IN BRIEF

PROFOR

E. PERGE, R. BEHAL APR 2018

FOREST-SWIFT— HIGH FREQUENCY DATA COLLECTION IN FORESTS

Poverty estimates are essential to track the progress/impact of programs and policies that target poverty itself. However, while forest areas are home to the poor, very little is known about poverty rates in these areas and the role forest resources could play to reduce poverty. To fill this information gap, Forest-SWIFT has been designed to measure (i) poverty incidence and (ii) forest dependence using a country-specific econometric model and questionnaire. Forest-SWIFT was piloted in forest villages in Turkey in 2017 using recently collected baseline data on forest income and activities. The pilot verified that Forest-SWIFT is cost-efficient, quick, and robust, but can also face certain challenges in implementation.

Why is it important to measure poverty in forest areas?

Poverty reduction features amongst the goals of the largest multilateral and bilateral development institutions such as the World Bank, the UN, the IMF, and even tops the list of the Sustainable Development Goals (SDGs). Even though measuring poverty is essential to track the poor and to measure the impact of policies and programs, poverty monitoring remains infrequent. Collecting poverty data can be time consuming and expensive especially when including representative samples of forest households, which are difficult to access.

BOX 1: DEFINITIONS AND MEASUREMENTS OF INDICATORS

Poverty Incidence	measured as the proportion of population living below per capita national or international poverty line. The per capita welfare aggregate includes food and non-food consumption, measured using either expenditure or income, depending on country definitions.
Forest Dependence	measured as the proportion of income derived from forest related activities. Forest related income includes sales from forest products, wages from forest activities, and payments for environmental services.

To circumvent the limitation of accessibility, poverty within forests has been assessed using either satellite imagery [1] [2] or case studies, resulting in metrics that are often not comparable to national measures of welfare. The particularities of forest households and their livelihoods often remain unacknowledged in the design and implementation of national policies for poverty reduction [3].

There is a clear need for an inexpensive and robust tool that can increase the frequency of data collection to track poverty and forest-dependence. Forest-SWIFT has been developed by PROFOR to serve this purpose, encouraging researchers, practitioners, and project managers to collect more and better evidence of poverty and forest dependence.

What is Forest-SWIFT?

Forest-SWIFT is a high-frequency data collection method developed to provide timely, quick, and accurate data on poverty and forest dependence. Forest-SWIFT is based on an existing high-frequency data collection tool called the Survey of Well-being via Instant Frequent Tracking (SWIFT). SWIFT is a methodology developed by the World Bank to track poverty between two rounds of comprehensive household surveys (LSMS-type surveys) [4][5]. As an extension to SWIFT, Forest-SWIFT estimates forest dependence in addition to poverty, using country-specific models for each indicator.

Forest-SWIFT uses baseline household data from surveys such as LSMS survey and the Forestry Modules, which include information on consumption and forest income, to identify the household characteristics that most explain these variables [6]. Both models assume a linear relationship between consumption/forest income y_h and their correlates x_h with a projection error u_h.

$$\ln y_h = a + \sum_{(k=1)}^{K} \beta_k * x_{kh} + u_h$$

Forest-SWIFT includes multiple steps to improve the ability of the formulas to estimate household income/consumption and forest income. Forest-SWIFT controls for issues linked to over-fitting – when a model performs well within the sample but poorly outside the dataset – by cross-validating the model [7]. The purpose of cross-validation in Forest-SWIFT is to identify the optimal level of significance in the model, which would balance the number of determinants and the goodness of fit across the sample. Cross-validation consists of two steps: (a) splitting the sample in n-folds and running the model in n-1 folds and testing it on the n^{th} fold, (b) running multiple models per fold, testing various thresholds of significance for model variables. This process entails adding variables to the model sequentially if they bring enough information, and simultaneously removing them if they do not. The optimal threshold for including variables in the final model is chosen by comparing the goodness of fit across all models generated per sub-sample. The optimal threshold is then applied to the whole sample, and the final model is generated using variables that meet the established criteria.

To ensure the quality and robustness of the models, Forest-SWIFT carries out two tests (if data are available): backward imputation and validity test. The former applies the final model to a previous round of data to check the stability of the model over time. The latter tests whether the error term follows a normal distribution using a simulation method developed by Elbers, Lanjouw and Lanjouw (2002, 2003).

After running all these tests, we have a set of determinants for welfare and forest income. These determinants are compiled into a short country-specific questionnaire. Data are collected using computer-assisted personal interviews (CAPI) application. The data are finally used to predict for forest-dependence and poverty.

How to implement Forest-SWIFT?

In this note, we describe the application of Forest-SWIFT in forest villages in Turkey. The availability of baseline datasets for both, poverty and forest dependence, made Turkey the ideal location for a pilot. Poverty data were available in the 2013 Household Budget Survey, which were subsequently standardized by the ECAPOV team at the World Bank.

Step 1 establishes comparability between both data sources i.e. compares the characteristics of rural population in the Household Budget Survey (HBS) 2013 to the forest dwelling population in the Socioeconomic Household Survey (SEHS) 2016.

We found that forest households are larger than rural households, but have fewer dependents (children and elders). Heads of forest households are more likely to be male and more educated. However, labor force participation of household heads and prime-aged adults is lower **(Table 1)**.

TABLE 1. COMPARING HBS 2013 AND SEHS 2016 DATA

	HBS 2013 (rural)	SEHS 2016
Household size	3.88	4.52
Dependence ratio	0.607	0.532
HEAD CHARACTERISTICS		
Age	52.08	53.35
Male	87.33%	96.48%
No school	19.70%	9.06%
Primary school	80.30%	90.06%
Employed	71.81%	65.48%
PRIME AGED ADULTS' CHARACTERISTICS		
No school	18.49%	9.85%
Primary school	81.51%	90.04%
Employed	61.01%	49.60%
Neither student, nor employed (15-29 yrs)	29.45%	21.16%
Unemployed	2.74%	48.89%
Labor force participation	61.01%	56.71%
Female labor force participation	38.67%	26.30%

Source: authors' computation using SEHS 2016 and HBS 2013. Weights applied. Note: All statistics are at the HH level. [a] SEHS 2016 collected this data at the community level

Step 2 constructs the consumption model using HBS 2013 data. The poverty rate among the HBS households was estimated at 35%, using a poverty line of $7 (in terms of 2011 PPP). The SWIFT model, which used the log of per capita consumption as the dependent variable, and a significance threshold of 0.005, selected 14 necessary variables out of a pool of 23, that would best estimate poverty for that population. SWIFT purposefully estimates the log transformation of the dependent variable to smooth asymmetries and normalize the distribution of the variable, making it easier to estimate.

Step 3 builds the forest-income model using SEHS 2016 data. The model estimates the log of per capita net forest income, and a forest-dependence headcount, using median per capita forest income as the threshold in place of a poverty line. The model only includes households who report non-zero, positive, forest-related income from any source, which represents around 60% of the original sample. The model had to be revised after the data collection since recent government interventions compromised the integrity of certain variables (e.g. freezers which were given for free before the 2017 referendum). The final model has a set of 25 explanatory variables, of which 10 were selected using a p-value 0.01.

What are the results from Forest-SWIFT?

Forest-SWIFT is ultimately a customized tool to estimate aggregates for poverty and forest-dependence, which results in a short questionnaire that only collects data on the necessary variables identified by the final models. The final questionnaire included a household roster and 20 questions on forest collection, wages, dwellings and assets. We conducted the survey in 100 out of the 202 villages surveyed in SEHS 2016. The collected data form an unbalanced panel dataset since we only re-surveyed 1000 households from the original sample. The sample is still representative of national forest villages, with households being self-weighted. The survey took place over a three-week period, during which enumerators spent less than 20 minutes with each household.

Running the consumption and forest-income models with the current survey data yielded estimates for the current survey population. Predicted consumption using Forest-SWIFT 2017 data is higher than consumption in 2013 (**Table 2**); poverty incidence falls to 23% when using a poverty line of $7 (in terms of 2011 PPP).

TABLE 2. POVERTY RATE AND CONSUMPTION PER CAPITA IN RURAL AND FOREST AREAS (2013 & 2017)

	HBS 2013 (original)	SWIFT 2017
Poverty rate (%)	34.9	23.2
Mean per capita Consumption	5,906	6,442

Source: authors' estimations using HBS and Forest-SWIFT data. Note: Consumption values in HBS 2013 and in Forest-SWIFT 2017 are all in Turkish Lira 2013

The average per capita forest-income was TL 3,407 in 2017 using 2016 prices. The sharp increase from the 2016 estimates is attributed to higher response rates in 2017. The estimates were not drastically compromised since 50.5% of forest households were still below median forest income **(Table 3)**.

TABLE 3. FOREST INCOME AND RATIO BELOW MEDIAN FOREST INCOME (2016 & 2017)

	SEHS 2016 (original)	Forest-SWFT 2017
Mean log per capita forest income	4.23	4.56
Mean per capita forest income	893	1,217
Ratio Below median forest income	0.5	0.455

Suggested formula for inverse natural log = Exp (m+ sigma (m)^2/2) Note: Forest incomes with SEHS 2016 and Forest-SWIFT 2017 are in Turkish Lira 2016

What are Forest-SWIFT limitations and advantages?

The first pilot of Forest-SWIFT in Turkey proved to have a steep learning curve. Developing Forest-SWIFT using the forest income data from Forestry Modules emphasized the importance of a clear definition for forest products and a clear questionnaire design. During model development, we found a number of agricultural products incorrectly classified as forest products, which lengthened the list of questions in the baseline about individual product use. These questions were asked for a total of 90 products. The low participation rate in forest product extraction is potentially linked to respondents' fatigue and low willingness-to-answer truthfully about activities involving each forest product. Better specification and selective questioning could result in more thorough reporting. The 2017 Forest-SWIFT survey, which included a more limited list of 9 forest products and fewer questions per product, recorded higher participation rates for forest product extractions.

In its initial phase, Forest-SWIFT is limited to estimating non-zero positive forest income since the original SWIFT model was developed to measure poverty through consumption and total income. While it is quite rare that households have zero or negative consumption or total income, households who participate in forest-related activities can have zero or negative net income because of higher costs or lower market sales. To improve accuracy, Forest-SWIFT estimates the log transformation of net positive forest income, since log transformations of zero and negative numbers do not exist. Consequently, we had to restrict our sample to households with positive forest-income, and can only obtain non-zero positive estimates for the SWIFT data.

These limitations do not undermine the advantages of Forest-SWIFT, or the importance of having frequent data on poverty and forest dependence. The SWIFT poverty rate is plausible when comparing it to Turkey's national poverty rate for 2016. Collecting such poverty data was completed over a short period of time and at a very low cost compared to traditional household survey data. In addition, the forest income data clearly identify the importance of forest products in the livelihoods of the poor, even though households with higher forest dependence have lower incomes.

Thanks to this Forest-SWIFT pilot, the resulting data for forest households can be used to explore more the characteristics of forest households, forest dependence and its causes, and channels through which forests can contribute to poverty reduction.

References

[1] Sunderlin, William D., Arild Angelsen, and Sven Wunder. 2004. "Forests and Poverty Alleviation." Bogor, Indonesia: Center for International Forestry Research.

[2] Sunderlin, William D., Sonya Dewi, Atie Puntodewo, Daniel Muller, Arild Angelsen, and Michael Epprecht. 2008. "Why Forests Are Important for Global Poverty Alleviation: A Spatial Explanation." *Ecology and Society 13 (2): 1–21.*

[3] Oksanen, Tapani, Brita Pajari, and Tomi Tuomasjukka. 2003. *Forests in Poverty Reduction Strategies: An Assessment of PRSP Processes in Sub-Saharan Africa*. Vol. EFI Proceedings. 47 vols. European Forest Institute. http://www.efi.int/files/attachments/publications/proc47_net.pdf.

[4] Ahmed F., C. Dorji, S. Takamatsu, and N. Yoshida. 2014. "Hybrid Survey to Improve the Reliability of Poverty Statistics in a Cost-Effective Manner" World Bank Policy and Research Paper 6909. World Bank, Washington D.C.

[5] Yoshida, N., R. Munoz, A. Skinner, C. Kyung-eun Lee, M. Brataj, W. Durbin and D. Sharma. 2015. "SWIFT Data Collection Guidelines version 2." World Bank, Washington D.C.

[6] Bakkegaard, R.Y., N.J. Hogarth, I.W. Bong, A.S. Bosselmann, and S. Wunder. 2017. "Measuring Forest and Wild Product Contributions to Household Welfare: Testing a Scalable Household Survey Instrument in Indonesia." *Forest Policy and Economics 84: 20–28.*

[7] Kuhn, Max and KjellJohnson. 2013. "Applied predictive modeling" Springer.

The Program on Forests (PROFOR) multi-donor partnership generates innovative, cutting-edge knowledge and tools to advance sustainable management of forests for poverty reduction, economic growth, climate mitigation and adaptation, and conservation benefits. Through its programs, PROFOR is advancing forest-smart development, which recognizes forests' significance for sustaining growth across many sectors, including agriculture, energy, infrastructure, and water.

UKaid from the British people

For further information, contact Emilie Perge, Economist, Poverty GP; eperge@worldbank.org