
The Middle Class in Tajikistan

Estimation, Dynamics, and Profile

Contents

I. Introduction	2
II. Previous Approach using TLSS Data	4
III. Updated Approach using HBS Data	5
IV. Data	6
V. Model	8
VI. Findings	10
VII. Profile	11
VIII. Conclusions	14

I. Introduction

Over the last decade, the countries in Central Asia have achieved remarkable progress in terms of economic growth and poverty reduction. Thanks to favorable energy prices the economic growth in oil exporting economies was high, spilling over to oil importing countries in the region. Between 2005 and 2015, poverty rates fell rapidly in Tajikistan, the Kyrgyz Republic, and Kazakhstan – the three countries with available micro-economic data, and similar pattern might be expected in Turkmenistan and Uzbekistan (two countries, where household data is not readily accessible). Growth in income from employment was the most important driver of poverty reduction. Generally, the benefits of economic growth were also widely shared. Consumption growth of the bottom forty percent was higher than the average growth rate in Kazakhstan and the Kyrgyz Republic, and since 2013, official estimates for Tajikistan also show improvements among population in the lower tail of distribution. The Gini coefficients for the Central Asian countries rank among the lowest in the world and have slightly declined in recent years.

Although, poverty has fallen throughout Central Asia, many people remain vulnerable to economic volatility. A large share of population has a consumption just above the relevant poverty line and at significant risk of falling below it. There is substantial churning in-and out of poverty among households, which points to low degree of resilience. In the context of increased economic volatility, country-wide or idiosyncratic shocks easily translate to reduction in welfare. From an economic perspective, exiting that state of vulnerability in the countries of the region could be viewed as a transition process to becoming the “middle class” society and highlights the process of reaching a threshold of economic stability associated with a low probability of falling back into poverty.

Country-led national development strategies increasingly focus on revamping growth models, emphasizing resilience, and building human capital as pillars of sustainable development. National strategies implicitly aim at expanding the middle class to harness its productivity, innovations, and growth potential. Despite growing interest in identifying the middle class in Central Asia, there has been little recent empirical work to define the middle class in the local context. The central policy question to promote shared prosperity is whether past poverty reduction episodes have been associated with expansion of middle class, or the observed welfare gains were only marginal/short-term, resulting in growing share of vulnerable population instead.

Previous work conducted by the World Bank on the middle class around the world, has underlined two important issues that are directly relevant to understanding the welfare dynamics in Central Asia. First, the definition of the middle class should be grounded in the local context. As much as poverty line is defined by national standards the definition of middle class would also need to be linked to national living standards. Second, building a thriving middle class requires building an environment conducive to improving productivity, enhancing resilience, and strengthening the agency of people. In other words, governments, policies and markets play vital role in shaping the process of becoming the middle-class society. Expanding the equitable access to finance, health, education and public services will improve the economic capacity and resilience of population, leading to growth in welfare across all income groups.

Context in Tajikistan

In Tajikistan, growing the middle class is a central aspiration of the Government, and one of the key commitments laid out in the country's National Development Strategy. The size of the middle class was most recently officially measured using the Tajikistan Living Standards Surveys 2007 and 2009. That year was the last time that the TLSS was conducted however, and from that point on, projections were used to estimate the size of the middle class.

This note extends the information available to policy makers in Tajikistan in two ways. First, it leverages data collected by the World Bank and other partners to estimate the size of the middle class using a definition and measurement approach that is consistent with the method adopted starting in 2009. This is an improvement on the existing projection estimates because it uses real data, which become increasingly imprecise over time and are vulnerable to mismeasurement. Second, this note lays out a method for defining and measuring the middle class going forward that is based on the Household Budget Survey (HBS). The HBS is collected by TajStat every quarter, and includes a full consumption module suitable to measuring the middle class. This note empirically estimates the threshold that separates the upper and lower middle classes from the poor and vulnerable groups using HBS data, and describes the process by which this exercise was completed. The note also profiles three economic groups: poor, vulnerable, and middle class. The profiles help to understand the characteristics and constraints faced by people in each category.

If adopted, the method of measuring and monitoring the middle class defined in this paper would have significant advantages over both the existing approach using projections (that are not based on observed data) and a related but alternative approach, which would require using infrequent surveys conducted by third parties (such as those used to extend the 2009 estimates reported in the following section). The advantages of instead using the HBS for these purposes is that it is a nationally owned data source, it is continuously collected, and it doubles as the source of the country's official poverty measure.

However, there are some drawbacks to using the HBS for these purposes as well. The design of the survey is such that the sample is concentrated in districts that participate in the survey, and due to the age of the sample (and the lack of documentation on the initial design) this can lead to some bias in the estimates. The questionnaire also uses an extended diary approach to gather information (with a one-month recall for many goods, such as food items), which experiments in Tajikistan have shown tends to lead to

undercounting of some types of consumption and expenditure. The HBS is also currently being updated by TajStat to address these and other considerations. This is an advantage in the sense that more and better data may be forthcoming, but also adds an additional question, because the method of measuring the middle class proposed here may need to be updated in the future if the survey instrument changes. These limitations should be considered in the discussion of whether adopting an approach based on the HBS is appropriate to the Government's purposes.

The note is organized in three sections: first, previous methods used to track the middle class in Tajikistan are presented. Section three describes the data, model, and the results of the estimation of middle class lines using the HBS. Section four provides profiles of the three groups: poor vs. vulnerable vs. the middle class, also looking at dynamics over time, along with remarks to frame the findings for policy focused discussion.

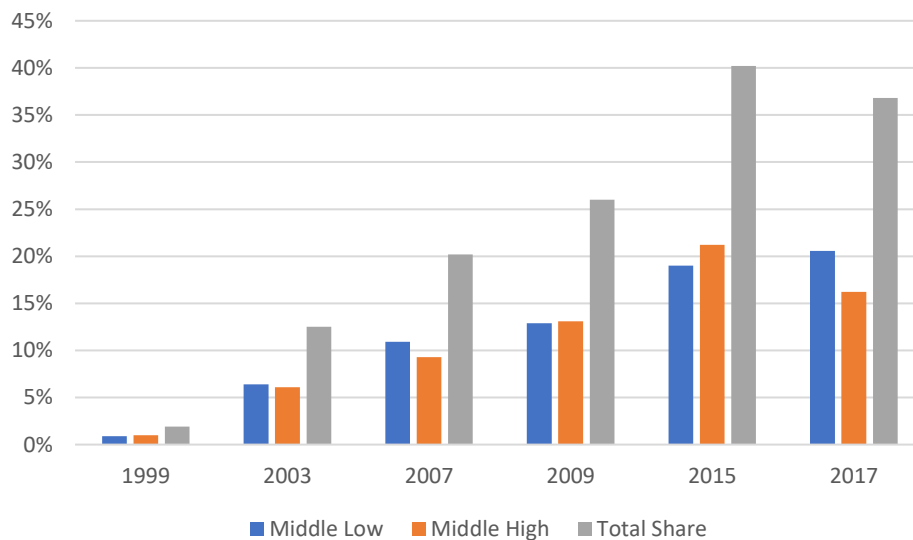
II. Previous Approach using TLSS Data

Previously, the middle-class lines for Tajikistan were estimated using data from the Tajikistan Living Standards Measurement Survey for 2007 and 2009. It is important to note that panel component of LSMS data was only a small subset of the whole sample, and available for only two years, which precluded updating the estimates from surveys not directly designed by Tajikistan's Statistical Agency.

Using the lines estimated in this effort and applying those to the most recent TLSS type surveys, the dynamics of middle class growth can be tracked. However, TajStat has not collected such surveys since 2009. But the World Bank and other partners have conducted two surveys with similar sampling and questionnaire designs that can be used for this purpose. The first was the Proxy Means Testing survey conducted in the spring of 2015 with 3000 households in 150 primary sampling units (PSUs). The second was conducted in late 2016/early 2017 with 3000 households for the purposes of monitoring Water, Hygiene, and Sanitation issues in Tajikistan. Each of these studies included a full consumption and expenditure model which allows for rough comparisons with the 2009 TLSS, after adjusting for consumer price changes using the official CPI.

Using these data, the overall trend is pointing to gradual expansion of middle class. However, it appears that the upper middle class reduced as a share of the population between 2015 and 2017. This may relate to the economic slowdown experienced in Tajikistan and the region in the wake the decline of oil prices in 2015, and the large reduction in remittance flows to Tajikistan that followed.

Figure 1: The Middle Class: TLSS-approach



Source: Staff estimates based on TLSS data of various years

III. Updated Approach using HBS Data

The approach adopted in section two has several limitations however. Due to the infrequent and contingent nature of third party surveys, such an approach cannot be reliably counted on to produce consistent estimates of the middle class in the future. This section described a method for defining the middle class that i) conforms to international best practice, and ii) uses data (the HBS) that are expected to be consistently available in the future.

There are many acceptable ways to define the middle class. Broadly, there are two distinct literatures arising from the disciplines of sociology and economics. In the economics tradition, the definition of the middle class focuses on finding the threshold(s) for inclusion based on either per capita income or the distribution of consumption expenditure. Accordingly, the middle class could be defined by:

- i) A relative threshold, that divides the income/consumption into deciles and considers, for example, the third and seven deciles as middle class or confining the middle class to values around the median of income (e.g. factor of 0.75 and 1.25 of median);
- ii) Self-reported perceptions of households as belonging to different economic classes, such as poor, lower middle class, upper-middle class, and rich;
- iii) An absolute threshold that defines the households as belonging the middle class based on: a) an international benchmark e.g. between \$10 to \$50 in international dollars per day per person in PPP terms, or b) estimating the threshold framing the method in the concept of economic security.

In the case of Tajikistan, the method grounded in the concept of economic vulnerability and security was adopted for two main reasons. First, the relative methods are somewhat arbitrary – there are no *apriori* factors that point that specific deciles could be considered middle class. Second, in the absence of suitable data on self-perceived status, the vulnerability approach is most appropriate in the local context, and allows estimation the threshold empirically, highlighting the economic security of households.

The motivating idea behind the concept of economic vulnerability is that households in the middle class need to be “free from poverty” and in that sense, are more economically secure and resilient. From an empirical perspective, the concept of vulnerability is defined in terms of the probability of a household falling into poverty. In this case, a standard threshold of less than ten percent probability of falling into poverty in the next period(s) was adopted. The key building blocks and steps to estimate the threshold based on the vulnerability approach are as follows:

Step 1 – Probability model defines vulnerability index. Based on panel data, predict household specific vulnerability to poverty based on set of independent variables X_{it} (which include location, demographic, education and other covariates associated with risk of poverty). If $poor_{i,t+1}$ describes vulnerability status, then:

$$p_{it} = E(poor_{i,t+1}|X_{it}) = F(X_{it}\beta_{it})$$

where, $poor_{i,t+1} = 1$ if poor for at least one period between $t=2$ and $t=T$

Step 2 – Consumption model isolates the structure of expenditure. If one predicts household expenditure on the same set of correlates X_{it} , and if Y_{it} is household per capita consumption in the initial period, then:

$$\ln Y_{it} = \alpha + X_{it}\gamma_{it} + \varepsilon_{it}$$

Step 3 – Identification of threshold for the middle class. Predicted expenditure for a fixed probability range, i.e. at /or below 10 percent of probability/risk of falling into poverty for upper middle-class line and at/below 30 percent risk for lower middle-class line.

IV. Data

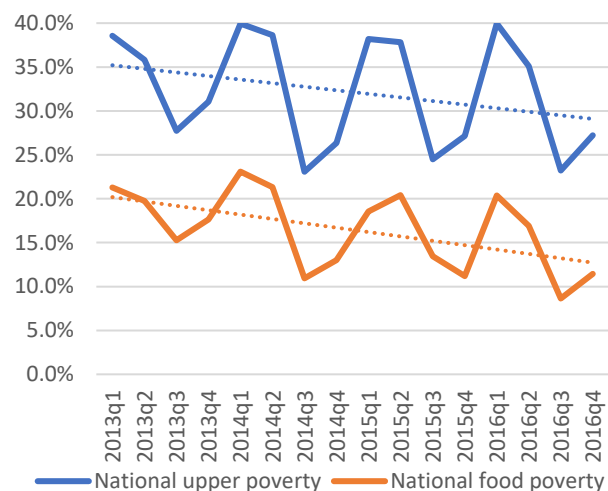
The national HBS is administered by Tajikistan’s National Statistics Office and used for official monitoring of poverty and the welfare of the population. A wide range of topics is collected on quarterly basis, ranging from demographics to agricultural production, for a sample of 3000 households. The Statistical Agency reports that few households were replaced over the last five years, and the HBS between 2013 and 2015 is used as a panel in the analyses that follow.

The World Bank’s technical assistance has supported the NSO on improving the HBS since 2014. This has resulted in gradual improvement of quality of data collected and analyses conducted using the survey microdata. However, the data have significant limitations due to an outdated sampling design and

questionnaire currently used. Consequently, the performance of the statistical models developed in section V have relatively low goodness-of-fit in comparison to gold-standard examples. However, as discussed above, one of the main benefits of using HBS versus other data (e.g. LSMS type surveys) is that the former is collected regularly and allows for constant updates and checking the stability of middle class lines. A large share of the HBS sample is designed as the panel, and thus readily amenable to estimation middle-class line based on vulnerability approach.

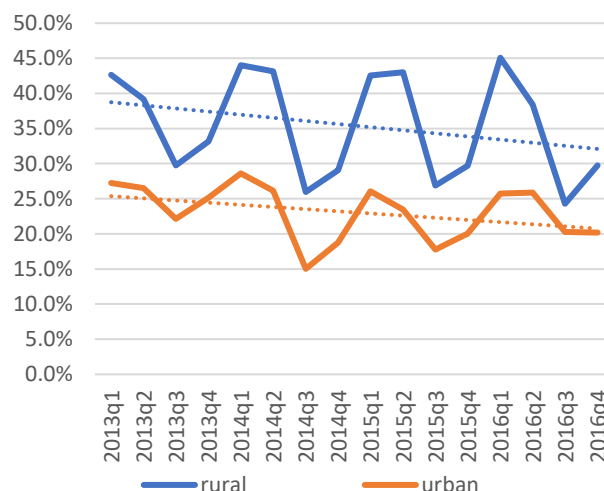
Although poverty rates have steadily fallen between 2013 and 2016, three out of ten people in Tajikistan are still considered poor. Poverty rates are highly seasonal, especially in rural areas where households are more depend on agricultural income. In addition, there are strong regional dimensions to poverty in Tajikistan: poverty is lowest in the capital of Dushanbe, and highest in Districts of Republican Subordination region.

Figure 2: National poverty rates, using upper and food poverty lines, by quarters (in percent to population)



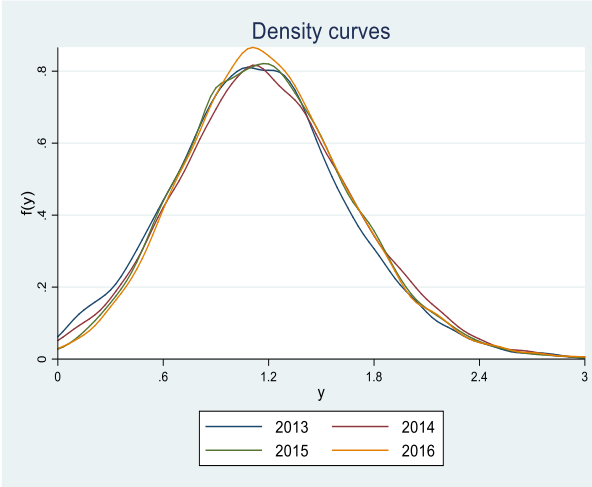
Source: Staff estimates based on TJK HBS data

Figure 3: National poverty rates, in urban and rural areas, by quarters (in percent to population)



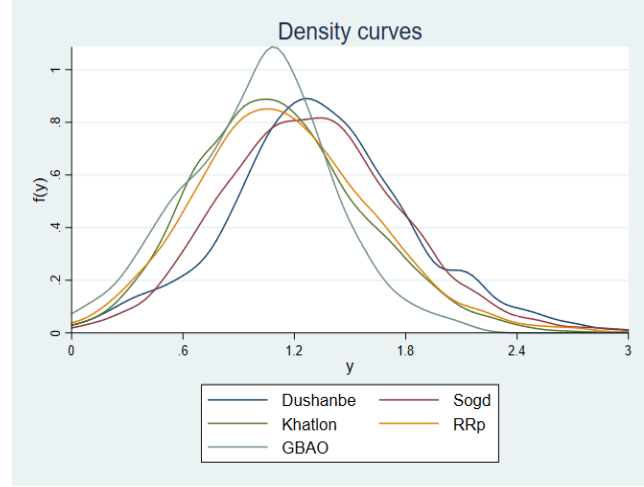
Note: Estimates based on re-calibrated population weights

Figure 4: Density curves of consumption per capita, per day, in logarithmic terms by years



Source: Staff estimates based on TJK HBS data

Figure 5: Density curves of consumption per capita, per day, in logarithmic terms by regions (2016)



Regardless of the season, the share of the poor who remain poor in the following period is high: two thirds of those who were poor in 2013 remained poor in 2015. The share of poor and non-poor who move between rankings, either fall into poverty or move out of poverty, is remarkably high. The high extent of churning points to low resilience to fluctuating economic conditions.

Table 1: Poverty transition matrix (by quarters), 2013-2015

Quarter 1	Non-poor	Poor	Quarter 2	Non-poor	Poor
Non-poor	76.13	23.87	Non-poor	74.5	25.5
Poor	38.95	61.05	Poor	40.08	59.92
Total	61.8	38.2	Total	62.17	37.83
Quarter 3	Non-poor	Poor	Quarter 4	Non-poor	Poor
Non-poor	85.61	14.39	Non-poor	84.72	15.28
Poor	49.24	50.76	Poor	46.52	53.48
Total	75.52	24.48	Total	72.85	27.15

Source: Staff estimates based on TJK HBS data

V. Model

Logistic and OLS models are estimated to derive the probabilities and predicted values of consumption corresponding to probabilities of falling into poverty. Both models include location, demographic, education, economic, and other variables related to living conditions of a household. These are generally variables which are correlated to poverty. It is important to note that that the objective of the models is not to directly explain the drivers of poverty, merely to explain its variation in a statistical sense, and thus

some degree of endogeneity and use of proxies is acceptable. But there is also some trade-off- inclusion of more variables result in better fit, but collinearity might bias the magnitude of coefficients. In the end, model includes parsimonious set of variables to avoid the bias, but covers important dimensions/correlates of poverty.

Table 2: Results of logistic and OLS regressions

Variables:	Logistic (Poverty status)		OLS (Consumption)	
	coef	se	coef	se
Sogd	0.143	-0.091	0.075***	-0.019
Khatlon	0.258***	-0.089	-0.046**	-0.019
RRS	0.145	-0.099	-0.069***	-0.023
GBAO	0.395***	-0.111	-0.150***	-0.023
HHH Single	0.255***	-0.078	-0.071***	-0.018
HHH Divorced	0.185**	-0.085	-0.039**	-0.018
HHH Secondary	0.475***	-0.067	-0.126***	-0.014
HHH Primary	0.517***	-0.163	-0.146***	-0.039
Urban	-0.017	-0.090	0.031*	-0.018
Household size	0.178***	-0.017	-0.041***	-0.004
Age of hh head	0.047***	-0.009	-0.012***	-0.002
Age hhh squared	-0.001***	0.000	0.001***	0.000
HH head male	0.074	-0.047	-0.000	-0.011
Change in hh size	0.121***	-0.012	0.001	-0.003
Change in # of children	-0.015	-0.020	0.011**	-0.004
Share of adults	-0.518**	-0.260	0.160***	-0.054
Number of dependents	0.049	-0.044	-0.008	-0.009
HH receives wage	0.038	-0.066	-0.045***	-0.014
HH receives agr. income	-0.988***	-0.088	0.275***	-0.017
HH receives soc. assist	-2.766**	-1.078	0.119	-0.093
HH receives remittance	0.225***	-0.050	-0.008	-0.011
HH with sewerage	0.522***	-0.184	-0.134***	-0.038
HH with hot water	-0.205	-0.355	0.236*	-0.123
HH with bath/showers	-0.673***	-0.171	0.163***	-0.036
Share of employed	-1.275***	-0.181	0.544***	-0.035
Constant	-2.301***	-0.315	5.666***	-0.068
Observations	11,904		11,904	
R-squared	0.110		0.254	

*** p<0.01, ** p<0.05, * p<0.1

Source: Staff estimates based on TJK HBS data

Generally, the direction of coefficients in the logistic regression are of expected sign, except for a small number of cases, such as access to sewerage and if household receives the wage for instance, but the sign becomes consistent in OLS regression.

VI. Findings

While there is significant noise in the estimation results, the correlation between probability of falling into poverty and consumption is strong. The results show that the relationship is not always linear – especially at the lower levels of probabilities. Around the mark of 10-15 percent risk of falling into poverty, the predicted consumption rises significantly.

Figure 6: Scatter plot of log of actual consumption and estimated probability of falling into poverty, 2013

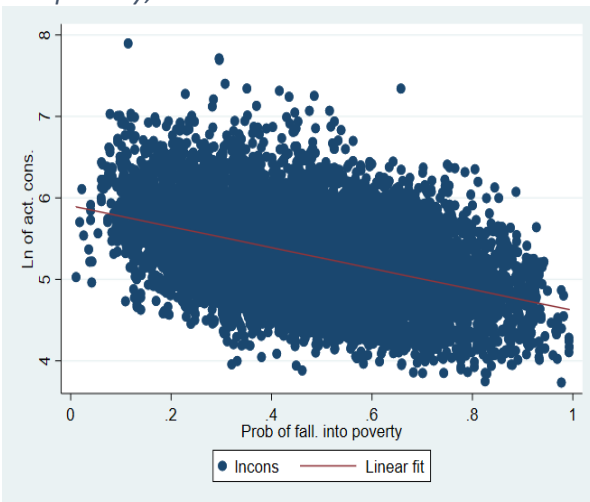
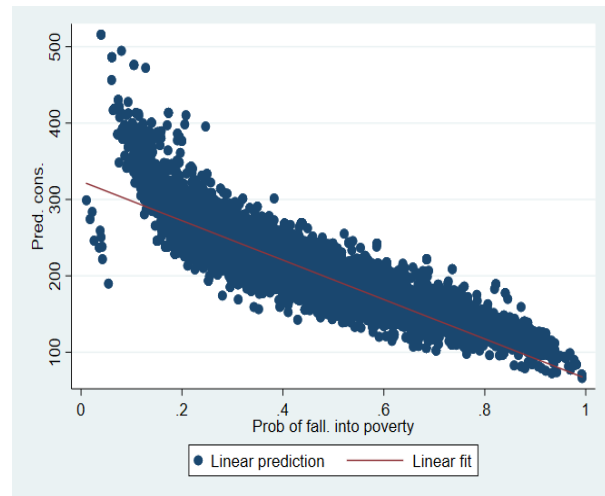


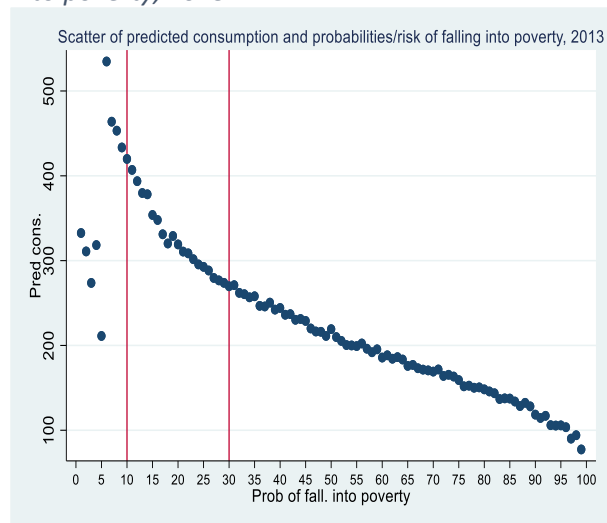
Figure 7: Scatter plot of predicted consumption and estimated probability of falling into poverty, 2013



Source: Staff estimates based on TJK HBS data

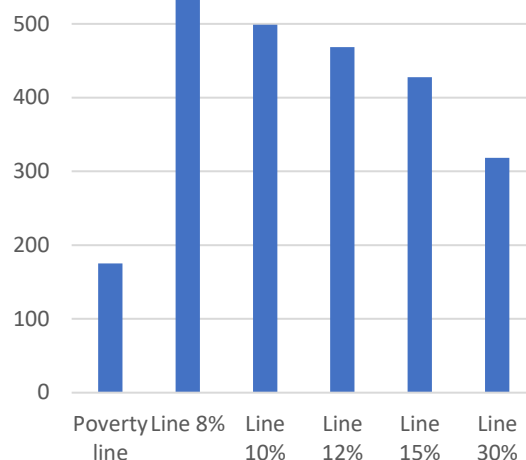
Based on the model, the threshold value that differentiates between the vulnerable and lower middle class is estimated at 318.35 TJK Somoni (TKS) per person per month, and for upper middle class at 498.71 TKS in 2016. The threshold value is 1.8 and 2.8 times higher than the value of poverty line respectively, which generally conforms to similar ratios in other countries. In absolute terms, the upper middle-class cut-off line for Tajikistan is only slightly lower than similarly estimated value of line for Kyrgyzstan, where the poverty line is also slightly higher, in comparable terms. As expected, the values of middle class lines are sensitive to the probability threshold level and the changes in poverty line. However, different specifications of the model, extended vs. parsimonious, result in similar values of the thresholds.

Figure 8: Scatter plot of predicted and estimated consumption for specific levels of risk to falling into poverty, 2013



Source: Staff estimates based on TJK HBS data

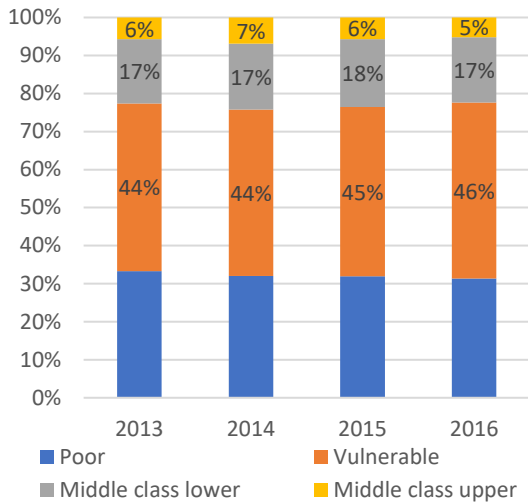
Figure 9: Values of cut-off lines at selected levels of probabilities (in LCU per month, 2016)



VII. Profile

At the estimated middle-class lines, the middle class is estimated to have slightly expanded through 2014, followed by a slight decline in 2015-2016. This indicates that the trend of declining *poverty* over recent years was not accompanied by an expanding middle class. The decline in poverty rates over the periods were essentially matched by an increasing share of vulnerable population. The share of the population who is not poor, but not yet middle class, is growing. The shares of lower and upper middle classes remained around 17 and 6 percent respectively of the total population, and like poverty, are also subject to strong seasonality. The findings point to fragility of poverty reduction episodes and substantiate the descriptive statistics suggesting a high degree of churning.

Figure 10: The trend and size of the middle class in Tajikistan, (2013-2016) based on HBS data (in percent to total population)



Source: Staff estimates based on TJK HBS data

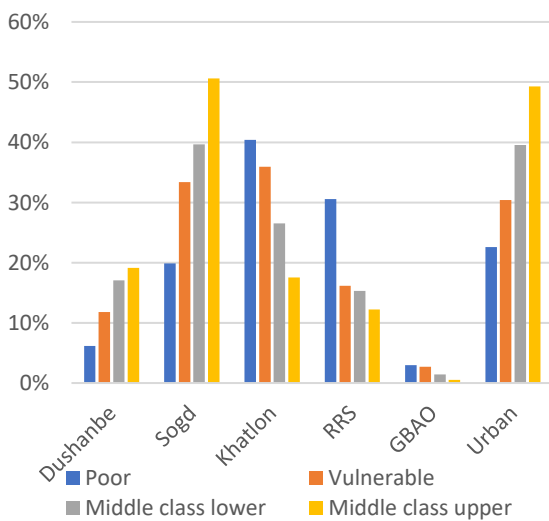
Figure 11: The trend and size of the middle class in Tajikistan, (2013-2016, by quarters) based on HBS data (in percent to total population)



Note: Estimates based on re-calibrated population weights

As with poverty, there is pronounced demographic and spatial pattern to the middle class. Interestingly, about half of all those who belong to middle class come from Sogd oblast, with Dushanbe and Khatlon accounting for another two-thirds. About half of the middle class comes from urban areas, while cities and towns account for only about one third of population. Middle class households are substantially smaller: they have fewer children, fewer female members, and smaller share of seniors.

Figure 12: Distribution of economic classes across regions, 2013 (in percent to total population in each economic group)



Source: Staff estimates based on TJK HBS data

Figure 13: Average number of household members, number of children and number of females across economic classes, 2013



As expected, household heads in middle class households are more educated, and are especially more likely to hold a tertiary degree. Also related to education, the share of adult members receiving wage in the middle-class households is significantly higher than vulnerable or poor households. The living conditions of middle class households are much better compared to other two economic classes; middle class has higher access to improved water, sewerage, bathrooms and telephone connection. Certainly, this infrastructure is correlated to urban location of middle class. However, there are also some amenities, like gas, district heat and hot water that do not substantially vary with economic class status. However, in Tajikistan, these amenities are not always functioning or available where one might expect to find them, due to the unavailability of gas for heating following the breakup of the regional network after independence.

Figure 14: Average shares of household heads with tertiary education; members above 65 years and members receiving the wage across economic classes, 2013

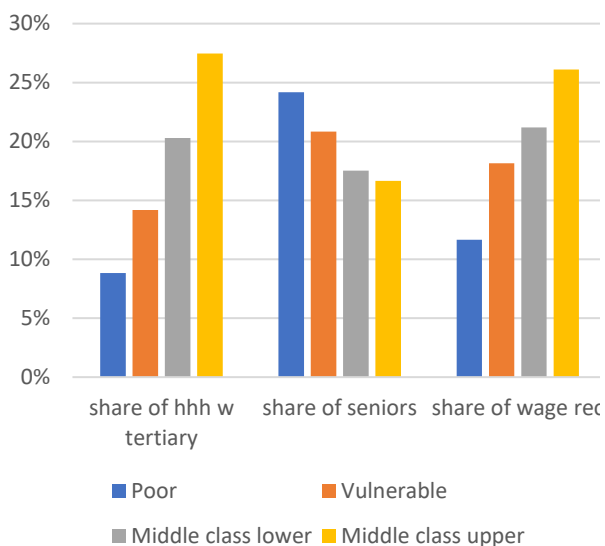
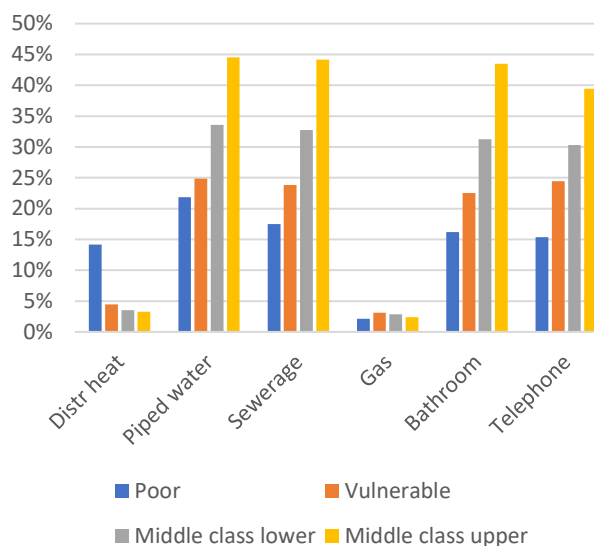


Figure 15: Average shares of households with access to district heat; piped water; sewerage; piped gas; bathroom and landline- telephone across economic classes, 2013



Source: Staff estimates based on TJK HBS data

Like in the case of poverty, there is pronounced seasonal churning in and out of the middle class. A small share of population is in the middle-class status in all four quarters of the year. The largest share of population staying persistently in the one specific class is vulnerable – those not poor, but not yet middle class.

VIII. Conclusions

Tajikistan currently uses a projection approach for estimating the size of the middle class, and set related objectives for the National Development Strategy. This note includes updates, the estimates from this approach leveraging data collected by the World Bank and other partners to estimate the size of the middle class using a consistent definition. The note also lays out a method for defining and measuring the middle class going forward that is based on HBS, while discussing some of the advantages and disadvantages of the various alternatives. The analysis then estimates the thresholds that separate the upper and lower middle classes from the poor and vulnerable groups using HBS data, and describes the process by which this exercise was completed.

The results of the estimation strategy highlight that although Tajikistan has grown its middle class over the long term, recent years have seen little progress. This is in contrast of poverty rates have slowly but steadily declined in recent years. The share of the population categorized as 'vulnerable' expanded, pointing to a high degree of vulnerability and low resilience of households to changing economic conditions. Generally, policies that tackle both poverty and vulnerability would lead to expanding middle class. As was found in the Systematic Country Diagnostic for Tajikistan, the development of human capital and promoting the private sector could lead to both higher stock and better returns to human capital and eventually form a basis for middle class growth.