ON THE DURABILITY OF NON-DURABLE GOODS: SOME EVIDENCE FROM U.S. TIME-SERIES DATA

by

Luigi Ermini

Working Paper No. 91-1
January, 1991

ABSTRACT

Recent empirical evidence from monthly data has suggested that consumption of non-durables and services may follow an integrated, first-order moving average process, or IMA(1,1), with a negative MA coefficient, rather than a random walk. This is precisely the generating mechanism of consumption implied by Hall’s model under durability of goods. In effect, non-durables (and services) should exhibit a degree of durability on a monthly or quarterly basis, as this category contains commodities with service lifes of up to three years. The paper tests this implication in a univariate framework, and estimates the average durability rates for the three National Income and Product Account categories of durables, non-durables and services. The rates for non-durables and services are found to be plausible while the rates for durables are found to be too low. The paper also reconciles through temporal aggregation the seemingly different results obtained with monthly and quarterly data.
ON THE DURABILITY OF NON-DURABLE GOODS:
SOME EVIDENCE FROM US TIME-SERIES DATA (*)

1. Introduction

Recent literature, based on evidence from monthly data, has suggested that consumption of non-durables and services may follow an integrated, first-order moving average process, or IMA(1,1), with a negative MA coefficient, rather than a random walk, or IMA(1,0) (Ermini [1989a]). As this alternative process is precisely the generating mechanism implied by Hall’s [1978] model suitably modified to account for durability of goods, the purpose of this paper is to investigate whether the durability of non-durable goods and services, possibly combined with the phenomenon of temporal aggregation, can rescue Hall’s model from well reported failures. Along this line, but with Japanese panel data and without temporal aggregation, Hayashi [1985] reported some encouraging results.

The potential relevance of durability is evident if one considers the standard practice in consumption theory of adopting a rigid differentiation between expenditures on goods and services that depreciate entirely within the consumer’s decision period (the "non-durable goods and services" category), and expenditures on goods that depreciate only partially (the "durable goods" category). This differentiation is clearly arbitrary, since the assignment of a specific commodity to either category depends in general on the length of the consumer’s decision period: the shorter this period, the greater the number of goods that fall into the category of durables. In practice we see that, on the one hand, economists are typically uninterested in determining the length of the decision period; this is customarily made to coincide with a quarter or a year, depending on the type of data used for empirical work. On the other hand, theoretical models and empirical works on consumption behavior are mostly focused on explaining expenditures on non-durables and services only, on the grounds, perhaps, that these two categories comprise the great part of consumers’ expenditures (about 75-80%). Expenditures on durable goods are then given a special treatment and explained in isolation.

The standard justification for this rigid differentiation arguably rests on the presumption that the non-durable data used in empirical work (mostly, quarterly time-series of the National Income and Product Account) relate to commodities that depreciate totally within the quarter. In fact, the Bureau of Economic Analysis, which

(*) For very helpful comments I thank John Cochrane, Clive Granger, Anthony Hall, David Hendry and Timo Teräsvirta.
prepares the NIPA data, defines the category of non-durables as comprising goods with an average service life of up to three years. Clothing and shoes, for example, are treated as non-durables (see Byrnes, Donahoe, Hook and Parker [1979] for details). It follows that the practice of building consumption models in which commodities depreciate entirely within the quarterly decision period, and of testing these models with quarterly NIPA data of non-durables (and services), may be grossly inappropriate: many non-durable goods last longer than the quarter in which they are purchased. A similar conclusion can be reached for some expenditures on services; think, for example, of dental services.

Table 1 reports for different decision periods the durability rate $\delta$, that is the fraction surviving after each period, of a commodity that depreciates exactly in three years (defining total depreciation as a 5% scrap value at the end of the three years). The table also reports the ranges of durability rates corresponding to the two NIPA categories of non-durables and durables, containing commodities with a full range of service lives respectively under and over three years. For example, with quarterly data one would expect an average durability rate between 0 and 0.779 for non-durables, and between 0.779 and 1 for durables.

The table also shows that the assumption of full depreciation within the decision period for non-durable goods is even less appropriate if one believes that consumers make decisions at intervals shorter than a quarter: the shorter the decision period, in fact, the greater the number of items contained in the NIPA categories of non-durables and services that do not depreciate entirely within the shorter decision period, and therefore the higher the average durability. In the limit, the practice of using NIPA consumption data of non-durables and services to test continuous-time consumption models becomes paradoxical: in these models decisions are assumed to be taken continuously and commodities are still assumed to depreciate entirely within the infinitesimally small decision period.

To establish the effect of this rigid differentiation on our understanding of consumption behavior is a fundamental issue, at the core of both economic theory (the way we model consumers behavior) and empirical economics (the effect of the non-durability assumption on test results). Surprisingly, the literature appears to be only sporadically aware of this problem, particularly the literature on financial economics, despite the rapid growth of interest and empirical research on continuous-time consumption-based asset pricing models.

Regarding the modelling aspect of the problem, a natural solution is to relax this rigid differentiation, and to model consumers behavior as if all commodities and
services were durable with a full spectrum of durability rates between zero and one (for an interesting example along this line, see Dunn and Singleton [1986]). Regarding the empirical aspect of this problem, it is crucial to establish whether the empirical rejection of well known consumption models could be explained by a significant degree of durability of non-durables and services.

The purpose of this paper is to investigate this empirical aspect of the durability issue, by focusing on Hall’s [1978] model of consumption behavior. The investigation seeks to reconcile three different pieces of empirical evidence. Firstly, it is well known in the literature that Hall’s model is rejected with quarterly data of non-durables and services (see Myron [1986] for a summary of test results). With US quarterly data on durables, Mankiw [1982] also rejects Hall’s model, suitably modified to account for the difference between consumption services and consumption expenditures, on the basis of a very implausible (negative, but very small) estimate of the durability rate for durables \(^1\). Secondly, but in contrast with this evidence, Hayashi [1985] fitted a similarly modified Hall’s model to Japanese panel data of quarterly consumption expenditures, and estimated very plausible durability rates for several categories of services and commodities traditionally classified as non-durables, thus implicitly suggesting that Hall’s model may not be rejected if non-durables and services are treated as durables. A similar suggestion also appears in Muellbauer [1983]. However, Hayashi also estimates a low durability rate for durables (0.38), which, compared with the admissible range of table 1, would again reject the modified Hall’s model for durables, as in Mankiw. Furthermore, Hayashi also estimates the durability rates of several categories of consumption with quarterly Japanese time-series data, obtaining a practically zero durability rate for durables (as in Mankiw) and positive but difficult to interpret durabilities for other categories.

Thirdly, Ermini [1989a] shows that monthly consumption follows an integrated first-order moving average process, IMA(1,1), with a negative MA coefficient, suggesting that, with or without temporal aggregation, consumption may indeed be generated by the representative agent as an IMA(1,1) rather than as a random walk. This alternative generating mechanism can be explained as an implication of the permanent income hypothesis by introducing transitory consumption into Hall’s model (Ermini [1990]). As shown below, it can also be explained by introducing into Hall’s model

---

\(^1\) Reversions of Mankiw’s results, however, have been obtained by introducing adjustment costs for durable goods (Bemanke [1985]), and non-separability between durables and non-durables (Startz [1989]).
the durability of goods. Incidentally, the possibility of this alternative generating mechanism for consumption bears some interesting consequences on addressing the equity-premium puzzle (Ermini [1991]), or on explaining positive risk premia in the term structure (Backus, Gregory and Zin [1989]).

The paper carries out the following exercises: (i) Hall’s model, modified to account for goods durability, is first tested with the three NIPA categories (durable, non-durables and services) of US quarterly data, with and without temporal aggregation. The latter case allows a comparison with both Mankiw’s and Hayashi’s results. The former case investigates the effect of temporal aggregation on the modified Hall’s model as a necessary bridge to compare the results from monthly data with the results from quarterly data. (ii) Hall’s modified model is then tested with US monthly data, again with and without temporal aggregation, and the results are compared with the previous ones. As some economists might dismiss empirical inference based on monthly data because of their noisiness, the paper also tests the modified Hall’s model under different hypotheses of measurement errors.

It is found that under temporal aggregation Hall’s modified model is not rejected in a univariate framework when tested with consumption data of non-durables and services; the estimated durability rates for these two categories are greater than zero and in line with the ranges of table 1. More importantly, and again in line with table 1, it is found that the smaller the consumers decision period, the higher are these rates. Thus, if consumers make decisions at intervals shorter than a quarter, they seemingly face a degree of durability for services and non-durables, as predicted by the way the NIPA data are constructed. On this grounds, economists should perhaps reconsider the practice of modelling consumption behavior as if all expenditures depreciated entirely within the decision period.

However, the paper also finds, similarly to Hayashi’s, that the estimated durability rate for durables, though greater than zero, is still much lower than the range indicated in table 1. So, whereas durability of goods and temporal aggregation apparently rescue Hall’s model for non-durables and services, these two factors are not sufficient to rescue Hall’s model for durables. This conclusion, in a sense, strengthens those research efforts, such as in Bernanke [1985], Lam [1989], Startz [1989] and Caballero [1990] among others, aiming at explaining consumption of durables in isolation through such additional factors as adjustment costs.

The paper is organized as follows. Section 2.1 tests Hall’s modified model with quarterly data and no temporal aggregation, and Section 2.2 with temporal aggregation. Section 3.1 tests Hall’s modified model with monthly data, with and without temporal
aggregation. As the assumption of no measurement error, commonly adopted for quarterly data, is generally considered unacceptable for monthly data, Section 3.2 discusses the case of monthly data affected by different types of measurement errors. Finally, Section 4 offers some concluding remarks.

As a final remark, the conclusions obtained with both quarterly and monthly data can be interpreted in isolation or as joint empirical evidence. If inference with monthly data is deemed unacceptable because of excessive noise, one can interpret in isolation the conclusions obtained with quarterly data under temporal aggregation. If one believes that monthly data contain relevant information, then the similar conclusions from the two data sets can be interpreted as a strengthening of the empirical evidence. However, one should be aware that even if pieces of empirical evidence separately may corroborate an economic theory, their joint existence may refute the same theory (see Ericsson and Hendry [1989] for a discussion).

2. Testing with quarterly data

Following Mankiw [1982] and Hayashi [1985], expenditures on a given category of goods follow an IMA(1,1) process with a negative moving average coefficient:

$$\Delta C_t = \alpha + u_t - \delta u_{t-1},$$

(1)

where $C_t$ is consumer expenditures, $u_t$ is a zero-mean white noise innovation process, $\alpha$ is a constant term, and $\delta$ is the (average) durability rate of the category of goods over the decision period, that is the fraction of goods that remains available for consumption after one period; hence, $0 \leq \delta \leq 1$. When $\delta = 0$, the goods depreciate entirely within the decision period, and Hall’s random walk model of consumer expenditures obtains. When $\delta = 1$, consumption follows a white noise process with a linear time trend.

The generating process (1) is readily derived by assuming that under durability of goods the flow of consumption services from a category of commodities in any period, $s_t$, does not depend only on the current expenditures for this category, $C_t$, but on past expenditures as well. That is, $s_t = \sum_{j=0}^{K} \delta_j C_{t-j}$, with $\delta_0 = 1$. By considering a utility function defined for each period over the sum of the flows of consumption services provided by each category of commodities/services, and using a derivation similar to Hall’s [1978], one readily obtains that the flow of consumption services for each category follows a random walk, $\Delta s_t = \alpha + u_t$. Replacing into this model the above distributed-lag relation between consumption services and consumption expenditures, one obtains a $K$-order autoregressive process in the first differences of consumption expenditures, or ARIMA($K$,1,0):
\[ \Delta C_t = \alpha - \sum_{j=1}^{k} \delta_j \Delta C_{t-j} + u_t \]  

(2)

This is the representation chosen by Hayashi (with \( k = 4 \)) to fit Japanese quarterly time-series data. Alternatively, one can approximate this relation with the more parsimonious geometric decaying Kyoyck representation, that is \( \delta_j = \delta^j \), to obtain the standard stock-flow relationship

\[ S_t - \delta S_{t-1} = C_t \]  

(3)

Replacing (3) into the random walk model for \( S_t \), the IMA(1,1) model (1) follows. As model (1) contains only one durability parameter instead of four in the Hayashi’s version of (2), it permits more direct inferences about goods durability and comparisons across consumption categories. For this reason, this paper chooses model (1) as the modified Hall’s model under durability of goods. The null hypothesis of the tests that follow is thus based on model (1) with \( \delta \) restricted to the values indicated in table 1 for each category of goods.

2.1 Testing with quarterly data and no temporal aggregation

Using quarterly data and selecting model (1) as the model under the null hypothesis is equivalent to adopting the unmentioned, but critical assumption that consumers make decision quarterly (no temporal aggregation effects). Table 2 reports the maximum likelihood estimates (standard errors in parenthesis) of model (1) fitted to per-capita, seasonally adjusted, constant dollars ’82, post-war quarterly data of all the NIPA categories of services, non-durables and durables, and their aggregation 2. This model was tested against the set of all nesting univariate ARIMA\((p,1,q+1)\) alternatives, with \( p + q \leq 3 \), and could not be rejected at the 5% level for all categories of data, and at the 1% level for the case of total consumption expenditures. (Incidentally, the residuals from all the models considered in this exercise passed the usual white noise tests.) Furthermore, the random walk model, or IMA(1,0) was also rejected at the 5% level against the IMA(1,1).

However, though the IMA(1,1) model is not rejected in a univariate framework, we still reject the modified Hall’s model on the basis of the implausible estimates of the durability rates for all categories. Comparing the second column of table 2 with table 1, the durability hypothesis is rejected in all cases: for durable goods, because the

\[ \text{Source: Citibank database 1989. For example, the series of per-capita non-durables was obtained (in Citibank notation) as GCN82xGYDPC8/GYDP82.} \]
hypothesis $\delta = 0$ cannot be rejected, nor can the hypothesis of a typically higher prior for the durability rate of durables be accepted; for services and non-durables, because the estimates are significantly negative. Thus, taking these results at face value would lead us to conclude that the modified Hall's model (1) is inconsistent with all categories of post-war US quarterly consumption data. Regarding durables, these results confirm the practically zero durability found by both Mankiw and Hayashi with time-series data. Regarding non-durables and services, these results confirm well known rejections of Hall's model with quarterly data. This conclusion, however, is reversed - at least for non-durables and services - if we introduce into the model the possibility that consumers make decisions more frequently than a quarter.

2.2 Testing with quarterly data under temporal aggregation

The reason why temporal aggregation might reverse the unfavorable results of the previous section rests on the property that under temporal aggregation an IMA(1,1) process with negative MA coefficient, like model (1), can be transformed into an IMA(1,1) process with positive (or zero) MA coefficient, as estimated in table 2.

Let $m$ be the sampling ratio, that is, the ratio between the period of data observation (in our case, a quarter) and the period over which expenditures are generated by the agent according to model (1). Without loss of generality, $m$ will be assumed to be an integer. Usually in empirical work $m$ is assumed equal to one: if the data is available at quarterly intervals, consumers are implicitly assumed to make consumption decisions at quarterly intervals. This implies, for our case, that the model to be considered under the null when testing with quarterly data is precisely model (1). As seen in the previous section, under this assumption model (1) is rejected. However, if consumers generate consumption according to model (1) but at intervals shorter than a quarter ($m > 1$), the corresponding observed quarterly consumption may follow a process similar to the estimated model of table 2, in which case model (1) would not be rejected.

The phenomenon of temporal aggregation is seldom taken into account in the economic literature, although its theoretical effects on the structure of time-series models are well known (for a summary, see especially Weiss [1984]), and despite some encouraging results and applications (see, among others, Christiano and Eichenbaum [1987], Christiano, Eichenbaum and Marshall [1991], Ermini [1988, 1989a, 1990, 1991], Grossman, Melino and Shiller [1987], and Breeden, Gibbons and Litzenburger [1989]).
In defense of this lack of interest, some economists argue that if a model does not fit the data well, the researcher’s attention should be directed to modifying assumptions with economic content, rather than bothering with such technical assumptions as, for example, \( m = 1 \). Apart from considerations of method (all the assumptions that make up a model are equally liable to cause rejection), this argument must be rejected for a very simple reason: the assumption \( m = 1 \) is not a "technical" assumption of consumption behavior. On the contrary, the timing of agents’ decisions is a fundamental feature of economic theory, as it is well known, for example, in monetary theory or in the theory of the firm with costs of adjustment. The lack of interest in the timing of consumers’ decisions, thus, is not sufficient reason to dequalify the issue to a "technicality" of the model.

Some economists might also argue that, precisely because of this lack of interest in the timing of agents’ decisions, to assume a value \( m > 1 \) is as arbitrary as to assume \( m = 1 \), unless some arguments in support of the plausibility of either case can be presented. Indeed, Hall’s model of consumption behavior offers a good example of how the underlying economic theory can be invoked to support the hypothesis \( m > 1 \). In fact, \( m > 1 \) is precisely an implication of life-cycle rational expectations models of consumption behavior of the representative agent with no costs of making or implementing decisions (for further discussion see, for instance, Ermini [1989b]). As this class of models, which Hall’s model belongs to, entails that consumption decisions are made and revised as often as the agent receives innovations of labor income, and as in advanced economies the greater part of labour income is received, with a good time-alignment, partly every month and partly every fortnight, it follows that the interval of consumption decisions consistent with this class of models should be on average a month or less.

The effects of temporal aggregation on the structure and parameters of IMA(1,1) processes are analyzed in Ermini [1989a]. Let (1) be the model of consumption expenditures generated by the agent at any interval shorter than a quarter. In line with most literature, this model is assumed to have no seasonality. Let \( \bar{C}_t \) be the quarterly aggregation of \( C_t \),

\[
\bar{C}_t = \sum_{j=0}^{m-1} C_{t-1-j}
\]

(the index \( t \) refers now to quarters). Under temporal aggregation the IMA(1,1) process (1) remains an IMA(1,1) process for all values of \( m \) greater than one; i.e. \( \bar{C}_t \) obeys the model.
\[ \Delta \bar{C}_t = \alpha_m + \nu_t - \delta_m \nu_{t-1} \] (4)

where \( \alpha_m \) and \( \delta_m \) are indexed on \( m \) to indicate that in general they depend on the sampling ratio. Further, let \( \rho = -\delta/(1+\delta^2) \) be the first-lag autocorrelation of the first differences \( \Delta \bar{C}_t \) of consumption generated by the agents as in (1). Under temporal aggregation the first-lag autocorrelation of \( \Delta \bar{C}_t \), \( \rho_m = -\delta_m/(1+\delta_m^2) \), becomes (Ermini [1989a]):

\[ \rho_m = \frac{(m^2-1) + 2(m^2+2)\rho}{2(2m^2+1) + 8(m^2-1)\rho} \] (5)

for all \( \rho \in (-0.5, 0.5) \). For \( \rho = 0 \), the process \( C_t \) is a white noise, and (5) does not apply. Setting \( \rho = 0 \), Working [1960] result on the effects of temporal aggregation on random walk processes obtain.

For any initial value of \( \rho \) between -0.5 and 0.25, \( \rho_m \) is monotonically increasing towards the limit of 0.25; for any initial value between 0.25 and 0.5, \( \rho_m \) is monotonically decreasing towards the same limit. Therefore, the estimated positive first-lag autocorrelation of quarterly consumption, \( \hat{\rho} = -\delta/(1+\delta^2) \) (third column of table 2), can be consistent under temporal aggregation with a negative first-lag autocorrelation of consumption generated by the agent, that is with a positive durability rate. Model (1) thus can be true with a durability rate \( \delta \) between 0 and 1 at intervals shorter than a quarter, and yet quarterly data may appear generated by an IMA(1,1) with positive first-lag autocorrelation.

A series of calculations indeed confirms this case. The last column of table 2 reports the estimated durability rates implied by the estimated quarterly first-lag autocorrelation \( \hat{\rho} \), for the case that agents make decisions monthly \( (m = 3) \). These durability rates are obtained from (5) simply considering that with \( m = 3 \), \( \hat{\rho} \) would correspond to \( \rho_3 \). Hence, (5) can be solved for \( \rho \). The durability rates for the original model (1) are then calculated from the quadratic expression \( \rho = -\delta/(1+\delta^2) \), by choosing the smallest of the two roots. The 95% confidence interval is also reported, obtained by applying the same procedure to the limits of the confidence interval of \( \hat{\rho} \).

The results of table 2 can be summarized as follows:

(i) as opposed to the previous section, under temporal aggregation the modified Hall's model cannot be rejected for the categories of non-durables and services, as the implied durability rates for \( m = 3 \) are all in line with the ranges of table 1. Additional calculations, not reported here, based on higher values of \( m \) confirmed the same result. Furthermore, the category of services appears to be less durable than the category of non-durables, as also found by Hayashi with Japanese panel data.
(ii) based on calculations with different values of the sampling ratio $m$, the shorter the consumers’ decision period, the higher the estimated durability rates, in line with the predictions of the theory (see also table 3).

(iii) compared with the ranges of table 1, the durability rates for durable goods are too low. Thus, similarly to the previous section, we would reject the modified Hall’s model for durables.

So, following table 2, if consumers make consumption decisions according to Hall’s modified model at monthly intervals, then they face an average durability rate over the monthly period of about 0.1 for services and 0.25 for non-durables. At the current level of aggregation of the NIPA data we cannot establish whether these values are too low or not. A natural question, though, is to ask what would the corresponding durability rates over the quarterly period be. As a first approximation, a durability rate for non-durables of 0.25 over the month would imply complete depreciation over the quarter, as $0.25^3 = 0.016$ (and similarly for services). This result, then, would contradict the argument that the NIPA category of non-durables contains goods with a service life of up to three years.

In fact, inferring the durability over the quarter from the durability over the month as $\delta^3$ may grossly underestimate the former. The durability of a category of goods over a given period is the average over each single good in the category, $\sum_{j=1}^{M} \alpha_j \delta_j$, where $\delta_j$ is the durability of good $j$ over the given period, $M$ is the number of goods in the category, and $\alpha_j, \sum_{j=1}^{M} \alpha_j = 1$, is the share of good $j$ in the category. Over $m$ such periods, the fraction of each good remaining is $\delta_j^m$, so that the average durability for the category is $\sum_{j=1}^{M} \alpha_j \delta_j^m$. Because of the convexity of the exponential function, clearly

$$\sum_{j=1}^{M} \alpha_j \delta_j^m \geq (\sum_{j=1}^{M} \alpha_j \delta_j)^m,$$

so that the true average durability of, say, non-durable-goods over the quarter may well be greater than the 0.016 previously calculated. For example, consider a category comprising two goods, with $\delta_1 = 0$ and $\delta_2 = 0.8$, and shares $\alpha_1 = \alpha_2 = 0.5$. Clearly, the average monthly durability rate is 0.4, while the true average rate over the quarter is 0.25, much greater than $0.4^3 = 0.06$. Unfortunately, the magnitude of the inequality in (6) cannot be assessed without knowledge of the goods composition in the category, and of their durabilities. As this argument is crucial to a further understanding of the
durability issue, a promising direction of research seems to rest on assessing the durability rates, or equivalently the service lives, of more finely disaggregated consumption data, similarly to what reported in the literature, for example, for producers’ durables (Bureau of Economic Analysis [1987]).

The same argument holds if the average durability rate for each good is replaced by the frequency distribution of its service life over the entire population. The same argument also holds for the case of durable goods. Thus, even if their estimated durability rate over the month is too low to corroborate Hall’s modified model for their case, their corresponding rate over the quarter would still be greater than the practically zero value estimated by Mankiw and Hayashi.

3. Testing with monthly data

As described in Byrnes, Donahoe, Hook and Parker [1979], quarterly and annual NIPA series on personal income and outlays are calculated from monthly data. Moreover, approximately 70% of monthly data are directly collected from monthly observations, with the remaining obtained by interpolation of information available at quarterly/annual frequency (as for the case of the information extracted from tax returns) and by extrapolation of past trends. However, though monthly series are the building block of macrodata, they are seldom used in macroeconometrics, in contrast, for example, with financial economics where high-frequency data (daily, weekly, monthly) are routinely used. The standard argument against the use of monthly data in macroeconometrics is that they are too noisy to make any sensible inference (as an example of this argument, see Backus, Gregory and Zin [1989], p. 390). The issue of noisiness, however, is not raised with quarterly data as the reduction of the noise variance to one third due to the aggregation of error-corrupted monthly data is deemed sufficient for the problem to disappear. Clearly, not using monthly data is a waste of potentially relevant information. On the other hand, to ignore the noisiness issue altogether when using monthly data (as, for example, in Dunn and Singleton [1986], and Eichenbaum, Hansen and Singleton [1988]), is similarly unacceptable as it neglects a possibly substantial bias due to measurement errors.

Unfortunately, little is known about these errors, and, as convincingly discussed for example in Eisner [1989], macroeconometrics would greatly benefit from addressing this issue rigorously. Several sources of errors can be identified in macrodata (see, among others, de Leeuw [1990], Canada System of National Accounts [1989], and Wilcox [1990]): sampling errors, errors due to the projection and distribution of quarterly/annual information over months, procedural errors (recording, tabulating,
seasonally adjusting, etc.), adjustment or conceptual errors, whereby what is measured is not exactly what is meant to be according to the theory. For some categories of data the statistical properties of sampling errors can be assessed; for example, for consumption data these errors appear to be strongly correlated at 4- and 12-months frequency, and to have an average size between 0.5 and 2% of consumption levels (Wilcox [1990] and the references therein). Regarding projection and distribution errors, a first approximation can be found in the magnitude of the periodic revisions of monthly data. However, the revisions from preliminary to final estimates only affects the last three months; thus, with a series of over 350 monthly data points, their effect is negligible. Less so for the "latest" revisions based on the benchmark input-output model built with Census’s data at a 5-years frequency. These revisions, of much greater magnitude than the others (de Leeuw [1990]), can affect up to the last 12 years of the overall sample. Finally, statistical discrepancy can be cautiously used as an indication of the overall size of measurement errors (Canada System of National Accounts [1989]).

Given the current state of knowledge, a suitable approach from the viewpoint of empirical inference is to attempt to isolate the effects of measurement errors from the implications of the theory by investigating the effects on the model test results of different conjectures about the nature and size of these errors. Once again, Hall’s model, in its simplicity, offers a good example of this approach. First, the results from monthly data under the assumption of no measurement errors are presented, and then two different conjectures about these errors are considered.

3.1 Monthly data with no measurement errors

Table 3 reports the maximum likelihood estimates of the IMA(1,1) model (1) fitted to per-capita, seasonally adjusted, constant dollars '82, monthly data of the same categories of table 2. Similarly to the case with quarterly data, the IMA(1,1) is not rejected against a battery a univariate ARIMA models. Thus, this model fits monthly data quite satisfactorily. For the combined category of services and non-durables, this fact was already recognized, and its implications discussed, in Ermini [1989a]. As the choice of model (1) under the null entails $m = 1$, e.g. that consumers make decisions

---

3 Source: Citibank database 1989. For example, the series of per-capita non-durables was obtained (in Citibank notation) as GMCN82xGMYDP8/GMYD82. Monthly data was treated as if the deseasonalization procedure X-11, adopted by the Bureau of Economic Analysis (see Byrnes, Donahoe, Hook and Parker [1979]), and the interpolation/extrapolation procedures mentioned above had no effect on the error autocorrelations.
monthly, the estimated values of \( \delta \) refer now to durability rates over the month. Not surprisingly, these estimates confirm the same conclusions of the quarterly case under temporal aggregation; namely, that Hall's modified model is not rejected for non-durables and services, that services are less durable than non-durables, and that the durability rate for durables is too low. The estimated rates for non-durables and services are in line with the ranges of table 1, but are higher than the values implied by quarterly data under temporal aggregation (last column of table 2). As discussed below, this discrepancy can be explained by the greater noisiness of monthly data, compared to quarterly data. Moreover, different levels of noise for different categories of commodities could also explain the lower spread of the monthly estimates compared to the quarterly estimates.

To investigate the effect of temporal aggregation, table 3 reports also the implied durability rates for the case \( m = 2 \) (decisions every fortnight) and \( m = 4 \) (decisions every week). These values are obtained with the same procedure described for the quarterly case, and applied to the monthly estimates. It is again seen that the shorter the consumers decision periods, the higher are the durability rates, as predicted by the theory. Finally, regarding the possibility of inferring from these estimates the corresponding durability rate over the quarterly period, the same argument based on inequality (6) applies.

3.2 Monthly data with measurement errors

Consider first the case of monthly data affected by additive measurement errors generated as a fixed proportion of consumption levels. Letting \( \bar{C} \) be the monthly aggregate of consumption expenditures generated according to the temporally aggregated model (4), the observed monthly series is \( \bar{C}^*, = (1 + \gamma \bar{C}) \), where \( \gamma \bar{C} \) is the additive measurement error. To simplify the analysis without losing much insight, assume \( \gamma \) to be an average non-stochastic fraction; say, 2% of consumption level. Then, the monthly estimates of table 3 relates now to the model

\[
\Delta \bar{C}^*_t = \alpha^* + u^*_t - \delta_m u^*_{t-1},
\]

where \( \alpha^* = (1 + \gamma) \alpha_m \), and \( u^*_t = (1 + \gamma) v_t \). It follows that the presence of proportional measurement errors affects only the drift \( \alpha_m \) and the variance of the innovations \( v_t \), but not the MA coefficient \( \delta_m \). Therefore, the estimated durability rates of table 3 would also be valid for this case. A similar conclusion holds under temporal aggregation as well.
Consider now the case of additive measurement errors, $e_t$, independent of consumption levels, e.g. $\Delta \tilde{C}^* = \Delta \tilde{C} = \tilde{C} + e_t$. This error specification is quite common in the literature, despite its inappropriateness when applied to integrated processes like consumption (the measurement error would disappear asymptotically as a percentage of consumption levels). Adopting this specification, from (4) we have

$$\Delta \tilde{C}^* = \alpha_m + v_t + \delta_m v_{t-1} + e_t - e_{t-1}$$

which is still an IMA(1,1) process if $e_t$ is also white noise, with $\text{var}(\Delta \tilde{C}^*) = (1+\delta_m^2) \sigma_v^2 + 2 \sigma_e^2$, and first-order autocorrelation of the differences $\Delta \tilde{C}^*$:

$$\rho_m = \frac{-\delta_m \sigma_v^2 + \sigma_e^2}{(1+\delta_m^2) \sigma_v^2 + 2 \sigma_e^2} = \frac{-\delta_m \text{SN} + 1}{(1+\delta_m^2) \text{SN} + 2},$$

where $\text{SN} = \sigma_v^2/\sigma_e^2$ is the signal-to-noise ratio, and $\sigma_e^2$ is the error variance.

One interesting implication of (8) is that observed consumption follows an IMA(1,1) process only if $e_t$ is a white noise process. If $e_t$ follows a richer ARMA process, then observed consumption would follow a richer ARIMA process than IMA(1,1). As the IMA(1,1) model fits the various NIPA categories of monthly data quite well, the white noise hypothesis for measurement errors appears to be quite plausible (see also Ermini [1990]). This conclusion can be consistent with the aforementioned evidence of non-zero error autocorrelations at various lags, as these autocorrelations would make a negligible contribution to the autocorrelation function of consumption for relatively small error sizes.

The presence of additive errors goes against the durability hypothesis. Intuitively, an additive white noise error reduces the first-lag autocorrelation of observed consumption, proportionally to the error variance. For sufficiently large error variance, or equivalently for sufficiently low signal-to-noise ratio, consumption can be generated by the agent with a positive first-lag autocorrelation (that is, with a negative $\delta_m$) and yet it may be observed by the economist with a negative first-lag autocorrelation (positive $\delta$). However, the conclusions obtained without measurement errors for non-durables and services can be retained if the error variance is relatively low. Different error variances for different commodity groups can also explain the relative sizes of durability rates of table 3. For example, a larger measurement error variance for non-durables compared to the measurement errors for services would be sufficient to explain why the two estimated durability rates are so close.

To isolate the effects of measurement errors from the implications of the theory, one procedure is to establish for each sampling ratio $m$ the minimum error variance necessary not to reject the non-durability hypothesis. As this hypothesis is not rejected
when the first-lag autocorrelation of consumption generated by the agent is strictly less than zero, the random walk generating mechanism for consumption is the upper limit. For $m = 1$, and $\rho^*_{m} = \hat{\beta} = -\hat{\delta}(1+\hat{\delta}^2)$, (9) can be solved for SN setting $\delta_m = 0$. For $m > 1$, (9) can be solved for SN setting $\delta_m$ at the value corresponding to the temporal aggregation of a random walk (see Working [1960]). The lower the sampling ratio the higher the chance to reject the durability hypothesis (in other words, the lower the sampling ratio, the lower the minimum variance). Also, for a given sampling ratio, the higher the error variance the higher the chance to reject the durability hypothesis. The last column of table 3 reports the maximum values $SN^*$ for each category of data, for the case that consumers make decisions monthly ($m = 1$). These values are calculated from (9) setting $\delta_m$ equal to zero and $\rho^*_{m}$ to the upper limit of the confidence interval for $\hat{\beta}$. For example, $SN^* \sim 2$ for non-durables implies that the durability hypothesis cannot be rejected for any error variance equal to or lower than 50% the variance of the expenditures innovations, $\sigma_e^2$. In light of what previously argued about the relative size of measurement errors, this condition seems to be plausibly satisfied.

4. Conclusions

It is found that under temporal aggregation Hall’s model, modified to account for the difference between consumption services and consumption expenditures, is not rejected in a univariate framework when tested with consumption data of non-durables and services; the estimated durability rates for these categories are greater than zero and in line with the ranges predicted by the way the NIPA series are constructed. More importantly, and again in line with these predictions, it is found that the smaller the consumers decision period, the higher are these rates. Thus, if consumers make decisions at intervals shorter than a quarter, they seemingly face a degree of durability for services and non-durables which are consistent with the modified Hall’s model. On this grounds, economists should perhaps reconsider the practice of modelling consumption behavior as if all expenditures were discarded entirely within the decision period.

However, the paper also finds, similarly to Mankiw [1982] and Hayashi [1985], that the estimated durability rate for durables, though greater than zero, is still much lower than the range indicated by the theory. So, whereas durability of goods and temporal aggregation apparently rescue Hall’s model for non-durables and services, these two factors are not sufficient to rescue Hall’s model for durables. This conclusion, in a sense, strengthens those research efforts that aim at explaining consumption of durables in isolation through additional behavioral and technological considerations.
The paper also shows that these conclusions are not affected by the presence of measurement errors, as long as the error variance is relatively small compared to the variance of income innovations. Finally, the paper shows how valuable the use of monthly data and the contribution of temporal aggregation can be in empirical work on consumption behavior.
References


Caballero R.J., 1990, Durable goods: an explanation for their slow adjustment, discussion paper series no. 507, dept. of economics, Columbia University.


Ericsson N.R., and D.F. Hendry, 1989, Encompassing and rational expectations: how sequential corroboration can imply refutation, discussion paper no. 354, Board of
Governors of the Federal Reserve System.


Ermini L., 1990, Effects of transitory consumption and measurement errors on tests of the permanent income hypothesis, discussion paper, dept. of economics, University of Hawaii at Manoa.


Wilcox D. W., 1990, What do we know about consumption?, working paper no. 107, Division of Research and Statistics, Board of Governors of the Federal Reserve System.

<table>
<thead>
<tr>
<th>decision period</th>
<th>durability rate for 3-year service life</th>
<th>range of durability (non-durables)</th>
<th>range of durability (durables)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>0.368</td>
<td>0. - 0.368</td>
<td>0.368 - 1.</td>
</tr>
<tr>
<td>1 quarter</td>
<td>0.779</td>
<td>0. - 0.779</td>
<td>0.779 - 1.</td>
</tr>
<tr>
<td>1 month</td>
<td>0.920</td>
<td>0. - 0.920</td>
<td>0.920 - 1.</td>
</tr>
<tr>
<td>1 fortnight 5</td>
<td>0.959</td>
<td>0. - 0.959</td>
<td>0.959 - 1.</td>
</tr>
<tr>
<td>1 week</td>
<td>0.979</td>
<td>0. - 0.979</td>
<td>0.979 - 1.</td>
</tr>
</tbody>
</table>

### TABLE 2

Model \( \Delta C_t = \alpha + \eta_t - \delta \eta_{t-1} \) fitted to quarterly data 7
(in parenthesis: standard errors / 95% confidence intervals)

<table>
<thead>
<tr>
<th></th>
<th>( \alpha )</th>
<th>( \delta ) (durability rate over the quarter)</th>
<th>( \beta = -\delta(1+\delta^2) )</th>
<th>( \delta ) at ( m = 3 ) (durability rate over the month)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DURABLES</td>
<td>7.871</td>
<td>-0.082</td>
<td>0.081</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td>(2.418)</td>
<td>(0.077)</td>
<td>(-0.062, 0.215)</td>
<td>(0.213, 0.569)</td>
</tr>
<tr>
<td>NON-DURABLES</td>
<td>7.916</td>
<td>-0.126</td>
<td>0.124</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>(2.018)</td>
<td>(0.075)</td>
<td>(-0.018, 0.251)</td>
<td>(-0.279, 0.451)</td>
</tr>
<tr>
<td>SERVICES</td>
<td>20.261</td>
<td>-0.192</td>
<td>0.185</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>(1.664)</td>
<td>(0.076)</td>
<td>(0.040, 0.307)</td>
<td>(-1, 0.381)</td>
</tr>
<tr>
<td>SERVICES + NON-DURABLES</td>
<td>28.177</td>
<td>-0.211</td>
<td>0.202</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(3.147)</td>
<td>(0.075)</td>
<td>(0.061, 0.319)</td>
<td>(-1, 0.353)</td>
</tr>
<tr>
<td>TOT. CONS.</td>
<td>36.050</td>
<td>-0.109</td>
<td>0.107</td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td>(4.741)</td>
<td>(0.077)</td>
<td>(-0.041, 0.243)</td>
<td>(-0.193, 0.476)</td>
</tr>
</tbody>
</table>

---

5 As a convention, here a fortnight is simply half a month, and a week a quarter of a month.

7 NIPA series of per-capita, quarterly, seasonally adjusted, constant dollars '82, from 1947-1 to 1989-1 (no. of observations = 169).
| TABLE 3 | Model $\Delta C_t = \alpha + u_t - \delta u_{t-1}$ fitted to monthly data $^9$  
(in parenthesis: standard errors) |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha$</td>
</tr>
<tr>
<td>DURABLES</td>
<td>3.252 (1.075)</td>
</tr>
<tr>
<td>NON-DURABLES</td>
<td>3.073 (0.737)</td>
</tr>
<tr>
<td>SERVICES</td>
<td>8.141 (0.603)</td>
</tr>
<tr>
<td>SERVICES + NON-DURABLES</td>
<td>11.22 (1.13)</td>
</tr>
<tr>
<td>TOT. CONS.</td>
<td>14.483 (1.873)</td>
</tr>
</tbody>
</table>

$^9$ NIPA series of per-capita, monthly, seasonally adjusted, constant dollars '82, from 1959 to 1989 (no. of observations = 366).
<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Durability rates $\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>decision period</td>
<td>durability rate for 3-year service life</td>
</tr>
<tr>
<td>1 year</td>
<td>0.368</td>
</tr>
<tr>
<td>1 quarter</td>
<td>0.779</td>
</tr>
<tr>
<td>1 month</td>
<td>0.920</td>
</tr>
<tr>
<td>1 fortnight $^5$</td>
<td>0.959</td>
</tr>
<tr>
<td>1 week</td>
<td>0.979</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Model $\Delta C_t = \alpha + \mu_t - \delta \mu_{t-1}$ fitted to quarterly data $^7$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(in parenthesis: standard errors / 95% confidence intervals)</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
</tr>
<tr>
<td>DURABLES</td>
<td>7.871 (2.418)</td>
</tr>
<tr>
<td>NON-DURABLES</td>
<td>7.916 (2.018)</td>
</tr>
<tr>
<td>SERVICES</td>
<td>20.261 (1.664)</td>
</tr>
<tr>
<td>SERVICES + NON-DURABLES</td>
<td>28.177 (3.147)</td>
</tr>
<tr>
<td>TOT. CONS.</td>
<td>36.050 (4.741)</td>
</tr>
</tbody>
</table>

$^5$ As a convention, here a fortnight is simply half a month, and a week a quarter of a month.

$^7$ NIPA series of per-capita, quarterly, seasonally adjusted, constant dollars '82, from 1947-I to 1989-I (no. of observations = 169).
| Table 3 | Model $\Delta C_t = \alpha + u_t - \delta u_{t-1}$ fitted to monthly data \(^9\)  
|         | (in parenthesis: standard errors)  
|---------|----------------------------------|---------------------------------|
|         | $\alpha$ | $\delta$ | $\delta$ at $m = 2$ | $\delta$ at $m = 4$ | $SN^*$  
|         |          | (durab.over month) | (durab.over fortnight) | (durab.over week) |  
| DURABLES | 3.252 (1.075) | 0.376 (0.048) | 0.620 | 0.795 |  
| NON-DURABLES | 3.073 (0.737) | 0.364 (0.049) | 0.601 | 0.780 | 2.03  
| SERVICES | 8.141 (0.603) | 0.260 (0.050) | 0.528 | 0.729 | 4.41  
| SERVICES + NON-DURABLES | 11.22 (1.13) | 0.253 (0.051) | 0.503 | 0.719 | 4.80  
| TOT. CONS. | 14.483 (1.873) | 0.254 (0.050) | 0.517 | 0.729 |  

\(^9\) NIPA series of per-capita, monthly, seasonally adjusted, constant dollars '82, from 1959 to 1989 (no. of observations = 366).