

NTTA Production in Time Units

Gretchen Donehower
NTA/CWW Time Use and Gender Workshop
Tuesday, July 15, 2014
East-West Center
Honolulu, Hawaii



But first...

1. Questions about NTA by sex
2. What were you able to accomplish yesterday?
3. Future problems/issues?
4. Please send me asap:
 - Population by age/sex
 - Single-sex, Male, and Female YL age profiles (smoothed and unsmoothed if possible)

Outline

1. Working with your time use data
2. Identifying NTTA activities
3. Estimation strategy
4. File preparation
5. Smoothing
6. Understanding results
7. Sensitivity testing

Know your time use survey

- How is the data structured?
- How are activities coded?
- Whose time use data do you have in the household?
- Do you need to merge different types of files (like activity files linked to demographic information files by a person ID number)
- Are there sample weights? What are they adjusting for?
- When were the surveys done? (Want to know what time of year because of seasonality)

Time use survey data cleaning

- The usual data cleaning activities
 - Variables are as described in the codebook?
 - Duplicate observations?
 - Unrealistic outliers?
 - Invalid codes?
 - Codes for missing data changed to missing or zeros as appropriate?

Time use survey data cleaning

- Time-based diagnostics and checks on your use of the data
 - Does total time sum to 24 hrs/day or all persons? (Or if you don't have all activities accounted for, is sum of accounted activities reasonable?)
 - If not, is error correlated with age and/or sex?
 - Do you observe time patterns as expected
 - more market work done on typical workdays
 - childcare marked as an activity in households with very young children
 - sleep is a reasonable amount on average by age

What to do if there are problems?

- Total time not equal to 24 hours?
 - Proportional adjustment for each individual if errors are not too large and your survey accounts for all possible activities
 - Can delete some observations with far more than 24 hours accounted for, and adjust the rest
 - OR look for specific outliers by individual or by type and see if problem can be isolated (eg observations from one province have large errors but others are fine)

Which activities are in NTTA?

- Find time use survey's coding scheme for activities, or list of activities with time variables in your survey
- Identify NTTA productive activities in coding scheme
 - Not already included in national accounts
 - So market production is NOT included
 - Unpaid family labor that generates market income is not included, even if the person who did the work did not get paid
 - Satisfies third party criterion
 - You could pay someone else to do it for you
 - Leisure, education, and many self-care activities do NOT satisfy this criterion
- Group NTTA productive activities into 11 categories, data permitting

Which activities are **NOT** in NTTA?

- Ones whose value is already in national accounts and NTA
 - Consumption of goods produced by household for own use are imputed in national accounts consumption measures
 - Returns to labor to produce these household-consumed and produced goods are imputed in mixed income

Groups of NTTA Activities

1. Cleaning
2. Laundry (includes sewing and clothing repair)
3. Cooking (food and drink preparation)
4. Household maintenance and repair
5. Lawn and garden care
6. Household management (incl. finances, scheduling, coordinating, and related telephone calls)
7. Pet care (not veterinary care)
8. Purchasing goods and services
9. Childcare **
10. Eldercare and care outside the home (includes volunteering) **
11. Travel (related to care activities and purchasing goods and services)

** Needs to be divided into variables for care in household versus outside household

Estimation Strategy (time units)

- Prepare files with production data and hh members
 - Need individual-level file with an observation for each household member and variables for time spent in household production activities
- [impute consumption... we will do that tomorrow]
- Calculate age-sex means of time spent in 11 categories (data permitting)
 - Keep care categories separated by location (in household vs. outside household) if possible, for use in consumption calculation
 - Use survey weights
- Smoothing
 - Examine smoothed profiles by looking at each one
 - Adjust spans so as not to “smooth over” a real pattern
 - care variables will often need smaller spans than other household activities or need to be smoothed in sections

Time use survey file preparation

- Need to create a file with one observation per household member and variables for time average daily time spent in each household production activity
 - Includes all household members, with unique household id for each household
 - Time respondents will have time values in those variables, non-respondents will have missing data code (“.” in Stata)
 - Non-time respondents are not necessary for production but are for consumption

File prep for diary surveys

Given time diary file and household roster file:

Time Diary Data (one obs per activity)

HHID	PID	Time	Act
1	1	30	101
1	1	240	201
1	1	60	201
1	1	90	101
1	1	360	301
1	1	240	401
2	1	60	101
2	1	90	101
2	1	120	201
2	2	30	401
2	2	60	302
2	2	60	302
2	2	90	301
2	2	360	401
2	2	240	401

Household Roster Data (one obs per person, including non-time respondents)

HHID	PID	Age	Sex
1	1	25	1
1	2	8	2
2	1	32	1
2	2	37	2
2	3	4	2
2	4	8	1
2	5	1	1

Time Diary Data (one obs/activity)

HHID	PID	Time	Act
1	1	30	101
1	1	240	201
1	1	60	201
1	1	90	101
1	1	360	301
1	1	240	401
2	1	60	101
2	1	90	101
2	1	120	201
2	2	30	401
2	2	60	302
2	2	60	302
2	2	90	301
2	2	360	401
2	2	240	401

Collapses to one obs per person/activity

HHID	PID	Time	Act
1	1	120	101 cleaning
1	1	300	201 cooking
1	1	360	301 childcare - hh
1	1	240	401 laundry
2	1	150	101 cleaning
2	1	120	201 cooking
2	2	90	301 childcare - hh
2	2	120	302 childcare - nhh
2	2	630	401 laundry

Reshapes to one obs per person (time respondents only)

HHID	PID	Act101	Act201	Act301	Act302	Act401
		clean	cook	chcarehh	chcarenhh	laundry
1	1	120	300	360	.	240
2	1	150	120	.	.	.
2	2	.	.	90	120	630

Replace activities respondent did not do with zeros

HHID	PID	Act101	Act201	Act301	Act302	Act401
		clean	cook	chcarehh	chcarenhh	laundry
1	1	120	300	360	0	240
2	1	150	120	0	0	0
2	2	0	0	90	120	630

Household Roster Data (one obs/person, incl. non-time resp)

HHID	PID	Age	Sex
1	1	25	1
1	2	8	2
2	1	32	1
2	2	37	2
2	3	4	2
2	4	8	1
2	5	1	1

Merges to one obs per person, including all household members

HHID	PID	Age	Sex	Act101	Act201	Act301	Act302	Act401
				clean	cook	chcarehh	chcarenhh	laundry
1	1	25	1	120	300	360	0	240
1	2	8	2
2	1	32	1	150	120	0	0	0
2	2	37	2	0	0	90	120	630
2	3	4	2
2	4	8	1
2	5	1	1



File prep for activity questions

With activity based questions, you start here instead of with the activity level file

Reshapes to one obs per person (time respondents only)

HHID	PID	Act101 clean	Act201 cook	Act301 chcarehh	Act302 chcarenhh	Act401 laundry
1	1	0	0	0	.	0
2	1	0	0	.	.	.
2	2	.	.	0	0	0



Replace activities respondent did not do with zeros

HHID	PID	Act101 clean	Act201 cook	Act301 chcarehh	Act302 chcarenhh	Act401 laundry
1	1	120	300	360	0	240
2	1	150	120	0	0	0
2	2	0	0	90	120	630



Household Roster Data (one obs/person, incl. non-time resp)

HHID	PID	Age	Sex
1	1	25	1
1	2	8	2
2	1	32	1
2	2	37	2
2	3	4	2
2	4	8	1
2	5	1	1



Merges to one obs per person, including all household members

HHID	PID	Age	Sex	Act101 clean	Act201 cook	Act301 chcarehh	Act302 chcarenhh	Act401 laundry
1	1	25	1	120	300	360	0	240
1	2	8	2
2	1	32	1	150	120	0	0	0
2	2	37	2	0	0	90	120	630
2	3	4	2
2	4	8	1
2	5	1	1

Unsmoothed Age Means

- Collapse file by age/sex, using survey weights, keeping number of observations:

```
gen obs=1
```

```
collapse (mean) cook clean chcare... (rawsum)
```

```
obs [w=wgt], by(age sex)
```

(need “obs” for smoothing routine)

Smoothing

- Current preferred method: R's "supsmu" (Friedman's super smoother)
 - Pros: good smoother and small span results come out pretty well
 - Cons: have to implement in R which can be confusing and clunky if you have not used R before
- Alternative: Stata's "lpoly"
 - Pros: stays within Stata so you don't have to mess with R
 - Cons: settings are very sensitive, can be hard to find the best settings, often need to manually adjust smoother at endpoints

`lpoly wtd_mean_variable age [if] [in] [#_of_observations] [, options]`

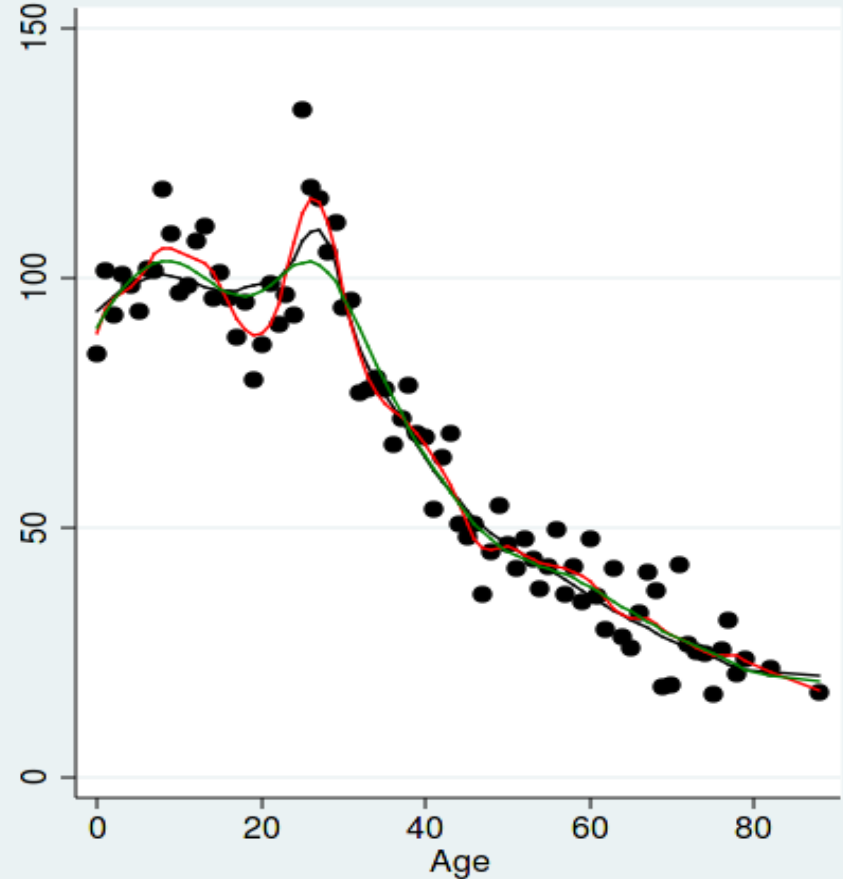
Options for controlling the smooth:

<code>kernel(kernel)</code>	kernel function ("gaussian" works best)
<code>bwidth(# varname)</code>	bandwidth
<code>degree(#)</code>	degree of polynomial smooth (so far 2 works best)

Comparing Smoothers

- Example (welfare payments to families with young children):

Unsmoothed versus Different Smoother



Understanding the results

- Can further cross-cutting reveal how overall patterns emerge?
 - Regional variation
 - SES
 - Different household structures
 - If you have more than one time use survey, are patterns stable over repeated cross-sections?
- Examine time use and gender differences across different countries

Sensitivity Tests

- Change methodology to evaluate how different results would be if you made different assumptions
- Most important one for production account in time units is including secondary activities
 - Preserve 24 hours in a day (i.e. no multitasking miracle)
 - Divide the time unit in half between the two activities