TVSEP Data User Workshop

Internet use and non-farm employment

Presenter: MSc. Nguyet Tran
Institute for Environmental Economics and World Trade
Leibniz University Hannover

Overview

I. Introduction to the topic, data use, econometric models, results

II. Preparing data sets

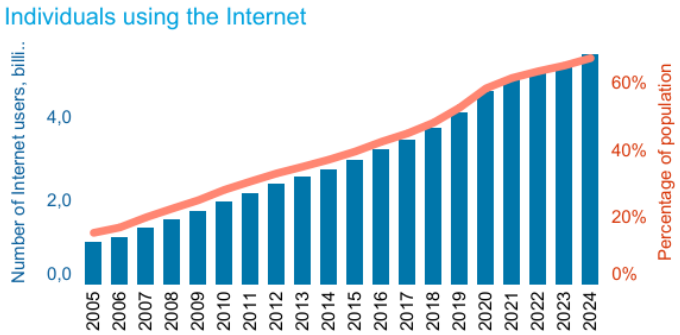
III. Regression

I. Introduction

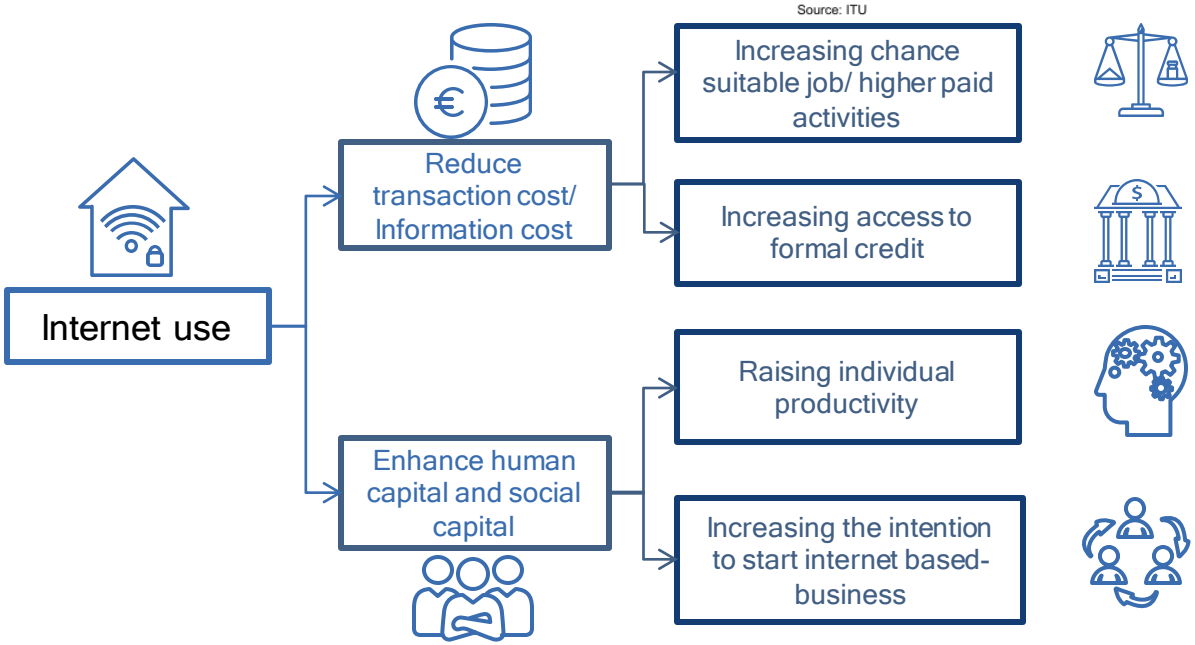
- **Promoting non-farm sectors is considered a key driver of ecomomic growth and poverty alleviation (UN, 2021)**
  - Providing income, food security, reducing poverty and promoting welfare (Hoang et al., 2014, 2001; Do et al., 2022)
- **Challenges for rural developing countries**
  - Rural members have access to **low-paid and unstable** non-farm activities (Brünjes & Diez, 2016)
  - **Information asymmetries** are prevalent in the labor markets in rural developing countries (Rajkhowa and Qaim, 2022)

Introduction

- **One potential channel is to expand internet access and use**
- 68 % (5.3 billion) people use internet globally (International Telecomunication Union, 2025)



Source: ITU

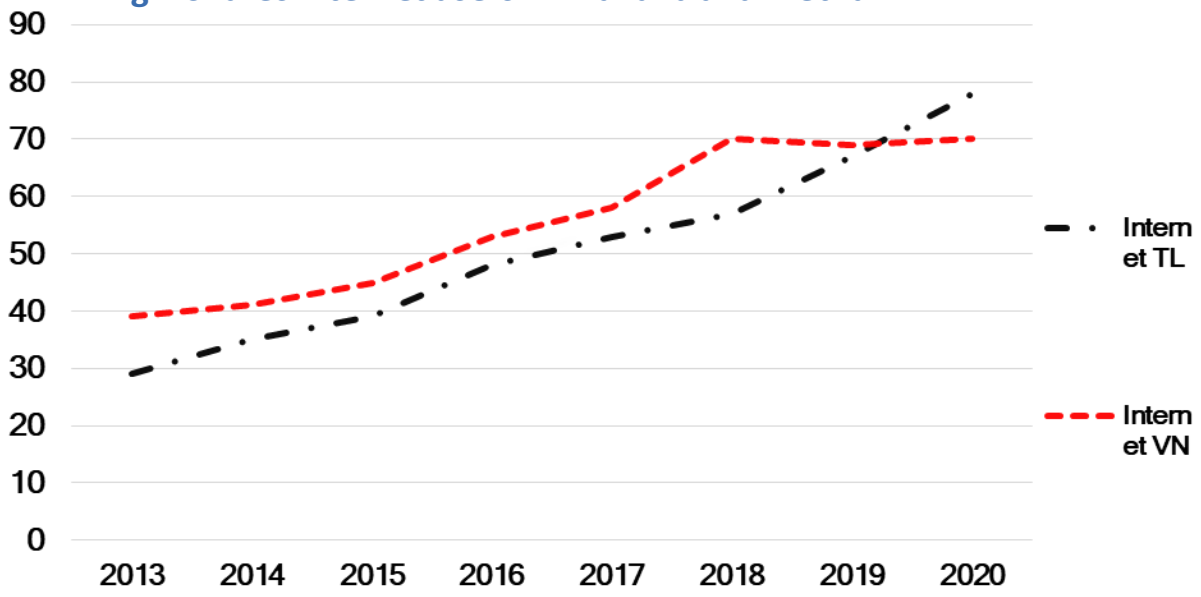


Sources: Self-formulation from Galperin&Viecens, 2017; Tan&Li, 2022

Why Thailand and Vietnam

- Internet access and use increases rapidly (WB data, 2023)
- Non-farm sectors are expanding (Amare et al., 2023)
- Studies on the effects of internet use mainly focus on agriculture (Nguyen et al., 2021; Nguyen et al., 2023; Kaila & Tarp, 2019)

Fg.1 Shares internet users in Thailand and Vietnam



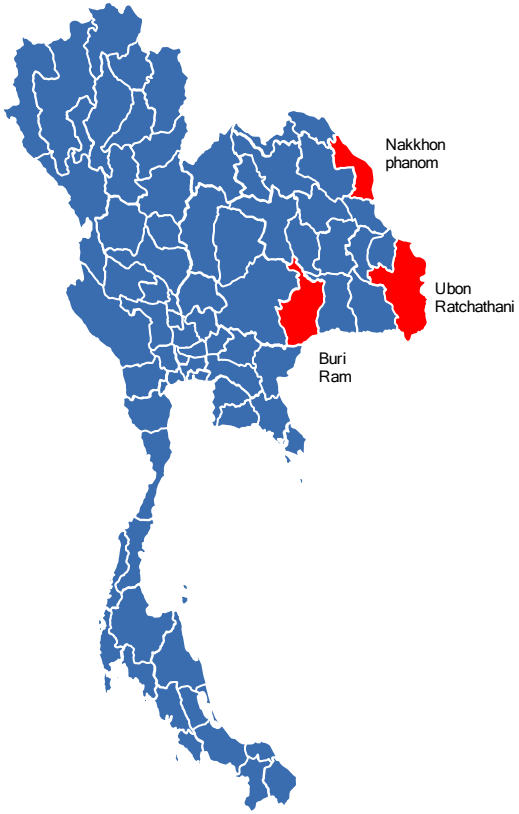
Source: World Bank Data, 2023

Objectives

- 1. To investigate the effects of internet use on non-farm employment
- 2. To examine the heterogeneity in the impact of internet on different population groups

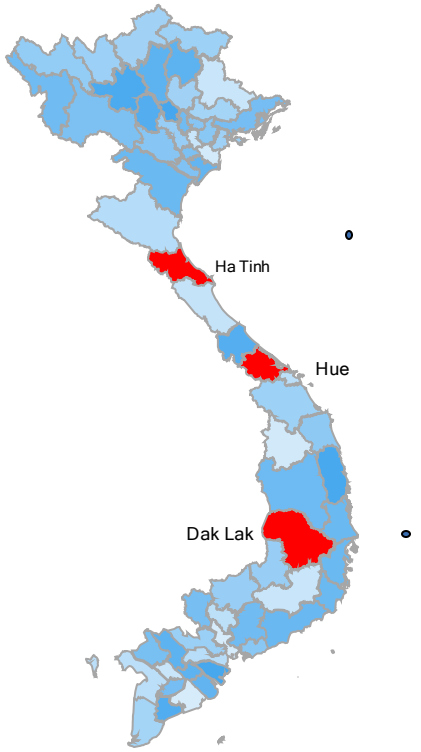
Data

Thailand



- Thailand Vietnam Social Economic Panel (TVSEP) covers about **4400 households in 440 villages** in **Thailand** and **Vietnam**
- This study employs **2 waves of 2016 and 2017 with 6044 observations**

Vietnam



Emprirical Strategies

$$NF_{ij} = \alpha_0 + \alpha_1 I_{ij} + \alpha_2 H_{ij} + \alpha_3 L_j + \alpha_4 P + \varepsilon_{ij} \qquad (1)$$

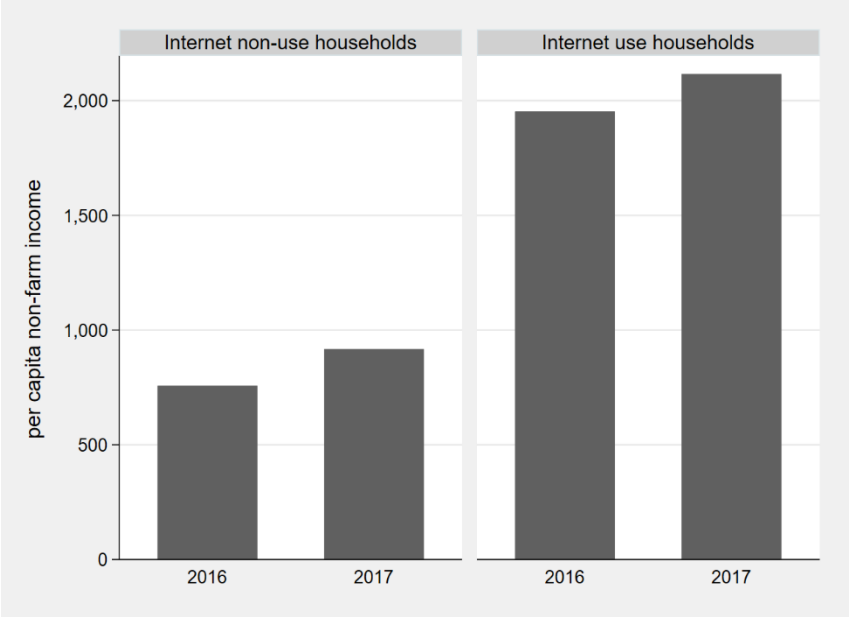
$$NF_{ij} = \beta_0 + \beta_1 I_{ij} + \beta_2 Inter + \beta_3 H_{ij} + \beta_4 L_j + \beta_5 P + \theta_{ij} \qquad (2)$$

- $\mathbf{NF_i}$: **outcome variables** (probability of NF participation, NF per capita income) of household *i* in village j
- $\mathbf{I_i}$: 1 if household uses internet for working purposes→ **endogenous variable** *(Nguyen et al., 2023)*
- $\boldsymbol{Inter}$: **interaction term** between internet use and some other variables
- $\boldsymbol{H_i}$: household control variables ; $\boldsymbol{L_j}$: local characteristics
- $\boldsymbol{P}$ and $\boldsymbol{C}$ are provincial dummies
- $\boldsymbol{\varepsilon_{ij}}$ *and* $\boldsymbol{\theta_{ij}}$ are error terms

→ Method **Heteroscedasticity-based Instrumental Variable** helps to generate **internal instrument variable** *(Lewbel., 2012; Baum & Lewbel, 2019)*

Data descriptives

Fg 2. Per capita non-farm income over years and internet use status



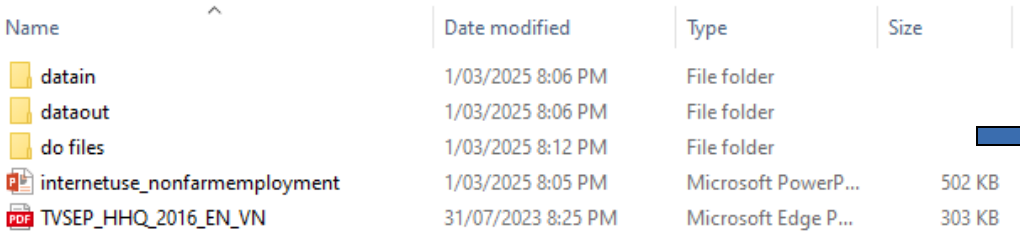
Source: own calculation based on TVSEP data

II. Preparing the data sets

- Follow the dropbox link:

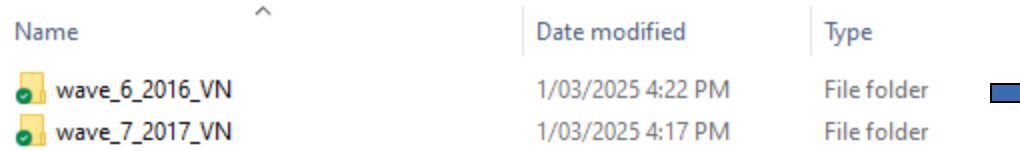
https://www.dropbox.com/scl/fo/dn6liq63c2p9ynuk3loyd/ANLUg8pdnUi3xvFn5mMcVqI?rlkey=m5nakgv7jihhx2ft1ldv5j14k&st=rk67i5pg&dl=0

Instruction for files



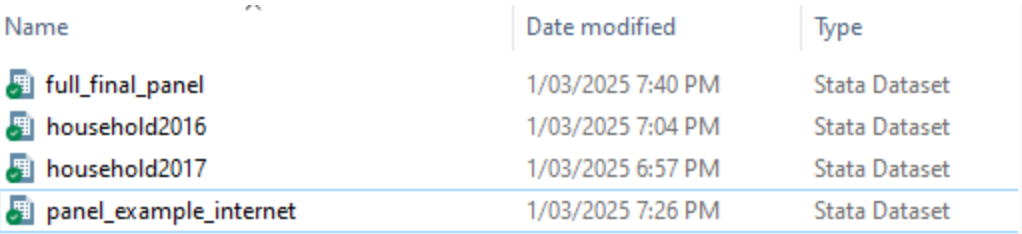
Name	Date modified	Type	Size
datain	1/03/2025 8:06 PM	File folder	
dataout	1/03/2025 8:06 PM	File folder	
do files	1/03/2025 8:12 PM	File folder	
internetuse_nonfarmemployment	1/03/2025 8:05 PM	Microsoft PowerP...	502 KB
TVSEP_HHQ_2016_EN_VN	31/07/2023 8:25 PM	Microsoft Edge P...	303 KB

Main Folders
"internet&non-farm"



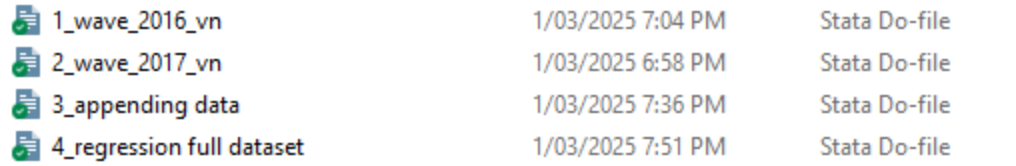
Name	Date modified	Type
wave_6_2016_VN	1/03/2025 4:22 PM	File folder
wave_7_2017_VN	1/03/2025 4:17 PM	File folder

Folder of "datain" contains original data (.dta)



Name	Date modified	Type
full_final_panel	1/03/2025 7:40 PM	Stata Dataset
household2016	1/03/2025 7:04 PM	Stata Dataset
household2017	1/03/2025 6:57 PM	Stata Dataset
panel_example_internet	1/03/2025 7:26 PM	Stata Dataset

Folder of "dataout" contains all generated data files (.dta)

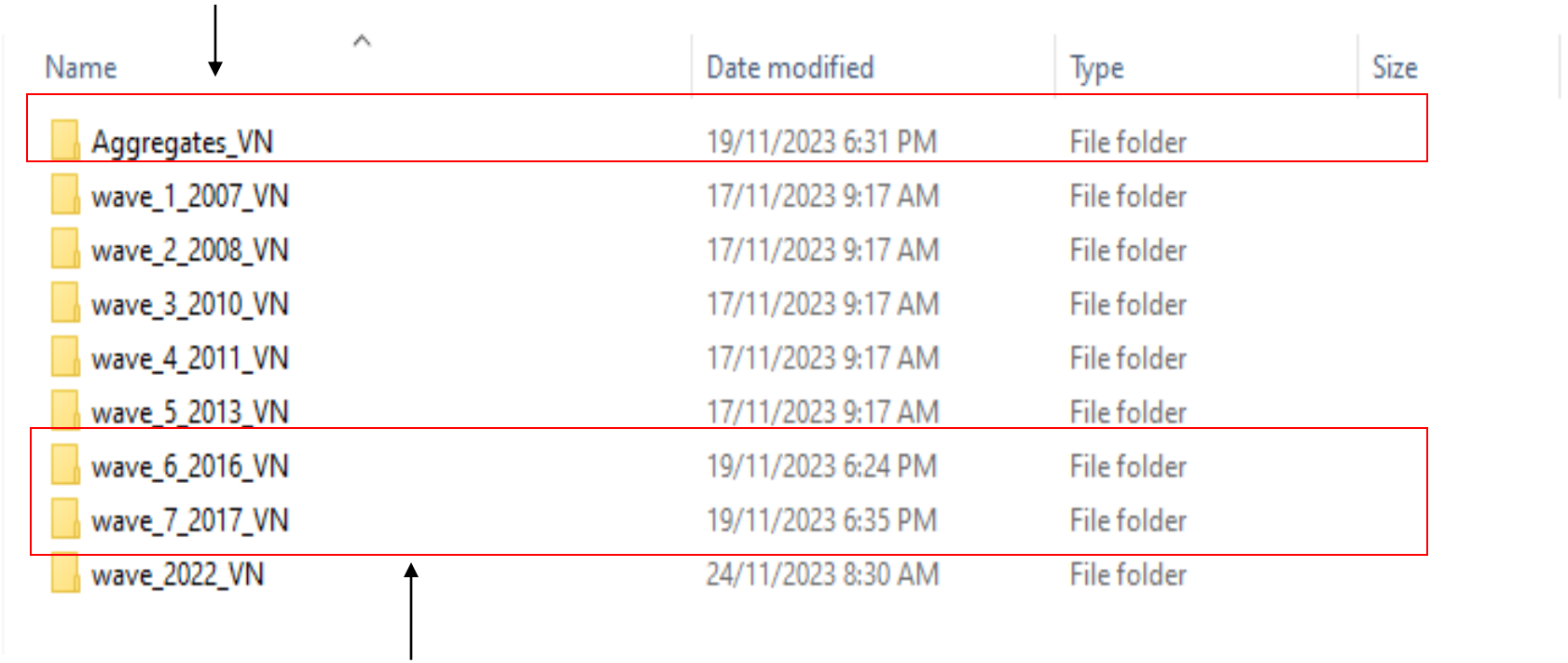


1_wave_2016_vn	1/03/2025 7:04 PM	Stata Do-file
2_wave_2017_vn	1/03/2025 6:58 PM	Stata Do-file
3_appending data	1/03/2025 7:36 PM	Stata Do-file
4_regression full dataset	1/03/2025 7:51 PM	Stata Do-file

Folder of
"do files"

Data folders

Step 1: Determining data folders

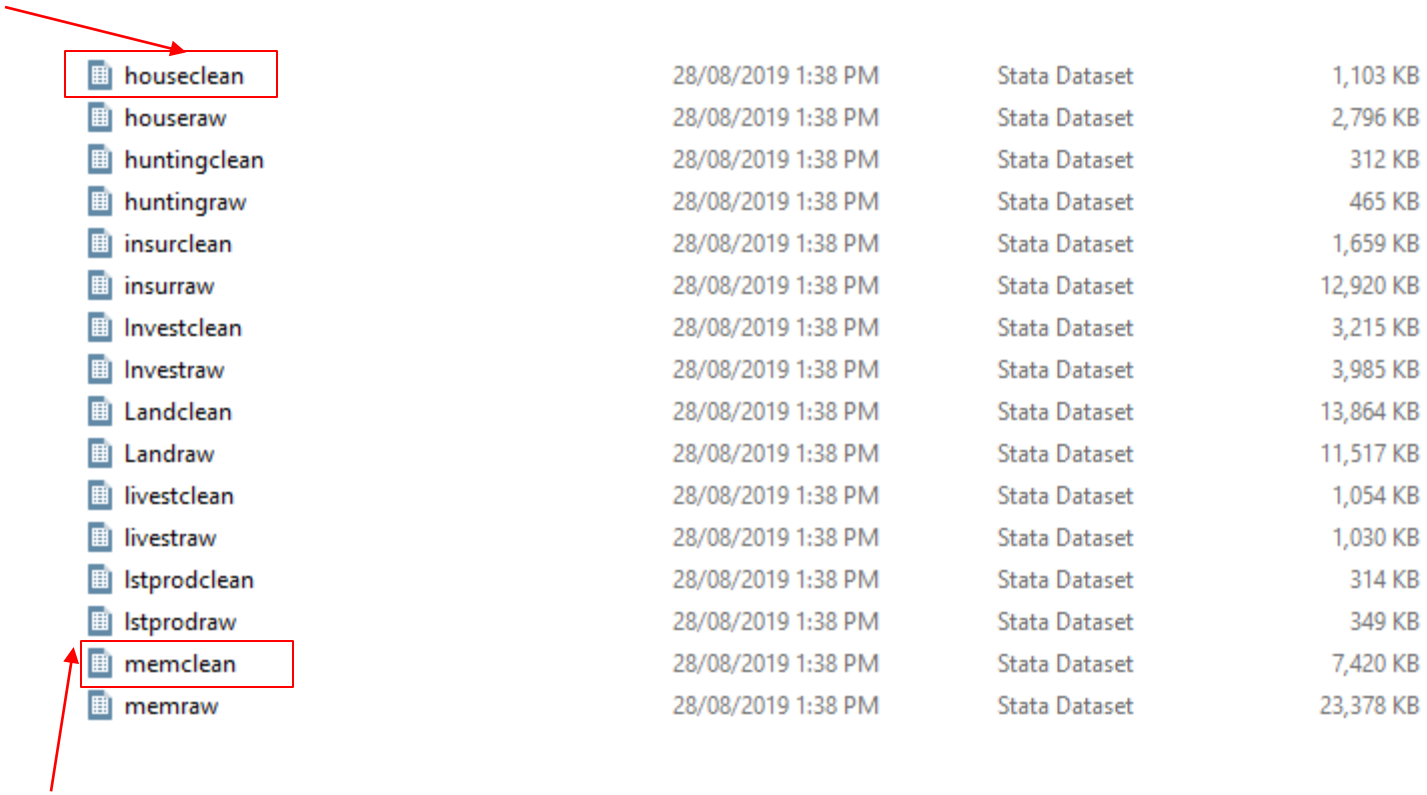


Name	Date modified	Type	Size
Aggregates_VN	19/11/2023 6:31 PM	File folder	
wave_1_2007_VN	17/11/2023 9:17 AM	File folder	
wave_2_2008_VN	17/11/2023 9:17 AM	File folder	
wave_3_2010_VN	17/11/2023 9:17 AM	File folder	
wave_4_2011_VN	17/11/2023 9:17 AM	File folder	
wave_5_2013_VN	17/11/2023 9:17 AM	File folder	
wave_6_2016_VN	19/11/2023 6:24 PM	File folder	
wave_7_2017_VN	19/11/2023 6:35 PM	File folder	
wave_2022_VN	24/11/2023 8:30 AM	File folder	

These folders were sent when you apply for tvsep data on the homepage

Extracting variables

Step 1.1. Locating dta file that contain our needed variables



houseclean	28/08/2019 1:38 PM	Stata Dataset	1,103 KB
houseraw	28/08/2019 1:38 PM	Stata Dataset	2,796 KB
huntingclean	28/08/2019 1:38 PM	Stata Dataset	312 KB
huntingraw	28/08/2019 1:38 PM	Stata Dataset	465 KB
insurclean	28/08/2019 1:38 PM	Stata Dataset	1,659 KB
insurraw	28/08/2019 1:38 PM	Stata Dataset	12,920 KB
Investclean	28/08/2019 1:38 PM	Stata Dataset	3,215 KB
Investraw	28/08/2019 1:38 PM	Stata Dataset	3,985 KB
Landclean	28/08/2019 1:38 PM	Stata Dataset	13,864 KB
Landraw	28/08/2019 1:38 PM	Stata Dataset	11,517 KB
livestclean	28/08/2019 1:38 PM	Stata Dataset	1,054 KB
livestraw	28/08/2019 1:38 PM	Stata Dataset	1,030 KB
lstprodclean	28/08/2019 1:38 PM	Stata Dataset	314 KB
lstprodraw	28/08/2019 1:38 PM	Stata Dataset	349 KB
memclean	28/08/2019 1:38 PM	Stata Dataset	7,420 KB
memraw	28/08/2019 1:38 PM	Stata Dataset	23,378 KB

Data folder name „datain“

Step 2. Coding variables

Section 9.2: Housing conditions

19	When your household uses the internet what is the major device used?		Code KK
20	If your household has a smartphone, what is the internet on the smartphone mainly used for?	a	Code LL
		b	Code LL
21	If household has other internet access than the smartphone, what is this access mainly used for?	a	Code LL
		b	Code LL

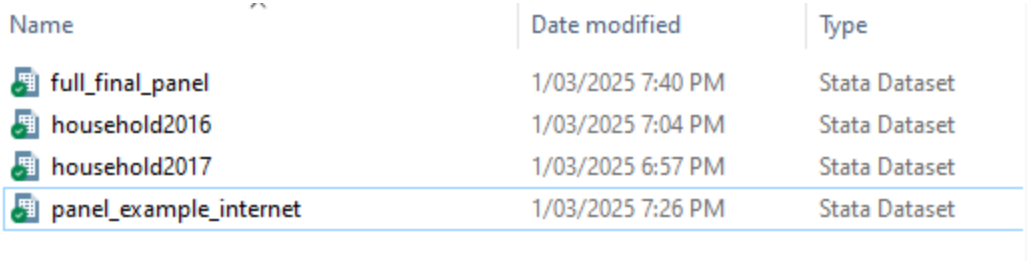
Code KK

1 smartphone
2 computer (PC and/or laptop)
3 tablet
4 device in internet café /shop
5 does not apply (no access)
90 others, please specify

Code LL

1 carry out financial transactions
2 used for entertainment
3 contact family member via email or messenger app
4 contact friends via email or messenger app
5 contact business partmers via email or messenger app
6 find information about job opportunities
7 offer information about jobs
8 search for information about the weather
9 find information about things to buy
10 offer information about own things to sell
11 receive medical/pharmaceutical information
12 learning or studying
13 Used for trading activities (buy-sell)
14 Household does not own such device
90 others, please specify

- Using do files which are named as “1_wave_2016_vn“ and “2_wave_2017_vn“
- The outcome data files will be saved in folder “dataout“



Name	Date modified	Type
full_final_panel	1/03/2025 7:40 PM	Stata Dataset
household2016	1/03/2025 7:04 PM	Stata Dataset
household2017	1/03/2025 6:57 PM	Stata Dataset
panel_example_internet	1/03/2025 7:26 PM	Stata Dataset

Main variables

Variable name	Measurement	Definitions
Internet use	Dummy	Whether household has access to the internet for productive purposes, Yes = 1; Otherwise = 0
Non-farm employment	Dummy	Whether household has at least one member participating in non-farm jobs, yes = 1; Otherwise = 0
Number of non-farm laborers	Quantity	Number of household members employed in non-farm sectors
Per capita non-farm income	PPP $	Total non-farm income per capita (divided by the number of household members)
A. Household characteristics		
Age	Years	Age of household head
Gender	Dummy	Gender of household head. Male = 1; Female = 0
Marital status	Dummy	If the household head is married. Yes = 1; Otherwise = 0
Mean education years	Years	The average years of education of all adult members in the household
Household size	Persons	The number of people living in the household and not leaving more than 180 days per year
Labor share	Percentage	The share of laborers in total household size
Total land value	PPP $	Total current land value
Number of shocks	Numbers	The total number of shocks (both idiosyncratic shocks and covariate shocks) that household experienced during the last 12 months
Ethnic minority	Dummy	If the household head belongs to an ethnic minority. Yes = 1; Otherwise = 0
B. Physical capital		
Number of phones	Quantity	The number of phones that household owns

Looking at the questionnaire

2.1: Household Members

Who is mainly in charge of financial or burea

1	2	14	15	16	
I.D. code	**Name/Nickname**	Main occupation between 5/15 - 4/16 according to time spent	Second occupation between 5/15 - 4/16 according to time spent	How many days did [NAME] stay in the household between 5/15 - 4/16? *if = 366, skip to next person*	r
		H	**H**	days	
01					
02					
03					
04					
05					
06					
07					
08					
09					
10					
11					
12					
13					
14					
15					

Code H

1 Engaged in own agriculture (including livestock and aquaculture)
2 Engaged in fishing, hunting or collecting
3 Non-farm self-employed
4 Casual off-farm labour in agriculture
5 Casual labour in non-agriculture
6 Permanently employed in agriculture
7 Permanently employed in non-agriculture
8 Government official
9 Housewife
10 Student/Pupil
11 Child below school age
12 Unemployed
13 Performing only occasional and light work

14 Monk
15 joined the army
17 Unable to work because of disability
18 Unable to work-other reasons
19 Taking care of disabled/impaired household members
20 no second occupation
90 Other, specify
98 no answer

Code AA

1 yes
2 no
98 no answer

Final data sets achievement

Step 2: Coding variables (cont)

By executing do files number 1 and 2, we obtain two datasets: "household2016.dta", "household2017.dta" with exactly the same generated variables in the folder "dataout ". Then we need to append datasets to achieve final panel.

Step 3: Appending datasets

- After generating variable for each year, we will append the data to get the final data sets
- → Do file is named as "3_appending data"

III. Regression

- Open the dataset „full_final_panel.dta“ in the folder “dataout “.
- This dataset contains 6044 observations of households who use internet and engaging in non-farm employment of two year 2016 and 2017
- Please execute the do file namely “4_regression full dataset“

Thank you for your attention!

References

- Amare, M., Parvathi, P., & Nguyen, T. T. (2023). Micro insights on the pathways to agricultural transformation: Comparative evidence from Southeast Asia and Sub-Saharan Africa. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, *71*(1), 69-87.
- Do, M. H., Nguyen, T. T., Halkos, G., & Grote, U. (2022). Non-farm employment, natural resource extraction, and poverty: evidence from household data for rural Vietnam. *Environment, Development and Sustainability*, 1-38.
- Galperin, H., & Fernanda Viecens, M. (2017). Connected for development? Theory and evidence about the impact of internet technologies on poverty alleviation. *Development Policy Review*, *35*(3), 315-336.
- International Telecommunication Union (ITU) World Telecommunication/ICT Indicators Database (accessed date 1st March 2025) https://www.itu.int/en/ITU-D/Statistics/pages/stat/default.aspx
- Hoang, T. X., Pham, C. S., & Ulubaşoğlu, M. A. (2014). Non-farm activity, household expenditure, and poverty reduction in rural Vietnam: 2002–2008. *World Development*, *64*, 554-568. https://doi.org/10.1016/j.worlddev.2014.06.027
- Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business and Economic Statistics*, *30*(1), 67–80. https://doi.org/10.1080/07350015.2012.643126
- Kaila, H., & Tarp, F. (2019). Can the Internet improve agricultural production? Evidence from Viet Nam. *Agricultural Economics*, *50*(6), 675-691.
- Nguyen, T. T., Nguyen, T. T., & Grote, U. (2023). Internet use and agricultural productivity in rural Vietnam. *Review of Development Economics*, *27*(3), 1309-1326.