

## Appendix A7

### **Biases and Shortcomings in the Construction of Provincial-level Absolute Prices**

Our findings are subject to a wide range of potential biases. One set of biases is related to the proper prices for rural vs. urban areas, and another set to the proper price index to obtain basket costs in other years.

#### *Problems with absolute prices*

The rural price level is likely to be *underestimated* for a number of reasons. Lacking rural prices for many consumer goods, we have adopted the retail prices in the 29 provincial capitals. Unless the distribution system is subsidized by the government, these consumer goods are likely to be more expensive in the countryside due to transportation costs. The implicit price of grain and the composite price of grain may be too low in 1990 when the marginal (market) price was significantly higher; but the same (low) prices as are used in the household survey living expenditures are also used in the income derivation, and to maintain comparability of the basket cost with income data, we used the available prices which underlie both expenditure and income data. Rural land prices, which presumably are included in construction costs, are likely to be too low in 1990, when the proper valuation of land was barely beginning in the urban areas.

On the other hand, the rural price level could be *overestimated* in comparison to urban areas. The price of energy may be lower in the countryside than in the 29 provincial capitals (whose prices we assume to hold for the whole province). For some foods products we have to use retail prices in the provincial capitals; these are likely to be too high, especially in the case of vegetables.

Under- or overestimation of prices in the countryside, compared to urban areas, does not matter for comparisons across rural areas, or for comparisons across urban areas, as long as the degree of bias is the same across all provinces. It does matter for direct rural-urban comparisons based on the joint basket.

Implicit pricing of staples in rural areas, as is done in the rural and joint basket in the paper, makes allowance for local substitution, which is desirable if the purpose is to price “grain” as an aggregate consumer good, but not if specific types of grains are to be priced all over the country. In urban areas, there may be no big difference, if all types of staples, arriving through a state distribution system, come at a similar price. In rural areas there could be big price differences among different types of staples, with the dominant type of staple much cheaper than other types. The use of composite agricultural procurement prices of grain to price the rural basket (as done in a table in an appendix) enforces the pricing of a specific mixed-grain sub-basket in each province.

Not all products included in the basket are a perfect choice—the choice was limited by data availability. For example, in the case of clothing, the basket is heavy on cloth in contrast to manufactured (finished) pieces of clothing.

The urban basket completely misses out on a number of products, namely those not captured in the urban monetary living expenditure survey. On the one hand, this means the

basket values are perfectly suitable to spatially deflate urban disposable income (which also does not take into consideration in-kind income). On the other hand, one may wonder what the impact of including *all* living expenditures on price levels would be; given what one could expect the amount of urban in-kind living expenditures to be, it would presumably be small.

### *Problems with CPIs*

The second big issue is the choice of price indices to obtain basket costs in other years. China's CPI, as the CPIs of other countries, is likely to be biased upward. Among the causes of this upward bias are substitution bias, outlet bias, weighting bias, and measurement problems for quality changes.<sup>1</sup> In the case of the U.S., two comprehensive estimates of the overall annual bias in the CPI were 1.1% (Michael Boskin et al., 1996) and 0.87% (David Lebow and Jeremy Rudd, 2003), with much of the correction due to quality changes.<sup>2</sup> As Gale Johnson (2001) argues, these biases are likely to be much more severe in China: "In fact, given the much higher growth rates of real per capita incomes in China than in the United States since 1978, the biases created by outdated weights and changes in the quality of products would be much greater in China." (p. 3)

The NBS reveals little information about how it calculates the CPI. Presumably, price subindices are Laspeyres price indices and aggregated into the overall rural, or urban, or total CPI using arithmetic means, i.e., substitution bias is not controlled for. No further details are available on the choice of outlets.

The expenditure weights are likely to be held constant over one decade, which at a time of rapid economic growth leads to large inaccuracies at the end of the decade. At a real growth rate of, for example, 7%, real expenditures double within a decade; in a developing country much of the increase in expenditures is likely to occur on non-foods, thus changing the relative weights of different product categories in the living expenditures. This may be even more true for urban than for (the poorer) rural areas.

---

<sup>1</sup> Substitution bias refers to lower-level substitution bias and upper-level substitution bias. Thus, the CPI does not take into account that consumers switch to slightly different specifications of a product (a close substitute) if the price of the close substitute falls (lower-level substitution bias); consumers may also substitute between product categories (upper-level substitution bias). The CPI ignores substitution because it uses base-year product specifications and base year expenditure weights. Outlet bias refers to the recent spread of supermarkets and discount stores which may be underrepresented in the original sample of outlets (and a switch to new outlets will not lead to a correction of the index if prices in the new outlet are linked to the old index). Weighting bias occurs if the expenditure weights used in the calculation of the CPI are incorrect. Quality change refers to changes in quality that are missed by price surveyors, or to the introduction of a completely new product if it enters the index without taking into account the improvement (reduction in living costs) it brings.

<sup>2</sup> For an extensive discussion of the CPI in the U.S. see the Winter 1998 and Winter 2003 issues of the *Journal of Economic Perspectives* with contributions in the Winter 1998 issue by, among others, Michael Boskin et al. (with a description of how the CPI was derived in the mid-1990s), Katharine Abraham, John Greenlees, and Brent Moulton (cautioning that the CPI may well *underestimate* inflation when it comes to quality issues), Robert Pollak (in particular on the democratic vs. plutocratic weights), Angus Deaton (cautioning against quality corrections and warning about the implications of the plutocratic expenditure weights), and in the Winter 2003 issue by Charles Schultze (reporting on the recommendations of the Committee on National Statistics of the National Academy of Sciences), Katherine Abraham (describing recent improvements to the CPI, such as a switch to geometric rather than arithmetic means in the compilation of many strata price subindices), and Jerry Hausman (with another evaluation of the biases in the CPI).

The U.S. CPI is in fact an urban CPI, covering approximately 87% of U.S. households. As of 2002, it covers 186 item strata in 87 urban areas; in the case of about 60% of the item strata, prices are aggregated into strata price subindices based on geometric means, while further aggregation towards the overall CPI is based on arithmetic means. Price information is collected on (in 1996) 71,000 goods and services at about 22,000 retail outlets.

Furthermore, as explained in the appendix on the derivation of price levels for years other than 1990, provincial CPIs incorporate provincial weights for the different product categories. Rather than obtain provincial prices of a nationwide uniform basket in years other than 1990, the further away from 1990 one goes, the more likely the derived provincial basket value reflects the local price of an increasingly localized (rather than nationwide) basket.

One big concern is new products, which the NBS may well first introduce to the CPI at their current prices, without being able to correct for large improvements that the new product may have in terms of the fundamental services they provide at a given price (which means the CPI should fall) or to correct for monopolistic pricing at the initial stages (which means the CPI should rise). On the other hand, given that the NBS has issued a nationwide uniform list of products to include in the provincial CPIs, and given that this list appears to not be updated for the same periods as specific expenditure weights are adopted, new products may well be incorporated into the CPI with a time lag of many years only.<sup>3</sup> (In the U.S., cellular phones, which took 18 years to be included, are the extreme example.) The CPI then is likely to overestimate inflation.

If the set of specific products on which prices were collected in the 1990s was determined in the late 1980s, there is also a fair chance that the selection from today's point of view is heavily biased towards traditionally planned products, which are likely to be basic consumer goods. The available absolute retail price data for 1990, used repeatedly in this paper, for example, do not cover telephones, computers, VCRs, or any other recent technological gadgets. If basic consumer goods are products with little scope for price changes, while the prices of technology products fall, the CPI overestimates inflation.

In our paper, given the focus on spatial deflation, CPI issues do not constitute a problem as long as the biases are proportionally (to base-year basket values) the same across all provinces, and, in the case of rural-urban comparisons, the same across rural and urban areas. The latter may not hold well if new products are priced differently across China's provinces and rural vs. urban areas.

Quality improvements may be most rapid in urban areas where new products appear first, implying that the overestimation of the CPI is most severe in urban areas (as Gale Johnson believes).<sup>4</sup> This, in turn, would imply that prices in urban areas in 2000 are overestimated. In a comparison of spatially deflated rural vs. urban incomes, we would correct the urban income downward (relative to the nominal income) too much. Quality differences between rural and urban areas are likely to be particularly pronounced in housing construction, and it is unclear whether the construction costs properly capture all quality differences.<sup>5</sup>

---

<sup>3</sup> In the U.S., approximately 30% of sample items whose prices the Bureau of Labor Statistics collects disappear every year due to natural attrition and need to be replaced; in two-thirds of these cases, a comparable product can be found, but in one-third of the cases a different-quality product has to be integrated. Furthermore, as new stores are rotated into the sample each year with a complete rotation over four years, 25% of the sample of items sold in retail outlets are automatically replaced; the BLS chooses outlets and then specific products in a specific outlet. (Charles Schultze, 2003, p. 12) In other words, the actual list of products on which prices are collected is changing continuously. No information is available on how China's NBS handles such changes; while product attrition and sample store rotation may all be less expansive in China, these problems cannot be avoided altogether.

<sup>4</sup> Gale Johnson (2001) also questions the accuracy of various rural price deflators in the early 1980s; these are likely to strongly underestimate the actual degree of inflation in the countryside.

<sup>5</sup> David Lebow and Jeremy Rudd (2003) elaborate on the numerous complications of deriving a proper housing price index in the U.S.

On the other hand, the urban spatial deflators are likely to be too small because urban in-kind income is likely to have fallen, relative to monetary income, over time. Reductions in in-kind income were probably offset by increases in monetary income. The CPI does not capture this drastic increase in prices from zero (or heavily subsidized low prices of, for example, communal services) to market prices. Urban income may have to be deflated much more than the CPI implies. Again, this does not matter for cross-province comparisons if the relative changes are the same across all provinces, but it matters for rural-urban comparisons.

One double-check on the validity of our derived year 2000 basket values is possible for the urban case, with a year 2000 urban basket and year 2000 urban prices. Using the limited price data available for 2000 in the *Price Yearbook* (also see appendix on sources of price data) and year 2000 quantity data from the *Urban Household Survey Yearbook* allows the construction of a year 2000 urban basket, priced at year 2000 urban prices. (Trying to rely as much as possible on absolute prices rather than unit values implies larger adjustment factors for the individual product categories than in the case of the 1990 basket.) The resulting provincial urban basket values of the year 2000 are highly correlated with our derived urban basket values of the year 2000 (1990 urban basket values multiplied by urban CPIs), with a correlation coefficient of 0.7758, significant at well below the 0.1% significance level. This suggests that the derivations in our paper are highly valid.

### **References relevant for this appendix and not listed in the paper**

- Abraham, Katharine G., John S. Greenlees, and Brent R. Moulton. "Working to Improve the Consumer Price Index." *Journal of Economic Perspectives* 12, no. 1 (Winter 1998): 27-36.
- Abraham, Katharine G. "Towards a Cost-of-Living Index: Progress and Prospects." *Journal of Economic Perspectives* 17, no. 1 (Winter 2003): 45-58.
- Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson. "Consumer Prices, the Consumer Price Index, and the Cost of Living." *Journal of Economic Perspectives* 12, no. 1 (Winter 1998): 3-26.
- Deaton, Angus. "Getting Prices Right: What Should Be Done?" *Journal of Economic Perspectives* 12, no. 1 (Winter 1998): 37-46.
- Hausman, Jerry. "Sources of Bias and Solutions to Bias in the Consumer Price Index." *Journal of Economic Perspectives* 17, no. 1 (Winter 2003): 23-44.
- Johnson, D. Gale. "Have the Urban-Rural Disparities Increased Since 1978 in China?" Mimeo, Office of Agricultural Economics Research, The University of Chicago, 8 November 2001.
- Lebow, David E., and Jeremy B. Rudd. "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature* XLI, no. 1 (March 2003): 159-201.
- Pollak, Robert A. "The Consumer Price Index: A Research Agenda and Three Proposals." *Journal of Economic Perspectives* 12, no. 1 (Winter 1998): 69-78.
- Price Yearbook. Zhongguo wujia nianjian* (China Price Yearbook). Beijing: Zhongguo wujia chubanshe. Various years.
- Schultze, Charles L. "The Consumer Price Index: Conceptual Issues and Practical Suggestions." *Journal of Economic Perspectives* 17, no. 1 (Winter 2003): 3-22.