



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Does Chinese Inflation Understate Cost of Living?

Jonathan Aaron Cook*

This version: May 2013
Preliminary

Abstract

The Chinese consumer price index slightly overstates the increase in the cost of living and thus understates the growth in real income. Over the period 1995 to 2002, we find that the annual growth in median urban household income deflated by the cost of living was 7%, which is close to the growth in real GDP during this time. When deflated by official CPI, the median urban household sees only a 3% increase in real income.

1 Introduction

A downward bias in the measurement of Chinese inflation would overstate the competitiveness of Chinese products abroad, overstate the magnitude of China's real income growth, and understate the profitability of foreign sellers in China. China experts have questioned the reliability of many of China's official statistics (for example, Rawski [2001], Gale [2002], and Young [2003]). This paper tests for bias in China's consumer price index (CPI) over the period 1995 to 2002. The period 1995 to 2002 is especially interesting because nominal gross domestic product (GDP) grew at 10 percent per year while inflation was only around 1 percent per year and food prices were decreasing by 1 percent per year.

This paper infers the true change in the cost of living from consumer behavior. Engel's law states that a consumer's food share of expenditure is decreasing in his income. The decline in food share of expenditure that, *ceteris paribus*, follows an increase in real income is the

*Economic Research Service, United States Department of Agriculture. The views expressed here do not necessarily reflect the views of ERS or the US Department of Agriculture. Email: JACook@ers.usda.gov

best established empirical regularity in economics (Houthakker 1987). Engel curves for food have been shown to be stable over time and across countries (Almås 2012). Food share of expenditure has long been used to estimate consumer welfare.

This paper's approach is based on Hamilton (2001) and Costa (2001) who use shifts in consumers' Engel curves for food to estimate U.S. CPI bias. This approach to estimating CPI bias does not assume a specific cause of bias. This approach has also been used to estimate CPI bias in developing countries, including post-reform Mexico and Brazil (Filho and Chamon 2012b). This approach has also been used for estimating bias in purchasing power parity (PPP) (Almås 2012).

After finding evidence of bias, we examine the potential sources of CPI bias. One potential bias arises from changes in the variety of available goods. Another problem arises from the difficulty of incorporating quality changes. Comparing average prices with the consumer price index can shed some light on the which cause is more likely. Using average prices for the period 2001 to 2011, we find that new-good bias is not a likely source of bias for Chinese CPI. Substitution bias and quality changes are likely to be responsible culprits.

In examining average prices, we focus on Chinese food prices. Food is an important component of consumer prices and accounts for about one third of the Chinese National Bureau of Statistics (NBS) consumer price index. This paper compares the NBS food price index with rural and urban food price indexes based on average rural and urban prices for specific foods. Our urban index shows an increase in average food prices of 129 percent over 2001 to 2011 while the official price index indicates an increase of only 67 percent. Greater increases in average prices suggests that new-good bias is not a likely source of upward CPI bias.

This paper contributes a measure of Chinese CPI bias and a measure of the growth of real household income for the median urban Chinese household. The relatively small bias gives credibility to the use of Chinese CPI for measuring real income or a CPI-based real exchange rate.

Related work includes Young (2003), who showed that when the GDP deflator is replaced with other price indexes, Chinese real economic growth reduces by 2.5 percent per year. Gong and Meng (2008) use an approach that is also based on Hamilton (2001) to estimate spatial price indexes for China, but do not estimate CPI bias.

In work that was unknown to this author at the time the analysis was completed, Filho and Chamon (2012a) also estimate Chinese CPI bias using the same approach as this paper and find similar results. Filho and Chamon use a different dataset that includes the years considered by this paper. Given the surprising result that the Chinese consumer price index slightly overstates

the increase in the cost of living, it is reassuring that a second study, which uses a different dataset, confirms the results.

In the next section, we estimate cost-of-living bias from consumer behavior. Section 3 describes potential sources of the bias in Chinese CPI. We test for new-good bias in Section 4. Section 5 concludes the paper.

2 CPI Bias from Engel Curves

This section closely follows Hamilton's (2001) work on U.S. CPI bias. We begin with the equation for a household's share of expenditure on food based on Deaton and Muellbauer's (1980) almost ideal demand system

$$\omega_{i,j,t} = \phi + \gamma(\ln P_{f,j,t} - \ln P_{n,j,t}) + \beta(\ln Y_{i,j,t} - \ln P_{j,t}) + \theta X_{i,j,t} + \mu_{i,j,t} \quad (1)$$

where $\omega_{i,j,t}$ is the food share of household i in province j at time t , $P_{f,j,t}$, $P_{n,j,t}$, and $P_{j,t}$ are true, unobserved prices for food, non-food, and all goods, $Y_{i,j,t}$ is income, and $X_{i,j,t}$ is a vector of household characteristics. We have data for two time periods, so that t equals either 0 or T . We observe official measures of inflation, $\Pi_{j,t}$, which are potentially subject to bias. The true inflation for all goods over the period from 0 to T is equal to

$$\frac{P_{j,T}}{P_{j,0}} = (1 + \Pi_{j,T})(1 + \bar{E})^T, \quad (2)$$

where \bar{E} is the average annual CPI bias. We can analogously define true inflation for food and non-food items as

$$\frac{P_{f,j,T}}{P_{f,j,0}} = (1 + \Pi_{f,j,T})(1 + \bar{E})^T$$

and

$$\frac{P_{n,j,T}}{P_{n,j,0}} = (1 + \Pi_{n,j,T})(1 + \bar{E})^T.$$

The equations above assume that bias in the measurement of food and non-food is the same and that bias does not differ by region. We replace the true prices in Equation (1) with observed inflation to find

$$\begin{aligned} \omega_{i,j,t} = & \phi + \gamma [\ln(1 + \Pi_{f,j,t}) - \ln(1 + \Pi_{n,j,t})] + \gamma [\ln P_{f,j,0} - \ln P_{n,j,0}] \\ & + \beta [\ln Y_{i,j,t} - \ln(1 + \Pi_{j,t})] - \beta t \ln(1 + \bar{E}) - \beta \ln P_{j,0} \\ & + \theta X_{i,j,t} + \mu_{i,j,t}. \end{aligned} \quad (3)$$

From Equation (3), we replace

$$\gamma [\ln P_{f,j,0} - \ln P_{n,j,0}] - \beta \ln P_{j,0}$$

with a dummy variable for province, denoted by D_j , and we replace $\beta t \ln(1 + \bar{E})$ with a dummy variable for year T , denoted by δ_T . Equation (3) is now

$$\begin{aligned} \omega_{i,j,t} = & \phi + \gamma [\ln(1 + \Pi_{f,j,t}) - \ln(1 + \Pi_{n,j,t})] + D_j \\ & + \beta [\ln Y_{i,j,t} - \ln(1 + \Pi_{j,t})] + \delta_T + \theta X_{i,j,t} + \mu_{i,j,t}. \end{aligned} \quad (4)$$

We can find cumulative CPI bias over T years from

$$\exp(-\delta_T/\beta) = (1 + \bar{E})^T$$

and find the average annual bias from

$$[\exp(-\delta_T/\beta)]^{1/T} = 1 + \bar{E}.$$

2.1 Data

We require a large repeated cross-sectional dataset that includes income, share of expenditure on food, and other household characteristics. It appears that there are only two datasets that contain the necessary information for estimating China's CPI bias: the Chinese Household Income Project Survey and the Urban Household Survey. This paper uses the Chinese Household Income Project Survey; Filho and Chamon (2012a) use the Urban Household Survey. Meng (2012) provides an excellent description of available Chinese data.

The Chinese Household Income Project (CHIP) 1995 and 2002 datasets provide these variables and also a number of household attributes (Riskin, Renwei, and Li 1995; Li 2002). CHIP provides a random sample of over 6,000 urban households in each survey year. While the survey covers both urban and rural households, we only use urban households as this analysis is more difficult for rural households who may consume vegetables that they have grown. These surveys cover the provinces Beijing, Shanxi, Liaoning, Jiansu, Zhejiang, Henan, Hubei, Guangdong, Sichuan, Yunnan, and Gansu.

For our measure of food expenditure, we do not include expenditure on cigarettes or alcohol. When cigarettes, alcohol, or both are include in food expenditure the results are not greatly affected. Our measure of expenditure does not include expenditure on housing. A large increase in home ownership has occurred between 1995 and 2002.

For our vector of household characteristics, $X_{i,j,t}$, we use the head of household's age, age squared, years of education, the size of the household, and dummy variables for whether or not the household owns a refrigerator, car, and motorcycle. Some variables are not include because the 1995 and 2002 surveys contain different questions. For example, the 2002 survey contains

information on food away from home, but the 1995 survey does not. Table 1 presents summary statistics for these variables.

Our identification requires data on inflation by region. We collect data on inflation for food and for all goods from NBS's Statistical Yearbooks. For our measure of the inflation of non-food, we begin by defining the price indexes

$$\frac{I_{j,t}}{I_{j,t-1}} = 1 + \Pi_{j,t}, \quad \frac{I_{f,j,t}}{I_{f,j,t-1}} = 1 + \Pi_{f,j,t}, \quad \text{and} \quad \frac{I_{n,j,t}}{I_{n,j,t-1}} = 1 + \Pi_{n,j,t}$$

and set these indexes equal to 100 in the base period, $I_{j,0} = I_{f,j,0} = I_{n,j,0} = 100$. We can now express the inflation of non-food items using the relationship

$$\frac{I_{j,t}}{I_{j,t-1}} = \frac{w_t I_{f,j,t} + (1 - w_t) I_{n,j,t}}{w_{t-1} I_{f,j,t-1} + (1 - w_{t-1}) I_{n,j,t-1}},$$

where w_t is urban expenditure on food in year t (taken from Chinese Statistical Yearbooks). Gong and Meng (2008) use regional unit values for the price of food and non-food in a regression equation that is similar to the one presented here in their study of regional price variation.

2.2 Estimate of Bias

Table 2 reports the estimates of Equation (4) for all observations then for more specific groups. Log of expenditure on cigarettes and alcohol is included as a regressor in Table 3. Since Engel's law only holds *ceteris paribus*, it is important to control for other factors that are changing over time. One approach is to focus on a more homogeneous group of households. It is important to remember that cost of living will differ with demographics and estimates of CPI bias for a specific group may not apply to all households. Here we focus on households with working-age head of household that are not too far in the tails of the income distribution. This will allow us to draw inferences regarding middle-class households.

We begin by removing households with food share expenditure less than 2 percent or greater than 80 percent. This is done to remove households in the tails of the income distribution and remove possible incorrectly entered data. Since the cutoff levels of food share are arbitrary, I use the same cutoff levels as Hamilton (2001). We further restrict our attention to households with head of household between 25 and 60 years old. This is done to focus on the more homogeneous group of households with heads that are working age. Next, we focus on three-person households. Over sixty five percent of the observations are three-person households with the next largest group being four-person households at sixteen percent.

We use expenditure in place of income. As noted by Costa (2001), using expenditure better accounts for permanent income. Estimates using income will also be biased by the change in savings rates during this time period. (An excellent description of the change in household

savings in available in Chamon and Prasad [2010].) A potential problem with using expenditure is that, when using expenditure on both the left and right hand side of the equation, measurement error may bias the estimates of our regression coefficients. A natural approach to overcoming this is to instrument log expenditure with log income. For reference, estimates of CPI bias using income are presented in Table 3.

As expected, households that own a refrigerator pay a smaller share of their expenditure on food. Households that own a car or motorcycle are also have a smaller percentage of their expenditure on food. The coefficient on log of real expenditure is similar to that found for Mexico (Filho and Chamon 2012b) and for the United States in early 1900s (Costa 2001). When we estimate Equation (4) on only homeowners or by sector of employment or by province, similar results are obtained.

This bias implies that the growth of real income has been faster than what is implied by official measures. Figure 2 presents densities of real household income for three-person households in 1995 and 2002. Unlike with the densities found from official inflation, we see that the mode household income has increased.

For the median household, real income has risen by about 7 percent per year. High levels of real income growth are found throughout the distribution, but the the distribution of income has become more right skewed. Tables 4 and 5 present the views on income inequality by income percentile. The income distribution is viewed as more fair by wealthier households, although a large percentage from all income groups view the income distribution as “not very fair” or “very unfair.” Since we are interested the change in welfare over time, it is interesting to look at self-reported happiness. Unfortunately, the 1995 survey does not contain information on view of income inequality or happiness so a comparison over time is not possible. Table 6 shows that the heads of wealthier households report that they are happier than heads of lower income households (as expected).

The remainder of this paper examines the causes of bias in Chinese CPI as a cost-of-living index.

3 Price Index Bias

We can only talk of bias in reference to some theoretical ideal index. We begin by assuming that there is a representative Chinese consumer. This representative consumer maximizes his utility function, $U(\mathbf{q}_t)$, by selecting a vector of quantities consumed, $\mathbf{q}_t = (q_{1,t}, q_{2,t}, \dots, q_{N_t,t})$, of each of the N_t goods available at time t . The consumer maximizes his utility in every time period given his budget constraint W_t . We begin by defining the consumer’s utility in the base

period as

$$\bar{u} \equiv \max_{\mathbf{q}_0} U(\mathbf{q}_0) \quad \text{subject to} \quad \mathbf{p}_0 \cdot \mathbf{q}_0 \leq W_0.$$

The consumer's expenditure function is defined as the minimum expenditure needed to obtain a level of utility for a given price vector,

$$m(\mathbf{p}_t, \bar{u}) \equiv \min_{\mathbf{q}_t} \mathbf{p}_t \cdot \mathbf{q}_t \quad \text{subject to} \quad U(\mathbf{q}_t) \geq \bar{u}.$$

We can now define a cost of living index as the change in expenditure needed to obtain the same level of utility,

$$\frac{m(\mathbf{p}_T, \bar{u})}{m(\mathbf{p}_0, \bar{u})}.$$

If we assume that the consumer obtains \bar{u} by consuming the same basket of goods in both periods, we can write the cost of living index as a Laspeyres index,

$$\frac{m(\mathbf{p}_T, \bar{u})}{m(\mathbf{p}_0, \bar{u})} = \frac{\mathbf{p}_T \cdot \mathbf{q}_0}{\mathbf{p}_0 \cdot \mathbf{q}_0}.$$

If at time T the consumer can achieve \bar{u} with a basket of goods that is less expensive than \mathbf{q}_0 , we say that the Laspeyres index suffers from "substitution bias." Substitution bias will cause CPI to overstate the increase in the cost of living and is likely to be a source of bias in Chinese CPI. Existing work on substitution bias in the United States' CPI (for example, Braithwait [1980] and Manser and McDonald [1988]) could provide a starting point for making Chinese CPI more robust to substitution of goods.

A further complication arises in that some items may not be available in both periods. A Laspeyres index only includes goods for which there are prices in both periods. We denote the number of items that are available in both month t and month $(t - 1)$ as G_t . The one-month change in the aggregate price index, P_t , is given by

$$\frac{P_t}{P_{t-1}} = \frac{\sum_{i=1}^{G_t} (p_{i,t} \times q_{i,t-1})}{\sum_{i=1}^{G_t} (p_{i,t-1} \times q_{i,t-1})}. \quad (5)$$

Setting the index equal to 100 in the base period, we find the value of the index in the next month by the use of Equation (5). This procedure is repeated to find the value of the index in following months.

It is common to have data on expenditure instead of on consumption. The formula for a Laspeyres index based on expenditure is found by replacing consumption in Equation (5) with

expenditure divided by price to find

$$\frac{P_t}{P_{t-1}} = \frac{\sum_{i=1}^{G_t} \left[p_{i,t} \times \left(\frac{e_{i,t-1}}{p_{i,t-1}} \right) \right]}{\sum_{i=1}^{G_t} \left[p_{i,t-1} \times \left(\frac{e_{i,t-1}}{p_{i,t-1}} \right) \right]} = \frac{\sum_{i=1}^{G_t} \left[\left(\frac{p_{i,t}}{p_{i,t-1}} \right) \times e_{i,t-1} \right]}{\sum_{i=1}^{G_t} e_{i,t-1}} \quad (6)$$

where $e_{i,t-1}$ denotes expenditure in the previous year.

3.1 New-Good Bias and Quality Change

The new-good problem is well known in literature on price indexes (Hausman 2003). This problem arises because, as new goods enter the market, the price differences between the existing goods and the new goods are difficult to incorporate into a price index. As low-cost stores like Wal-Mart proliferated in the United States during the 1990s, unused price changes between existing food sellers and these new grocery stores biased consumer prices upward (Hausman and Leibtag 2009).

Reinsdorf (1993) estimated the magnitude of this bias by comparing the change in the average price of specific food categories with the change in a price index for these the same categories. Reinsdorf showed that average prices were increasing by less than the consumer price index. A new-good bias is created if the average price of entrants is less than the average price of existing goods. Comparing average prices and the consumer price index provides a way of checking if new-good bias is the cause of the slightly downward bias in Chinese CPI.

China's recent development has changed the variety of goods that are available for purchase. Increases in regulation enforcement have increased costs of production. As a result, some lower-priced food items are no longer available. New suppliers that enter the market must comply with the new food safety laws, and thus have higher costs. Although higher priced, new varieties may be higher quality. A consumer requires less consumption of a higher quality variety to achieve the same level of utility. When both prices and quality are increasing, the effect on cost of living could be positive or negative. To see how this change in the composition of suppliers affects a consumer price index, consider the following example of the prices of two varieties of a good over three time periods:

Time	Variety A	Variety B
1	\$1	-
2	\$1	\$2
3	-	\$2

The above example reflects that as safety standards increase some existing low cost sellers, Variety A, are forced to exit the market. New sellers that enter the market, Variety B, will

charge a higher price than the sellers that are leaving the market. A price index over these three time periods, which can be calculated using the equations above, shows zero inflation. This is because the difference in prices between the two items is not considered in the computation of the index. If we look at the change in average prices (from \$1 to \$2), we see a price increase of 100% over these three periods.

If Variety *A* and Variety *B* are of equal quality to the consumer, then the consumer's cost of living has increased by 100%. If, due to increased quality, consuming one unit of Variety *B* gives the representative consumer the same happiness as consuming two units of Variety *A*, then there is no change in the consumer's cost of living. The cost of living will decrease if the quality of Variety *B* is more than twice that of Variety *A*.

In this section, we compare the change in average price with the change in price index to estimate bias in the index that results from new goods. There are other factors that affect average prices, for example migration. The next section describes the effect of migration on average prices and cost of living.

3.2 Migration

Migration affects both average prices and cost of living. Changes in average prices due to migration may cause the change in average prices to overstate our estimate of the change in average prices due to new goods. This is a limitation of using our results on average prices as evidence against new-good bias.

This section shows that the effect of migration on the cost of living is greater than the change in average prices. Our earlier estimate of CPI bias used only urban households and thus cannot capture the effect of migration on average cost of living. Since average prices are increasing faster than the consume price index, it is possible that the consumer price index does not reflect changes in cost of living that result from migration. This affect is difficult to capture empirically.

Regional migration in China has been increasing (Fan 2005). Although the residential permit, *hukou*, system still limits movement, migration control has lessened since the 1980s. Movement to higher price regions for employment opportunities increases the average cost of living faced by a Chinese worker. Changes in cost of living due to migration are not reflected in CPI.

Regional difference in income are long standing and, since the 1990s, have been growing (Jian et al. 1996). During the period 1995 to 2000, over 19 million Chinese migrated form Western and Central China to the more prosperous East (Fan 2005). These migrants have greater employment opportunities, but also face higher prices than they would in the less prosperous regions.

We can understand how regional movement affects the cost of living through a simple ex-

ample. Suppose that the price of a food item is \$1 in Region A and \$10 in Region B . Over time, prices in our fictitious regions are not changing, but there is movement from Region A to Region B . Consider a period of migration that begins with two thirds of the population was in Region A and one third in Region B and ends with one third are in Region A and two thirds in Region B . Over this period, the average price of the food item has increased from \$4 to \$7. A Laspeyres price index, as calculated in Equation (5), shows no inflation. For the one third of the population that migrated, the food item is ten times more expensive.

To see how this bias affects our indexes more formally, let us stay with the setting of two regions and some migration between the regions. Let $\mathbf{p}_{A,t}$ and $\mathbf{p}_{B,t}$ denote prices at time t in regions A and B . To ease notation, consider a continuous population normalized to a mass of one. The measure of the population remaining in Region A is denoted as μ_A , and the analogous population remaining in Region B is denoted as μ_B . The migrant population is of measure μ_M . We define individual j 's expenditure function as above, now denoted with a subscript as m_j , rather than invoking a representative consumer. With a representative consumer, the change in the representative consumer's cost of living is equal to the average change in cost of living. The change in the average cost of living is now given by

$$\mu_A \int_i \frac{m_i(\mathbf{p}_{A,t}, \bar{u}_A)}{m_i(\mathbf{p}_{A,0}, \bar{u}_A)} di + \mu_B \int_j \frac{m_j(\mathbf{p}_{B,t}, \bar{u}_B)}{m_j(\mathbf{p}_{B,0}, \bar{u}_B)} dj + \mu_M \int_k \frac{m_k(\mathbf{p}_{B,t}, \bar{u}_M)}{m_k(\mathbf{p}_{A,0}, \bar{u}_M)} dk. \quad (7)$$

In contrast, a Laspeyres price index, which does not consider regional price differences, gives us a measure of

$$(\mu_A + \mu_M) \int_i \frac{m_i(\mathbf{p}_{A,t}, \bar{u}_A)}{m_i(\mathbf{p}_{A,0}, \bar{u}_A)} di + \mu_B \int_j \frac{m_j(\mathbf{p}_{B,t}, \bar{u}_B)}{m_j(\mathbf{p}_{B,0}, \bar{u}_B)} dj. \quad (8)$$

The migration bias of a Laspeyres price index is approximated by the difference of Equations (7) and (8). This measure of bias ignores nonpecuniary benefits of living in one region over another. The magnitude of the bias will depend on the proportion of migrants and the scale of regional price differences. If a constant basket of goods is consumed in both regions, the bias equals

$$\mu_M \left(\frac{(\mathbf{p}_{A,t} - \mathbf{p}_{B,t}) \cdot \mathbf{q}_{A,0}}{\mathbf{p}_{A,0} \cdot \mathbf{q}_{A,0}} \right). \quad (9)$$

A version of Equation (9) could be estimated using NBS microdata.

Before comparing changes in average prices with official measures of inflation, we need to see how migration affects average prices relative to the cost of living. The following proposition tells us that we should expect the change in average price to understate the change in the cost of living due to migration.

Proposition 1 *When inflation in region B is not too much larger than inflation in region A , the change in average price is less than the change in the cost of living.*

Proof. The change in average prices is given by

$$\frac{\mu_A \mathbf{P}_{A,t} \cdot \mathbf{Q}_{A,0} + \mu_B \mathbf{P}_{B,t} \cdot \mathbf{Q}_{B,0} + \mu_M \mathbf{P}_{B,t} \cdot \mathbf{Q}_{A,0}}{(\mu_A + \mu_M) \mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0} + \mu_B \mathbf{P}_{B,0} \cdot \mathbf{Q}_{B,0}}. \quad (10)$$

The true change in the cost of living is

$$\mu_A \frac{\mathbf{P}_{A,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}} + \mu_B \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{B,0}}{\mathbf{P}_{B,0} \cdot \mathbf{Q}_{B,0}} + \mu_M \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}}. \quad (11)$$

Let us define inflation in region A as

$$\Pi_A \equiv \frac{\mathbf{P}_{A,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}},$$

inflation in region B as

$$\Pi_B \equiv \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{B,0}}{\mathbf{P}_{B,0} \cdot \mathbf{Q}_{B,0}},$$

and the initial difference in price levels as

$$R \equiv \frac{\mathbf{P}_{B,0} \cdot \mathbf{Q}_{B,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}} > 1.$$

We can now write the change in average price as

$$\left(\frac{\mu_A}{\mu_A + \mu_B + R\mu_M} \right) \Pi_A + \left(\frac{R\mu_B}{\mu_A + \mu_B + R\mu_M} \right) \Pi_B + \left(\frac{\mu_M}{\mu_A + \mu_B + R\mu_M} \right) \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}}$$

and the change in the cost of living as

$$\mu_A \Pi_A + \mu_B \Pi_B + \mu_M \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}}$$

Subtracting the change in average price from the change in cost of living yields

$$\mu_A \Pi_A \left(\frac{\mu_A + \mu_B + R\mu_M - 1}{\mu_A + \mu_B + R\mu_M} \right) + \mu_B \Pi_B \left(\frac{\mu_A + \mu_B + R\mu_M - R}{\mu_A + \mu_B + R\mu_M} \right) + \mu_M \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}} \left(\frac{\mu_A + \mu_B + R\mu_M - 1}{\mu_A + \mu_B + R\mu_M} \right).$$

A little algebra shows that the above difference is positive as long as

$$\Pi_B < \left(\frac{\mu_A}{\mu_B} \right) \Pi_A + \left(\frac{\mu_M^2}{\mu_B [1 - \mu_M]} \right) \frac{\mathbf{P}_{B,t} \cdot \mathbf{Q}_{A,0}}{\mathbf{P}_{A,0} \cdot \mathbf{Q}_{A,0}}.$$

■

4 Measuring Average Price Change

4.1 Data

To construct our indexes of average prices we gather average rural free market prices from the NBS's Yearbook of Rural Price Survey and average urban prices from the China Price Information Center, National Development and Reform Commission. To weigh the various food items in our indexes, we use rural household per capita consumption by food category and urban expenditures by food categories, both from NBS's Statistical Yearbooks. The next section describes the construction of the indexes and how the weights are utilized.

4.2 Constructing Indexes

For both our urban and rural data, we have prices for items and weights for broader categories. Table 7 provides a list of categories and food items for the rural price index. Food items within a category receive equal weight. Table 8 provides the categories and food items for the urban index. For urban expenditure, a single expenditure is given for meat, which includes beef, pork, and mutton. We weight beef, pork, and mutton by the percent urban expenditure on meat times the percentage of rural meat expenditure on each item. For example, if 30% of urban expenditure was on meat and 90% of rural meat expenditure was on pork, then we weigh urban pork prices by 27%. This weighting is done to ensure items receive weights that beef and mutton do not receive disproportionately large weights.

We estimate Equation (5) using rural prices and consumption. We estimate Equation (6) using urban prices and expenditure.

There are a few instances of missing price observations. Missing price data is common in the Bureau of Labor Statistic's (BLS) International Price Program. We follow the general approach used by the BLS in handling missing prices. When the price of an item is missing at time t , that item is not used in constructing the index at time t . The missing price is replaced with the last observed price for that item. Feenstra and Diewert (2001) showed that this type of imputation does not lead to a long-run bias in an index.

4.3 Comparing Average Prices and Inflation

Figure 3 compares our urban and rural food price indexes with the OECD food price index. The urban index shows that average food prices increased by 129% over the period 2001 to 2011, whereas the OECD index shows an increase of only 67%. The rural index show an increase of 104% in average food prices over 2002 to 2011 compared with 71% in the OECD index.

We present the change in the price index from the same month in the previous year in Figure 4. The NBS now only releases food price inflation in this year-on-year format. Average prices show far more volatility than the NBS index.

Average prices show more inflation and are more volatile. One shortcoming of the approach taken in this section is that the lack of publicly available inflation for specific categories forces us to construct an index of average price changes. Another shortcoming is that the NBS measure of food price inflation contains more food items than our average price index. It could be that average prices are only collected for the items with the most volatile prices.

If new goods were the cause of the Chinese CPI found in Section 2, it should be that average prices were decreasing. We do not find evidence that would support new-good bias as a cause

of CPI bias. Increased quality of goods could cause average prices to increase faster than the cost of living. Migration also causes average prices to increase.

5 Conclusion

Over the period 1995 to 2002, the increase in the cost of living was less than the change in the consumer price index. The annual increase in median household real income was around 7% during this period. This increase is larger than what is found when median household income is deflated by official CPI. While the distribution of income has become more disperse, there have been significant income increases throughout the distribution.

A possible reason for the bias in Chinese CPI is that, as goods have increased quality, CPI has not reflected those quality changes. Another possibility is that consumers have substituted away from the goods that experienced the largest price increases.

References

- Almås, I. (2012), “International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food,” *American Economic Review*, 102, 1093–1117.
- Braithwait, S. D. (1980), “The Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes,” *The American Economic Review*, 70, pp. 64 – 77.
- Chamon, M. D. and Prasad, E. S. (2010), “Why Are Saving Rates of Urban Households in China Rising?” *American Economic Journal: Macroeconomics*, 2, pp. 93–130.
- Costa, D. L. (2001), “Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves,” *Journal of Political Economy*, 109, pp. 1288–1310.
- de Carvalho Filho, I. and Chamon, M. (2012a), “Consumption Based Estimates of Chinese Growth,” *Unpublished*.
- de Carvalho Filho, I. and Chamon, M. (2012b), “The Myth of Post-Reform Income Stagnation: Evidence from Brazil and Mexico,” *Journal of Development Economics*, 97, 368 – 386.
- Deaton, A. and Muellbauer, J. (1980), “An Almost Ideal Demand System,” *The American Economic Review*, 70, pp. 312–326.
- Fan, C. (2005), “Interprovincial Migration, Population Redistribution, and Regional Development in China: 1990 and 2000 Census Comparisons,” *The Professional Geographer*, 57, pp. 295 – 311.
- Feenstra, R. and Diewert, E. (2001), “Imputation and Price Indexes: Theory and Evidence from the International Price Program,” *BLS Working Papers, Working Paper 335*.
- Gale, F. (2002), “China’s Statistics: Are They Reliable?” in *China’s Food and Agriculture: Issues for the 21st Century*, eds. F. Gale, F. Tuan, B. Lohmar, H.-H. Hsu, and B. Gilmour, Economic Research Service, U.S. Department of Agriculture.
- Gong, C. H. and Meng, X. (2008), “Regional Price Differences in Urban China 1986-2001: Estimation and Implication,” *IZA Discussion Paper No. 3621*.
- Hamilton, B. W. (2001), “Using Engel’s Law to Estimate CPI Bias,” *The American Economic Review*, 91, 619 – 630.
- Hausman, J. (2003), “Sources of Bias and Solutions to Bias in the Consumer Price Index,” *The Journal of Economic Perspectives*, 17, pp. 23 – 44.
- Hausman, J. and Leibtag, E. (2009), “CPI Bias from Supercenters: Does the BLS Know that Wal-Mart Exists?” in *Price Index Concepts and Measurement*, eds. W. E. Diewert, J. S. Greenlees, and C. R. Hulten, pp. 203 – 231, University of Chicago Press.
- Houthakker, H. (1987), “CPI Bias from Supercenters: Does the BLS Know that Wal-Mart Exists?” in *The New Palgrave: A Dictionary of Economics*, eds. J. Eatwell, M. Milgate, and P. Newman, vol. 2, pp. 39 – 98, MIT Press.
- Jian, T., Sachs, J. D., and Warner, A. M. (1996), “Trends in regional inequality in China,” *China Economic Review*, 7, pp. 1 – 21.
- Li, S. (2002), “Chinese Household Income Project, 2002,” Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

- Manser, M. E. and McDonald, R. J. (1988), "An Analysis of Substitution Bias in Measuring Inflation, 1959-85," *Econometrica*, 56, pp. 909 – 930.
- Meng, X. (2012), "Labor Market Outcomes and Reforms in China," *The Journal of Economic Perspectives*, 26, pp. 75–101.
- Rawski, T. G. (2001), "What is happening to China's GDP statistics?" *China Economic Review*, 12, 347 – 354.
- Reinsdorf, M. (1993), "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index," in *Price Measurements and Their Uses*, eds. M. Foss, M. Manser, and A. Young, pp. 227 – 258, University of Chicago Press.
- Riskin, C., Renwei, Z., and Li, S. (1995), "Chinese Household Income Project, 1995," Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
- Young, A. (2003), "Gold into Base Metals: Productivity Growth in the Peoples Republic of China during the Reform Period," *Journal of Political Economy*, 111, pp. 1220–1261.

Descriptive Statistics				
	(i)	(ii)	(iii)	(iv)
<u>Food Expenditure</u>				
5th percentile	0.322	0.321	0.314	0.311
50th	0.412	0.411	0.401	0.396
95th	0.512	0.508	0.496	0.485
<u>Log of Household Income</u>				
5th percentile	8.822	8.831	8.853	8.884
50th	9.707	9.710	9.713	9.710
95th	10.679	10.682	10.686	10.667
<u>Log of Expenditure</u>				
5th percentile	8.622	8.640	8.671	8.708
50th	9.505	9.510	9.525	9.531
95th	10.455	10.458	10.467	10.463
<u>Head of Household Age</u>				
5th percentile	31	31	30	30
50th	46	46	44	41
95th	68	68	58	56
<u>Household Size</u>				
5th percentile	2	2	2	3
50th	3	3	3	3
95th	5	5	4	3
<u>Head of Household Years of Education</u>				
5th percentile	5	5	6	7
50th	11	11	11	11
95th	16	16	16	16
<u>Refrigerator</u>				
Mean	0.798	0.801	0.810	0.835
<u>Cars</u>				
Mean	0.007	0.007	0.008	0.008
<u>Motorcycles</u>				
Mean	0.108	0.108	0.116	0.121
Total Observations	13,335	13,199	10,704	6,993
Observations for 1995	6,771	6,197	5,348	3,366
Observations for 2002	6,564	6,315	5,356	3,627

Table 1: Descriptive statistics for variables used for estimating Engel curves for food. Food does not include alcohol or cigarettes. The variables refrigerator, cars, and motorcycles are dummy variables. Column (i) presents the results when the regression is estimated with all observations. Column (ii) includes households with food share between 2 and 80% of expenditure. Column (iii) further reduces the observations to households with head of households age 25 to 60. Column (iv) further restricts to 3-person households. Author's calculations from the 1995 and 2002 Chinese Household Income Project dataset.

Estimating Engel Curves from Equation (4)								
	<i>(i)</i>		<i>(ii)</i>		<i>(iii)</i>		<i>(iv)</i>	
	<u>OLS</u>	<u>IV</u>	<u>OLS</u>	<u>IV</u>	<u>OLS</u>	<u>IV</u>	<u>OLS</u>	<u>IV</u>
Log of Expenditure	-0.162 (.003)	-0.144 (.003)	-0.161 (.002)	-0.140 (.003)	-0.162 (.002)	-0.139 (.004)	-0.164 (.003)	-0.144 (.004)
Relative Price Change	-0.005 (.030)	0.007 (.030)	-0.001 (.028)	0.013 (.029)	-0.002 (.030)	0.013 (.030)	0.008 (.036)	0.022 (.036)
2002 Dummy	-0.046 (.007)	-0.042 (.007)	-0.043 (.006)	-0.038 (.007)	-0.036 (.007)	-0.030 (.007)	-0.032 (.008)	-0.027 (.008)
Head of Household Age	0.001 (.001)	0.001 (.001)	0.002 (.001)	0.001 (.001)	-0.003 (.001)	-0.003 (.001)	0.002 (.002)	0.002 (.002)
Head of Household Age Squared	0.000 (.000)	0.000 (.000)	0.000 (.000)	0.000 (.000)	0.000 (.000)	0.000 (.000)	0.000 (.000)	0.000 (.000)
Head of Household Years of Education	-0.002 (.000)	-0.003 (.001)	-0.002 (.000)	-0.003 (.000)	-0.002 (.000)	-0.003 (.000)	-0.001 (.000)	-0.002 (.000)
Household Size	0.016 (.001)	0.014 (.001)	0.017 (.001)	0.014 (.001)	0.017 (.001)	0.016 (.001)		
Refrigerator	0.006 (.003)	-0.000 (.003)	0.008 (.003)	0.001 (.003)	0.006 (.003)	-0.001 (.003)	0.008 (.004)	.002 (.004)
Car	-0.002 (.011)	-.009 (.011)	0.003 (.010)	-0.005 (.010)	0.003 (.010)	-0.005 (.011)	0.002 (.012)	-0.004 (.012)
Motorcycle	-0.007 (.003)	-0.009 (.003)	-0.008 (.003)	-0.010 (.003)	-0.011 (.003)	-0.013 (.003)	-0.008 (.004)	-0.010 (.004)
Cumulative Bias	-24.57% (3.27)	-25.21% (3.57)	-23.26% (3.14)	-23.90% (3.43)	-19.83% (3.46)	-19.60% (4.20)	-17.58% (4.28)	-17.32% (4.84)
Annual Bias	-3.95% (.56)	-4.06% (.65)	-3.71% (.59)	-3.83% (.65)	-3.11% (.61)	-3.07% (.67)	-2.72% (.69)	-2.70% (.81)

Table 2: Dependent variable is food share of expenditure. Food does not include alcohol or cigarettes. Regression also contains dummy variables for province. Robust standard errors for coefficients and bootstrapped standard errors for bias in parenthesis. Column *(i)* presents the results when the regression is estimated with all observations. Column *(ii)* includes households with food share between 2 and 80% of expenditure. Column *(iii)* further reduces the observations to households with head of households age 25 to 60. Column *(iv)* further restricts to 3-person households.

Estimating Engel Curves				
	(i)	(ii)	(iii)	(iv)
Log of Expenditure	-0.144 (.004)	-0.140 (.003)	-0.140 (.004)	-0.145 (.004)
Log of Expenditure on Cigarettes and Alcohol	0.002 (.000)	0.002 (.000)	0.002 (.000)	0.002 (.001)
Relative Price Change	-0.011 (.030)	-0.009 (.030)	-0.008 (.031)	0.002 (.037)
2002 Dummy	-0.046 (.007)	-0.042 (.007)	-0.035 (.007)	-0.032 (.009)
Head of Household Age	0.001 (.001)	0.001 (.001)	-0.004 (.001)	0.001 (.002)
Head of Household Age Squared	0.000 (.000)	0.000 (.000)	0.000 (.000)	0.000 (.000)
Head of Household Years of Education	-0.003 (.000)	-0.003 (.000)	-0.003 (.000)	-0.002 (.000)
Household Size	0.014 (.001)	0.013 (.001)	0.017 (.001)	
Refrigerator	0.001 (.003)	0.001 (.003)	-0.000 (.003)	0.002 (.004)
Car	-0.005 (.012)	-0.000 (.010)	-0.000 (.011)	-0.005 (.013)
Motorcycle	-0.009 (.003)	-0.010 (.003)	-0.013 (.003)	-0.010 (.004)
Cumulative Bias	-27.30% (3.43)	-26.05% (3.65)	-22.05 % (3.84)	-19.87% (4.58)
Annual Bias	-4.45% (.66)	-4.43% (.64)	-3.50% (.70)	-3.11% (.82)

Table 3: Dependent variable is food share of expenditure. Food does not include alcohol or cigarettes. Regression also contains dummy variables for province. Robust standard errors for coefficients and bootstrapped standard errors for bias in parenthesis. Column (i) presents the results when the regression is estimated with all observations. Column (ii) includes households with food share between 2 and 80% of expenditure. Column (iii) further reduces the observations to households with head of households age 25 to 60. Column (iv) further restricts to 3-person households.

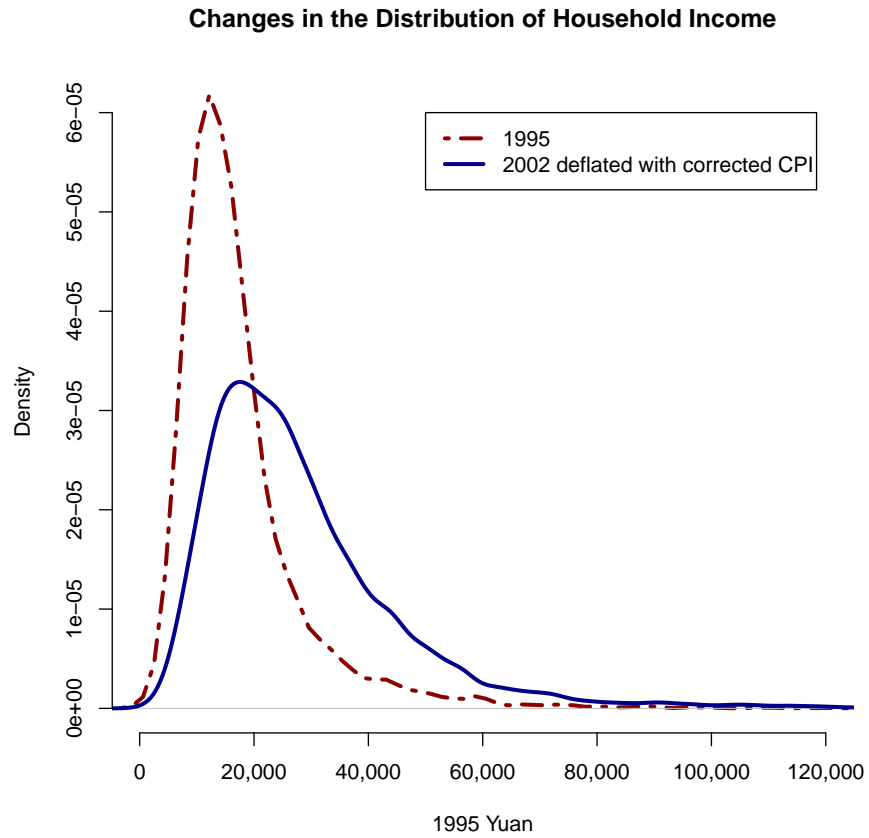


Figure 1: Density plots for urban household income. The red dotted line is for 1995; the solid blue line is for 2002 deflated with bias-corrected CPI.

Opinion of Income Distribution in City by Income Percentile				
	Very Fair	Fair	Not Very Fair	Very Unfair
Bottom 25%	0.17%	9.20%	48.21%	36.46%
Next 25%	0%	12.23%	48.74%	32.83%
Above Middle 25%	0.50%	10.83%	52.67%	29.33%
Top 25%	0.97%	14.72%	54.69%	23.30%

Table 4: Percentage of head of households who responded “Very fair,” “Fair,” “Not very fair,” or “Very unfair” to the question “Do you think the current situation on income distribution is fair? (in your city).” Only three-person households with head of household between 25 and 60 and food share of expenditure between 2 and 80 percent were included. Author’s calculations from the 2002 Chinese Household Income Project dataset.

Opinion of Income Distribution in Country by Income Percentile				
	Very Fair	Fair	Not Very Fair	Very Unfair
Bottom 25%	0.51%	8.01%	44.29%	40.89%
Next 25%	0.17%	11.73%	46.73%	35.51%
Above Middle 25%	0.83%	10.33%	48.83%	33.83%
Top 25%	1.78%	11.97%	53.24%	25.40%

Table 5: Percentage of head of households who responded “Very fair,” “Fair,” “Not very fair,” or “Very unfair” to the question “Do you think the current situation on income distribution is fair? (Countrywide).” Only three-person households with head of household between 25 and 60 and food share of expenditure between 2 and 80 percent were included. Author’s calculations from the 2002 Chinese Household Income Project dataset.

Overall Happiness of Head of Households by Income Percentile					
	Very Happy	Happy	So-So	Not Very Happy	Not Happy At All
Bottom 25%	4.60%	34.24%	35.09%	19.76%	5.11%
Next 25%	6.70%	43.89%	34.84%	11.56%	1.84%
Above Middle 25%	7.00%	49.83%	34.17%	6.83%	1.83%
Top 25%	7.12%	57.28%	30.10%	3.56%	1.12%

Table 6: Percentage of head of households who responded “Very happy,” “Happy,” “So-so,” “Not very happy,” or “Not at all happy” to the question “Generally speaking, do you feel happy?” Only three-person households with head of household between 25 and 60 and food share of expenditure between 2 and 80 percent were included. Author’s calculations from the 2002 Chinese Household Income Project dataset.

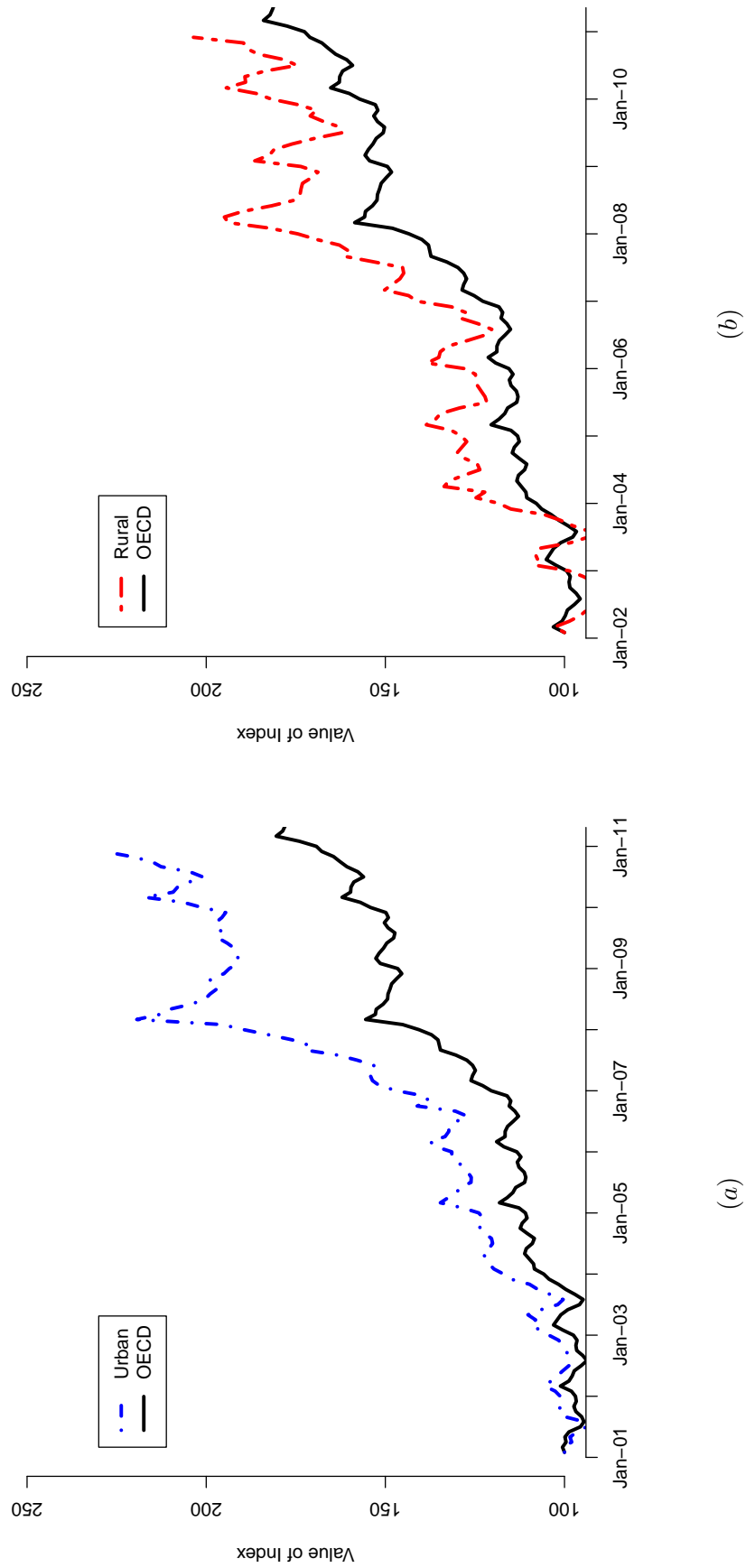


Figure 2: Comparing the OECD index of Chinese food prices (based on NBS data) with our indexes of average prices. In (a), we plot our urban index and the OECD index. In (b), we plot our rural index and the OECD index.

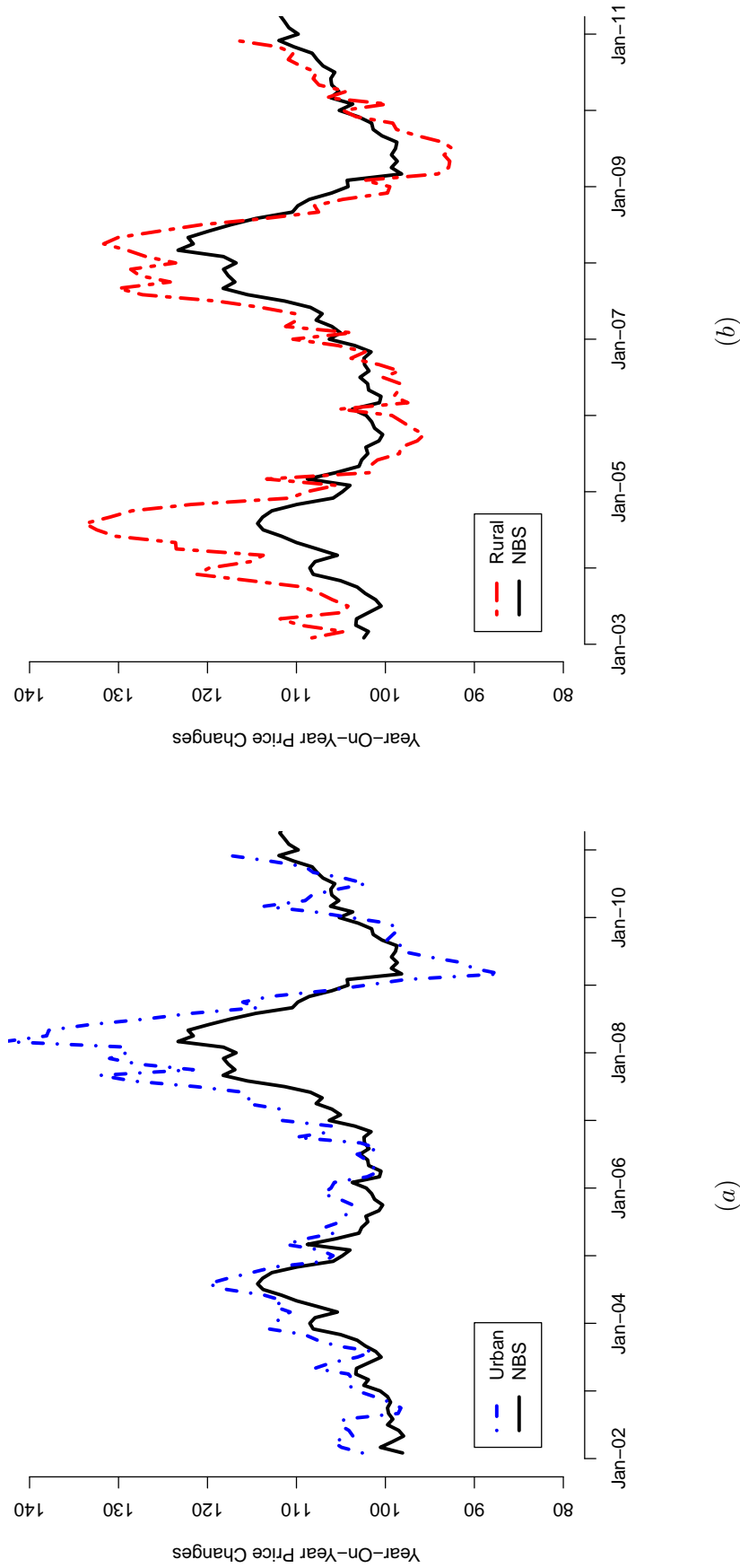


Figure 3: Comparing year-on-year inflation from the NBS index with our indexes of average prices. In (a), we plot year-on-year differences for our urban index and for the NBS index. In (b), we plot year-on-year differences for our rural index and for the NBS index.

<u>Pork</u>	<u>Rice</u>	<u>Vegetables</u>
- Pork	- Indica	-Corn
	- Quality Indica	- Cabbage
<u>Beef</u>	- Japonica	- Cucumber
- Beef	- Milled Inca	- Tomato
	- Milled Quality Indica	- Bell Pepper
<u>Mutton</u>	- Milled Japonica	- Kidney Bean
- Mutton		- Rapeseed
	<u>Wheat</u>	
<u>Eggs</u>	- Wheat	<u>Fruit</u>
- Eggs	- Wheat Flour	- Fuji Apple
		- Banana
<u>Aquatic</u>	<u>Soybean</u>	- Orange
- Grass Carp	- Soybean	
- Carp		<u>Edible Oil</u>
- Silver Carp		- Peanut Oil
-Hairtail Fish		- Rapeseed Oil
		- Soybean Oil

Table 7: Food items and categories used in the construction of the rural index. The bold items are categories and each category is weighted by rural per capita consumption. Price data is used for the non-bold items.

<u>Grain</u>	<u>Meat: Pork</u>	<u>Fresh Vegetables</u>
-Corn Meal	-Pork Lean	- Cucumber
- Flour: Premium	- Fresh Pork	- Radish
- Flour: Standard		- Eggplant
	<u>Meat: Beef</u>	- Tomato
<u>Oils and Fats</u>	-Beef Bone In	- Potato
- Rpaee Oil		- Carrot
- Soybean Oil	<u>Meat: Mutton</u>	- Green Pepper
- Peanut Oil	- Mutton, No Bone	- Chili Pepper
- Blended Salad Oil	- Mutton Bone In	- Head Cabbage
		- Okra (doujiao)
<u>Poultry</u>	<u>Condiments</u>	- Garlic
- Live Chicken	- Vinegar	- Leeks
- Chicken (baitiaoji)	- Soy Sauce	- Celery
		- Dabaicai
<u>Eggs</u>	<u>Sugar</u>	- Rape
- Eggs	- Sugar	
	- White Sugar	<u>Milk</u>
<u>Fruit</u>	- Brown Sugar	- Milk
- Orange		
- Apple: Fuji		<u>Starches</u>
- Banana		- Rice: Early Indica
- Watermelon		- Rice: Late Indica
		- Rice: Japonica

Table 8: Food items and categories used in the construction of the urban index. The bold items are categories and each category is weighted by urban per capita expenditure. Price data is used for the non-bold items.