vol. 1. Amsterdam: Elsevier. 803–887.

Cochrane, J. H. 2005. *Asset Pricing*. Revised Edition. Princeton, New Jersey: Princeton University Press.

Dunn, K. B. and K. J. Singleton. 1986. "Modeling the term structure of interest rates under non-separable utility and durability of goods". *Journal of Financial Economics*. 17.

Eichenbaum, M. S. and L. P. Hansen. 1987. "Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data". NBER Working Paper 2181.

Epstein, L. G. and S. E. Zin. 1991. "Substitution, risk aversion, and the temporal behaviour of consumption and asset returns: An empirical analysis". *Journal of Political Economy*. 99.

Fama, E. F. and K. R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds". *Journal of Financial Economics*. 33: 3–56.

Gospodinov, N., R. Kan, and C. Robotti. 2014. "Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors". *Review of Financial Studies*. 27(7): 2139–2170.

Haan, W. J. den and A. T. Levin. 1998. "Vector Autoregressive Covariance Matrix Estimation". San Diego, manuscript.

Hansen, L. P. 1982. "Large Sample Properties of Generalised Method of Moments Estimators". *Econometrica*. 50(4): 1029–1054.

Lustig, H. and A. Verdelhan. 2007. "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk". *American Economic Review*. 90(Mar.).

Mebane Jr, W. R., J. S. Sekhon, *et al.* 2011. "Genetic optimization using derivatives: the rgenoud package for R". *Journal of Statistical Software*. 42(11): 1–26.

Peñaranda, F. and E. Sentana. 2015. "A unifying approach to the empirical evaluation of asset pricing models". *Review of Economics and Statistics*. 97(2): 412–435.

Stock, J. H. and J. H. Wright. 2000. "GMM with weak identification". *Econometrica*. 68(5): 1055–1096.

Yogo, M. 2006. "A Consumption-Based Explanation of Expected Stock Returns". *The Journal of Finance*. 61(2).