# README FILE FOR: Consumption behaviours and poverty traps: A case study of detailed diary data from rural poor smallholder households in Uganda

## Daniel Gregg, David Just, and Daniel Hill

## Last update: 24 May 2024

## Repository ID: [openicpsr-203761](https://www.openicpsr.org/openicpsr/workspace?goToPath=/openicpsr/203761&goToLevel=project)

## Manuscript ID: XXX

This replication package contains the data, code and supplementary materials for the manuscript titled *Consumption behaviours and poverty traps: A case study of detailed diary data from rural poor smallholder households in Uganda.* The files and folders contained in this package include:

* **Sticky\_consumption\_EXP\_INC\_MAIN.R** reproduces all the main results for the manuscript. The script performs all the necessary steps for the analysis, calling necessary functions and data saved in the data and utilities folder and saving results in the results folder. The script can be executed when through the statistical program R, with the dependent packages downloaded.
* **Sticky\_consumption\_EXP\_INC\_LOW.R** replicates the analysis conducted in the MAIN script detailed above, but for a ‘low’ performance expected income model. This is a robustness check to ensure the results are stable for different specifications of expected and unexpected income, and the results are presented in the appendix and are saved in the result folder with \_LOW.
* **Sticky\_consumption\_EXP\_INC\_HIGH.R** replicates the analysis conducted in the MAIN script detailed above, but for a ‘HIGH’ performance expected income model. =
* **Data**  contains all the necessary data for the analysis. More details are provided below under the data availability section.
* **Utilities** contains all the necessary functions to run the analysis scripts. More details are provided below in the utilities section.
* **Results** contains all results printed by the main analysis scripts. These results are also presented in the manuscript for the main model, as well as the appendix for the high and low expected income models. More details are provided below.
* **Diary survey example.docx** provides an example of the diary tool used to collect data on income and consumption on a weekly basis. More details on this are provided in the Data availability section.
* **Appendix.pdf** provides supplementary results for the manuscript, including high and low specification of expected (and unexpected income) and more details and results for the wealth simulations.

## Data availability

Within the data folder, two data files are available and allow for the paper replication:

* ***Data.Rda*** – This is a de-identified version of the diary data collected for each individual, for a total of 442 individuals. *Data.Rda* contains the respondent ID (*SURVEY\_ID),* a week identifier *(WEEK),* and detailed income and expenses for the individual in the given week. .*RDA* files can be opened by the R statistical program.
* ***bg\_data.Rda*** – This is a de-identified version of the background information collected for each individual, for a total of 442 individuals. Relevant for this study and included in the replication repository, *bg\_data.Rda* contains village, parish, and subcounty information, as well as respondent IDS (*SURVEY\_ID)* to map this information to the diary data. Village, parishes, and sub-county information were renamed to deidentify the data. .*RDA* files can be opened by the R statistical program.

The data used in this paper were obtained from a high-frequency diary which recorded detailed weekly income and consumption at a household level for 442 households across 24 weeks. The sample was selected randomly from a list of households involved in a program focused on coffee value chain research in the Kapchorwa region of eastern Uganda. The original list we obtained this sample from was itself generated as a stratified random sample of 1,600 households engaged in selling coffee within the Kapchorwa region.

Each household was trained on the use of the diary tool which involved tallying income and consumption for each week. An example of this diary is provided in this replication repository. The diary was designed using images of both expenditure categories/income sources and monetary values with entries recorded as a tally allowing participants with low literacy/numeracy skills to easily participate. Participating households were regularly visited (multiple times each week) and were asked to record earnings and consumption at the end of each day. This took households typically less than 5 minutes each day.

The timing of the diary collection began at the end of the coffee harvest in November 2018 until April 2019. This period coincides with the end of the coffee harvesting season (running until December), where on farm labor opportunities are higher, and the beginning of a period of lower demand for labor with harvest of seasonal crops falling outside of the collection period. The majority of households are dependent on labor income where opportunities are often sporadic and unpredictable, and from regularly drawing down on savings accumulated from the coffee harvest period. The timing and study population of this survey is therefore well suited for the analysis of behavioral consumption responses from income changes where, for most households, changes in income are transitory and we expect that there are relatively low levels of planned income in comparison to cohorts in more developed areas or in comparison to periods of high labor demand.

This project was part of a broader project (*Developing Value Chain Innovation Platforms to Improve Food Security in East and Southern Africa (VIP4FS)*), approved by the University of Adelaide Human Research Ethics Committee (Approval number: H-2016-251, starting from 1 November 2016 to 31 July 2019)

## Computing requirements

The main analysis scripts are executable in the statistical program R, version 4.2.0. Please refer to the R documentation to see the relevant requirements for your computer and your organisation.

* <https://cran.rstudio.com/bin/windows/base/>
* <https://cran.rstudio.com/bin/macosx/>
* <https://cran.rstudio.com/bin/linux/>

Thes dependent packages for the analysis should be installed before executing the code, and include:

* *here -* <https://cran.r-project.org/web/packages/here/vignettes/here.html>. Please refer to the documentation of *here* especially to ensure the working directories are set up correctly for your needs.
* *plm -* [*https://cran.r-project.org/web/packages/plm/index.html*](https://cran.r-project.org/web/packages/plm/index.html)
* *ggplot2 -* [*https://cran.r-project.org/web/packages/ggplot2/index.html*](https://cran.r-project.org/web/packages/ggplot2/index.html)

## Replication steps

For replication of the manuscript *Consumption behaviours and poverty traps: A case study of detailed diary data from rural poor smallholder households in Uganda:*

1. Install R and the dependent packages for the analysis script.
2. Ensure the here() package is set up according to the documentation such that it calls the correct working directory.
3. Run the Sticky\_consumption\_EXP\_INC\_MAIN.R script.
4. Check the results saved in the results folder.
5. For the robustness checks on the expected income estimation, run the Sticky\_consumption\_EXP\_INC\_LOW.R and Sticky\_consumption\_EXP\_INC\_HIGH.Rscripts.

## List of utilities

The utilities folder contains necessary functions for the analysis. All functions are documented but more details are provided below:

|  |  |
| --- | --- |
| **Function** | **Description** |
| All\_simulation.R | Runs the asset dynamic simulations combining permanent income hypothesis, rule of thumb, and keeping up with the Joneses. |
| asset\_dynamics.R | Generates asset dynamic charts from the simulation results |
| estimateModels.R | Estimates all models after establishing the analysis dataframe, and saves the outputs for packaging as csv files. |
| expected\_income\_linear\_high.R | Estimates the expected and unexpected income series. The middle function is called within the Sticky\_consumption\_EXP\_INC\_HIGH.R script. |
| expected\_income\_linear\_low.R | Estimates the expected and unexpected income series. The middle function is called within the Sticky\_consumption\_EXP\_INC\_LOW.R script. |
| expected\_income\_linear\_middle.R | Estimates the expected and unexpected income series. The middle function is called within the Sticky\_consumption\_EXP\_INC\_MAIN.R script. |
| get\_outmat\_ROT\_08\_12\_20.R | Packages the Rule of Thumb model outputs for a .csv file output. |
| get\_outmat\_HABIT\_08\_12\_20.R | Packages the Habit formation model outputs for a .csv file output. |
| get\_outmat\_JONES\_08\_12\_20.R | Packages the Keeping up with the Joneses model outputs for a .csv file output. |
| getDWforPGM.R | Manual Durbin Watson tests for mean group estimators. |
| getFSTATforPGM.R | Manual F statistic tests for mean group estimators. |
| getplmfitted.R | Manual predicted and residuals series calculator for Mean group estimators. |
| KUJ\_simulation.R | Runs the asset dynamic simulations combining permanent income hypothesis with credit constraints and keeping up with the Joneses. |
| PIH\_credit\_constraints.R | Runs the asset dynamic simulations combining permanent income hypothesis with credit constraints. |
| ROT\_simulation.R | Runs the asset dynamic simulations combining permanent income hypothesis and rule of thumb. |
| runSimulations.R | Main script to execute asset dynamic simulations. Parameterises all behavioural coefficients, income, asset, and consumption series |

## List of results outputs.

The utilities folder contains necessary functions for the analysis. All functions are documented but more details are provided below:

|  |  |
| --- | --- |
| **Results file** | **Description** |
| “causality results\_HIGH.csv" | Results output for the causality tests with the 'high' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "causality results\_LOW.csv" | Results output for the causality tests with the 'low' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "causality results\_MAIN.csv" | Results output for the causality tests with the central model for expected and unexpected income. Presented in the manuscript. |
| "data summary.csv" | Summary of main data series |
| "Habits results full sample\_HIGH.csv" | Habit formation model results with the 'high' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "Habits results full sample\_LOW.csv" | Habit formation model results with the 'low' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "Habits results full sample\_MAIN.csv" | Habit formation model results with the central model for expected and unexpected income. Presented in the manuscript |
| "income persistence results\_HIGH.csv" | Results output for the tests of income persistence for the 'high' performance specification of expected and unexpected income. Presented in the appendix of the manuscript |
| "income persistence results\_LOW.csv" | Results output for the tests of income persistence for the 'low' performance specification of expected and unexpected income. Presented in the appendix of the manuscript |
| "income persistence results\_MAIN.csv" | Results output for the tests of income persistence with the central model for expected and unexpected income. Presented in the manuscript. |
| "Jones results full sample\_HIGH.csv" | Keeping up with the Joneses model results with the 'high' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "Jones results full sample\_LOW.csv" | Keeping up with the Joneses model results with the 'low' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "Jones results full sample\_MAIN.csv" | Keeping up with the Joneses model results with the central model for expected and unexpected income. Presented in the manuscript |
| "predicted income regression\_HIGH\_PREDICTION\_POWER.csv" | Regression results determining expected and unexpected income series. This is the high-performance model with more explanatory variables, which assumes more income is 'expected' and less income (the residual) is unexpected |
| "predicted income regression\_LOW\_PREDICTION\_POWER.csv" | Regression results determining expected and unexpected income series. This is the low performance model with fewer explanatory variables, which assumes less income is 'expected' and more income (the residual) is unexpected |
| "predicted income regression\_MIDDLE\_PREDICTION\_POWER.csv" | Regression results determining expected and unexpected income series. This is the central estimate used in the manuscript. |
| "RoT results full sample\_HIGH.csv" | Rule of Thumb model results with the 'high' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "RoT results full sample\_LOW.csv" | Rule of Thumb model results with the 'low' performance specification for expected and unexpected income. Presented in the appendix of the manuscript |
| "RoT results full sample\_MAIN.csv" | Rule of Thumb model results with the central model for expected and unexpected income. Presented in the manuscript |
| "Simulation\_KUJ with asymmetry.jpg" | Simulation model results - keeping up with the joneses assuming asymmetry |
| "Simulation\_KUJ with symmetry.jpg" | Simulation model results - keeping up with the joneses assuming symmetry |
| "Simulation\_PIH with credit constraints.jpg" | Simulation model results for permanent income hypothesis with credit constraints |
| "Simulation\_ROT and KUJ with asymmetry.jpg" | Simulation model results - rule of thumb and keeping up with the joneses assuming symmetry |
| "Simulation\_ROT and KUJ with symmetry.jpg" | Simulation model results - rule of thumb and keeping up with the joneses assuming asymmetry |
| "Simulation\_ROT with asymmetry.jpg" | Simulation model results - rule of thumb assuming symmetry |
| "Simulation\_ROT with symmetry.jpg" | Simulation model results - rule of thumb assuming asymmetry |

## Variable names and descriptions

### Data.Rda

|  |  |
| --- | --- |
| **Variable** | **Description** |
| "SURVEY\_ID | Respondent ID |
| "WEEK\_COUNT" | Number of weeks observed for respondent |
| "DATE\_FROM" | Date income and consumption observations collected from |
| "DATE\_TO" | Date income and consumption observations collected to |
| "WEEK" | Week identifier |
| "MONEY\_EARNT\_PICKING" | Money earned from picking coffee cherries in the given week (in UGX) |
| "MONEY\_EARNT\_COFF\_SALES" | Money earned from selling coffee cherries in the given week (in UGX) |
| "MONEY\_EARNT\_PARCHMENT" | Money earned from selling coffee parchment in the given week (in UGX) |
| "MONEY\_EARNT\_OTHER\_WORK" | Money earned from other businesses or waged work in the given week (in UGX) |
| "MONEY\_EARNT\_GOVT\_NGO" | Money earned from government sources or NGOs in the given week (in UGX) |
| "MONEY\_EARNT\_PARTNER" | Money provided to individual from their partner in the given week (in UGX) |
| "MONEY\_EARNT\_FRIEND" | Money provided to individual from a friend in the given week (in UGX) |
| "MONEY\_EARNT\_SAVINGS" | Money retrieved from savings in the given week (in UGX) |
| "MONEY\_EARNT\_LOANS" | Money earned from loan repayments back to individual in the given week (in UGX) |
| "GAMBLING" | Money earned from gambling in the given week (in UGX) |
| "MONEY\_WENT\_PARTNER" | Money that was given to individual's partner |
| "MONEY\_WENT\_EXPENSES" | Money that was spent on any expenses |
| "MONEY\_WENT\_SAVINGS" | Money that went to savings |
| "MONEY\_WENT\_STOLEN" | Money that was stolen |
| "MONEY\_WENT\_FRIENDS" | Money given to friends |
| "MONEY\_WENT\_LOAN" | Money that was used to repay loans |
| "EXPENSES\_FOOD" | Expenditure on food items in the given week (in UGX) |
| "EXPENSES\_FARM\_TOOLS" | Expenditure on food tools in the given week (in UGX) |
| "EXPENSES\_FARM\_INPUTS" | Expenditure on food inputs in the given week (in UGX) |
| "EXPENSES\_HOME\_CONSUMABLES" | Expenditure on home consumables (i.e. light bulbs, paper, etc) in the given week (in UGX) |
| "EXPENSES\_EDUCATION" | Expenditure on education (school fees, materials, uniform) in the given week (in UGX) |
| "EXPENSES\_ALCOHOL\_CIGARETTES | Expenditure on alcohol and cigarettes in the given week (in UGX) |
| "EXPENSES\_HOUSELOAN\_RENT" | Expenditure on mortgages or rent in the given week (in UGX) |
| "EXPENSES\_GAMBLING" | Expenditure on gambling in the given week (in UGX) |
| "EXPENSES\_TRANSPORT" | Expenditure on transport in the given week (in UGX) |
| "EXPENSES\_OTHER\_LOAN" | Expenditure on other loans in the given week (in UGX) |
| "EXPENSES\_OTHER" | Other expenditure in the given week (in UGX) |

### Bg\_data.Rda

|  |  |
| --- | --- |
| **Variable** | **Description** |
| ID | ID assigned for bg\_data frame |
| ID\_IN\_SURVEY | ID assigned in diary dataframe to map background data to diary observations |
| VILLAGE | Village of respondent. De-identified. |
| PARISH | Parish of respondent. De-identified |
| SUBCOUNTY | Subcounty of respondent. De-identified |
| NUM\_IN\_HOUSEHOLD | Number of individuals within the respondent’s household |
| VAL\_ASSETS\_SAVINGS | Value of savings available to the individual (in UGX) |
| VAL\_ASSETS\_BODA | Value of Bodas (motorbikes) available to the individual (in UGX) |
| VAL\_ASSETS\_BICYCLE | Value of bicycles available to the individual (in UGX) |
| VAL\_ASSETS\_TV | Value of TVs available to the individual (in UGX) |
| VAL\_ASSETS\_SMARTPHONE | Value of smartphones available to the individual (in UGX) |