## Kazakhstan

**Sources:**
Atkinson and Micklewright 1992, Tables UE5, UE6, and UI3 from the Statistical Appendix
Alexeev and Gaddy 1993
Milanovic 1998
World Bank Poverty Monitoring Database 2002
Deiniger & Squire, World Bank 2004
Transmonee 2004
World Bank, World Development Indicators 2004

The documentation relies heavily on Atkinson and Micklewright (1992)

**Surveys:**
**Family Budget Survey 1988, 1989, 1990**
The Family Budget Survey, used by Atkinson and Micklewright (1992), Alexeev and Gaddy (1993) and Milanovic (1998), was a survey of families of persons employed in the state sector and of families of collective farmers (covering the whole USSR). Families have been selected by sampling individuals at their place of work. Sampling appears to have operated principally on a quota basis, with quotas for economic sectors, industrial branches and, within these, republics and oblasts. Beneath this level enterprises were selected on the basis of average age level and within each enterprise individuals were chosen according to their skill level and wage. Once included in the sample, an individual (and his or her family) was asked to participate until they left the enterprise or retired; retirement did not lead to automatic exclusion from the sample but was usually associated with a family dropping out from the survey. Replacement occurred only when a household dropped out from the survey. Participating households were monitored by the survey throughout the year. There were interviews with the whole household twice every month with diary records being maintained continuously. At inception in its post-war form in 1951, the sample size was about 51 000 families, in 1969, 62 000 and in 1988, 90 000. A family was defined as relatives who share a common budget. Approx. one third of the oblasts were not represented in the survey, high wage heavy industrial branches were over-represented in the selection of enterprises and less-skilled workers and those outside the direct production process were less likely to be selected. State farm workers seem also to have been under-represented but the rural population as a whole over-represented. When the sample size was expanded in 1988 it was specifically to correct problems of regional and branch representation.

For the above mentioned reasons the FBF sample was unrepresentative: families of those employed in the co-operative or private sectors and those not employed were in general excluded, old-age pensioners were heavily under-represented since they were originally excluded, the probability of selection was proportional to the number of working members since the sampling unit was the worker and, the panel nature of the survey biased the sample towards families of elder persons with long service records.

No figures of non-response are available but apparently the response rates were very high due to material and moral incentives to participate. Where non-response did occur, the household concerned was substituted with another household with similar observable characteristics.

When reporting results, figures have been adjusted for the over-sampling of collective farm families.

Annual gross family income was collected, including all money income from employment and from social security benefits. Reported cash income from sales of agricultural produce was included as was the value at state prices of agricultural production for self-consumption. The value of benefits in-kind from the employers such as meals and transport was included. Apparently benefits in-kind from the state are not included. Respondents' information on earnings and pensions were checked with employer records.

In Atkinson and Micklewright (1992) the reference period was a year but the data has been divided by twelve in the tables. The authors had to interpolate in order to arrive at values for the mean, median, and the Gini coefficient. To do this the authors used the program INEQ written by F.A. Cowell. Alexeev and Gaddy (1993) used a simple nonparametric technique based on the Kolmogorov-Smirnov test to fit the data to a lognormal distribution. They report that the estimates for 1990 are more reliable than those for 1988 as incomes were grouped only into 5 categories in 1988 but 7 in 1990.

**Earnings Surveys 1981, 1986, 1989**

Two sources of data have been used for earnings by Atkinson and Micklewright (1992). The first is Goskomstat 100% census of enterprises, held periodically since 1956, and from 1976 every five years (this is called the March Census by the authors). Enterprises were obliged to provide information on earnings for all their employees in a number of discrete earnings bands. The second source relates to data for 1984 and 1989. This is a survey of households of state sector and collective farm employees (this is called the March Household Survey by the authors). The information from this survey should be comparable with the first one. The March census covered only persons working in state enterprises, whereas those about 10% working on collective farms or private agricultural plots were excluded. Thos employed by "social organisations", usually taken to mean the Communist Party and its close affiliates, were also excluded. The earnings data used from the household survey 1989 have been selected to include only those employed in state enterprises and farms (i.e. excluding collective farm workers). Employees not working a full month were excluded and only first jobs were considered. Part-time workers were included but are excluded from the tables provided by the authors.

The earnings concept is gross earnings for the month in question. The monthly bonus and the monthly value of any quarterly bonuses were included but annual bonuses or any other rewards based on a period of more than three months were excluded. No account was taken of the value of income in kind provided by enterprises. The reference period

was gross earnings in March for both surveys. Atkinson and Micklewright (1992) had to interpolate in order to arrive at values for the mean, median, and the Gini coefficient. To do this the authors used the program INEQ written by F.A. Cowell.

**Family Budget Survey 1993**

This survey is used by Milanovic (1998). It is reported not to be representative.

**Living Standards Measurement Survey 1996**

This survey is used by Deininger & Squire (2004). It was counducted by the Sigma-institute in Berlin with technical assistance from the World Bank. The survey covered 2000 households and was carried out in July 1996. The income concept seems quite comprehensive and should be recorded net of taxes. A long list of items were asked for past month and include earnings, interests, dividends, sell of property, rents, a long list of transfers, alimony and help from relatives. In-kind components are included but imputed rent and apparently also home production are excluded (home production was asked about in the expenditure part but does not seem to be included in the income aggregate). Incomes are missing for approx. 11% of the households.

**Data from Transmonee**

Only a secondary source was used. The estimates from 1989 originate from Flemming & Micklewright (1999). They use the same methodology and survey as Atkinson and Micklewright (1992) but have apparently access to more detailed tables.