# SURVEY-BASED ESTIMATES OF BIASES IN CONSUMER PRICE INDICES DURING TRANSITION: EVIDENCE FROM ROMANIA

RANDALL K. FILER[*] and JAN HANOUSEK[**]

May 2001

Abstract

*Mismeasurement of inflation is likely to be more severe in a transition economy than in a more stable environment. Comparisons of self-reported changes in economic welfare with changes in incomes suggest that official Romanian inflation measures may be overstated by between 100 and 300 percent at a reported annual inflation rate of 40 per cent.*

Abstrakt

*Dá se očekávat, že v období přechodu od plánované k tržní ekonomice budeme pozorovat daleko větší chyby v měření inflace než je tomu v zemích se stabilní tržní ekonomikou. Výsledky založené na porovnání změn v blahobytu se změnami příjmů naznačují, že oficiální hodnoty inflace v Rumunsku jsou výrazně vychýleny, přičemž vychýlení se pohybuje mezi 100 až 300 procenty při oficiální míře inflace 40 procent.*

JEL Codes: C8, E3, P2

---

[*] Professor of Economics, Hunter College and The Graduate Center, The City University of New York and CERGE-EI, a joint workplace of Charles University and The Academy of Sciences of the Czech Republic, e-mail: rfiler@hunter.cuny.edu;

**Citibank Associate Professor of Economics, CERGE-EI, corresponding author: CERGE-EI, Politických vězňů 7, 111 21 Prague 1, Czech Republic, e-mail: jan.hanousek@cerge.cuni.cz. Both Filer and Hanousek are Research Associates of the William Davidson Institute. This research was supported in part by grants P97-8119-R from the PHARE/ACE program of the European Union and II/75831 from the Volkswagen Foundation.

It is generally believed that biases in traditional measures of consumer prices may be especially great in transition economies (see Filer and Hanousek, 2000). All the well-understood reasons why conventional indices may overstate inflation should be exacerbated in the transition. Among these are massive changes in relative prices leading to consumer substitutions, substantial improvements in quality in even the most mundane goods, and the rapid proliferation of new, low-cost distribution channels.

Nordhaus (1998) and Krueger and Siskind (1998) have used consumer's self-reported changes in economic well-being to assess the possible extent of biases in measures of inflation in the United States. The intuition is that if inflation measures are biased (say, upwards), some consumers will report themselves to be better off when their real income adjusted for the official inflation rate will appear to have declined. An appropriate, unbiased inflation rate would be one at which only consumers who report themselves as worse off experienced an apparent decline in real income. Krueger and Siskind suggest estimating the regression:

$$Y = \alpha + \beta Z$$

where Y is the net percentage of families who report an improvement in their financial situation (i.e. those reporting an improvement minus those reporting a decline) and Z is the net percentage whose real income actually increased. Under the assumption that measures of Y are accurate (especially the absence of money illusion) and that the deflator used to calculate Z is unbiased, $\alpha$ should equal zero and $\beta$ should equal 1. Krueger and Siskind perform a grid search over various possible CPI deflators (and associated implied biases in the official CPI), and suggest that the value maximizing the *P*-value for the joint test of these restrictions is an appropriate selection for the true, unbiased CPI. For the U.S. they find that this value is close to, and may even be slightly greater than, the official inflation rate.

Although this is a clever strategy, it is of limited usefulness in a transition economy. First,

estimating the proposed regression requires several years of data.[1]  Second, there is an implicit

assumption that the coefficients are stable over time (i.e. that the extent of bias in the CPI is

constant).  In the context of transition, where inflation rates are both high and highly variable, and

where economies have moved from seriously distorted to relatively normal over the course of a

decade in which local statistical offices have become increasingly sophisticated, such an

assumption of stability is inherently questionable.

Fortunately, it is possible to develop a similar technique that uses only a single year's data.

Below we outline such a methodology and apply it to measures of inflation in 2000 in Romania.

Romania lends itself to such analysis because among the Central European countries with

reasonably advanced statistical offices,[2] it retains a high  inflation rate.  Officially, consumer prices

in January 2001 were 39.9 per cent higher than in January 2000.  Monthly rates of increase in the

CPI during 2000 ranged from a low of 1.8 per cent in March, May and August to a high of 6.0

per cent in July.[3]

Suppose a sample of individuals is asked their current monthly household income and their

monthly income one year ago.  For a given inflation rate $r$, denote by  $r^+$ families whose real

income has risen and by $r^-$ those whose real income has fallen.  Similarly, denote as $s^+$ those who

report that their financial situation has improved and $s^-$ as those who report a worsening in their

situation.  Table 1 is a 2 x 2 contingency table where misclassifications are denoted as $n^{+-} + n^{-+}$.  The total number of such misclassifications, denoted $n = n^{+-} + n^{-+}$ will vary as the inflation

---

[1]Krueger and Siskind use 27 years between 1968 and 1995.

[2]In the context of European Union accession, Romania's National Institute of Statistics
has received considerable technical assistance from the EU's PHARE program including
various "twinning" arrangements with West European statistical agencies.

[3]Taken from National Bank of Romania, <u>Monthly Bulletin</u>, January 2001, Table 2.
Original data provided by National Institute of Statistics and Economic Studies of Romania.

rate $r$ changes. The value of $r$ that minimizes misclassifications seems a reasonable candidate for the "true" or unbiased inflation rate.

**Table 1**

**Contingency Table for Misclassification Analysis**

| | Reported Well-Being Risen $s^+$ | Reported Well-Being Fallen $s^-$ |
|---|---|---|
| Real Income Risen $r^+$ | $n^{++}$ | $n^{+-}$ |
| Real Income Fallen $r^-$ | $n^{-+}$ | $n^{--}$ |

One advantage of this technique is that it does not require any information on actual incomes, only their year to year change. We asked a random probability sample of 1513 Romanian households[4] the following two questions:

1)  A household's income can change for many reasons. A member of the household may get a new job or be promoted, or someone may lose their job or retire. Someone could start working after leaving school or maternity leave or someone may add or leave a part-time job or start or end some private business. Someone in the household may become eligible for or lose government benefits. Considering all of these AND any other changes you can think of, by what PERCENTAGE was your household's income for the year 2000 greater or less than for 1999? Please note that we do NOT want you to tell us actual incomes, only the percentage change (record increases as + and decreases as -).

---

[4]The questions were added for us to the March 2000 Omnibus Survey conducted by the market research firm Mercury. We were referred to this firm through corporate supporters of CERGE-EI who use them for survey work in Romania. In addition to the responses to our questions we were provided with full demographic data on respondents and weights enabling adjustment of sample respondents to national norms by age, gender, region and income level.

2)      Of course, prices also changed last year, some by more than others. Considering everything together, including the change in your family's income, changes in prices and changes in what you bought, overall how would you say that your standard of living in 2000 has changed, compared with 1999? We are only concerned with your actual current economic condition, not any changes in your personal situation, the country as a whole, or your beliefs about the future.  Using this definition, please tell me how was your standard of living in 2000 compared to 1999?

      1.      A great deal better
      2.      Somewhat better
      3.      About the same
      4.      Somewhat worse
      5.      A great deal worse

Table 2 shows that a majority of Romanians believed that their financial situation deteriorated during 2000.  Almost half, however, said that their situation either improved or remained about the same over the past year.  Interesting, among those who said their situation remained about the same, the average increase in income was only 7.3 percent, far below the official inflation rate of 39.9 per cent.  Indeed, even among those who reported that their situation improved, the average increase in income (26.3 percent) was well below the official inflation rate. This strongly suggests that Romanians themselves do not feel that they need an income increase as large as the supposed increase in the CPI to maintain their financial situation.

**Table 2.**

**Average Income Change by Reported Change in Financial Well-Being**

| Financial Situation in 2000 | Number of Observations | Weighted No. of Observations | Average Income Change | Weighted Average Income Change |
|---|---|---|---|---|
| A Great Deal Better Than 1999 | 51 | 50 | 48.9% | 49.7% |
| Somewhat Better Than 1999 | 254 | 255 | 21.4% | 21.7% |
| About the Same as 1999 | 425 | 415 | 6.8% | 7.3% |
| Somewhat Worse Than 1999 | 484 | 486 | -9.1% | -8.4% |
| A Great Deal Worse Than 1999 | 299 | 307 | -21.1% | -20.2% |
| TOTAL | 1513 | 1513 | 0.9% | 0.5% |

The pattern of beliefs regarding changes in economic well-being is consistent with general patterns in transition. Younger respondents (under age 45) are substantially more likely to report an improvement in their economic situation (21 percent) than those 45 to 60 (15 per cent) or those over 60 (8 percent). Similarly, urban residents are more likely to report an improvement in their condition (23 per cent) than rural residents (17 per cent), while the highest frequency of respondents reporting an improvement occurs in the capital city, Bucharest (34 per cent).

Given these results, we performed a grid search over various possible "true" inflation rates ranging from 0 per cent to 40 per cent (slightly over the "official" rate of 39.9 per cent). There is considerable heaping of responses to the question asking for nominal income changes at even multiple of 5 per cent. Over 96 percent of respondents gave an answer that was a multiple of 5

per cent.  We therefore limited our grid search to these round number inflation rates.[5]  Table 3 presents the results of this exercise.  Consider the results in the first row.  The numbers show that 303.95 respondents (weighted sample) reported that their financial situation was better in 2000 than in 1999 and would have had a higher real income if the inflation rate were 0 percent while 555.40 respondents reported themselves as worse off in 2000 and would have had a lower real income if the inflation rate were 0 per cent.  We regard these respondents as "correctly classified" for a 0 per cent inflation rate.[6]  Similarly the 0.98 respondents who reported themselves as better off but who would have had a lower real income with 0 per cent inflation and the 237.76 respondents who reported themselves as worse off but who would have had a higher real income if inflation were 0 per cent are considered "incorrectly classified."  Our goal is to pick the inflation rate that minimizes the number of such incorrect classifications.

It is apparent in the table that the lowest number of misclassifications occurs in the range between 5 and 20 per cent, with the lowest actual number of misclassifications at a rate of 10 per cent, only slightly higher than the 7.3 per cent inflation rate that would be consistent with the nominal income changes of those who reported their financial situation as about the same.[7]

---

[5]A search over all inflation rates yields similar results with very small changes in misclassifications between multiples of 5 per cent.  Results for all inflation levels are available at http://home.cerge-ei.cz/hanousek/romania.

[6]The analysis of correct and incorrect classification is based on those who reported their financial situation as "better" or "worse" in 2000 than in 1999 and ignores those who reported their situation as about the same.  In the grid search, we must, therefore, decide how to classify those whose reported nominal income change exactly matches the inflation rate being used.  We have classified those with nominal income change equal to the assumed inflation rate as correctly classified if they report their financial situation as "somewhat better" or "somewhat worse" and incorrectly classified if they report their situation as "a great deal better" or "a great deal worse."  Recalculation assuming that all of these cases are incorrectly classified does not change the results.

[7]At first the existence of local minima at both 10 per cent and 20 per cent might seem somewhat odd. Recall, however, that a respondent who reported a nominal income change of 10 per cent will be deemed correctly classified if he reported his financial situation as either

## Table 3
## Relationship between Reported Well-Being and Changes in Real Income for Various Inflation Levels

| ASSUMED INFLATION | Correctly classified | | | Incorrectly classified | | |
|---|---|---|---|---|---|---|
| | BETTER | WORSE | TOTAL | BETTER | WORSE | TOTAL |
| 0 | 303.95 | 555.40 | 859.35 | 0.89 | 237.76 | 238.66 |
| 5 | 281.86 | 603.59 | 885.45 | 22.98 | 189.57 | 212.56 |
| 10 | 259.95 | 651.59 | 911.54 | 44.90 | 141.57 | 186.47 |
| 15 | 199.04 | 692.25 | 891.29 | 105.81 | 100.92 | 206.72 |
| 20 | 177.26 | 722.53 | 899.79 | 127.58 | 70.64 | 198.22 |
| 25 | 129.55 | 738.18 | 867.73 | 175.29 | 54.99 | 230.28 |
| 30 | 100.20 | 759.05 | 859.24 | 204.65 | 34.12 | 238.77 |
| 35 | 65.76 | 770.60 | 836.36 | 239.08 | 22.57 | 261.65 |
| 40 | 59.24 | 778.16 | 837.40 | 245.60 | 15.01 | 260.61 |

Note: All outcomes were weighted in order to get fully representative sample of Romanian population.

We have tried various modifications of our specification. In particular, results are not changed when we adopt a loss function that weights misclassification by 2 categories (reported one's self as a great deal better off when real income has declined) as twice as bad a failure as misclassification by only one category. Neither are they changed when individuals whose real income was unchanged are considered misclassified if they reported themselves as either better or worse off. Results do change slightly if we include those who reported there economic well-being as "about the same." Including these respondents reduces the misclassification minimizing apparent inflation rate to zero per cent for the year 2000. We prefer a more cautious conclusion and will base further discussion on the results underlying Table 3.[8]

---

"somewhat better" or "somewhat worse" when analyzing a 10 per cent assumed inflation rate. Thus, the number of correct classifications will depend in part on the number of observations at that rate. There were 207 respondents who reported a 10 per cent increase in their nominal income and 121 who reported a 20 per cent increase. This contrasts with only 62 respondents who reported a 15 per cent increase in their nominal income.

[8]Results from all of the alternative specifications are available at http://home.cerge-ei.cz/hanousek/romania.

While it is clear from Table 3 that many lower inflation rates predict individuals' self-report of changes in their financial well-being better than the official rate of approximately 40 per cent, Table 3 does not establish whether this difference in predictive accuracy is statistically significant.   Table 4 presents Z statistics and associated probabilities from a nonparametric test of whether the assumed inflations rates that minimize misclassifications (10 per cent and 20 per cent) are significantly better at predicting self reports of changes in economic well-being than various alternative inflation rates.  (See the appendix for a formal presentation of the test).

**Table 4**

**Tests of Significance of Differences in Predictive Accuracy for Various Inflation Rates**

|  | Predictive Accuracy of 10% Rate | Predictive Accuracy of 20 % Rate |
|---|---|---|
| Tested Against 0% Rate | 2.57 (p=.01) | 1.96 (p=.05) |
| Tested Against 5% Rate | 3.12 (p=.001) | 0.96 (p=.169) |
| Tested Against 10% Rate | n.a. | 0.98 (p=.171) |
| Tested Against 15% Rate | 2.01 (p=.022) | 0.96 (p=.119) |
| Tested Against 20% Rate | 0.95 (p=.171) | n. a. |
| Tested Against 25% Rate | 2.97 (p=.001) | 4.03 (p=.000) |
| Tested Against 30% Rate | 3.20 (p=.000) | 3.80 (p=.000) |
| Tested Against 35% Rate | 4.25 (p=.000) | 5.02 (p=.000) |
| Tested Against 40% Rate | 4.10 (p=.001) | 4.73 (p=.000) |

The pattern of results are quite clear.  For both of the suggested true inflation rates, there is no significant difference when compared with each other or with alternative low possible

inflation rates.  On the other hand, each suggested rate is significantly better at predicting consumers' self-reports of the change in their economic well being than any possible alternative inflation rate of 25 per cent per year or more.  This suggests that while there is considerable uncertainty as to the actual "best" inflation rate, we can confidently rule out any possibility greater than 25 per cent per year, including the official rate of 40 per cent per year.

Table 5 presents further confirmation that an inflation rate in the range of 10 to 20 percent is consistent with Romanian's self-perception of changes in their economic well-being.  While for the entire sample, substantially more Romanians report that their situation worsened than improved during 2000, the self-reports of changes in economic well-being for those who reported a 10 per cent increase in their nominal incomes were symmetric around a modal value of no change in their economic status (with 29 percent reporting an improvement and 28 per cent a worsening.  For those with a 20 per cent increase in nominal incomes, the modal respondent reported an improvement in their economic well-being.

**Table 5**

**Distribution of Reported Change in Economic Well-Being**

**for Various Nominal Income Changes**

| Financial Situation in 2000 | Number of Observations for Full Sample | Number of Observations for Those Reporting a 10% Change in Income | Number of Observations for Those Reporting a 20% Change in Income |
|---|---|---|---|
| A Great Deal Better Than 1999 | 51 (3.4%) | 4 (1.9%) | 2 (1.7%) |
| Somewhat Better Than 1999 | 254 (16.8%) | 56 (27.1%) | 45 (37.2%) |
| About the Same as 1999 | 425 (28.9%) | 89 (43.0%) | 42 (34.7%) |
| Somewhat Worse Than 1999 | 484 (32.0%) | 37 (17.9%) | 23 (19.0%) |
| A Great Deal Worse Than 1999 | 299 (19.8%) | 21 (10.1%) | 9 (7.4%) |
| TOTAL | 1513 | 207 | 121 |

These results are quite striking. If we accept the upper bound on our range of uncertainty as to the true inflation rate, 20 per cent a year, official rates for the year 2000 in Romania are twice what Romanians themselves say would be required in the way of an income increase to enable them to maintain their standard of living. At the lower bound (5 percent or even the 8 percent annual increase in income actually received by those Romanians who say their standard of living was about the same year-to-year), official inflation rates are *five times* actual rates.

Of course it is possible that self-reports of economic well-being are themselves biased and that Romanians are for some reason over-reporting their economic status. Such would be the case if there were severe money illusion among respondents, who noticed increases in their nominal income but not in prices. We suspect, however, that true price increases of 40 per cent in a given year would be hard for consumers to miss. Indeed, given the extensive documentation

11

in the Romanian press of the problem of inflation, it might be expected that respondent's would overestimate the problem caused by rising prices rather than suffer from money illusion. Thus, we are left with the conclusion that the official measure of increases in consumer prices in Romania for the year 2000 substantially overstates the true rate of inflation faced by Romanian consumers.

The fact that the true inflation rate in Romania appears to be substantially lower than the official rate has a number of important policy implications. To the extent that it is desired that wages and public assistance payments increase in order to compensate recipients for losses due to inflation, our findings suggest that these increases can be far smaller than previously supposed and yet still leave citizens feeling that they have maintained their standard of living. This will obviously reduce strains on tight local budgets and enable growth-promoting reductions in tax rates. Furthermore, if true inflation rates are substantially lower than official statistics suggest, there may well be room for regional central banks to adopt a more accommodating monetary policy, again with implications for investment and rates of growth.

**References**

Antoch, J. and Hanousek, J. (1999) "A specification test for discrete choice models," CERGE-EI Discussion Papers Series.  Prague, November 1999, no. 32.

Anderson, E.B. (1994) *The Statistical Analysis of Categorical Data (3rd ed.)*, Springer-Verlag, Berlin.

Filer, R. K. and Hanousek, J.  (2000)  "Output Changes and Inflationary Bias in Transition," *Economic Systems*, 24:3 pp: 285-294.

Gibbons, J. D. (1997)  *Nonparametric Methods for Quantitative Analysis (3rd ed.)*, American Sciences Press, Columbus, Ohio.

Hanousek, J. (2001) "Note on Specification Tests of (Binary) Choice Models: A Non-Parametric Approach," *Central European Journal of Operation Research*, forthcoming

Krueger, A. B. and Siskind, A. (1998) "Using Survey Data to Assess Bias in the Consumer Price Index," *Monthly Labor Review, 121:4 pp: 25*-33.

Nordhaus, W.  (1998) "Quality Changes in Price Indexes," *Journal of Economic Perspectives*, 12:1 pp: 59-68.

Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*, New York: Wiley.

Zvára, K. (1995) "True level of the tests about the parameters in the binomial distribution," *Bulletin of the Czech Statistical Society,* 6  pp. 27-31.

**Appendix**

**Nonparametric Test of Significance of Differences in Predictive Accuracy**

The following nonparametric test is used to determine whether there is a significant difference in the accuracy of two different inflation rates (See Antoch and Hanousek, 1999 and Hanousek, 2001 for more details). The joint distribution of correct and incorrect predictions of economic status is summarize by the following contingency table:

| | Inflation Rate 2 Predicts Correctly | Inflation Rate 2 Predicts Incorrectly | $\sum$ |
|---|---|---|---|
| Inflation Rate 1 Predicts Correctly | $n_{11}$ | $n_{12}$ | $n_{1x}$ |
| Inflation Rate 1 Predicts Incorrectly | $n_{21}$ | $n_{22}$ | $n_{2x}$ |
| $\sum$ | $n_{x1}$ | $n_{x2}$ | $n$ |

The hypothesis of whether or not one method dominates the other in terms of predictive accuracy is equivalent to testing for symmetry in the contingency table.

The frequencies in contingency tables represent a sample from the multinomial distribution with four categories and related probabilities equal to $p_{11}$, $p_{12}$, $p_{21}$, and $p_{22}$. The test of the null hypothesis $H_0$: $p_{11} = p_{22}$ is a test of whether the probability of correct prediction is the same for both models. In the 2 x 2 framework, this is equivalent to $H_0$: $p_{12} = p_{21}$. Using well-known results for the multinomial distribution, we obtain the conditional distribution of $n_{12}$ and $n_{21}$ given $n_{11}$ and $n_{22}$:

$$Pr(n_{12}, n_{21} \mid n_{11}, n_{22}) = \frac{(n - n_{11} - n_{22})!}{n_{12}! \, n_{21}!} \left( \frac{p_{12}}{1 - p_{11} - p_{22}} \right)^{n_{12}} \left( \frac{p_{21}}{1 - p_{11} - p_{22}} \right)^{n_{21}} \qquad (1)$$

For a fixed $n$, $n_{11}$ and $n_{22}$, it must be the case that $n - n_{11} - n_{22} = n_{12} + n_{21}$, (analogously $1 - p_{11}$

$- p_{22} = p_{12} + p_{21}$). Under the validity of the null hypothesis $H_0$, $p_{12} = p_{21}$, and, therefore,

$p_{12}/(p_{12} + p_{21}) = p_{21}/(p_{12} + p_{21}) = \frac{1}{2}$.

Hence, (1) reduces to:

$$Pr(n_{12}, n_{21} \mid n_{11}, n_{22}) = \binom{n_{12} + n_{21}}{n_{12}} \left(\frac{1}{2}\right)^{n_{12} + n_{21}}, \tag{2}$$

which is a binomial distribution with parameters $n_{12} + n_{21}$ and $\frac{1}{2}$. Then the test statistic

$$\chi^2 = \frac{[n_{12} - (n_{12} + n_{21})/2]^2}{(n_{12} + n_{21})/2} + \frac{[n_{21} - (n_{12} + n_{21})/2]^2}{(n_{12} + n_{21})/2} = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} \tag{3}$$

under the null hypothesis has an asymptotically $\chi^2$ distribution with one degree of freedom. The

proof follows from the central limit theorem and the distribution of quadratic forms (see,

Anderson (1994), Serfling (1980), among others).[9] It should be noted that we would get the

identical test statistic using the central limit theorem approximation of the binomial distribution.

Indeed,

$$Z = \frac{\hat{p}_{12} - \hat{p}_{21}}{\sqrt{\hat{p}(1 - \hat{p})\left(\dfrac{1}{n_{1x}} + \dfrac{1}{n_{x1}}\right)}} = \frac{\dfrac{n_{12}}{n_{1x}} - \dfrac{n_{21}}{n_{x1}}}{\sqrt{\dfrac{(n_{12} + n_{21})(n_{11} + n_{22})}{n^2}\left(\dfrac{1}{n_{1x}} + \dfrac{1}{n_{x1}}\right)}} \sim N(0, 1). \tag{4}$$

Because $Z^2 = \chi_1^2$, the test statistics described by (3) and (4) are equivalent.[10]

---

[9]Although the $\chi^2$ test formally holds only asymptotically, it is well-established that this approximation works very well if $(n_{12} + n_{21}) \geq 8$ (see Gibbons, 1997, among others).

[10]In the literature several modifications of the test statistics (3) and (4) have been suggested based on the arcsin transformation stabilizing the variance. However, it has been shown in Monte Carlo simulations that these arcsin transformations for sample sizes $\leq 100$ cannot be used in practice. For extensive simulation study, see Zvára (1995).