GARBAGE IN - ECONOMETRICS OUT?

by

Luigi Ermini*
Fred Joutz**

*University of Hawaii, Honolulu
**Washington University, Washington, DC

Garbage In - Econometrics Out?

by Luigi Ermini - University of Hawaii, Honolulu

and Fred Joutz - George Washington University, Washington DC

March 1993

How faithful should one be to the purist's view that in econometric work one should never look at the data before testing, lest the test outcome be prejudiced (or lest one is prepared to use a much larger confidence level than the usual 5%)? If the purist advice is followed, how believable is their contention that ultimately data "problems" will be detected by the proper use of diagnostics? Indeed, from a purist's perspective, the chemist's rule that one should always make sure that glassware is properly cleaned before laboratory experiments appears laughable, but is it? The following brief story provides some material suggesting that perhaps the chemist may have the last laugh.

The authors of this note set out to retrieve NIPA consumption data from an available Citibank database, to carry out empirical work on monthly consumption and its various components. They retrieved the seasonally adjusted, constant 82-dollar series (in Citibase notation) GMCN82, GMCS82 and GMCD82, and all related components. As it happens, nowadays data can be fed from the Citibase tape directly to an econometric software package, making the purist's view even more compelling. As the research project was to extend Ermini's [1992] work on the durability of services and non-durable goods to the NIPA disaggregated components of consumption, the autocorrelogram of the various series was estimated, and Ermini's model tested. Various diagnostic checks did not reveal anything "irregular", so that, under a strict interpretation of the purist's approach, we had our results and could proceed with writing the empirical and concluding sections of the paper.

However, out of habit - one of the two authors was previously a chemical engineer - time plots of the various series were also produced. Two plots in particular caught the authors' attention: figure 1 which reports GMCHO8 (monthly expenditures for household operations), and figure 2 which reports GMCN82 (monthly consumption for non-durables) and some of its components. The former exhibits an unlikely structural break; the latter has the total smaller than its components. Evidently, something was wrong with this data set.

An inquiry to the database issuer alerted us about the availability of a new data series based on the benchmark year 1987 instead of 1982. Comparing pairs of series revealed that

most series of the 1982 data set were consistent with the 1987 data set. For example, figures 3-6 report total monthly consumption of services (GMCS), monthly consumption of electricity and gas (GMCHU), monthly consumption of transportation services (GMCST), and monthly consumption of gasoline and oil (GMCNG) respectively. The suffix "8" refers to 82-series, the suffix "q" to 87-series. Clearly, all these series are consistent pairwise, the differences being related to a rescaling of the corresponding price indexes.

As suspected, however, some series of the 1982 data set exhibited obvious inconsistencies. Compare figure 7, which reports the new monthly 87-series of expenditures for household operations (GMCHOQ), with the corresponding 82-series of figure 1. More troublesome is the comparison of monthly consumpion for electricity (figure 8 vs figure 9 for GMCHE), of monthly consumption of natural gas (figure 10 vs figure 11 for GMCHG), and of monthly consumption of fuel oil and coal (figure 12 vs figure 13 for GMCOF): the long-term trending movement of the 82-series are opposite to the long-term trending movement of the 87-series! (In fact, the former series look more like price indices than consumption series.) Finally, the monthly 87-series of total consumption of non-durables confirmed the add-up problem mentioned above; furthermore, comparing figures 14 and 15 also revealed different shapes and trends not justifiable by a change of benchmark years.

The central issue here is not to try to establish the origin of these irregularities [1]. Rather, the relevant point is that looking at data plots of the 1982 data set turned out to be a crucial step in making sure that we were dealing with the "right" data set. One may ask whether these irregularities would be revealed by proper diagnostics. We doubt it and, of course, now it is too late to answer this question (the answer would be prejudiced)! We all agree that the validity of an experiment depends on the quality and consistency of the data. Standard practice in econometrics condones the removal of such outliers as, for example, negative personal income in panel studies or recordings with unclassified values. So the "look at the data first" advice is often followed by the experimenter, at times in an indirect and disguised way.

The real question - to date unanswered - is how far can one go with this advice. Leamer [1978] refers to *excessive* prior examination of the data as a form of specification search instigated by the data - or Sherlock Holmes inference. As Holmesian search may reveal anomalies in the realized series that suggest an alteration in the hypothesized model, Leamer proposes to penalize this prior investigation. In practice, it is not clear how such penalties would be determined. On the contrary, other econometricians advocate the policy of letting

---

1. Our initial downloading of the data occurred in April 1992. At that time - and still now - the CITIBASE files included both the 82 and the 87 data set. We retrieved our data from the 82 data set for consistency with Ermini [1992], without knowing that CITIBASE no longer defends the reliability of this data set. We checked the 1982 series using the May 1st, 1991 release of CITIBASE - i.e. prior to the publication of the 1987 data set - and found none of the irregularities described above.

the data "speak for itself" without penalty as a way to provide a valuable benchmark against which specifications of economic conjectures can be properly tested (for example, Spanos [1986], Hendry [1989]). So, the debate still rages. In the interim, by first looking at the data, the two authors managed to avoid an embarassing inference from bad data.

References

CITIBASE, 1991, Citibank Economic Database (machine-readable magnetic data file). 1946-Present, New York, Citibank, N.A.

Ermini L., 1992, "On the durability of non-durable goods: Some evidence from US time series data", *Economics Letters*, 39, 135-141.

Hendry D.F., 1989, *Lectures on Econometric Methodology*, mimeo, Oxford University

Leamer E.E., 1978, *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley, New York.

Spanos A., 1986, *Statistical Foundations of Economic Modelling*, Cambridge University Press.
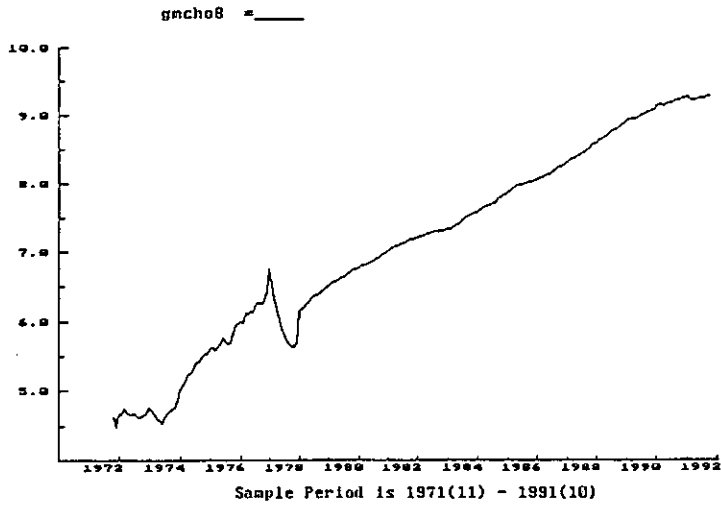
Figure 1
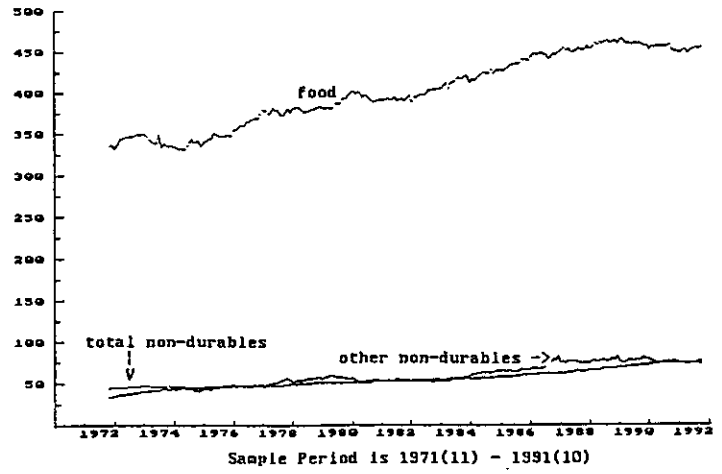
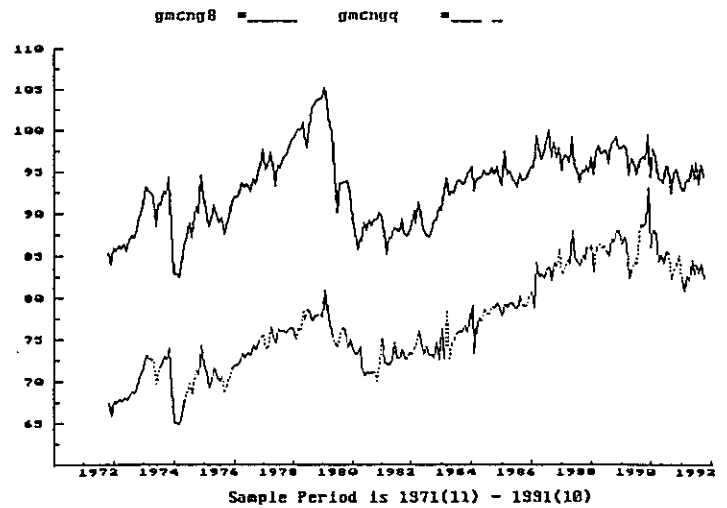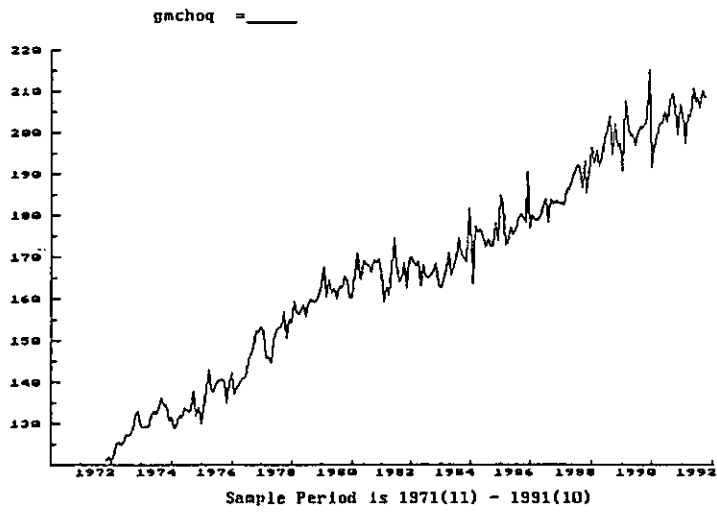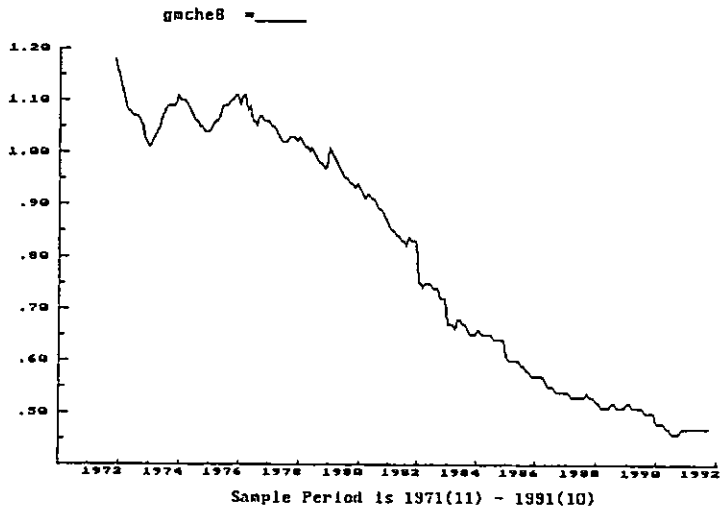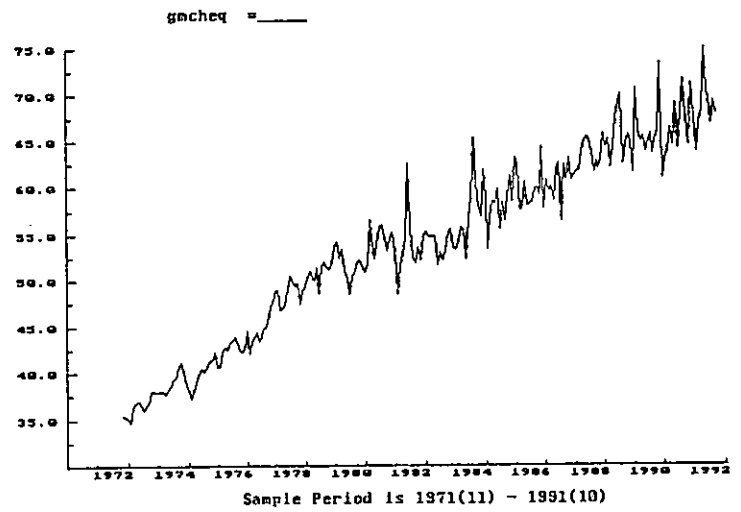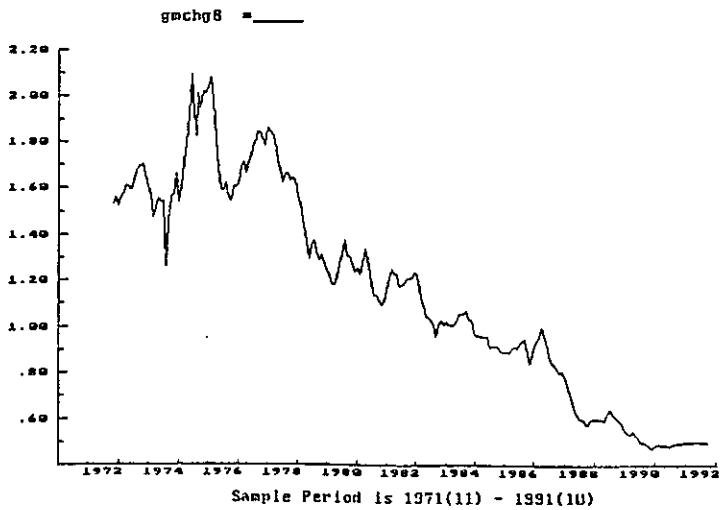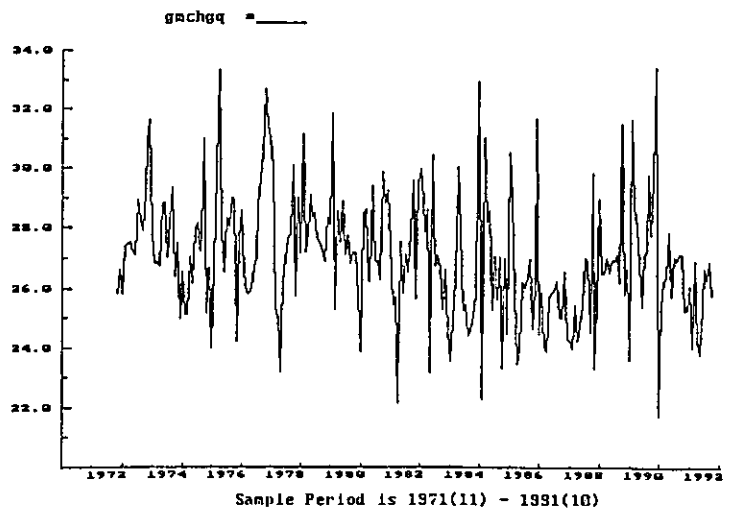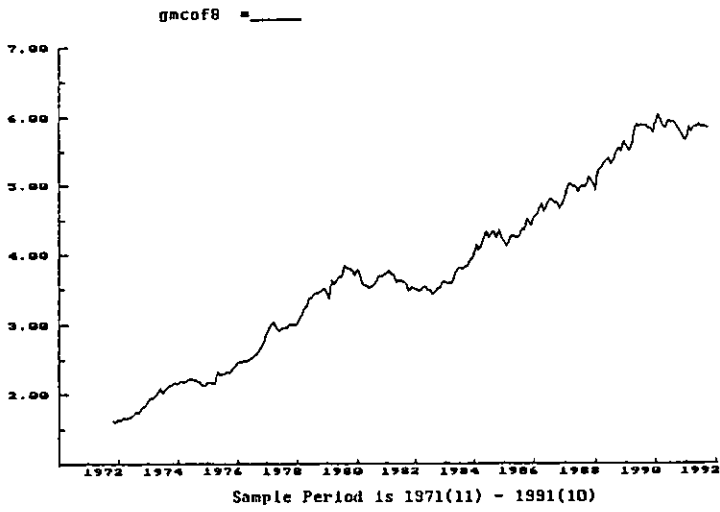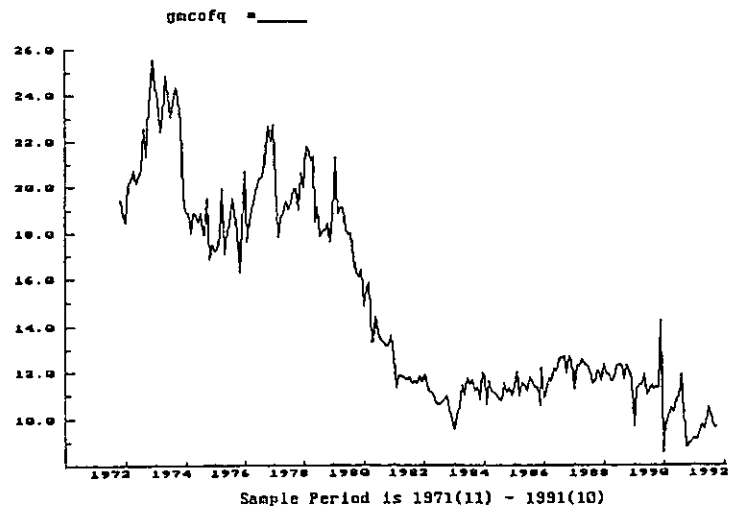

Figure 2
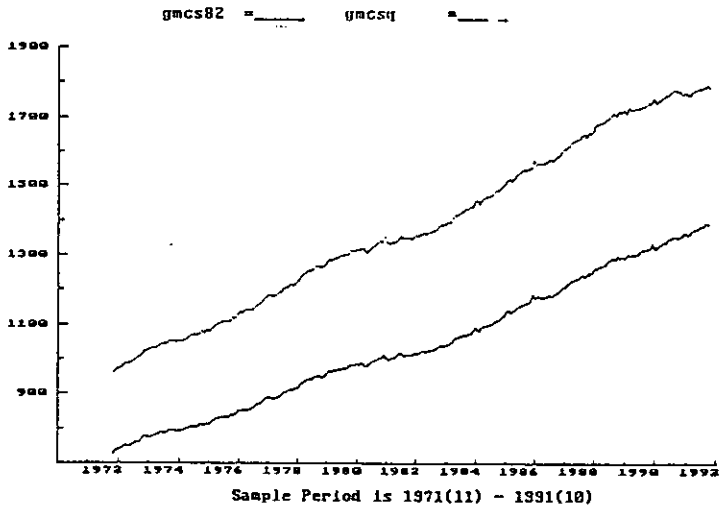
Figure 3
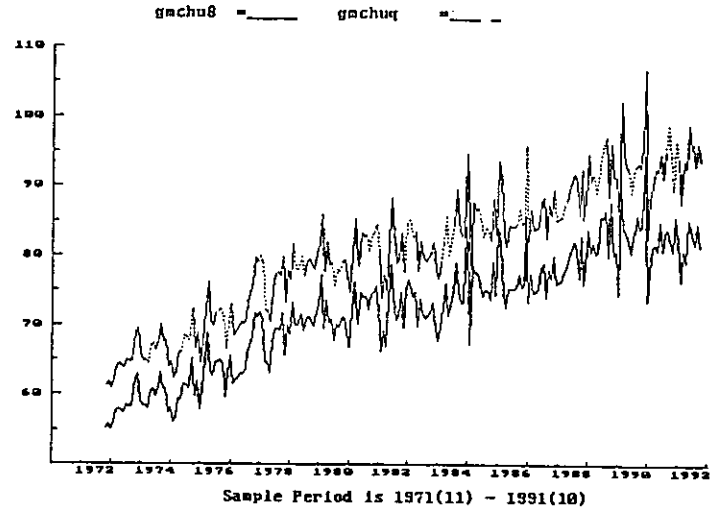
Figure 4
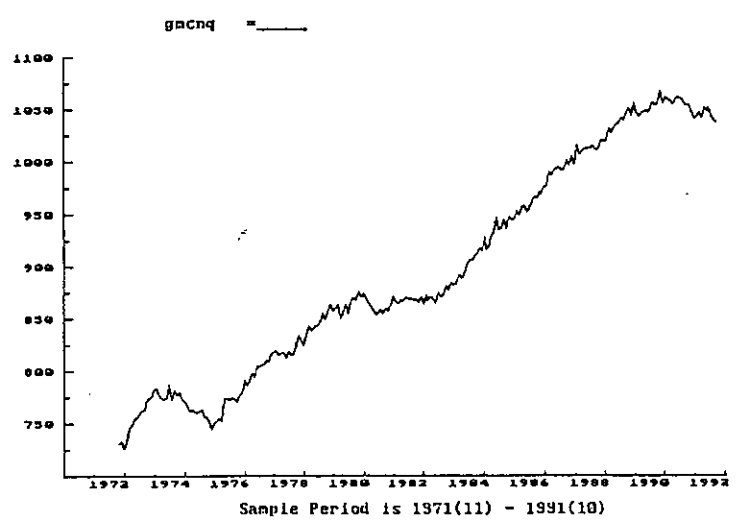

Figure 5

Figure 6

gmchoq =_____

Sample Period is 1971(11) - 1991(10)

Figure 7

gmche8 =_____

Sample Period is 1971(11) - 1991(10)

Figure 8

gmcheq =_____

Sample Period is 1971(11) - 1991(10)

Figure 9

gmchg8 =_____

Sample Period is 1971(11) - 1991(10)

Figure 10

gmchgq =_____

Sample Period is 1971(11) - 1991(10)

Figure 11

gmcof8 =___

Sample Period is 1971(11) - 1991(10)

Figure 12

gmcofq =___

Sample Period is 1971(11) - 1991(10)

Figure 13

gmcn82 =___

Sample Period is 1971(11) - 1991(10)

Figure 14

gmcnq =___

Sample Period is 1971(11) - 1991(10)

Figure 15