Draft Only
January, 2004

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.

Acknowledgements:
We are grateful for assistance from Olga Antimonova, John Bryant, Geua Boe-Gibson, Daniel Berkowitz, and Konstantin Gluschenko and to helpful comments from Betsy Brainerd, Valery Lazarev, Tom Mroz, and seminar audiences at Canterbury, Cornell, Duke, Michigan, Williams and the Northeast Universities Development Consortium Conference.

JEL: C43, E31, O12, P36
Keywords: CPI bias, Living standards, Measurement error, Prices, Russia, Transition

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

---

[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, the number of retail stores and public catering establishments doubled, from 455,000 in 1992 to 935,000 by 1999 (Spulber, 2003, 323). While many of the new outlets were small (with fewer than 30 workers) large, foreign-owned, hyper-stores such as IKEA also entered 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).

crisis on the physical well-being of the Russian population, using six measures of nutrition – 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, 2002a). 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

[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.

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).

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 (2001) argues, inconsistency between the trends in

---

[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.

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.[10] The food budget share for this household would be 43.9% 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 48.4% in 1998. In fact, households with CPI-deflated total expenditures of 1000 roubles in November 1998 had food budget shares of only 44.4%. Thus, when viewed from the standard of their budget shares, Russian households in November 1998 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.[11]

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:[12]

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

---

[10] Values are in terms of the new roubles, introduced from 1998, where one new rouble=1000 old roubles. A value of 1900 roubles would place this household at the 83rd percentile of the expenditure distribution in 1996.

[11] 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.

[12] 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.

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), $\mathbf{X}$ is a vector of 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:

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

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

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

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,

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

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

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

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:

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

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

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

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:

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

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:

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

if either γ or $(1\text{-}r)$ 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.[13] 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:

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

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:

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

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 the 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,

[13] 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.[14] These households were located in 21 survey sites in 16 different *Oblasts*.[15] The second phase spans the period 1994 through 2001, with six rounds of data collected from approximately 4000 households.[16] 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 the monetary value of production for own consumption. Because it is mainly food that is self-produced, any attempt by us to value self-produced items is likely to affect the food budget share and the Engel curve estimates. This sensitivity to imputation procedures is most likely to affect rural households, so these households are excluded from the analysis.[17] This sample restriction should not diminish the relevance of the results because the prices for the CPI are collected from towns and cities, so urban households seem to be the relevant sample. Moreover, urban households account for 77% of the sample and population.

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.[18] Analysis by Stillman and Thomas (2002) suggests that attrition related to observable characteristics is not a serious concern in these data.[19] 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.

[14] 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.

[15] 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*.

[16] 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.

[17] See below for an analysis of accounting for the so-called "Dacha production", which is gardening by urban households in their weekend country cottages.

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

[19] 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. These restrictions are similar to those employed by Hamilton and Costa, and are designed to provide a more homogeneous sample with higher data quality. This sample selection does not appear to influence the pattern of results (robustness tests are reported below). Control variables include real total expenditures, relative food price changes, demographic, educational and employment characteristics, indicators of dwelling characteristics, and regional and time dummies.[20] Two variants of the total expenditure variable are used; one that includes all items enumerated by the survey and one that excludes purchases of durables. The model also includes the budget share for 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 54% in Phase II.[21] The average food share fell by 2 percentage points between Rounds 1 and 4 in Phase I and by 9 percentage points between Rounds 5 (late 1994) and 10 (late 2001) of Phase II, despite declines in CPI-deflated total consumption in both phases. Several other factors, such as relative food prices, employment rates and household composition also varied over the sample period, so regression analysis is needed to untangle the effect of these changes on food budget shares.

The estimation is carried out using both OLS and Instrumental Variables, because of concerns about measurement error in the total expenditures variable. measurement error bias in the total expenditures coefficient, *ß* would flow through into the estimate of CPI-bias (see equation 9). Household income is used as an instrument because this variable is collected independently of the total expenditures variable.

Phase I results

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. This fall is most apparent in the IV results, which are the preferred ones because the Hausman test indicates some inconsistency in the OLS estimates.

[20] 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.
[21] These two averages are not comparable because of differences in the samples and the questionnaires.

Relative to the base period (July-Oct, 1992), the food share is about two percentage points lower in Round 2 and three points lower in Round 4 (Nov-Dec, 1993), conditional on the other covariates. But in contrast to the expected pattern with CPI-bias, the food share in Round 3 (June-July, 1993) is higher than in the base period. Seasonality is a possible culprit for this pattern, because the Household Budget Survey also shows higher average food shares in the June quarter than in any other quarter.

Seasonality interferes with the measurement of CPI bias, but to the extent that Round 1 and Round 4 have some overlapping months, a tentative estimate can be made. Combining the three percentage point fall in the conditional food budget share from Round 1 to Round 4 with movement in the national food-nonfood inflation rate (using an estimate of $\bar{g} = 0.19$ which is derived using an approach described below), the application of equation (11) suggests a cumulative CPI bias of approximately 0.33 between July 1992 and the end of 1993.[22] This implies an average monthly bias of about 2 percentage points per month, during a period when the average monthly change in the CPI was about 20 percentage points. This estimate should not be regarded as definitive because of the short time period for 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.

Phase II results

Table 2 contains the estimates of the food Engel curves for the phase II RLMS data. The key result is that relative to the base period (Nov-Dec, 1994), the food share is 1, 4, 8, 10, and 11 percentage points lower in the subsequent survey rounds, conditional on the other covariates. All of these changes, except for the fall in the food share from Round 5 to Round 6 are statistically significant. Thus, there has been continual downward drift in the food Engel curve, a specific example of which was illustrated in Figure 3 for 1996 and 1998. By comparing the four columns of regression results, it is clear that this drift in the food Engel curve is not affected very much by the particular definition of household expenditures or by the estimation method. In contrast to Phase I, there is no significant difference between the IV and OLS results.

Sensitivity analyses reported in Table 3 all point to the same result, that there is an unexplained decline in the food budget share of between 10 and 13 percentage points between 1994 and 2001. The first sensitivity analysis in Panel A uses the national-level model, where regional effects and the regional variation in relative food prices is excluded (equation (10)). Some analysts might favour this model because the relative food price effect is not very precisely estimated in Table 2. Regardless, the pattern of the period dummy variables is very similar to what was previously estimated, although these now measure not just CPI bias but also the effect on budget shares of intertemporal variation in the measured national inflation rate for food relative to nonfood. Augmenting this model with a quadratic expenditure term does not change the dummy variable coefficients showing the downward drift in the food Engel curve, and in fact the quadratic terms are not statistically significant in the IV results.[23]

[22] If the OLS estimates are used, the cumulative bias is between 0.56 and 0.64, with the higher figure from the estimates that exclude durables purchases.

[23] The Hausman test result of $F_{(2,24)}$=290 suggests that the IV results should be used.

The analyses in Table 3B consider two other modelling choices, whether to use sampling weights and whether to focus just on two-adult households.[24] In neither case does it make any difference to the results. The unweighted decline in conditional food budget shares is slightly larger than the weighted decline, so ignoring the weights would strengthen any claims about the magnitude of the CPI bias. Similarly, using all household types, and adding dummy variables to control for the different categories, gives slightly larger coefficient values than when the results are just for two-adult households.

The sensitivity of the results to extreme values in the sample is considered in Table 3C, and once again the basic pattern is confirmed. First, the sample was trimmed to include only those households with food shares in the 0.02-0.90 interval. This is a fairly deep cut compared with the rules used by Hamilton and Costa, yet it has relatively little effect.[25] The magnitude of the dummy variables for each time period is reduced but so too is the coefficient on deflated expenditures, ß and because the CPI bias is estimated as the ratio of these two coefficients, there is no net effect on the bias estimates. Removing households with low (and possibly mis-measured) levels of expenditures, here defined as less than 600 roubles per month, has only a small impact on the pattern of the coefficients.

The panel structure of the data is exploited in Table 4, which contains the results of including household fixed effects in the regressions. Compared with the cross-sectional results, the dummy variables for most time periods are somewhat larger in absolute value, so the introduction of the fixed effects acts to slightly raise the CPI bias estimates. A comparison of the results in 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 (6120 household-round observations). The sample in Table 2 includes the new households who moved into sample dwellings (7753 household-round observations). The similarity of the two sets of results suggests that attrition is not affecting the estimates.

VII) Analysis

The estimation results indicate a persistent and substantial downward drift in the food Engel curves. We attribute this drift to unmeasured growth in real expenditures. We have no reason to believe that the nominal expenditure estimates from the RLMS are becoming increasingly understated, so in turn, this mis-measurement of real expenditures is attributed to CPI bias. If the assumptions underlying equation (9) are satisfied, the cumulative CPI bias after each round of the RLMS is found by dividing the coefficient on the dummy variable for the round by the income coefficient: $1-\exp(-d_t/b)$. The average monthly bias can be found by dividing the difference between cumulative bias estimates by the number of months separating them.

Over the seven years from November 1994 to November 2001, the cumulative bias in the Russian CPI is estimated to lie somewhere between 0.64 and 0.87, depending on the estimation method, sample, and expenditure definition used (Table 5). This range corresponds to an average monthly bias of between 0.8 and 1.0 percentage points. The lower values come from the IV estimates but only one of the four Hausman tests was statistically significant, so there is no

[24] Deaton (1997) discusses the debate about whether survey sampling weights should be used in regressions.

[25] Hamilton (2001) excluded households if their food share exceeded 80%, which is over four times higher than the mean food share in his sample. In contrast, the 90% threshold used here is less than twice the mean food share.

strong reason for favouring these lower values. Amongst the OLS results, the most precise measure of bias comes from the cross-sectional estimate of a cumulative bias of 0.80, with $\hat{s} = 0.04$, when durable purchases are excluded from the expenditure definition. Thus, our preferred estimate of the average monthly bias is 0.9 percentage points, which compares with a monthly CPI inflation rate of 3.2 percentage points over 1994-2001. In other words, almost one third of the measured rise in the cost of living can be attributed to CPI bias.

The level of CPI bias appears to be falling, both absolutely and relative to the inflation rate. Between Rounds 5 and 8 (late 1994 to late 1998) the average bias was 1.5 percentage points per month, while the monthly inflation rate was 4.2 percentage points.[26] But between Rounds 8 and 10 (late 1998 to late 2001) the average bias was just 0.3 points per month, with a monthly inflation rate of 1.9 percentage points. This fall in the bias seems plausible because it is earlier in the period where there was the greatest volatility in prices, which would contribute to commodity substitution bias. It was also in the immediate aftermath of price liberalisation when prices between state-run and private stores diverged the most, giving more scope for outlet bias if the price surveys by Goskomstat failed to keep pace with changing consumer shopping patterns. The improvements made by Goskomstat in 1998, such as the extension from 280 to 400 items in the basket, also are likely to have contributed to a reduction in the bias.

One concern with the results in Table 5 is that the substantial change in relative food prices illustrated in Figure 1b is controlled for only imprecisely because the food price coefficient, γ in Tables 2 and 4 are surrounded by wide standard errors. So as a sensitivity analysis, a value for γ was derived from an equation with no price data (this equation is reported in the first column of Table 3A).[27] Equation (11) was then used to combine this derived estimate, $\bar{g}$ with information on the aggregate movement in the relative price of food in order to retrieve estimates of CPI bias from the coefficients of a model without regional effects and without relative food prices. The cumulative bias estimates ranged from 0.70 to 0.76, and were only 0.04 points below the estimates from the model with food prices included. Thus, we doubt that uncertainty about the size of γ greatly affects the results. In terms of the other sensitivity analyses reported in Table 3, the cumulative bias estimates for Round 10 range from 0.64-0.75, with a mean value of 0.70. Once again, this is quite close to the preferred values reported in Table 5.

What are the implications of this CPI bias for assessments of the trend in living standards during Russia's transition to a market economy? Figure 4 displays the trend in real per capita GDP, which shows the 'usual' story that GDP in 2001 was only three-quarters of its 1991 value.[28] To the extent that GDP statistics measure welfare, this suggests that transition was associated with a precipitous drop in the Russian population's standard of living. But if the value of real Household Final Consumption Expenditure is adjusted for the effect of CPI bias, a rather

[26] This monthly average of 1.5 points lends some credence to our earlier estimate, from the RLMS Phase I data, of a two percentage point monthly bias between July 1992 and the end of 1993.

[27] The specific steps were to first use the method proposed by Frisch (1959) to get an own-price elasticity from the food budget share, w_i the income elasticity of food demand, h and the 'flexibility of money' $e_{ii} = (1/w)h_i(1 - w_i h_i) - w_i h_i$, where ω is -4.2, based on the relationship used by Lluch *et. al.* (1977) of $w \approx -36 X^{-0.36}$, 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. The resulting value for e_{ii} of -0.56 was then used to derive an estimate of γ=0.19, noting that for equation (1), the own-price elasticity is: $e_{ii} = -1 + (g - ab)/w$ where a is the share of food in the overall price index.

[28] The estimates are in local currency units, series NY.GDP.MKTP.KN from the *World Development Indicators*.

different picture emerges. Just adjusting consumption, and leaving all other components unchanged, real per capita GDP would be 20% higher in 2001 if it is assumed that there was no CPI bias prior to 1994 (that is, just using the Engel curve results from Phase II of RLMS). Allowing for bias back to 1992, by also using the results from Phase I of RLMS, real per capita GDP in 2001 is 30% higher than its officially reported level, and has returned to the level experienced in 1991.

VIII) Corroborating Evidence

Our results suggest that the official inflation figures considerably overstate rises in the cost of living in Russia and contribute to an overly pessimistic view of declining living standards during the transition. Some corroborating evidence for these conclusions comes from data in two other parts of the RLMS questionnaires. The first concerns the ownership of durable goods, while the second uses subjective questions on changes in economic welfare.

Increasing ownership rates for consumer durables are not consistent with the prolonged decline in real household consumption that is indicated by the official statistics. According to the RLMS figures, the proportion of urban households owning a VCR rose steadily from only 20% in 1994 to 53% in 2001. Less dramatic but equally steady rises in ownership rates are indicated for color TVs, and cars and trucks (Figure 5). Yet between 1994 and 1998, the official figures suggest that real per capita GDP fell by 11%. There is a high income elasticity of durables ownership, so this falling real GDP seems inconsistent with the improvements in living standards indicated by rising ownership rates for household durables.[29]

Self-reported changes in economic welfare in Russia are more consistent with the bias-adjusted CPI than with the official CPI. In Round 9 (2000) the RLMS asked: ‘How did you and your family live five years ago compared to how you live now?’ and respondents were able to answer using a 5-point scale that ranged from ‘lived much better’, through ‘lived the same as now’ to ‘lived much worse’. Time series of similar questions have been used by Nordhaus (1998) and Krueger and Siskind (1998) to measure CPI bias in the U.S., while Filer and Hanousek (2002) use a cross-section of similar data to measure CPI bias in Romania. The idea is to look for the deflator that gives the closest association between the subjective, self-rated welfare change, and an objective measure of welfare change based on deflated per capita expenditures.

When the CPI is used to deflate per capita expenditures, a majority (59%) of those who rated themselves as being better off in 2000 than in 1995 are in the group for whom the CPI indicates a *fall* in real per capita expenditures (Table 6a). Similarly, the group who indicated no change in their living standards are divided 63:37 between those with apparent falls in CPI-deflated expenditure and those with rises. When the bias-adjusted CPI is used to deflate per capita expenditures, a (slight) majority of those who indicate that they were better off in 2000 also have a rise in the objective welfare measure, while the 55:45 split for those indicating ‘no change’ is closer to the 50:50 split that would be expected (Table 6b). In terms of statistical significance, two measures of association for categorical data – Cramer’s V and the chi-squared statistic – indicate significantly better fit (p=0.04) between the objective and subjective welfare changes

[29] In cross-sections for the first (1994) and last (2001) years in the sample, unit increases in log household expenditures are associated with 26 percentage point and 28 percentage point increases in the probability of the household owning a VCR. These values come from an IV estimation of a linear probability model.

when the bias-adjusted CPI is used.[30] Taking the comparisons in Table 6 one step further, a grid search was carried out to find the CPI that best reconciles the subjective report of welfare change with the objective change in deflated per capita total expenditures. Both Cramer's *V* and the chi-squared statistic are maximized by a CPI that equals 379 in 2000 (1995=100). This is rather closer to the value of the bias-adjusted CPI than it is to the official CPI (Figure 6).

IX) Other Explanations for Drifting Food Engel Curves

The unexplained fall in the food budget share of Russian households points to an understatement of real income, which we attribute to CPI bias. But there are other possible causes of this trend, including greater own-production of food by urban households, the so-called 'Dacha production'. If Dacha production has increased over time, there may be a spurious fall in our measure of the food share. To check this possibility, RLMS data on consumption from self-produced potatoes were evaluated. Potatoes are the main food grown by urban households.

The available evidence on urban household's own-production of food does not suggest any bias in our previous estimates (Table 7). First, between 1995 and 2001 there been a continuous decline in the proportion of urban households consuming self-produced potatoes. Second, when consumption values are imputed from the reported quantities (using the median unit value for potato purchases), and compared to the total value of consumption, the apparent budget share for self-produced potatoes has fallen by two-thirds. Finally, augmenting the food budget share to include self-produced potatoes has no effect on the level or pattern of cumulative CPI bias.

Another concern relates to how much faith can be placed on the results for an Engel curve that is estimated over a period that includes a major economic crisis (following the August 1998 financial crisis). McKenzie (2001) has shown that after the 1994 Mexican Peso crisis, budget shares for food staples increased by more than a pre-crisis food Engel curve would predict. In contrast, shares for non-staple foods and semi-durables like clothing fell by more than the Engel curve would predict. However, such adjustments are less apparent amongst Russian households (Table 8). Whilst there was a small rise in the budget share for staple food in 1998, which was accommodated by a fall in the non-staple food share, there was no reduction in the budget shares for semi-durables.[31] Moreover, if we simply remove all observations from 1998 (Round 8) from our sample, and re-estimate CPI bias, the cumulative bias estimates by 2001 are exactly the same as those reported previously in Table 5.

X) 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

[30] Cramer's *V* statistic tests the null hypothesis of no association between the row variable (the objective measure of welfare change) and the column variable (the subjective measure). For 2×2 tables, $-1<V<1$, and $0<V<1$ otherwise. The chi-squared statistic is based on a comparison of the observed joint frequency with the frequency that would exist if the row and column variables were statistically independent. See Agresti (1984) for discussion of both statistics.

[31] There are 16 items included as staples: white bread, black bread, macaroni products, rice/cereals, cabbage, potatoes, beets/carrots, onions/garlic, vegetable oil, flour, salt/spices, tea, milk, margarine and sugar. These were identified based on their low expenditure elasticity of demand, in both rounds 5 and 10.

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 growth to be understated.

We find an average CPI bias of about two percentage points per month during Phase I of the RLMS (1992-93) and about one percentage point per month in Phase II (1994-2001). Greater confidence can be attached to the results from Phase II. 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. Even just allowing household final consumption to be deflated with bias, while assuming that the other components of GDP are unbiased, we find that the level of real per capita GDP in 2001 may be understated by up to 30% compared with using a bias-corrected deflator (Figure 4). Combined with other adjustments, such as for growth in the unofficial economy and the reduction in wasteful production (Shleifer and Treisman, 2003), the real value of GDP may in fact be rather larger than it was at the beginning of transition. Even the official figures show that household consumption collapsed less than investment and government spending, so after correction for CPI-bias, a rise in average household living standards during the transition seems highly likely.

The widely reported 'Russian mortality crisis' (Shkolnikov, *et al.*, 1998) may seem inconsistent with this claim of rising average living standards. Between 1990 and 2001, male life expectancy in Russia dropped from 63.8 years to 58.6 years, with most of the decline in the early period.[32] Yet there are at least two reasons for doubting that this decline is a symptom of falling living standards during the transition. First, life expectancy was already declining prior to the transition, falling by four years for males between the mid-1960s and the mid-1980s (Eberstadt, 1995). Campaigns to reduce alcohol consumption briefly interrupted this trend but after these ended life expectancy fell again to the level predicted from the earlier trend (Andreev, 2001). Second, if the higher mortality reflected greater poverty, it would be expected to strike economically vulnerable groups like the young and the elderly but instead most of the increased mortality occurred among working age Russian males (Shleifer and Treisman, 2003).

Our results also contribute to the puzzle raised by Stillman and Thomas (2002) about the contrast between the stability of nutritional indicators and the size of the shock to real incomes in 1998. While the 1998 crisis still shows up in the bias-adjusted data, as an 8% deviation from the trend in per capita GDP, it is preceded by some years of growth, rather than the decline that is apparent in the official data. Thus it is less surprising that households were able to cope with the 1998 shock and buffer their nutrition than it would have been if they had exhausted their resources in dealing with the previous seven year decline in living standards that the official data indicate.

[32] We attempted to use the Phase II RLMS data to see if the trends in life expectancy and death rates in the survey were consistent with the aggregate evidence. However, the sample is too small to have enough deaths (roughly 100 per survey round) to analyse either mortality change or the determinants of mortality.

References

Agresti, A. 1984. *Analysis of Ordinal Categorical Data*. New York: Wiley.

Andreev, E. 2001. "Did the economic crisis cause the mortality fluctuation in Russia in the 1990s?" Paper presented at the 24th IUSSP Conference, Salvador de Bahia, Brazil.

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-xx.

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. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: Johns Hopkins University Press.

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.

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

Eberstadt, N. 1995. *The Tyranny of Numbers: Mismeasurement and Misrule* AEI Press, Washington DC.

Filer, R., and Hanousek, J. 2002. "Survey-based estimates of biases in Consumer Price Indices during transition: evidence from Romania" *Journal of Comparative Economics* 30(3): 476-487.

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(3): 411-434.

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.

Krueger, A. and Siskind, A. 1998. "Using survey data to assess bias in the Consumer Price Index" *Monthly Labor Review* 121(4): 25-33.

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):

McKenzie, D. 2001. "The consumer response to the Mexican Peso crisis" Northeast Universities Development Consortium Conference, Boston University.

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

Nordhaus, W. 1998. "Quality changes in price indexes" *Journal of Economic Perspectives* 12(1): 59-68.

Shkolnikov, V., Cornia, G., Leon, D., and Meslé, F. 1998. "Causes of the Russian mortality crisis: evidence and interpretations" *World Development* 26(11): 1995-2011.

Shleifer, A., and Treisman, D. 2003. "A normal country" *Working Paper* 10057, National Bureau of Economic Research.

Spulber, N. 2003. *Russia's Economic Transitions: From Late Tsarism to the New Millennium* Cambridge: Cambridge University Press.

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. *Transition – the First Ten Years: Analysis and Lessons for Eastern Europe and the Former Soviet Union* Washington: The World Bank.

World Bank. 2002a. "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
% change over previous month
40
35
30
25
20
15
10
5
0
-5
Jan-92
Jan-93
Jan-94
Jan-95
Jan-96
Jan-97
Jan-98
Jan-99
Jan-00
Jan-01
Jan-02
Jan-03

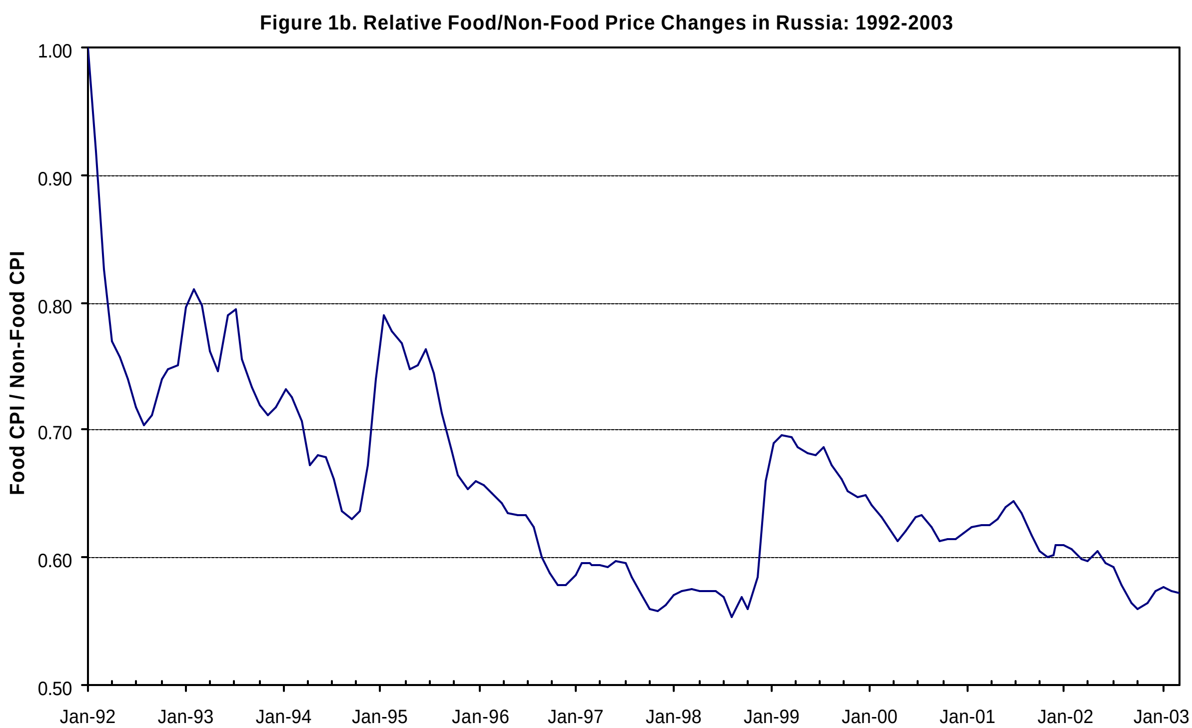


Figure 1b. Relative Food/Non-Food Price Changes in Russia: 1992-2003
Food CPI / Non-Food CPI
1.00
0.90
0.80
0.70
0.60
0.50
Jan-92
Jan-93
Jan-94
Jan-95
Jan-96
Jan-97
Jan-98
Jan-99
Jan-00
Jan-01
Jan-02
Jan-03

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

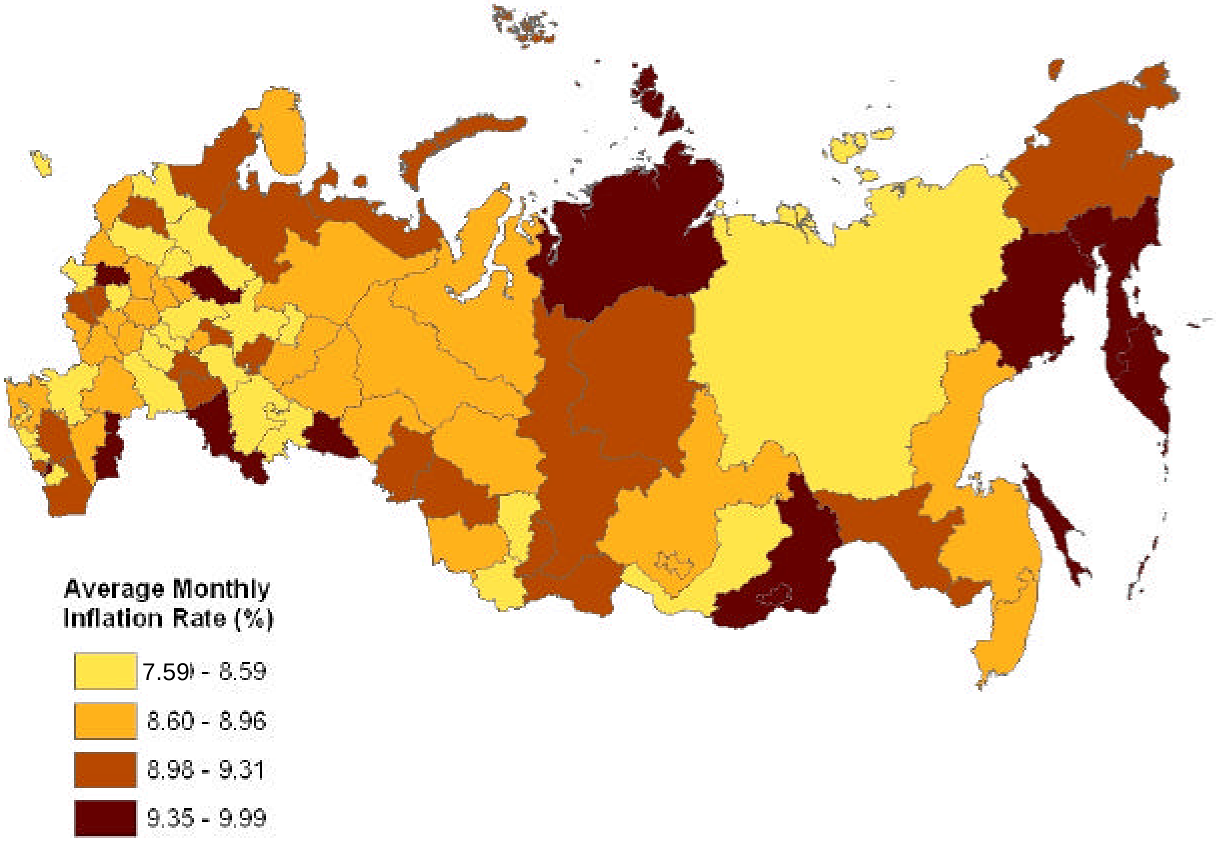


Figure 3: Food Engel Curves for 1996 and 1998

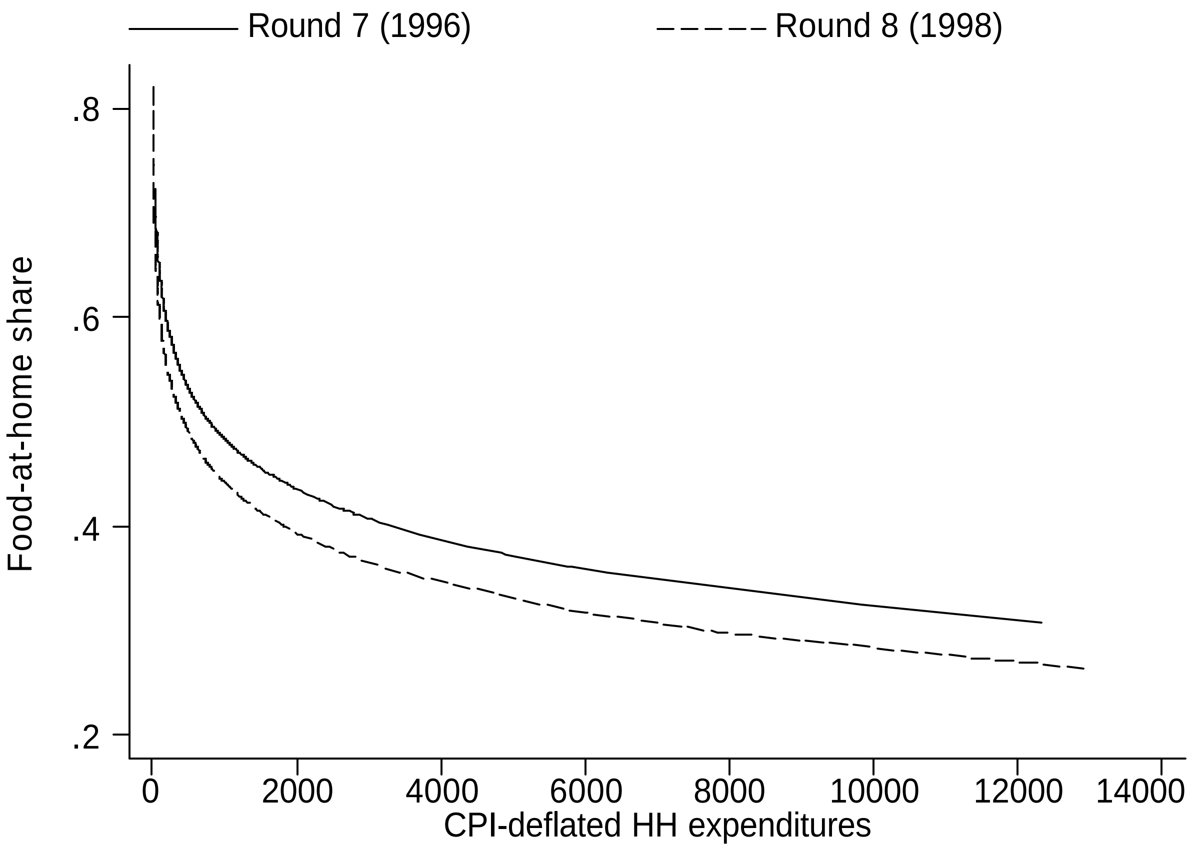
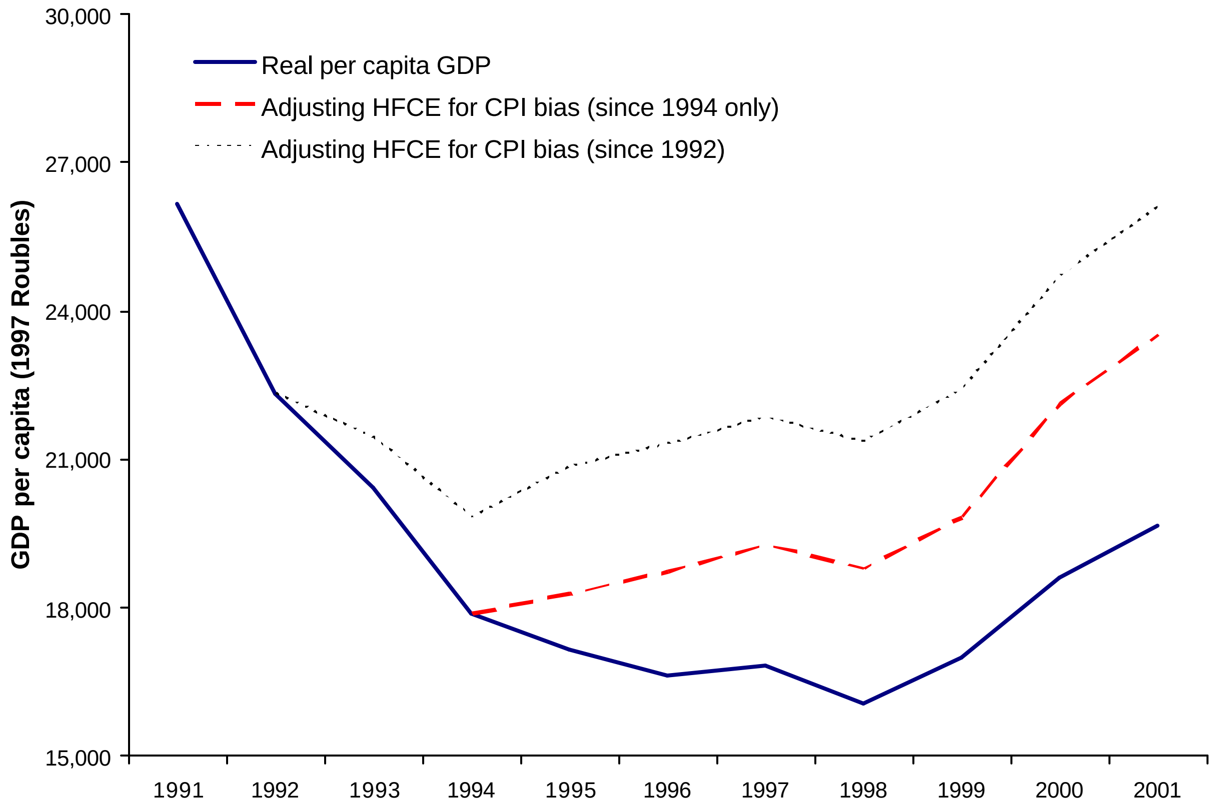


Figure 4: Effect of CPI Bias on Estimates of Real per capita GDP in Russia



Note: HFCE = Household Final Consumption Expenditure

Figure 5: Ownership Rate for Certain Household Durables
(urban, 2-adult families)

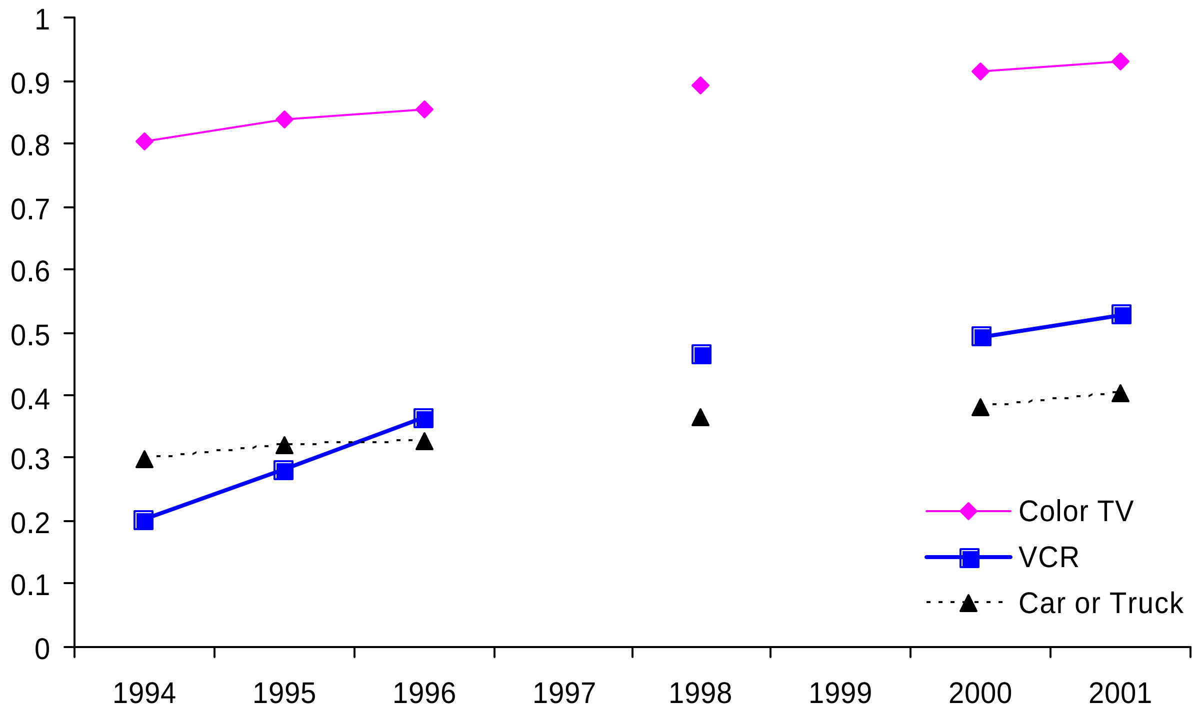


Figure 6: Grid Search Results for Russian Consumer Price Inflation Rate That Reconciles Objective and Subjective Welfare Changes

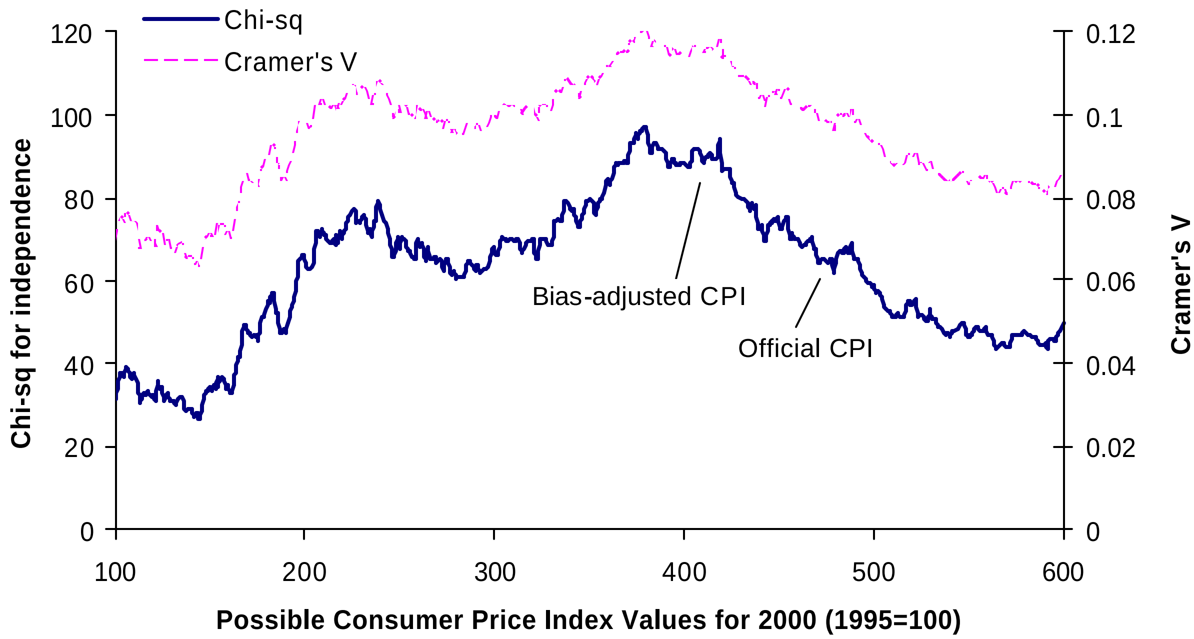


Table 1: Descriptive Statistics and Econometric Results for Food Engel Curve Estimated from Phase I RLMS

	Descriptive Statistics			Engel Curve Regression Results			
				OLS Estimates		IV Estimates[a]	
	Mean[b] (std. dev)	Round 1 Mean	Round 4 Mean	All expenditures	Excluding durables	All expenditures	Excluding durables
Budget share for food at home	0.581 (0.217)	0.581	0.561				
ln (real total expenditure)	9.101 (0.791)	9.119	9.071	-0.045 (5.25)**	-0.038 (4.55)**	-0.112 (6.56)**	-0.099 (5.60)**
ln (household size)	1.103 (0.305)	1.115	1.101	-0.032 (1.64)	-0.040 (2.09)+	-0.001 (0.04)	-0.011 (0.47)
% of household = 2 years old	0.021 (0.077)	0.026	0.018	0.134 (2.16)*	0.173 (2.87)*	0.049 (0.70)	0.092 (1.32)
% of HH 3-14 year old boys	0.094 (0.152)	0.094	0.093	0.170 (5.85)**	0.181 (6.53)**	0.127 (4.05)**	0.142 (4.60)**
% of HH 3-14 year old girls	0.086 (0.147)	0.090	0.085	0.159 (4.80)**	0.167 (5.43)**	0.125 (3.50)**	0.136 (4.00)**
% of HH 15-17 year old boys	0.022 (0.076)	0.022	0.024	0.112 (3.58)**	0.123 (3.78)**	0.099 (3.47)**	0.111 (3.71)**
% of HH 15-17 year old girls	0.023 (0.077)	0.023	0.022	0.042 (1.26)	0.055 (1.55)	0.037 (1.14)	0.049 (1.43)
Age of household head	45.186 (11.686)	44.598	45.698	0.003 (3.76)**	0.002 (3.17)**	0.003 (3.59)**	0.003 (3.17)**
Age of spouse	42.783 (12.618)	42.134	43.320	0.003 (3.05)**	0.003 (3.40)**	0.002 (1.83)+	0.002 (2.25)*
Head has tertiary education	0.245 (0.430)	0.287	0.216	-0.003 (0.60)	-0.005 (0.89)	0.007 (1.19)	0.005 (0.85)
Spouse has tertiary education	0.230 (0.421)	0.271	0.199	-0.015 (2.74)*	-0.013 (2.65)*	-0.008 (1.33)	-0.007 (1.40)
Head is working	0.715 (0.452)	0.764	0.678	-0.007 (0.96)	-0.004 (0.55)	0.001 (0.12)	0.003 (0.34)
Spouse is working	0.694 (0.461)	0.725	0.670	0.004 (0.45)	0.004 (0.43)	0.011 (1.28)	0.010 (1.15)
% of budget on food out of home	0.025 (0.066)	0.030	0.022	-0.427 (7.73)**	-0.453 (8.78)**	-0.443 (7.20)**	-0.463 (8.25)**

Round 2 (Jan-Mar, 1993)	0.240	-0.015	-0.021	-0.018	-0.023
	(0.427)	(1.31)	(1.86)+	(1.57)	(2.02)+
Round 3 (June-July, 1993)	0.243	0.019	0.015	0.024	0.020
	(0.429)	(2.38)*	(1.83)+	(2.79)*	(2.32)*
Round 4 (Oct 1993-Jan 1994)	0.242	-0.031	-0.032	-0.031	-0.033
	(0.428)	(3.72)**	(3.93)**	(3.34)**	(3.46)**
Constant		0.772	0.737	1.378	1.281
		(9.03)**	(8.99)**	(8.69)**	(7.97)**
R^2		0.130	0.117	0.075	0.071
F-test (time dummies=0)		11.76**	13.11**	14.74**	15.70**
F-test (instrument = 0 in first stage regression)				183.1**	167.4**
F-test (Hausman test for consistency of OLS)				28.07**	18.87**

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*=8416.
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.
[a] ln (real total expenditure) is treated as the endogenous variable, with ln(real total household income) as the instrument.
[b] 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			
				OLS Estimates		IV Estimates[a]	
	Mean[b] (std. dev)	Round 5 Mean	Round 10 Mean	All expenditures	Excluding durables	All expenditures	Excluding durables
Budget share for food at home	0.536 (0.226)	0.565	0.475				
ln (real total expenditure)	12.910 (0.878)	13.116	12.970	-0.081 (11.29)**	-0.073 (10.21)**	-0.113 (6.80)**	-0.103 (5.84)**
ln relative food price[c]	0.007 (0.123)	0.060	-0.036	0.041 (1.14)	0.032 (0.87)	0.032 (0.87)	0.021 (0.58)
ln (household size)	1.077 (0.301)	1.081	1.074	0.017 (1.16)	0.013 (0.92)	0.026 (1.57)	0.022 (1.37)
% of household = 2 years old	0.019 (0.073)	0.022	0.017	0.138 (4.73)**	0.155 (4.86)**	0.115 (3.58)**	0.132 (3.79)**
% of HH 3-14 year old boys	0.078 (0.140)	0.083	0.068	0.111 (4.21)**	0.108 (4.20)**	0.106 (3.80)**	0.104 (3.83)**
% of HH 3-14 year old girls	0.078 (0.140)	0.087	0.063	0.091 (3.27)**	0.087 (3.17)**	0.082 (2.90)**	0.080 (2.89)**
% of HH 15-17 year old boys	0.021 (0.074)	0.021	0.022	0.079 (2.31)*	0.068 (1.81)+	0.086 (2.59)*	0.074 (2.04)+
% of HH 15-17 year old girls	0.021 (0.075)	0.017	0.025	0.014 (0.36)	0.010 (0.25)	0.012 (0.31)	0.006 (0.17)
Dummy: detached dwelling	0.089 (0.284)	0.081	0.092	-0.022 (1.64)	-0.021 (1.67)	-0.024 (1.54)	-0.023 (1.54)
Age of household head	45.886 (12.212)	44.689	47.301	0.003 (3.26)**	0.003 (3.37)**	0.003 (3.23)**	0.003 (3.35)**
Age of spouse	43.649 (13.345)	42.379	45.076	0.002 (2.34)*	0.002 (2.13)*	0.002 (2.39)*	0.002 (2.15)*
Head has tertiary education	0.264 (0.441)	0.236	0.286	-0.027 (4.45)**	-0.028 (4.65)**	-0.017 (2.14)*	-0.019 (2.29)*
Spouse has tertiary education	0.278 (0.481)	0.269	0.283	-0.009 (1.66)	-0.009 (1.62)	-0.004 (0.62)	-0.005 (0.70)

Head is working	0.669	0.684	0.687	-0.010	-0.010	-0.005	-0.006
	(0.470)			(1.72)+	(1.88)+	(0.74)	(0.86)
Spouse is working	0.652	0.652	0.681	0.001	0.000	0.007	0.005
	(0.476)			(0.31)	(0.06)	(1.10)	(0.91)
% of budget on food out of home	0.041	0.046	0.041	-0.478	-0.503	-0.474	-0.497
	(0.086)			(16.40)**	(17.57)**	(17.13)**	(18.34)**
Round 6 (Oct-Nov, 1995)	0.184			-0.006	-0.012	-0.009	-0.013
	(0.387)			(0.60)	(1.16)	(0.84)	(1.27)
Round 7 (Oct-Nov, 1996)	0.169			-0.035	-0.040	-0.044	-0.048
	(0.375)			(4.38)**	(4.85)**	(4.83)**	(5.11)**
Round 8 (Nov-Dec, 1998)	0.154			-0.080	-0.087	-0.098	-0.103
	(0.361)			(8.49)**	(9.28)**	(7.57)**	(8.15)**
Round 9 (Oct-Nov, 2000)	0.145			-0.102	-0.110	-0.112	-0.118
	(0.352)			(11.02)**	(11.51)**	(12.57)**	(12.92)**
Round 10 (Oct-Nov, 2001)	0.148			-0.108	-0.115	-0.115	-0.120
	(0.355)			(13.05)**	(14.94)**	(15.62)**	(17.56)**
Constant				1.388	1.311	1.799	1.697
				(16.58)**	(15.59)**	(8.79)**	(7.78)**
R^2				0.267	0.254	0.258	0.246
F-test (time dummies=0)				46.2**	59.9**	70.4**	87.9**
F-test (region dummies=0)				16715**	17702**	78332**	63482**
F-test (instrument = 0 in first stage regression)						57.8**	58.1**
F-test (Hausman test for consistency of OLS)						2.59	2.29

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*=7753. 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.

[a] ln (real total expenditure) is treated as the endogenous variable, with ln(real total household income) as the instrument.

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

[c] 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]				Using Quadratic Expenditures			
	OLS Estimates		IV Estimates[b]		OLS Estimates		IV Estimates[b]	
	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables
ln (real total expenditure)	-0.079	-0.071	-0.102	-0.093	0.750	0.805	0.390	0.442
	(13.19)**	(12.09)**	(6.17)**	(5.61)**	(8.33)**	(8.52)**	(0.95)	(1.06)
[ln real total expenditure]^2					-0.032	-0.034	-0.019	-0.021
					(8.86)**	(9.00)**	(1.22)	(1.31)
Time dummy variables								
Round 6 (Oct-Nov, 1995)	-0.007	-0.012	-0.009	-0.013	-0.008	-0.012	-0.009	-0.013
	(0.64)	(1.17)	(0.80)	(1.22)	(0.72)	(1.22)	(0.85)	(1.24)
Round 7 (Oct-Nov, 1996)	-0.039	-0.043	-0.047	-0.049	-0.039	-0.043	-0.047	-0.049
	(3.95)**	(4.17)**	(4.50)**	(4.58)**	(4.19)**	(4.38)**	(4.55)**	(4.62)**
Round 8 (Nov-Dec, 1998)	-0.082	-0.089	-0.096	-0.101	-0.074	-0.081	-0.092	-0.097
	(7.93)**	(8.62)**	(7.12)**	(7.77)**	(7.43)**	(8.06)**	(6.34)**	(6.78)**
Round 9 (Oct-Nov, 2000)	-0.103	-0.111	-0.111	-0.117	-0.103	-0.112	-0.111	-0.118
	(10.00)**	(10.44)**	(11.23)**	(11.57)**	(10.12)**	(10.63)**	(11.28)**	(11.69)**
Round 10 (Oct-Nov, 2001)	-0.112	-0.118	-0.117	-0.122	-0.114	-0.119	-0.118	-0.123
	(10.82)**	(12.21)**	(11.88)**	(13.17)**	(10.78)**	(12.26)**	(11.92)**	(13.24)**
R^2	0.246	0.234	0.242	0.230	0.277	0.268	0.267	0.258
F-test (time dummies=0)	34.68**	44.57**	51.66**	65.47**	31.86**	41.42**	48.31**	60.36**

Note: Absolute value of t-statistics in parentheses corrected for cluster effects but not stratification; * significant at 5%; ** significant at 1%; + significant at 10%.
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 ln (real total expenditure) and food-away-from-home share.
[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.
[b] ln (real total expenditure) is treated as the endogenous variable, with ln(real total household income) as the instrument.

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

	Ignoring Sampling Weights				Using All Households Rather Than Just Two-Adult Households[a]			
	OLS Estimates		IV Estimates[b]		OLS Estimates		IV Estimates[b]	
	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables
ln (real total expenditure)	-0.080	-0.072	-0.103	-0.093	-0.082	-0.073	-0.121	-0.112
	(13.32)**	(12.08)**	(6.46)**	(5.87)**	(11.49)**	(10.34)**	(9.45)**	(8.33)**
Time dummy variables								
Round 6 (Oct-Nov, 1995)	-0.007	-0.012	-0.009	-0.013	-0.010	-0.016	-0.013	-0.018
	(0.63)	(1.15)	(0.79)	(1.21)	(1.09)	(1.84)+	(1.44)	(2.05)+
Round 7 (Oct-Nov, 1996)	-0.041	-0.045	-0.049	-0.051	-0.044	-0.050	-0.053	-0.058
	(4.19)**	(4.38)**	(4.81)**	(4.85)**	(6.57)**	(7.39)**	(6.82)**	(7.60)**
Round 8 (Nov-Dec, 1998)	-0.084	-0.090	-0.097	-0.102	-0.091	-0.099	-0.113	-0.119
	(8.31)**	(8.78)**	(8.02)**	(8.56)**	(11.22)**	(12.26)**	(10.85)**	(12.13)**
Round 9 (Oct-Nov, 2000)	-0.105	-0.112	-0.112	-0.118	-0.109	-0.116	-0.121	-0.126
	(9.69)**	(10.12)**	(10.95)**	(11.25)**	(12.72)**	(13.15)**	(14.60)**	(15.24)**
Round 10 (Oct-Nov, 2001)	-0.114	-0.119	-0.119	-0.123	-0.114	-0.122	-0.122	-0.128
	(11.09)**	(12.14)**	(12.14)**	(13.08)**	(13.16)**	(15.91)**	(14.55)**	(17.41)**
R^2	0.247	0.234	0.243	0.230	0.267	0.256	0.253	0.241
F-test (time dummies=0)	41.90**	50.43**	62.74**	72.25**	46.27**	65.56**	53.61**	75.13**

Note: Absolute value of t-statistics in parentheses corrected for cluster effects but not stratification; * significant at 5%; ** significant at 1%; + significant at 10%.
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 ln (real total expenditure) and food-away-from-home share.
[a] Dummy variables for couples with and without children, for mixed families and for non-family households are added to the model (the excluded category is sole occupants). The sample size increases to 10,466 with the additional types of households included.
[b] ln (real total expenditure) is treated as the endogenous variable, with ln(real total household income) as the instrument.

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

	Excluding Households With Food Shares Outside 0.02-0.90[a]				Excluding Households With Monthly Expenditures < 600 Roubles[b]			
	OLS Estimates		IV Estimates[b]		OLS Estimates		IV Estimates[b]	
	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables
ln (real total expenditure)	-0.077	-0.069	-0.084	-0.086	-0.086	-0.077	-0.108	-0.111
	(13.17)**	(12.07)**	(6.34)**	(6.21)**	(14.26)**	(12.97)**	(6.11)**	(5.97)**
Time dummy variables								
Round 6 (Oct-Nov, 1995)	-0.005	-0.004	-0.006	-0.005	-0.008	-0.006	-0.009	-0.008
	(0.51)	(0.37)	(0.58)	(0.51)	(0.71)	(0.58)	(0.85)	(0.75)
Round 7 (Oct-Nov, 1996)	-0.034	-0.033	-0.039	-0.039	-0.041	-0.039	-0.048	-0.048
	(3.61)**	(3.38)**	(4.01)**	(4.00)**	(4.18)**	(3.93)**	(4.63)**	(4.62)**
Round 8 (Nov-Dec, 1998)	-0.075	-0.069	-0.080	-0.079	-0.072	-0.067	-0.081	-0.079
	(8.17)**	(7.38)**	(7.20)**	(7.08)**	(6.78)**	(6.23)**	(6.45)**	(6.33)**
Round 9 (Oct-Nov, 2000)	-0.090	-0.086	-0.093	-0.092	-0.104	-0.100	-0.110	-0.109
	(9.35)**	(8.95)**	(10.15)**	(10.06)**	(9.91)**	(9.58)**	(11.06)**	(10.96)**
Round 10 (Oct-Nov, 2001)	-0.097	-0.095	-0.099	-0.099	-0.112	-0.110	-0.116	-0.115
	(9.56)**	(9.31)**	(10.57)**	(10.26)**	(10.75)**	(10.53)**	(11.74)**	(11.37)**
R^2	0.219	0.204	0.221	0.203	0.261	0.246	0.259	0.236
F-test (time dummies=0)	27.97**	25.37**	40.91**	39.04**	32.58**	30.33**	47.04**	44.75**

Note: Absolute value of t-statistics in parentheses corrected for cluster effects but not stratification; * significant at 5%; ** significant at 1%; + significant at 10%.
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 ln (real total expenditure) and food-away-from-home share.
[a] Removes 360 observations.
[b] This is equivalent to US$22 per household per month. This exclusion removes 173 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			
				OLS Estimates		IV Estimates[a]	
	Mean[b] (std. dev)	Round 5 Mean	Round 10 Mean	All expenditures	Excluding durables	All expenditures	Excluding durables
Budget share for food at home	0.544 (0.228)	0.566	0.487				
ln (real total expenditure)	12.890 (0.872)	13.118	12.858	-0.073 (15.58)**	-0.062 (13.21)**	-0.128 (4.58)**	-0.109 (3.70)**
ln relative food price[c]	0.019 (0.120)	0.060	-0.013	0.049 (1.35)	0.027 (0.77)	0.046 (1.22)	0.025 (0.69)
ln (household size)	1.065 (0.305)	1.088	1.028	-0.005 (0.18)	-0.003 (0.11)	0.018 (0.60)	0.014 (0.46)
% of household = 2 years old	0.016 (0.067)	0.023	0.007	0.194 (2.84)**	0.182 (2.70)**	0.141 (1.89)+	0.140 (1.92)+
% of HH 3-14 year old boys	0.075 (0.138)	0.085	0.056	0.084 (1.56)	0.073 (1.37)	0.071 (1.29)	0.061 (1.13)
% of HH 3-14 year old girls	0.074 (0.138)	0.088	0.050	0.193 (3.56)**	0.175 (3.27)**	0.169 (2.95)**	0.159 (2.84)**
% of HH 15-17 year old boys	0.022 (0.076)	0.021	0.026	0.058 (1.13)	0.041 (0.81)	0.060 (1.15)	0.042 (0.82)
% of HH 15-17 year old girls	0.021 (0.075)	0.018	0.028	0.187 (3.64)**	0.180 (3.55)**	0.186 (3.50)**	0.181 (3.45)**
Dummy: detached dwelling	0.097 (0.296)	0.079	0.126	-0.009 (0.37)	-0.012 (0.47)	0.001 (0.05)	-0.002 (0.07)
Age of household head	47.003 (12.077)	44.541	50.382	0.005 (1.36)	0.004 (1.24)	0.006 (1.74)+	0.006 (1.70)+
Age of spouse	44.881 (13.276)	42.225	48.440	0.001 (0.23)	0.000 (0.12)	0.001 (0.45)	0.001 (0.41)
Head has tertiary education	0.248 (0.432)	0.235	0.249	-0.009 (0.56)	-0.010 (0.63)	-0.002 (0.13)	-0.004 (0.23)
Spouse has tertiary education	0.272 (0.445)	0.267	0.254	0.011 (0.75)	0.007 (0.50)	0.017 (1.08)	0.013 (0.88)

Head is working	0.654	0.688	0.647	-0.012	-0.015	-0.008	-0.011
	(0.476)			(1.37)	(1.69)+	(0.80)	(1.14)
Spouse is working	0.658	0.652	0.684	0.003	0.002	0.006	0.004
	(0.474)			(0.31)	(0.21)	(0.64)	(0.46)
% of budget on food out of home	0.038	0.046	0.035	-0.496	-0.517	-0.489	-0.510
	(0.084)			(14.22)**	(15.21)**	(13.67)**	(14.54)**
Round 6 (Oct-Nov, 1995)	0.192			-0.003	-0.008	-0.012	-0.015
	(0.394)			(0.33)	(0.96)	(1.21)	(1.60)
Round 7 (Oct-Nov, 1996)	0.165			-0.038	-0.045	-0.057	-0.061
	(0.371)			(2.81)**	(3.33)**	(3.62)**	(3.97)**
Round 8 (Nov-Dec, 1998)	0.143			-0.093	-0.099	-0.138	-0.139
	(0.350)			(4.01)**	(4.33)**	(4.45)**	(4.53)**
Round 9 (Oct-Nov, 2000)	0.134			-0.104	-0.112	-0.139	-0.144
	(0.341)			(3.18)**	(3.45)**	(3.82)**	(4.03)**
Round 10 (Oct-Nov, 2001)	0.122			-0.119	-0.125	-0.148	-0.153
	(0.327)			(3.11)**	(3.32)**	(3.65)**	(3.85)**
Constant				1.283	1.199	1.868	1.670
				(5.28)**	(4.99)**	(4.66)**	(4.05)**
R^2				0.223	0.210	0.220	0.205
F-test (time dummies=0)				5.50**	5.76**	5.50**	5.16**
F-test (fixed effects=0)				1.85**	1.86**	1.79**	1.82**
F-test (instrument = 0 in first stage regression)						131.5**	117.4**
F-test (Hausman test for consistency of OLS)						3.89*	2.57

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

The sample is 1774 households who were surveyed in Round 5, with 6120 observations on those households. The excluded time dummy is for Round 5 (Nov-Dec, 1994). Each equation also includes 1773 household-level 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.

[a] ln (real total expenditure) is treated as the endogenous variable, with ln(real total household income) as the instrument.

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

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

[d] The *F*-tests have numerator degrees of freedom of 5 for the test of the time dummies, 1773 for the fixed effects, and 1 for the first stage instrument and Hausman test. The denominator degrees of freedom are 4325.

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

	Cross-Sectional Estimates				Panel Fixed Effects Estimates			
	OLS Estimates		IV Estimates		OLS Estimates		IV Estimates	
	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables	All expenditures	Excluding durables
Round 5 (Nov-Dec, 1994)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Round 6 (Oct-Nov, 1995)	0.074	0.149	0.079	0.122	0.039	0.127	0.089	0.131
	(0.118)	(0.117)	(0.088)	(0.089)	(0.118)	(0.124)	(0.065)	(0.072)
Round 7 (Oct-Nov, 1996)	0.351	0.422	0.325	0.375	0.407	0.513	0.361	0.430
	(0.061)	(0.060)	(0.049)	(0.051)	(0.110)	(0.106)	(0.074)	(0.087)
Round 8 (Nov-Dec, 1998)	0.624	0.696	0.580	0.632	0.718	0.796	0.660	0.720
	(0.046)	(0.044)	(0.038)	(0.044)	(0.089)	(0.075)	(0.067)	(0.074)
Round 9 (Oct-Nov, 2000)	0.713	0.779	0.627	0.683	0.759	0.834	0.663	0.732
	(0.041)	(0.039)	(0.058)	(0.064)	(0.109)	(0.088)	(0.098)	(0.102)
Round 10 (Oct-Nov, 2001)	0.735	0.795	0.637	0.689	0.802	0.866	0.686	0.755
	(0.048)	(0.043)	(0.066)	(0.071)	(0.104)	(0.083)	(0.108)	(0.110)
Average bias per month	0.9%	0.9%	0.8%	0.8%	1.0%	1.0%	0.8%	0.9%

Note: Based on coefficient estimates reported in Tables 2 and 4. Standard errors in ( ) robust to heteroscedasticity and cluster effects.

Table 6a: Comparison of subjective evaluation of welfare change with change in real per capita expenditures, using the CPI as the deflator[a]

	Subjective welfare change "How did you and your family live five years ago compared to now?"					
Real per capita expenditures in 1995:	Much better	Somewhat better	Same as now	Somewhat worse	Much worse	Total
Higher than in 2000	973	1,411	1,398	458	183	4,423
Lower than in 2000	376	643	834	332	125	2,310
Total	1,349	2,054	2,232	790	308	6,733

[a]Cramer's $V = 0.1009$; Chi-square = 68.5 (significant at $p < 0.0005$).

The CPI has a value of 485 in 2000 (1995=100).

Table 6b: Comparison of subjective evaluation of welfare change with change in real per capita expenditures, using the bias-adjusted CPI as the deflator[a]

	Subjective welfare change "How did you and your family live five years ago compared to now?"					
Real per capita expenditures in 1995:	Much better	Somewhat better	Same as now	Somewhat worse	Much worse	Total
Higher than in 2000	889	1,282	1,233	390	153	4,023
Lower than in 2000	460	772	999	400	155	2,710
Total	1,349	2,054	2,232	790	308	6,733

[a]Cramer's $V = 0.1158$; Chi-square = 90.2 (significant at $p < 0.00005$).

The bias-adjusted CPI has a value of 411 in 2000 (1995=100).

Table 7. The Effects of Own-Production of Potatoes

	Proportion Consuming from Own Production	Budget Share of Self-Produced Potatoes [a]	Cumulative Bias Estimates	
			Excluding Own-Production	Including Own-Production[a]
Round 5 (Nov-Dec, 1994)	0.513	0.010	0.000	0.000
Round 6 (Oct-Nov, 1995)	0.517	0.008	0.074	0.086
			(0.118)	(0.112)
Round 7 (Oct-Nov, 1996)	0.478	0.006	0.351	0.358
			(0.061)	(0.058)
Round 8 (Nov-Dec, 1998)	0.470	0.007	0.624	0.625
			(0.046)	(0.045)
Round 9 (Oct-Nov, 2000)	0.461	0.005	0.713	0.712
			(0.041)	(0.040)
Round 10 (Oct-Nov, 2001)	0.436	0.003	0.735	0.734
			(0.048)	(0.046)

[a] Imputed values are derived by applying the average unit values for purchases to the production quantities reported by households.

Table 8. Budget Shares Before, During and After the 1998 Financial Crisis

	Staple Food[a]	Non-Staple Food	All Food	Clothing	Education, Health, Recreation etc
Round 7 (Oct-Nov, 1996)	0.225	0.320	0.545	0.083	0.082
Round 8 (Nov-Dec, 1998)	0.241	0.301	0.542	0.083	0.095
Round 9 (Oct-Nov, 2000)	0.204	0.290	0.494	0.104	0.093
Average (1994 - 2001)	0.209	0.326	0.535	0.086	0.084

[a] Staple food includes white bread, black bread, macaroni products, rice/cereals, cabbage, potatoes, beets/carrots, onions/garlic, vegetable oil, flour, salt/spices, tea, milk, margarine and sugar.