CPI Bias and Real Living Standards in Russia During the Transition

John Gibson
Department of Economics
University of Waikato
jkgibson@waikato.ac.nz

Steven Stillman
Labour Market Policy Group
New Zealand Dept of Labour
steven.stillman@lmpg.dol.govt.nz

Trinh Le
Department of Economics
University of Canterbury
Trinh.Le@canterbury.ac.nz

Abstract

The Consumer Price Index (CPI) is calculated and closely monitored in almost every country. In addition to the importance of measured inflation to macroeconomic policy, the CPI is also used to deflate monetary measures of household living standards (including the updating of poverty lines). Yet there are well-known concerns that the CPI overstates the true increase in the cost of living, causing real economic growth and the growth in real living standards to be understated (Boskin and Jorgenson, 1997, Hausman, 2003). Although most attempts to measure CPI bias have focused on developed countries, and especially the U.S., the problem is likely to be even more serious in developing and transition economies. In these countries, large price shocks are more likely, causing consumer substitution away from the items in a fixed basket. Market liberalisation in these countries is also likely to shift consumer shopping patterns away from the outlets where prices are surveyed and may improve the quality of goods following expanded access to imports. The inability of a CPI to capture these effects contributes to its bias as a measure of the cost of living.

In this paper we measure CPI bias for Russia, using ten rounds of data from the Russian Longitudinal Monitoring Survey, covering the period 1992-2001. We follow the recently introduced method for measuring CPI bias that uses Engel's Law (Costa, 2001; Hamilton, 2001). Given that food's budget share is inversely related to household real income, by controlling for movements in relative prices and household characteristics, it is possible to infer changes in real incomes from movements in the share of food. Russian households have experienced large reductions in income and expenditure during the transition period, with a particularly severe decline occurring in the fall of 1998. Yet in contrast to the collapse in expenditure in 1998, nutritional status appears to be very resilient to variation in household resources (Stillman and Thomas, 2002). A number of factors may contribute to this contrast between income decline and nutritional stability, and in this paper we assess the role that bias in the deflators used to create the real income and expenditure series may play.

Acknowledgements: We are grateful to assistance from Olga Antimonova, Geua Boe-Gibson, Daniel Berkowitz, and Konstantin Gluschenko for assistance with the data.

JEL: C43, E31, O12
Keywords: CPI bias, Living standards, Measurement error, Prices, Russia

I. Introduction

The Consumer Price Index (CPI) is calculated and closely monitored in almost every country. In addition to the importance of measured inflation to macroeconomic policy, the CPI is also used to deflate monetary measures of household living standards. Yet there are well-known concerns that the CPI overstates the true increase in the cost of living, causing real economic growth and the growth in real living standards to be understated (Boskin and Jorgenson, 1997, Hausman, 2003). Although most attempts to measure CPI bias have focused on developed countries, and especially the U.S., the problem is likely to be even more serious in developing and transition economies. In these countries, large price shocks are more likely, causing consumer substitution away from the items in a fixed basket. Market liberalisation in these countries is also likely to shift consumer shopping patterns away from the outlets where prices are surveyed and may improve the quality of goods following expanded access to imports. The inability of a CPI to capture these effects contributes to its bias as a measure of the true cost of living.

In this paper we measure CPI bias for Russia, which is a country that has faced several large price shocks during the last decade. The most severe of these shocks occurred in the last two quarters of 1998. During this crisis, the monthly inflation rate approached 40% and real GDP declined by between 10-15% in two consecutive quarters. The magnitude of this decline in GDP is similar to that experienced in the US during the first year of the Great Depression and in Indonesia during their 1998 economic crisis. Even with its strong post-crisis performance, the Russian economy at the end of century appeared to be only three quarters of the size it had been seven years before, at least according to official statistics based on deflators like the CPI.

In addition to these price shocks, there have been fundamental changes in the economic structure in Russia and it is doubtful that the Russian CPI has been able to accurately track changes in the true cost of living during this tumultuous period. Amongst the many adjustments are the change from planned to market prices in 1992, changes in the nature of goods available to Russian households,[1] changes in retail structure and consumer purchasing patterns,[2] and changes in the degree of market integration.[3] Moreover, the degree and speed of price liberalisation also varied across Russian cities, and during the early stages of the transition the national market became more fragmented rather than more integrated (Gluschenko, 2003). It would be surprising if any country's CPI could keep pace with such remarkable changes, especially when change was occurring at different rates within the country.

While there are the usual macroeconomic reasons for interest in CPI bias, development microeconomists also may find Russia an interesting case, even if they have an inherent scepticism about macro statistics. First, emerging evidence from Russia and other countries suggests that individuals and households are very resilient in the face of major economic upheavals. For example, Stillman and Thomas (2002) examine the effect of the 1998 economic crisis on the physical well-being of the Russian population, using six measures of nutrition –

[1] The ratio of imports to GDP rose from an average of 19% in 1989-90 to 26% in 1999-2000. Moreover, the source of the imports also changed, with those sourced from the former Soviet republics (CIS countries) falling from 32% of the total in 1996 to 22% in 2002 (IMF, 2003). These changes are likely to have altered the quality of consumer goods available to the Russian population.

[2] For example, large, foreign-owned, hyper-stores such as IKEA can now be found in Russia.

[3] For example, it was three years after prices were liberalised during the 1992 economic reforms before food prices in state-run stores resembled closely the prices in private retail outlets (Berkowitz et al., 1998).

gross energy intake, two dimensions of diet quality, adult BMI, and for children, weight for height and height for age. In contrast to the collapse in expenditure in 1998, nutritional status appears to be very resilient to variation in household resources.[4] Similarly, in the Indonesian economic crisis, some measures of health and nutrition such as child height, actually improved despite the large fall in real GDP (Strauss et al., 2002). These findings are in contrast to earlier claims about the impact of stabilization on the health and nutrition of vulnerable groups (Helleiner, et al, 1991). To the extent that the newer evidence on the resilience of households is correct, the welfare costs associated with 'shock therapy' policies might be less severe than previously thought. But before such a conclusion can be drawn, we want to assess the role that bias in the deflators used to measure real income may play in this contrast between income decline and nutritional stability. In other words, perhaps the reason that nutrition didn't deteriorate so much in Russia is because the economic crisis really wasn't so bad as the deflated income variables imply. Indeed, there is a small literature that argues that the whole cost of the economic transition from communism has been overstated (Aslund, 2001), in part because the official statistics are based on potentially biased deflators.

The second microeconomic reason for interest in CPI bias concerns the measurement of poverty in Russia. The estimated headcount poverty rate for Russia in 1998 varies from 7% to 49%, due to different choices of poverty line, welfare indicator and survey used for the analysis (World Bank, 2002). To improve the consensus on poverty in Russia, two Russian ministries (the statistics agency, Goskomstat, and the Ministry of Labor and Social Development) and two donors (DFID and the World Bank) have begun a major project to improve the measurement of poverty in Russia. One output from this project has been a new poverty line, set for the year 2000 with a proposal for it to be updated using consumer price indices (Kakwani and Sajaia, 2003, p. 15). But to the extent that there is bias in the CPI, this bias will simply transfer into the measurement of the trend rate of poverty reduction. This issue is also relevant to other countries where either the CPI or a variant of it, such as a price index faced by low-income workers, is used to update poverty lines. In these countries, measurement biases in the CPI will affect debates about the rate of poverty reduction, as appears to have happened in India (Deaton and Tarozzi, 2000; Deaton, 2003).[5]

The final motivation for the paper is that Russia provides a new testing ground for a recently developed method of measuring CPI bias. The method introduced by Hamilton (2001) and Costa (2001) is much simpler than previous approaches, such as those of the Boskin Commission, which try to measure each particular type of bias for every component of the CPI. Interestingly, the simpler, 'Engel method' of Hamilton and Costa gives an estimate of CPI bias for the U.S. that is very similar to that obtained from the very detailed and much different approach of the Boskin Commission. Comparisons are needed for other countries to see whether this equivalence also holds and Russia is one of the few developing or transition economies where there are existing estimates of CPI bias that can serve as a benchmark (Bessonov, 1998).

[4] This resilience of nutritional indicators is not because of some ability of Russian households to smooth their expenditures during times of crisis. Stillman (2001) uses changes in oil prices as instruments for exogenous shocks to income and concludes that there is an almost one-to-one correspondence between exogenous shocks to income and changes in the non-durable expenditure of Russian households.

[5] The issue is especially salient for Russia because official poverty measurements are based on a household budget survey of almost 50,000 households per quarter. This large sample means that even small temporal differences in poverty rates are likely to appear statistically significant. Hence, any non-sampling errors, such as those due to bias in the deflator used for updating the poverty line, could have an unusually sharp impact on policy debates.

II. The Russian CPI and Russian Inflation

Each month, the Russian statistical authority (Goskomstat) collects prices on 400 representative goods and services from 350 towns and cities. This exercise covers every capital city of the 89 regions, with the other towns chosen by taking a representative sample of remaining urban areas. Currently, the prices are collected on the 23rd-25th day of each month, although weekly collections were done during the rapid inflation from 1992-96 and immediately after the financial crisis in August 1998. The slowing down of inflation since 1998 also led to a demand for a more precise tracking of overall price changes, which was met by increasing the number of goods and services whose prices are collected, from 280 previously to 400. Approximately one-half of the items are industrial goods, and the remainder are split between foods and services. Rented housing is excluded from the index. If an item cannot be priced at a surveyed store, town, or region, imputation is made by substituting an appropriate item from a neighbouring store, town or region. The prices are collected from a variety of different enterprises, including state-run, municipal, and private ones, as well as from urban markets. In total, the price collection covers 30,000 retail outlets.[6]

The monthly price changes are then aggregated for each of the 89 regions, where the weights are based on the structure of household expenditures for the region in the previous year. These expenditure estimates come from the Household Budget Survey, which surveys 49,000 households every quarter. The national monthly CPI is then calculated as a weighted average of the regional indices, where each region's weight is proportional to its population (Gluschenko, 2001). In addition to the overall CPI, indexes are also calculated for the three major groups: foods, industrial goods, and services. Despite the commodity and regional detail available, most attention is paid to the moments in the national average CPI.

The movements in the Russian CPI over the last decade are shown in Figure 1a. For the first two years after prices were liberalised (1992-93) the monthly inflation rate averaged 21%. Over the next two years the monthly average fell to only 9% and in the 1996-July 1998 period inflation seemed to be under control, with an average monthly increase of only 1.2%. However, the August 1998 financial crisis triggered a new bout of price rises, with the monthly inflation rate spiking at 38% in September 1998. In addition to increases in the overall price level, inflation in Russia has also been accompanied by a large shift in relative prices (Fig 1b). Food has become cheaper relative to non-food, especially in the early months of liberalisation in 1992, from early 1995 until just after the August 1998 crisis, and again since mid 1999. The fall in the relative price of food has been driven especially by increases in the prices of services.

Many socio-economic characteristics show marked regional differences in Russia but inflation appears to be an exception. According to the map in Figure 2, regions with the highest inflation rates are found equally in the east and the west. The regions with lower inflation rates are also spread about the country. More formally, an analysis of covariance indicated no difference between the 11 territories (aggregations of the 89 regions) in their average monthly inflation rates over the 1992-2002 period.

[6] Details in this paragraph come from Goryacheva (1999), Gluschenko (2001) and from the metadata provided by the OECD: http://www.mimas.ac.uk/macro_econ/oecd/Doc/SDRUSUk.htm.

III. CPI Bias and Methods of Bias Measurement

The Russian CPI is a Laspeyres index, which finds the cost of purchasing a fixed basket in a base period and the cost of buying the same basket in the present. Compared to the much-debated CPI in the United States, the Russian index has several positive features, such as the large sample used to obtain the expenditure weights and the frequent updating of the weights. Nevertheless, this type of index is known to produce a number of biases, compared to the conceptual standard of a true cost of living index (Hausman, 2003). In particular, because consumers may substitute away from higher priced goods (and outlets), while a Laspeyres index continues measuring the price of the higher priced items (from the original outlets), the CPI will be an upwardly biased estimate of changes in the true cost of living.[7] While this commodity substitution bias is typically thought of as contributing no more than one-fifth of the total CPI bias in developed countries, it may contribute more in transition economies where price shocks are larger.[8]

Estimates of commodity substitution bias exist for Russia, and these suggest that over the 1992-96 period, the official CPI overstated the rise in prices in Russia by 35% (Bessonov, 1998). However, evidence on the contribution that other sources, such as *outlet bias*, *quality change* and *new products* make to the total bias in the Russian CPI is unavailable.[9] It is because of the difficulty of isolating and measuring each individual source of bias that we adopt a different approach. Our approach gives reduced form estimates of the overall bias in the CPI, inferred from movements in food Engel curves over time. We provide the intuition for the approach here and a more formal treatment is given in Section IV.

The approach follows Costa (2001) and Hamilton (2001). Both authors estimate bias in the U.S. CPI by invoking Engel's Law, which states that food's budget share is inversely related to household real income. According to Houthakker (1987):

> "of all the empirical regularities observed in economic data, Engel's Law is probably the best established; indeed it holds not only in the cross-section data where it was first observed, but has often been confirmed in time-series analysis as well."

Provided we can control for movements in relative prices and household characteristics, it should be possible to "infer" changes in real incomes from movements in the share of food. For example, if the budget share of food were seen to be falling over a given period, yet CPI-deflated income had not risen commensurately, we would have circumstantial evidence that real income may be understated, perhaps because of bias in the price deflator. In other words, we are looking for 'drift' in the Engel curve, after all incomes have supposedly been put on a common temporal basis by deflating them by the CPI. As Costa (2000, p.2) argues, inconsistency between the

[7] Conversely, a Paasche index based on the current basket of goods, gives an underestimate of changes in the true cost of living. The geometric average of the Laspeyres and Paasche indexes (i.e., a Fisher index) is unbiased but not practical because statistical agencies cannot update the basket of goods instantaneously.

[8] The Boskin Commission estimated commodity substitution bias of 0.15 percentage points out of a total annual bias of 1.1 percentage points in the U.S. This was comparable to the *outlet bias* of 0.1 percentage points and smaller than the *formula bias* of 0.25 percentage points and the bias due to *quality change* and *new products* of 0.6 percentage points. Estimates are mostly similar for other developed countries, except for the formula bias, which relates to the way that individual price quotations are aggregated. In the U.S. they are aggregated using the arithmetic average of ratios (a.k.a. the Carli index) which produces a higher average price change than does either the ratio of averages (a.k.a. the Dutot index) or the geometric mean of the price ratios (a.k.a the Jevons index). Formula bias is less important in many other developed countries, which use either the Jevons or the Dutot index (Ducharme, 1997).

[9] Gluschenko (2001) points out that the Russian CPI also overstates inter-spatial price level differences.

trends in food budget shares and trends in real income can be attributed to changes in the relative price of food, demographic changes or to bias in the CPI. In the U.S., the rise in CPI-deflated income was able to explain only 1.5 points of the 4.5 percentage point fall in the food budget share from 1974-91. Relative price changes and other variables accounted for another 0.6 percentage points of the decline, leaving about 2.5 percentage points of the food-share decline to be explained by CPI bias (Hamilton, 2001). Bias was also present in other periods, and suggested that a considerable revision of economic history may be needed. For example, despite the Great Depression, real per capita personal expenditure actually rose by 0.6% per annum between 1919 and 1935 once account is taken of CPI bias (Costa, 2001).

For the case of Russia the method can best be illustrated by considering two cross-sections of household budgets, centered on November 1996 and November 1998. In the two years between these two cross-sections the Russian CPI rose almost 90%, due in large part to the August 1998 financial crisis. Consider a household with monthly total expenditures of 1900 roubles in 1996 and having otherwise average values of the characteristics described in Table 2 below. The food budget share for this household would be 44% according to the food Engel curve illustrated in Figure 3. Holding everything constant except for price level changes, this household two years later would have a real expenditure level of only 1000 roubles (in November 1996 prices). Hence, to the extent that the CPI measures the true cost-of-living for this household, it should retreat up the Engel curve to have a food budget share of 48% in 1998. In fact, households with CPI-deflated total expenditures of 1000 roubles in November 1998 had food budget shares of only 44%. Thus, when viewed from the standard of their budget shares, Russian households acted as if they are significantly better off than their CPI-deflated income would indicate.

IV) Empirical Methodology

In this section we describe the estimating framework used by Hamilton (2001) to infer CPI bias from a food Engel curve estimated on different years of cross-sectional micro data. This framework covers both the case when geographic and temporal variation in prices of food and non-food is available and when it is not. The advantage of food as an indicator good is that its low income elasticity makes its budget share sensitive to the mismeasurement of income, whereas goods with income elasticities close to one will have budget shares that are unchanged through time even if income growth is mismeasured. Food is also a non-durable, implying that expenditures in one period cannot provide a flow of consumption in another, and is likely to be separable from other goods in consumers' utility functions.[10]

The method starts with the Leser-Working form of the Engel curve, where the budget share is a linear function of the logarithm of real household income and a relative price term:[11]

$$w_{i,j,t} = f + g\left(\ln P_{F,j,t} - \ln P_{N,j,t}\right) + b\left(\ln Y_{i,j,t} - \ln P_{j,t}\right) + \mathbf{X}'q + u_{i,j,t} \quad (1)$$

where $w_{i,j,t}$ is the budget share of food for household *i* in region *j* and time period *t*, $P_{F,j,t}$, $P_{N,j,t}$, and $P_{j,t}$ represent the true but unobserved prices of food, nonfood, and all goods, *Y* is the household's total income (which is here measured by total expenditure), **X** is a vector of

[10] Hamilton (2001) shows that to decompose food and non-food expenditures into a price and a quantity index requires assuming additive separability of food and non-food in consumers' utility functions and homotheticity in the subutilities of food and non-food. If these conditions are met, CPI bias in such goods as computers will not affect food's budget share through any complementarities or substitutabilities.

[11] This functional form provides the basis of the Almost Ideal Demand System of Deaton and Muelbauer (1980). Results when a quadratic in log income is used are also described below.

individual household characteristics and u is the residual. The true cost of living is treated as a geometric weighted average of the prices of food and nonfood:

$$\ln P_{j,t} = a \ln P_{F,j,t} + (1-a)\ln P_{N,j,t} \quad (2)$$

and it is assumed that prices of a good G (either food, nonfood, or all goods) are measured with error,

$$\ln P_{G,j,t} = \ln P_{G,j,0} + \ln(1+\Pi_{G,j,t}) + \ln(1+E_{G,t}). \quad (3)$$

In equation (3), $\Pi_{G,j,t}$ represents the cumulative percentage increase in the CPI-measured price of good G from period 0 to period t and $E_{G,t}$ is the period-t percent cumulative measurement error in the cost-of-living index since the base period. By inserting equation (3) into (2), it is apparent that,

$$\ln(1+E_t) = a \ln(1+E_{F,t}) + (1-a)\ln(1+E_{N,t}) \quad (4)$$

Assuming that CPI bias does not vary geographically, inserting equations (2), (3) and (4) into equation (1) gives:

$$\begin{aligned} w_{i,j,t} = f &+ g\left[\ln(1+\Pi_{F,j,t}) - \ln(1+\Pi_{N,j,t})\right] \\ &+ b\left[\ln Y_{i,j,t} - \ln(1+\Pi_{j,t})\right] + \mathbf{X}'q \\ &+ g\left[\ln(1+E_{F,t}) - \ln(1+E_{N,t})\right] - b\ln(1+E_t) \\ &+ g\left(\ln P_{F,j,0} - \ln P_{N,j,0}\right) - b\ln P_{j,0} + u_{i,j,t}. \end{aligned} \quad (5)$$

An empirical version of equation (5) can be estimated if a database can be constructed from a time-series of cross-sectional household expenditure surveys and a temporal and cross-sectional CPI for food, non-food and all consumption:

$$\begin{aligned} w_{i,j,t} = \hat{f} &+ g\left[\ln(1+\Pi_{F,j,t}) - \ln(1+\Pi_{N,j,t})\right] \\ &+ b\left[\ln Y_{i,j,t} - \ln(1+\Pi_{j,t})\right] + \mathbf{X}'q \\ &+ \sum_{t=1}^{T} d_t D_t + \sum_{j=1}^{J} d_j D_j + u_{i,j,t} \end{aligned} \quad (6)$$

where D_t is a dummy variable equal to 1 in period t, D_j is a dummy equal to 1 for region j, d_t and d_j are their coefficients, and $\hat{f}$ is the intercept from equation (5), plus the coefficients of the omitted time and region dummies. The time dummy variables are crucial to the measurement of CPI bias because

$$d_t = g\left[\ln(1+E_{F,t}) - \ln(1+E_{N,t})\right] - b\ln(1+E_t) \quad (7)$$

and if equation (7) is written in terms of the cumulative bias in the CPI for all goods, $\ln(1+E_t)$, and if it is assumed that the relative bias between food and nonfood is constant across years, then:

$$\ln(1+E_t) = \frac{d_t}{-b - \dfrac{g(1-r)}{1-a(1-r)}}. \quad (8)$$

In other words, the bias can be identified up to an unknown parameter, r, which is the ratio of CPI bias in food to nonfood, and also depends on a, which is food's share in the cost-of-living index. Hamilton (2001) notes that equation (8) can be reduced to:

$$\ln(1+E_t) \approx \frac{-d_t}{b} \quad (9)$$

if either γ or (1-r) is close to zero. In other words, equation (9) is likely to hold if either relative price movements are unimportant to food demand or if CPI-bias in food and nonfood is equal. If instead, the price index for food is less badly biased (r<1), which seems plausible due to the measurement difficulties with items like computers, then equation (9) *understates* the bias. Thus, a lower bound for cumulative percentage CPI bias at period t is given by a simple ratio of estimated coefficients from equation (6), $1-\exp(-d_t/b)$.

When cross-sectional variation in relative food prices is unavailable, equation (6) cannot be estimated because there is no way to identify the parameter on food prices, γ.[12] Simply using temporal movements in an aggregate price index for food relative to nonfood will not work because this period-by-period variation will be perfectly correlated with the time dummy variables, D_t so the model could not be estimated. The specification that must be used when cross-sectional variation in food prices is unavailable is:

$$w_{i,t} = \hat{f} + b\left[\ln Y_{i,t} - \ln\left(1+\Pi_t\right)\right] + \mathbf{X}'q + \sum_{t=1}^{T} d_t D_t + u_{i,t}. \qquad (10)$$

The dummy variables in equation (10) measure not just the CPI bias of equation (7) but also the effect on budget shares of intertemporal variation in the measured inflation rate for food relative to nonfood. Hence, the cumulative percentage CPI bias at time t is calculated from:

$$1-\exp\left\{\frac{d_t - \bar{g}\left[\ln\left(1+p_{F,t}\right) - \ln\left(1+p_{N,t}\right)\right]}{-b}\right\} \qquad (11)$$

where $\bar{g}$ has to be obtained from outside of the estimated parameters for equation (10).

In the Russian context, regionally disaggregated data are available for the food and non-food inflation rates, so equations (6) and (9) provide the basic framework, following the approach of Hamilton (2001) of using food and non-food inflation rates rather than price levels to identify γ. However, we also use the no-regional-price variation approach described by equations (10) and (11) as a cross-check on the results.

V) Data

To estimate equation (6) we use data from the Russian Longitudinal Monitoring Survey (RLMS), which is an on-going longitudinal household survey designed and collected by Barry Popkin and his colleagues at Carolina Population Center, University of North Carolina, in collaboration with colleagues at the Russian Academy of Sciences and the Russian Institute of Nutrition. This survey is designed to be nationally representative and has been widely used to study demographic, economic and health-related topics during Russia's transition to a market economy (Mroz and Popkin, 1995; Lokshin and Ravallion, 2000). We also use the monthly CPI for food, industrial goods and services that is calculated for each of the 89 regions of Russia, and the overall CPI that is calculated nationally for the combined total of all goods and services.

The RLMS has operated in two phases, each with their own samples and data collection instruments. The first phase operated almost continuously between July 1992 and February 1994,

[12] Hamilton (2001) uses cross-sectional variation in inflation *rates*, rather than price levels, to identify γ from data for 25 major urban areas in the U.S.

with four rounds of data collected from approximately 6,700 households.[13] These households were located in 21 survey sites in 16 different *Oblasts*.[14] The second phase spans the period 1994 through 2001, with six rounds of data collected from approximately 4000 households.[15] The sampling for the second phase was based on a division of Russia into 38 strata, with one primary sampling unit (PSU) chosen from each strata. Several secondary sampling units were chosen within each PSU, giving approximately 160 survey sites from more than 30 different *Oblasts*. These selections appear to be representative because an analysis of covariance indicated no difference in average monthly inflation rates between the regions containing RLMS phase II sites and other regions. However, to accommodate the changed sample, plus changes in the questionnaire, our analysis is carried out separately for each of the two phases of the RLMS.

Two other features of the RLMS also affect the analysis. First, neither phase collected extensive information on production for own consumption. Because it is mainly food that is self-produced, it is likely that the food budget share for farm households will be biased downwards and this may affect the Engel curve estimates. To guard against this, rural households are excluded from the analysis. However, since urban households make up 77% of the RLMS sample and the Russian population, this sample restriction should not diminish the relevance of the results. Moreover, because the prices for the CPI are collected from towns and cities, urban households seem to be the relevant sample. Even after this restriction, the sample still covers 18 sites in phase I and 33 PSUs in phase II and the sample for each phase still has approximately 18,000 household-round observations.

The second survey feature is that the sampling frame for RLMS is a set of dwellings which are intended to be representative of the Russian population in the early 1990s. For cost reasons, the survey does not attempt to follow individuals or households who move from the sample dwelling. Instead, any new household member or new household living at the sample dwelling is included in the sample in each wave. The sample will remain representative of the underlying population assuming new entrants are exchangeable with movers.[16] Analysis by Stillman and Thomas (2002) suggests that attrition related to observable characteristics is not a serious concern in these data.[17] Since the Engel curve method for measuring CPI bias does not require the use of true panel data, and can be applied to repeated cross-sections (for example, Costa, 2001), we initially ignore the panel characteristics of the data in our analyses. But as a further check on the robustness of the results, the models are re-estimated using household fixed effects, exploiting the panel structure of the data.

---

[13] The second and subsequent survey rounds began in Dec 1992, May 1993 and October 1993. However, most interviews were conducted in August-October 1992; January-March 1993; June-July 1993; and November-December 1993.

[14] Russia's 89 regions are called either a *republic* (if it is a national autonomy), a *krai* (if it has a small scale national autonomy called *okrug* within its borders), or an *oblast*.

[15] Surveys were conducted in 1994, 1995, 1996, 1998, 2000, and 2001 (waves 5 through 10, respectively). A full project description is available at www.cpc.unc.edu/rlms which provides sampling procedures, survey instruments and field protocols. Surveys in phase II are conducted in the late Fall of each year with most of the interviews in the following months: November and December, 1994; October and November, 1995; October and November, 1996; November and December, 1998; October and November, 2000; October and November, 2001.

[16] See Thomas, Frankenberg, and Smith (2001) for a discussion of the likely implications of this assumption.

[17] Heeringa (1997) provides more information on attrition in RLMS and further discusses its overall representativeness.

VI) Estimation Results

Equations (6) and 10 are estimated for a sample of two-adult families, with or without children, where the adults are between 21-75 years old. Households whose food shares are outside the interval 0.02-0.95 are excluded. These restrictions are similar to those employed by Hamilton and Costa, and are designed to provide a more homogeneous sample with higher data quality. Control variables include real total expenditures, relative food price changes, demographic, educational and employment characteristics, indicators of dwelling characteristics (as proxies for wealth), and regional and time dummies.[18] Three different variants of the total expenditure variable are used; one that includes all items enumerated by the survey, one that excludes durables, and one that only includes food, clothes, fuel and services (following Stillman and Thomas, 2002). The model also includes the share of consumption devoted to food out of the home. This form of consumption is not part of the dependent variable because it is assumed that restaurant meals are not perfect substitutes for food-at-home. Ideally, the substitution possibilities between restaurants and home cooking should be captured by including the relative price of restaurant meals but the available price index is not available. Therefore, we follow Costa and Hamilton in using the budget share for restaurant meals as an explanatory variable, in place of the required price.

A description of the dependent and explanatory variables is contained in Table 1 (for Phase I) and Table 2 (Phase II). To show how food shares, prices, income and household characteristics have changed over time, the beginning and end-period averages of the variables are reported in addition to the full-sample average. The dependent variable, which is the share of consumption devoted to food at home, averages 58% in phase I and 53% in Phase II.[19] The average food share fell by 2 percentage points between Rounds 1 and 4 in Phase I and by 8 percentage points between Rounds 5 (late 1994) and 10 (late 2001) of Phase II, despite stagnation in CPI-deflated total consumption. The decline in the real price of food is also apparent in Table 2.

Table 1 contains the results of estimating equation (10) with the Phase I RLMS data. No attempt is made to estimate the more general equation (6) that uses regional relative food prices because for Phase I we lack the geographical identifiers needed to match the regional CPIs to the primary sampling units of the survey. The negative coefficient on deflated total consumption indicates that food budget shares fall as households become richer, which is precisely why food is used as the indicator good here. Relative to the base period (July-Oct, 1992), the food share is lower in Round 2 (albeit usually insignificantly so) and in Round 4 (Nov-Dec, 1993). But in Round 3 (June-July, 1993) the food share is higher than in the base period. Seasonality is a likely culprit for this pattern, making it difficult to interpret any CPI bias. However, to the extent that Round 1 and Round 4 occur in fairly similar periods of the year, an estimate of the average level of bias can be derived from the fact that the food budget share is three percentage points lower in Round 1 than in Round 4. If this is combined with movements in the national food-nonfood inflation rates (using an estimate of $\bar{g} = 0.18$ which is derived using an approach described below), the application of equation (11) suggests a cumulative CPI bias of approximately 0.6, which implies an average monthly bias of about 3% between July 1992 and January 1994. During the same period, the average monthly change in the CPI was about 20%. However, this estimate should

[18] We previously included dummies for the gender and ethnic minority status of the household head but these variables always had small and statistically insignificant coefficients. The indicator for separate dwellings is not included in the model for Phase I because details about the dwelling are not available from Round 3 of the survey.

[19] These two averages are not comparable because of differences in the samples and the questionnaires.

not be regarded as definitive because of the short time period covered by Phase I of the RLMS and the imperfect synchronisation of the survey rounds in the same period each year. These problems are much less apparent in Phase II of RLMS.

Table 2 contains the estimates of the food Engel curves for the phase II RLMS data. The key result is that the year dummy variables are all negative and statistically significant. Moreover, with the exception of the restricted definition of expenditures in Round 10, each dummy variable is larger (i.e., more negative) than in the preceding survey round. Thus, there has been an almost continual leftwards drift in the food Engel curve, as illustrated in Figure 3 for 1996 and 1998. By comparing the three columns of regression results, it is clear that this leftward drift in the food Engel curve is not affected very much by the particular definition of household expenditures.

The effect of relative food price inflation is not very precisely estimated in Table 2, so the alternative estimation strategy based on equation (10) was also used. Table 3 contains the results of estimating the national-level model, where regional effects and the regional variation in relative food price inflation is excluded. The pattern of the period dummy variables is very similar to what was previously estimated, although these variables now measure not just CPI bias but also the effect on budget shares of intertemporal variation in the measured inflation rate for food relative to nonfood. A variety of sensitivity analyses done with the national-level model are also reported in Tables 3A and 3B. The results are not sensitive to whether the survey sampling weights are used, to whether quadratic expenditure terms are included, and to the removal from the sample of households with especially low (and possibly mis-measured) expenditures.

The panel structure of the data is exploited in Table 4, which contains the results of including household fixed effects in the regressions. The dummy variables for each time period are now reduced somewhat in magnitude, especially when the restricted definition of expenditures is used. However the coefficient on deflated expenditures, *ß* is also reduced and because the CPI bias is estimated as the ratio of these two coefficients the introduction of the fixed effects has no net effect on the bias estimates. A comparison of the results derived from Table 4 with those from Table 2 also gives an indirect indication of the lack of sensitivity of the bias estimates to sample attrition. The sample for the fixed effects estimates is restricted to the households from Round 5 that were present in subsequent rounds (6011 household-round observations). The sample in Table 2 includes the new households who moved into sample dwellings (7800 household-round observations). The similarity of the two sets of results suggests that attrition is not affecting the estimates.

VII) Analysis

The CPI bias estimates that are derived from the coefficients in Table 2, using the equation: $1-\exp(-d_t/b)$ are presented in the first three columns of Table 5. Over the seven years from November 1994 to November 2001, the cumulative bias in the Russian CPI is estimated as 0.76 $((\hat{s} = 0.04),$ with an average bias of 0.9% per month. One possible weakness of these estimates is that the substantial change in relative food prices illustrated in Figure 1b is controlled for only imprecisely because $\hat{g} = 0.04 \pm 0.04$. We therefore experimented with various other methods of estimating the relative food price coefficient. First, we replaced the *oblast* data on the food/non-food inflation *rates* with a simple measure of the *level* of relative food prices. The measure used was the ratio of the food poverty line to the non-food poverty line, both of which have been calculated for nine broad regions by Kakwani and Sajaia (2003). The poverty line refers to the year 2000, so we restricted our estimates to rounds 9 and 10 of the RLMS data, covering 2000-01. The coefficient was estimated as 0.167, albeit with a wide standard error of 0.30.

Our second approach was to derive a value for γ by working backwards from estimates of the own-price elasticity of food demand, e_{ii} noting that for equation (1), the own-price elasticity is:

$$e_{ii} = -1 + \frac{\gamma - \alpha \beta}{w} \qquad (12)$$

where α is the share of food in the overall price index. We are unaware of any recent literature with econometric estimates of e_{ii} for food in Russia.[20] However, we were able to derive an estimate using the method proposed by Frisch (1959) for additive demand systems, where the own-price elasticity depends only on the food budget share, the income elasticity of food demand, η and the 'flexibility of money' ω:

$$e_{ii} = \frac{1}{\omega}\eta_i(1 - w_i\eta_i) - w_i\eta_i . \qquad (13)$$

The results of the model in the first column of Table 3 were used to calculate η, along with a value of -4.2 for ω.[21] The resulting values of e_{ii} and γ were -0.56 and 0.19. The estimates or CPI bias obtained using equation (11) and the derived estimate of γ are presented in the middle columns of Table 5. These bias estimates are almost identical to the ones previously estimated.

The final columns of Table 5 contain the bias estimates derived from the household-level fixed effects model. The temporal pattern of bias is somewhat different in these estimates but the end result is the same; by late 2001 the cumulative bias in the Russian CPI was of the order of 0.75, implying an average monthly bias of about 0.9-1.0%.

VIII) Conclusions and Implications

In this paper we have estimated Engel functions for the food budget share of Russian households, based on data from 1992 to 2001 from the Russian Longitudinal Monitoring Survey. Taste changes are unlikely over such a short period of time, so after allowing for changes in relative prices and demographic changes, we would expect households at the same real income to have a similar share of food in their budgets. In fact, we find that the share of food declines continuously over time, holding CPI-deflated incomes constant. One possible cause of this is that the CPI has overstated the rise in the cost of living and hence caused real income to be understated.

We find that the average bias in the CPI was over 3% per month during Phase I of the RLMS (1992-93) and approximately 1% per month during Phase II (1994-2001). The degree of bias varied considerable through time and was greatest in the beginning of the period. The cumulative effect of this bias causes a substantial understatement of the growth performance of the Russian economy during the transition. If the CPI is used as a deflator, the level of real per capita GDP in 2001 may be understated by up to 35% compared with using a bias-corrected deflator (Figure 4).

---

[20] Calibrated elasticities are available from the databases for CGE models, such as the GTAP model. In this model, a weighted average of the values for grains, meat and livestock, dairy and other food gives an own-price elasticity of food demand for the former Soviet Union of -0.12 (Dimaranan, et al, 2002).[20] However, we view this as less reliable than our own estimates because it is only based on income elasticities (whose sources are not revealed) with price elasticities derived from additive demand systems.

[21] This is based on the relationship used by Lluch *et. al.* (1977) of $\omega \approx -36X^{-0.36}$, where X is GNP per capita in 1970 U.S. dollars, which we estimate to between \$300 and \$700 for Russia over the 1992-2001 period.

Even with the adjustment for CPI bias, the effect of the August 1998 financial crisis shows clearly as a 16-20% deviation from the trend GDP level. Therefore, the contrast between the stability of nutritional indicators and the size of the shock to real incomes that is described by Stillman and Thomas (2002) does not seem to be explained by bias in the deflators used to measure real income. Thus, it really may be the case that households are more resilient than previously thought, in the face of macroeconomic shocks.

References

Aslund, A. 2001. "The Myth of Output Collapse After Communism" *Working Paper* No. 18, Carnegie Endowment for International Peace, Washington DC.

Berkowitz, D., DeJong, D., and Husted, S. 1998. "Quantifying Price Liberalization in Russia" *Journal of Comparative Economics* 26(4): 735-760.

Bessonov, V. 1998. "On Biases in Evaluations of Rise in Russian Consumer Prices" *Economical Journal of the High School of Economics* 2(2): 31-38 [cited in Gluschenko, 2001].

Boskin, M. and Jorgenson, D. 1997. "Implications of Overstating Inflation for Indexing Government Programs and Understanding Economic Progress" *American Economic Review* 87(2):89-93.

Costa, D. 2001. "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves" *Journal of Political Economy* 109(6): 1288-1310.

Deaton, A. 2003. "Prices and Poverty in India, 1987-2000" *Economic and Political Weekly* January 25, 362-368.

Deaton, A. and Tarozzi, A. 2000. "Prices and Poverty in India." *Mimeo* Research Program in Development Studies, Princeton University.

Deaton, A., and Muellbauer, J. 1980 "An Almost Ideal Demand System" *American Economic Review* 70(3): 312-326.

Dimaranan, B., McDougall, R., and Hertel, T. 2002. "Behavioral parameters" in Dimaranan and McDougall (eds) *Global Trade, Assistance, and Production: The GTAP Data Base*, Center for Global Trade Analysis, Purdue University, Chapter 20.

Ducharme, L. 1997. *Bias in the CPI; Experiences from Five OECD Countries*. Ottawa: Statistics Canada.

Frisch, R. 1959. "A Complete System for Computing all Direct and Cross Demand Elasticities in a Model with Many Sectors" *Econometrica* 27(2): 177-196.

Gluschenko, K. 2001. "Biases in the Russian Regional CPIs" *mimeo* Institute of Economics and Industrial Engineering, Siberian Branch of the Russian Academy of Sciences

Gluschenko, K. 2003. "Market Integration in Russia During the Transformation Years" *Economics of Transition* 11(2): xx-xx.

Goryacheva, I. 1999. "Methodology for Measuring Inflation in the Russian Federation" Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, 3-5 November.

Hamilton, B. 2001. "Using Engel's Law to Estimate CPI Bias" *American Economic Review* 91(3): 619-630

Hausman, J. 2003. "Sources of Bias and Solutions to Bias in the Consumer Price Index" *Journal of Economic Perspectives* 17(1): 23-44.

Heeringa, S. 1997. "Russia Longitudinal Monitoring Survey Sample Attrition, Replenishment, and Weighting: Rounds V-VII." *University of Michigan Institute for Social Research* (March).

Helleiner, G., Cornia, G., and Jolly, R. (1991) "IMF Adjustment Policies and the Needs of Children" *World Development* 19(12): 1823-34.

Houthakker, H. (1987) "Engel's Law" in J.Eatwell, M.Milgate and P.Newman (eds.) The New Palgrave Dictionary of Economics. Vol. 2 (London: McMillan), pp. 143-144.

International Monetary Fund. (2003) "Russian Federation: Statistical Appendix" *IMF Country Report* No. 03/145, May 2003.

Kakwani, N., and Sajaia, Z. (2003) "Poverty Line in Russia: Issues, Methodology and Recommendations" *mimeo* University of New South Wales and the World Bank.

Lluch, C., Powell, A., and Williams, R. (1977) *Patterns in Household Demand and Saving*, Oxford: Oxford University Press for the World Bank.

Lokshin, M., and Ravallion, M. (2000) "Welfare Impacts of the 1998 Financial Crisis in Russia and the Response to the Public Safety Net" *Economics of Transition* 8(2):

Mroz, T., and Popkin, B. 1995. "Poverty and the Economic Transition in the Russian Federation" *Economic Development and Cultural Change* 44(1):1-31.

Stillman, S. 2001. "The Response of Consumption in Russian Households to Economic Shocks." *IZA Discussion Paper* 411 (December).

Stillman, S., and Thomas, D. 2002. "The Effect of Economic Crises on Nutritional Status: Evidence from Russia" *mimeo*

Strauss, J., Beegle, K., Dwiyanto, A., Herawati, Y., Pattinasarany, D., Satriawan, E., Sikoki, B., Sukandi, and Witoelar, F. 2002. "Indonesian Living Standards Three Years After the Crisis: Evidence From the Indonesia Family Life Survey." Paper presented at the Northeast Universities Development Consortium Conference, Williams College.

Thomas, D., Frankenberg, E. and Smith, J. 2001. "Lost but Not Forgotten: Attrition and Follow-Up in the Indonesia Family Life Survey." *Journal of Human Resources* 36(3): 556-92.

World Bank. 2002. "Poverty in Russia" *Country Assistance Strategy of the World Bank Group for the Russian Federation* (Report No. 24127-RU) Attachment 3, pp. 43-45.

Figure 1a: Russian Monthly CPI Inflation Rate: 1992-2003

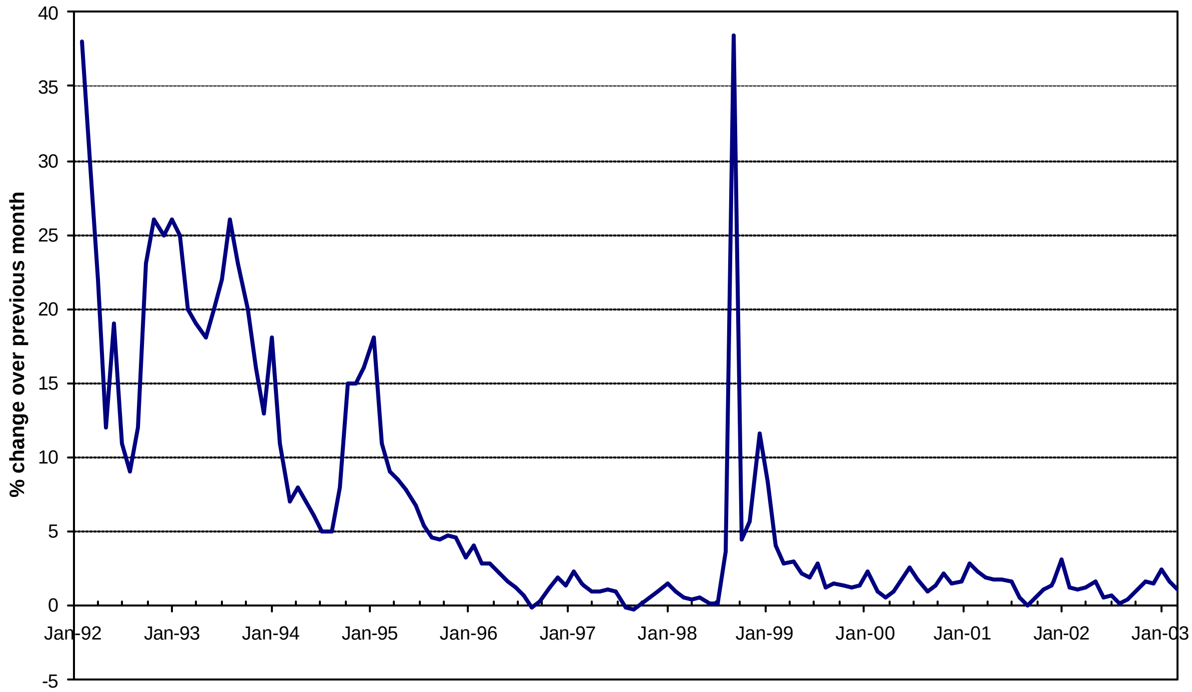

Figure 1b. Relative Food/Non-Food Price Changes in Russia: 1992-2003

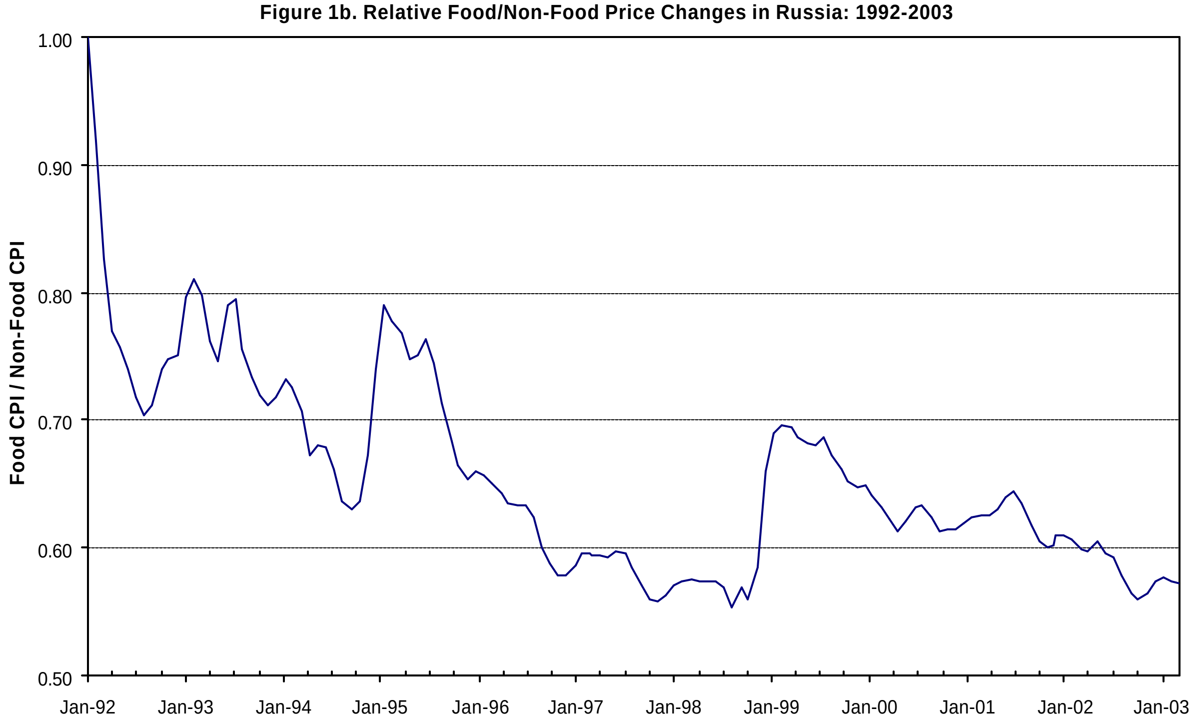

Figure 2: Average Monthly Inflation Rate (1992-2002) in Russian Regions

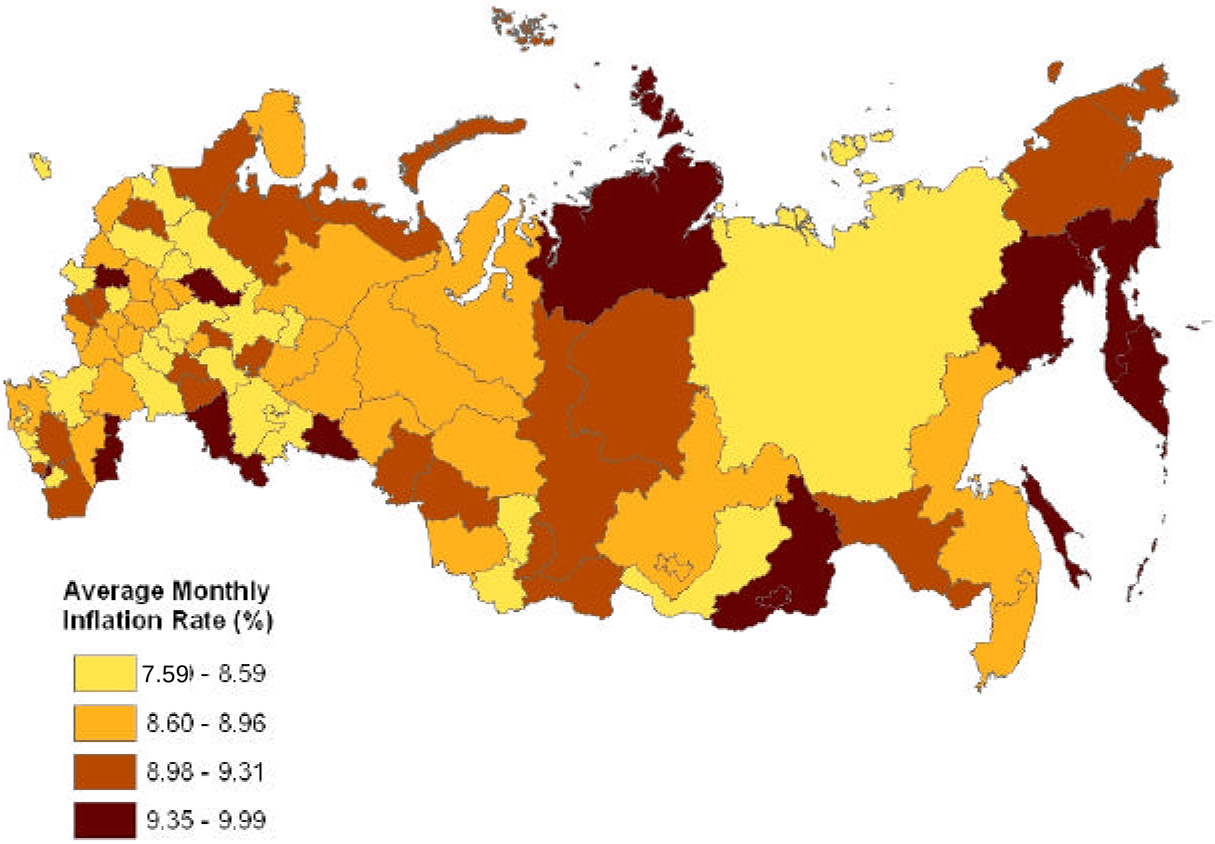

Figure 3: Food Engel Curves for 1996 and 1998

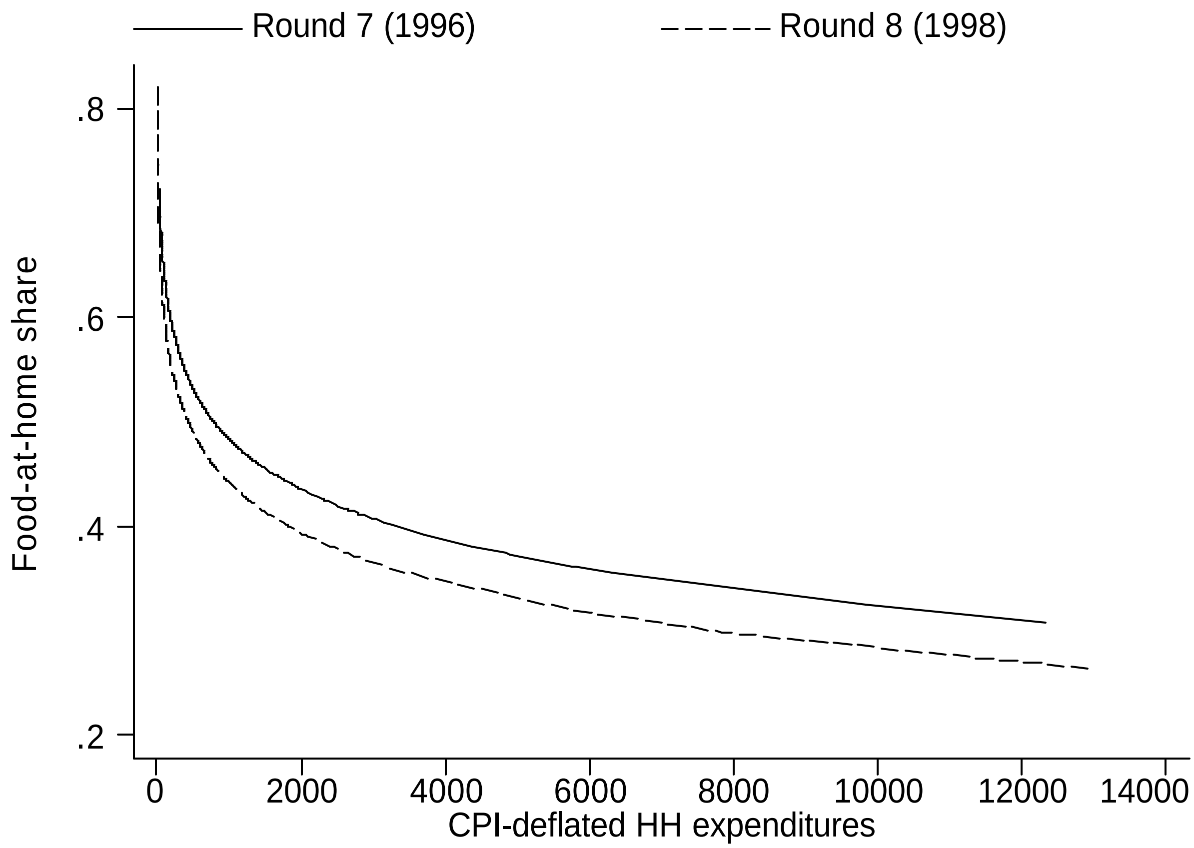

Figure 4: Effect of CPI Bias on Estimates of Real GDP in Russia

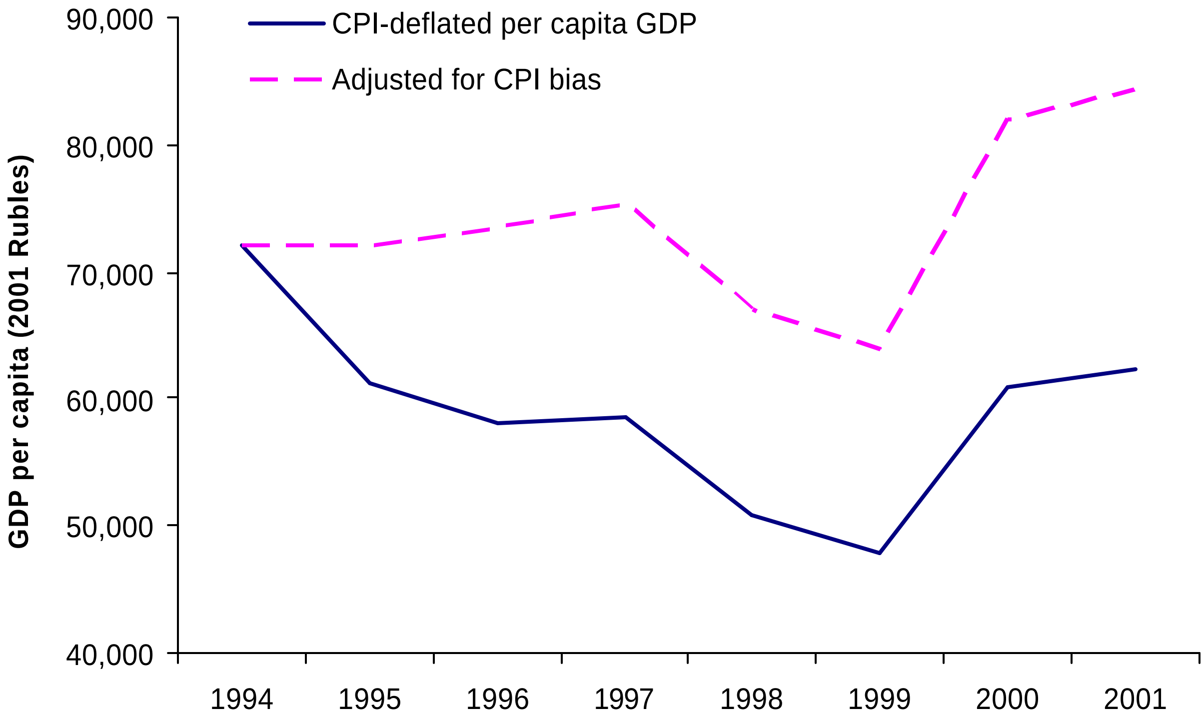

Table 1: Descriptive Statistics and Econometric Results for Food Engel Curve Estimated from Phase I RLMS

	Descriptive Statistics			Engel Curve Regression Results *Expenditure definition*		
	Mean[a] (std. dev)	Round 1 Mean	Round 4 Mean	All expenditures	Excluding durables	Restricted expenditures
Budget share for food at home	0.581	0.580	0.561			
	0.217					
ln (real total expenditure)	9.097	9.117	9.073	-0.045	-0.038	-0.035
	0.794			(5.15)**	(4.51)**	(4.58)**
ln (household size)	1.103	1.114	1.101	-0.032	-0.040	-0.082
	0.305			(1.66)	(2.13)*	(4.33)**
% of household = 2 years old	0.022	0.026	0.018	0.139	0.178	0.229
	0.078			(2.33)*	(2.98)**	(3.74)**
% of HH 3-14 year old boys	0.094	0.095	0.094	0.168	0.179	0.213
	0.152			(6.01)**	(6.73)**	(8.05)**
% of HH 3-14 year old girls	0.086	0.090	0.084	0.160	0.169	0.206
	0.147			(5.04)**	(5.71)**	(7.05)**
% of HH 15-17 year old boys	0.022	0.022	0.024	0.110	0.121	0.150
	0.076			(3.69)**	(3.89)**	(6.45)**
% of HH 15-17 year old girls	0.023	0.023	0.022	0.046	0.058	0.069
	0.077			(1.36)	(1.61)	(1.81)+
Age of household head	45.445	44.847	45.965	0.003	0.002	0.003
	11.971			(2.71)*	(2.30)*	(3.40)**
Age of spouse	42.508	41.899	43.037	0.003	0.003	0.003
	12.279			(3.10)**	(3.44)**	(5.13)**
Head has tertiary education	0.248	0.289	0.217	-0.011	-0.012	0.003
	0.432			(1.63)	(1.83)+	(0.41)
Spouse has tertiary education	0.227	0.269	0.197	-0.005	-0.005	0.005
	0.419			(1.03)	(0.91)	(0.90)
Head is working	0.713	0.760	0.675	-0.011	-0.008	-0.004
	0.452			(1.65)	(1.30)	(0.57)
Spouse is working	0.694	0.726	0.671	0.008	0.008	0.005
	0.461			(1.04)	(1.03)	(0.63)
% of budget food out of home	0.025	0.030	0.022	-0.422	-0.448	-0.549
	0.066			(7.61)**	(8.59)**	(11.88)**
Round 2 (Jan-Mar, 1993)	0.241			-0.013	-0.019	-0.023
	0.427			(1.11)	(1.68)	(2.15)*
Round 3 (June-July, 1993)	0.243			0.021	0.016	0.008
	0.429			(2.48)*	(1.93)+	(1.12)
Round 4 (Oct 1993-Jan 1994)	0.242			-0.030	-0.032	-0.031
	0.428			(3.48)**	(3.74)**	(3.85)**
Constant				0.772	0.740	0.785
				(8.93)**	(8.91)**	(9.33)**
R^2				0.13	0.12	0.15
F-test (time dummies=0)				12.02**	13.60**	11.64**

Note: Absolute value of t-statistics in parentheses corrected for cluster effects but not stratification; * significant at 5%; ** significant at 1%; + significant at 10%. *N*=8476. The excluded time dummy is for Round 1 (Jul-Oct, 1992).
The various expenditure definitions affect the dependent variable (food-at-home budget share), and ln (real total expenditure) and food-away-from-home share. The restricted definition of expenditures includes only food, clothes, fuel and services.
[a] For the expenditure and food share variables, the descriptive statistics are for "all expenditures" definition.

Table 2: Descriptive Statistics and Econometric Results for Food Engel Curve Estimated from Phase II RLMS

	Descriptive Statistics			Engel Curve Regression Results *Expenditure definition*		
	Mean[a] (std. dev)	Round 5 Mean	Round 10 Mean	All expenditures	Excluding durables	Restricted expenditures
Budget share for food at home	0.529 (0.221)	0.555	0.474			
ln (real total expenditure)	12.916 (0.878)	13.128	12.975	-0.082 (12.06)**	-0.074 (10.96)**	-0.064 (9.35)**
ln relative food price[b]	-0.382 (0.328)	-0.181	-0.465	0.042 (1.06)	0.033 (0.83)	0.028 (0.65)
ln (household size)	1.080 (0.300)	1.086	1.075	0.020 (1.43)	0.016 (1.21)	-0.032 (1.91)+
% of household = 2 years old	0.019 (0.073)	0.022	0.017	0.135 (4.60)**	0.150 (4.75)**	0.257 (7.80)**
% of HH 3-14 year old boys	0.078 (0.141)	0.084	0.068	0.107 (4.09)**	0.104 (4.07)**	0.161 (6.69)**
% of HH 3-14 year old girls	0.078 (0.140)	0.089	0.063	0.089 (3.19)**	0.085 (3.08)**	0.163 (5.79)**
% of HH 15-17 year old boys	0.021 (0.074)	0.021	0.022	0.080 (2.24)*	0.069 (1.76)+	0.097 (2.16)*
% of HH 15-17 year old girls	0.022 (0.075)	0.018	0.025	0.020 (0.52)	0.015 (0.41)	0.037 (1.23)
Dummy: detached dwelling	0.086 (0.281)	0.079	0.092	-0.028 (2.07)*	-0.027 (2.08)*	-0.042 (3.79)**
Age of household head	46.053 (12.442)	44.639	47.620	0.003 (3.23)**	0.003 (3.24)**	0.003 (4.13)**
Age of spouse	43.208 (12.895)	41.809	44.766	0.002 (2.07)*	0.002 (1.79)+	0.003 (3.72)**
Head has tertiary education	0.371 (0.483)	0.774	0.285	-0.022 (3.95)**	-0.023 (4.06)**	-0.020 (3.49)**
Spouse has tertiary education	0.362 (0.481)	0.787	0.286	-0.009 (1.72)+	-0.008 (1.64)	-0.009 (1.33)
Head is working	0.679 (0.467)	0.695	0.698	-0.010 (1.64)	-0.010 (1.81)+	-0.006 (1.41)
Spouse is working	0.653 (0.476)	0.658	0.673	0.007 (1.60)	0.005 (1.34)	0.001 (0.23)
% of budget food out of home	0.042 (0.087)	0.047	0.041	-0.459 (16.43)**	-0.485 (17.54)**	-0.645 (25.06)**
Round 6 (Oct-Nov, 1995)	0.183 (0.387)			-0.021 (1.98)+	-0.027 (2.57)*	-0.020 (2.02)+
Round 7 (Oct-Nov, 1996)	0.169 (0.375)			-0.050 (5.71)**	-0.055 (6.10)**	-0.033 (3.47)**
Round 8 (Nov-Dec, 1998)	0.154 (0.361)			-0.094 (10.45)**	-0.102 (11.28)**	-0.064 (7.76)**
Round 9 (Oct-Nov, 2000)	0.147 (0.354)			-0.111 (11.74)**	-0.121 (11.93)**	-0.096 (10.20)**
Round 10 (Oct-Nov, 2001)	0.150 (0.357)			-0.117 (12.92)**	-0.125 (14.45)**	-0.090 (9.18)**
Constant				1.404	1.332	1.282

	(18.09)**	(16.93)**	(16.66)**
R^2	0.26	0.24	0.32
F-test (region dummies=0)	27034**	74370**	4520**
F-test (time dummies=0)	44.05**	52.13**	25.16**

Note: Absolute value of t-statistics in parentheses corrected for cluster effects but not stratification; * significant at 5%; ** significant at 1%; + significant at 10%. *N*=7807. The excluded time dummy is for Round 5 (Nov-Dec, 1994). Each equation also includes 25 regional fixed effects. The various expenditure definitions affect the dependent variable (food-at-home budget share), and ln (real total expenditure) and food-away-from-home share. The restricted definition of expenditures includes only food, clothes, fuel and services.

[a] For the expenditure and food share variables, the descriptive statistics are for "all expenditures" definition.

[b] In terms of inflation rates rather than price levels.

Table 3A. Sensitivity of Key Coefficients to Changes in Model Specification: Phase II RLMS

	Excluding Regional Effects[a]			Ignoring Sampling Weights			Quadratic Expenditure		
	Expenditure definition			*Expenditure definition*			*Expenditure definition*		
	All expenditures	Excluding durables	Restricted expenditures	All expenditures	Excluding durables	Restricted expenditures	All expenditures	Excluding durables	Restricted expenditures
ln (real total expenditure)	-0.079	-0.071	-0.062	-0.080	-0.071	-0.061	0.741	0.796	0.471
	(13.40)**	(12.48)**	(10.74)**	(13.48)**	(12.43)**	(10.81)**	(8.02)**	(8.22)**	(5.68)**
[ln real total expenditure]^2							-0.032	-0.034	-0.021
							(8.56)**	(8.71)**	(6.27)**
Time Dummy Variables									
Round 6 (Oct-Nov, 1995)	-0.021	-0.027	-0.018	-0.019	-0.025	-0.017	-0.022	-0.028	-0.019
	(1.85)+	(2.41)*	(1.87)+	(1.64)	(2.20)*	(1.67)	(2.00)+	(2.55)*	(1.99)+
Round 7 (Oct-Nov, 1996)	-0.053	-0.058	-0.034	-0.054	-0.058	-0.034	-0.054	-0.058	-0.035
	(5.21)**	(5.41)**	(3.30)**	(5.20)**	(5.38)**	(3.21)**	(5.54)**	(5.76)**	(3.53)**
Round 8 (Nov-Dec, 1998)	-0.096	-0.104	-0.064	-0.095	-0.102	-0.063	-0.089	-0.097	-0.061
	(9.58)**	(10.47)**	(7.65)**	(9.51)**	(10.08)**	(6.98)**	(8.88)**	(9.70)**	(7.26)**
Round 9 (Oct-Nov, 2000)	-0.112	-0.121	-0.097	-0.112	-0.120	-0.095	-0.113	-0.123	-0.099
	(10.60)**	(10.97)**	(9.43)**	(10.13)**	(10.51)**	(8.55)**	(10.64)**	(11.15)**	(9.87)**
Round 10 (Oct-Nov, 2001)	-0.121	-0.128	-0.091	-0.121	-0.127	-0.089	-0.124	-0.130	-0.094
	(10.38)**	(11.80)**	(7.91)**	(10.47)**	(11.54)**	(7.61)**	(10.20)**	(11.65)**	(8.25)**
R^2	0.24	0.22	0.31	0.24	0.22	0.31	0.27	0.26	0.32
F-test (time dummies=0)	34.45**	43.16**	22.11**	40.26**	47.40**	24.21**	30.27**	38.22**	21.84**

Note: Absolute value of t-statistics in parentheses corrected for cluster effects but not stratification; * significant at 5%; ** significant at 1%; + significant at 10% . *N*=7807. Each model includes background coefficients and intercepts that are not reported. The excluded time dummy variable is for Round 5 (Nov-Dec, 1994).
The various expenditure definitions affect the dependent variable (food-at-home budget share), and two right-hand side variables (real total expenditure and food-away-from-home). The restricted definition of expenditures includes only food, clothes, fuel and services.
[a] Equation (10) where the time dummy variables capture the effect of variation over time in the inflation rate for food relative to non-food and where no regional intercepts and no relative food price is included in the specification.

Table 3B. Sensitivity of Key Coefficients to Changes in Sample Specification: Phase II RLMS

	Excluding Households With Monthly Total Expenditures < 200 Rubles[a]			Excluding Households With Monthly Total Expenditures < 400 Rubles[b]			Excluding Households With Monthly Total Expenditures < 600 Rubles[c]		
	Expenditure definition			*Expenditure definition*			*Expenditure definition*		
	All expenditures	Excluding durables	Restricted expenditures	All expenditures	Excluding durables	Restricted expenditures	All expenditures	Excluding durables	Restricted expenditures
ln (real total expenditure)	-0.081	-0.073	-0.063	-0.083	-0.075	-0.065	-0.086	-0.078	-0.067
	(14.27)**	(13.42)**	(11.30)**	(14.37)**	(13.52)**	(11.63)**	(14.23)**	(13.34)**	(11.30)**
Time Dummy Variables									
Round 6 (Oct-Nov, 1995)	-0.021	-0.027	-0.018	-0.020	-0.026	-0.018	-0.020	-0.026	-0.018
	(1.84)+	(2.40)*	(1.86)+	(1.77)+	(2.31)*	(1.80)+	(1.77)+	(2.30)*	(1.72)+
Round 7 (Oct-Nov, 1996)	-0.053	-0.058	-0.034	-0.053	-0.057	-0.034	-0.053	-0.058	-0.034
	(5.24)**	(5.44)**	(3.32)**	(5.15)**	(5.33)**	(3.29)**	(5.22)**	(5.38)**	(3.23)**
Round 8 (Nov-Dec, 1998)	-0.094	-0.102	-0.062	-0.090	-0.098	-0.060	-0.084	-0.092	-0.057
	(9.28)**	(10.15)**	(7.44)**	(8.65)**	(9.47)**	(6.95)**	(8.06)**	(8.91)**	(6.83)**
Round 9 (Oct-Nov, 2000)	-0.112	-0.121	-0.097	-0.111	-0.120	-0.096	-0.111	-0.120	-0.097
	(10.61)**	(10.97)**	(9.46)**	(10.59)**	(10.92)**	(9.31)**	(10.42)**	(10.70)**	(9.14)**
Round 10 (Oct-Nov, 2001)	-0.121	-0.128	-0.091	-0.121	-0.128	-0.091	-0.120	-0.126	-0.089
	(10.35)**	(11.75)**	(7.90)**	(10.27)**	(11.65)**	(7.86)**	(10.19)**	(11.53)**	(7.76)**
R^2	0.24	0.23	0.31	0.24	0.23	0.31	0.25	0.24	0.31
F-test (time dummies=0)	34.45**	43.16**	22.11**	31.50**	39.73**	20.24**	31.29**	39.41**	20.33**

Note: See Table 3A.

[a] Equivalent to approximately US$7 per household per month. This restriction removes 16 observations.

[b] Equivalent to approximately US$15 per household per month. This restriction removes 75 observations.

[c] Equivalent to approximately US$22 per household per month. This restriction removes 168 observations.

Table 4: Descriptive Statistics and Econometric Results for Fixed Effects Food Engel Curve Estimated from Households Present in the First Round of Phase II RLMS

	Descriptive Statistics			Engel Curve Regression Results *Expenditure definition*		
	Mean[a] (std. dev)	Round 5 Mean	Round 10 Mean	All expenditures	Excluding durables	Restricted expenditures
Budget share for food at home	0.529 (0.221)	0.555	0.474			
ln (real total expenditure)	12.916 (0.878)	13.128	12.975	-0.075 (15.81)**	-0.064 (13.44)**	-0.046 (10.59)**
ln relative food price[b]	-0.382 (0.328)	-0.181	-0.465	0.046 (1.27)	0.024 (0.66)	0.011 (0.33)
ln (household size)	1.080 (0.300)	1.086	1.075	0.003 (0.11)	0.005 (0.18)	-0.048 (1.81)+
% of household = 2 years old	0.019 (0.073)	0.022	0.017	0.199 (2.93)**	0.183 (2.74)**	0.253 (3.97)**
% of HH 3-14 year old boys	0.078 (0.141)	0.084	0.068	0.088 (1.66)+	0.077 (1.46)	0.135 (2.69)**
% of HH 3-14 year old girls	0.078 (0.140)	0.089	0.063	0.196 (3.65)**	0.177 (3.33)**	0.253 (5.02)**
% of HH 15-17 year old boys	0.021 (0.074)	0.021	0.022	0.070 (1.38)	0.053 (1.06)	0.080 (1.68)+
% of HH 15-17 year old girls	0.022 (0.075)	0.018	0.025	0.188 (3.67)**	0.180 (3.54)**	0.172 (3.58)**
Dummy: detached dwelling	0.086 (0.281)	0.079	0.092	0.010 (0.39)	0.006 (0.25)	0.016 (0.65)
Age of household head	46.053 (12.442)	44.639	47.620	0.005 (1.36)	0.004 (1.25)	0.003 (0.85)
Age of spouse	43.208 (12.895)	41.809	44.766	-0.001 (0.40)	-0.001 (0.49)	0.000 (0.07)
Head has tertiary education	0.371 (0.483)	0.774	0.285	-0.008 (0.79)	-0.007 (0.71)	-0.013 (1.31)
Spouse has tertiary education	0.362 (0.481)	0.787	0.286	0.012 (1.20)	0.011 (1.11)	0.002 (0.25)
Head is working	0.679 (0.467)	0.695	0.698	-0.016 (1.78)+	-0.018 (1.99)*	-0.019 (2.28)*
Spouse is working	0.653 (0.476)	0.658	0.673	0.004 (0.51)	0.002 (0.24)	-0.017 (2.01)*
% of budget food out of home	0.042 (0.087)	0.047	0.041	-0.479 (13.84)**	-0.500 (14.84)**	-0.645 (23.36)**
Round 6 (Oct-Nov, 1995)	0.183 (0.387)			0.002 (0.14)	-0.004 (0.36)	-0.006 (0.57)
Round 7 (Oct-Nov, 1996)	0.169 (0.375)			-0.033 (2.16)*	-0.040 (2.64)**	-0.026 (1.85)+
Round 8 (Nov-Dec, 1998)	0.154 (0.361)			-0.084 (3.50)**	-0.091 (3.84)**	-0.048 (2.12)*
Round 9 (Oct-Nov, 2000)	0.147 (0.354)			-0.087 (2.62)**	-0.095 (2.90)**	-0.071 (2.28)*
Round 10 (Oct-Nov, 2001)	0.150 (0.357)			-0.099 (2.57)*	-0.107 (2.80)**	-0.070 (1.92)+
Constant				1.368 (5.67)**	1.280 (5.36)**	1.220 (5.40)**
R^2						
F-test (fixed effects=0)				1.79**	1.80**	1.82**

F-test (time dummies=0)	5.71**	5.99**	1.67**

Note: Absolute value of heteroscedastically robust t-statistics in parentheses; * significant at 5%; ** significant at 1%; + significant at 10%.

The sample is 1761 households who were surveyed in Round 5, with 6011 observations on those households. The excluded time dummy is for Round 5 (Nov-Dec, 1994).

The degrees of freedom for the tests of all fixed effects are zero 1760,4229, and for the time dummies are 5,4229.

The various expenditure definitions affect the dependent variable (food-at-home budget share), and ln (real total expenditure) and food-away-from-home share. The restricted definition of expenditures includes only food, clothes, fuel and services.

[a] For the expenditure and food share variables, the descriptive statistics are for "all expenditures" definition.

[b] In terms of inflation rates rather than price levels.

Table 5. Estimates of Cumulative CPI Bias in Russia, 1994-2001

	Cross-Sectional Estimates						Panel Fixed Effects Estimates		
	Including Regional Variation in Relative Food Price Inflation			Using National Variation in Relative Food Price Inflation			Including Regional Variation in Relative Food Price Inflation		
	Expenditure definition			*Expenditure definition*			*Expenditure definition*		
	All expenditures	Excluding durables	Restricted expenditures	All expenditures	Excluding durables	Restricted expenditures	All expenditures	Excluding durables	Restricted expenditures
Round 5 (Nov-Dec, 1994)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Round 6 (Oct-Nov, 1995)	0.225	0.306	0.268	0.219	0.301	0.240	-0.022	0.062	0.124
	(0.100)	(0.099)	(0.110)				(0.155)	(0.165)	(0.202)
Round 7 (Oct-Nov, 1996)	0.453	0.526	0.404	0.240	0.307	0.076	0.356	0.465	0.436
	(0.058)	(0.056)	(0.082)				(0.131)	(0.127)	(0.175)
Round 8 (Nov-Dec, 1998)	0.682	0.750	0.628	0.619	0.693	0.522	0.675	0.760	0.645
	(0.040)	(0.037)	(0.046)				(0.104)	(0.090)	(0.172)
Round 9 (Oct-Nov, 2000)	0.742	0.805	0.776	0.716	0.782	0.747	0.688	0.776	0.787
	(0.039)	(0.036)	(0.044)				(0.139)	(0.117)	(0.146)
Round 10 (Oct-Nov, 2001)	0.760	0.817	0.753	0.732	0.789	0.702	0.734	0.812	0.778
	(0.044)	(0.039)	(0.054)				(0.138)	(0.114)	(0.175)
Average bias per month	0.9%	1.0%	0.9%	0.9%	0.9%	0.8%	0.9%	1.0%	0.9%

Note: Based on coefficients estimated in Tables 2, 3 and 4. Standard errors in ( ) robust to heteroscedasticity and cluster effects.