mates remain valuable documentation of the existence and likely range of outlet bias in the CPI for the products covered by this study.

Conclusion

Showing that a problem exists in the CPI is usually much easier than developing a workable and accurate solution for it. (That is, of course, the way things should be—we would expect all the easy problems to have been solved by now!) Alternatives to the assumption that the differences in price for items sold side-by-side in the same market are a measure of the value of their differences in quality are not easy to implement. In the case of outlet substitution bias, estimating quality adjustments for outlets from CPI samples is especially difficult because prices from different stores often represent varieties of differing quality levels. The diversity of varieties in the CPI follows from the need to obtain representative samples of the varieties purchased by consumers.

In the mid-1960s BLS asked Edward Denison to provide expert advice on improving the CPI. One of his remarks was that ideally prices should be collected from households rather than from stores, so the prices that are actually paid could be reflected in the index. Of course, he added, this would never be practical. Now the authors of the current chapter have used a unique data set to do just that. In doing so, they have provided important new evidence on the possible magnitude of outlet substitution bias in one component of the CPI.

References

Betancourt, R. R., and D. Gautschi. 1992. The demand for retail products and the household production model: New views on complementarity and substitutability. *Journal of Economic Behavior and Organization* 17 (2): 257–75.

———. 1993. Two essential characteristics of retail markets and their economic consequences. *Journal of Economic Behavior and Organization* 21 (3): 277–94.

Griliches, Z., and I. Cockburn. 1994. Generics and new goods in pharmaceutical price indexes. *The American Economic Review* 84 (5): 1213–32.

Fisher, F., and Z. Griliches. 1995. Aggregate price indexes, new goods, and generics. *Quarterly Journal of Economics* 110 (1): 229–44.

Comment Mick Silver

This excellent chapter addresses the important issue of outlet substitution bias. There is much in the methodology that is to be commended. The con-

Mick Silver is a senior economist in the Statistics Department of the International Monetary Fund. The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

cern of this comment lies with the empirical finding that the BLS CPI-U food at home inflation is too high by about 0.32 to 0.42 percentage points. The problem this comment draws attention to is the reliability of the finding given the data source employed, as opposed to econometric issues. It is first worth distinguishing between two types of scanner data: the consumer panel data used in this study that use handheld bar code scanners, as opposed to diaries to record purchases and retail bar code electronic point-of-sale (EPOS) scanner data.

Scanner Data

The study used ACNeilsen Homescan data. Such data arise from a consumer panel regularly scanning the bar codes on their shopping basket purchases. The data benefit from the attachment of the demographic characteristics of the household to the purchases and a history of purchase behavior so that, for example, repeat purchases can be identified as well as switching patterns. The coverage and representativity of the panel, and its consistency over time is restricted to, and relies upon, the efficacy of the selection of panel members, their attrition, and replacement. If it is true that the sample members are selected in a manner that they are more price sensitive than other members of the population, then the estimates of outlet substitution effects will be overstated.

A second type of scanner data is bar code point-of sale (POS) scanner data. Such data are compiled from the scanned transactions at the point of sale and have an impressive coverage of transactions. In a period of, say, a month, the quantity of sales can be aggregated of a particular variety of a product, and its transaction price summed and a unit value calculated. The summation can be over outlet-types and the product variety codes can be linked to files that contain detailed product characteristics for each variety. Such data can cover the vast majority of transactions and are unlikely to be subject to selectivity bias. However, such data cannot identify the demographics of each purchaser nor their purchase history as they substitute between outlets. Therein lays the advantage of the consumer panel data.

The chapter[1] describes the data panel used in the study as consisting of a subsample of transaction level data from the Fresh Foods Homescan Panel, ". . . approximately 61,500 randomly selected households across the US [. . .] randomly recruited to join the panel using sampling techniques to ensure household representation for demographic variables such as household income, family composition, education, and household location." The chapter notes that the panel is ". . . subject to turnover from normal attrition or adjustments to demographic targets necessitated by Census revisions." Also, that "Households lost through attrition are replaced with others having

1. I use the NBER working paper version of the chapter as the most recent available: http://www.nber.org/papers/w10712.

similar key characteristics". Emphasis is given to the fact that the sample is geographically dispersed and demographically balanced so that the sample matches the U.S. population as closely as possible.

There are two issues of concern. The first is the representativity of the sample of households due to nonresponse (self-selectivity) bias and the second, the attrition rate and replacement policy.

Nonresponse Selectivity Bias

The authors have noted that the sample selected is a random one, that each household, within the practicalities of such things, will have an equal chance of selection. There is not, to the knowledge of author of this comment, nor referenced in the chapter, any information on sample design, but we take this on trust. The method used to recruit and maintain panel members is again, not to our knowledge, documented. Our understanding is that Homescan recruit their members by first mailing the sampled householders and asking if they are willing to take part in a (regular) survey of spending in return for coupons and product information. Those who respond form the potential sample. All (or a sample) of those who complete the forms regularly and well (to some standard) are then selected for membership of the panel. We know little of the reward structure for being a panel member. Such members are likely have lower search costs, be better informed about prices, and be more price sensitive. We have no information on the nonresponse rate—the number of recruits to a panel over the number of mailings sent out to households. There are not unreasonable grounds to believe that this may be very high. The sample was post-stratified according to a number of Census-based demographic and geographic targets. However, such post-stratification is of course not in itself sufficient to remove selectivity bias. The sample comprises those households in each stratum who are more price conscious in the sense that they respond to calls for filling out forms and recording their purchases for coupons and rewards and stay with it. Findings of high price sensitivity and substitution behavior remain open to the charge that they are artifacts of the selected data.

Of course the loose description of these methods may be false and the nonresponse rate may be very low. But when similar comments were made by the author to a group of researchers who use the data there was no disagreement with the essentials of the point previously made.[2]

2. http://www.farmfoundation.org/projects/documents/ScannerDataWorkshopSummaries2_000.pdf. A Workshop on the Use of Scanner Data in Policy Analysis, Economic Research Service, USDA and the Farm Foundation, Washington DC, June 2003. The website includes summaries of the papers where shortcomings of Homescan data, as well as its very real benefits, are highlighted. In particular, see Helen Jensen, "Demand for Enhanced Foods and the Value of Nutritional Enhancements of Food" and J. Michael Harris, "Properties of Scanner Data."

Attrition and Replacement

The study uses a subset of the data to represent household purchases of food for at-home consumption. The study was over the period 1998 to 2001 and the Homescan sample used included 12,000 households in 1998 and 1999, increasing to 15,000 in 2000 and 2001, but was restricted to households that participated in the panel for at least ten out of the twelve months per year. This reduced the sample size in 1998, 1999, 2000, and 2001 to 7,624, 7,124, 7,523, and 8,216 households, respectively; by about one-third in 1998 and 1999 and one-half in 2000 and 2001. This reinforces the predisposition of the sample to the price conscious shoppers. Of the price sensitive households recruited only the more committed ones remained in the sample.[3]

Yet the sample may be further biased since there were in fact only 9,501 unique households over the four years, of which only ". . . 5,247 households participated for all four years, 1,877 households participated for three years, and 2,377 households were one year participants." Thus, of those households with over ten to twelve months of membership who constituted this self-selected sample, whose purpose is to reflect the purchase patterns of a representative consumer, over a half (5,247/9,501) were price-conscious households who had chosen to monitor, in return for the incentives, their shopping behavior *for at least four years.*

There should always be in economic analysis an awareness that what goes around in biased sample selection comes around in biased results. But all of the above remains a suspicion about possible bias and, by the nature of the data, one difficult to test. Some such testing can be carried out. For example, data are available by duration of panel membership and the estimation of substitution effects by duration of panel membership will give some insights into possible bias, though there would remain the problem of determining the selectivity bias from the initial self-selectivity decision. The purpose of the comment is of course only to draw attention to such possible selectivity bias and is not to negate the usefulness of the chapter's contribution in drawing further attention to outlet-substitution bias and providing a methodological basis for analyzing its effects.

3. J. Michael Harris makes a similar point: "In the HomeScan data set only 12,000 households reported both UPC and random weight purchases. However, if you restrict the sample to households present in the data for 10 of 12 months in 1999, only purchases for 7,195 households are available. Indeed, it is clear that all households are not present in the purchase data for every month. This situation can potentially create estimation problems for researchers and can magnify the censoring problem, especially when individual products are examined." (p. 25 of website, see footnote 2.).