How Americans' Time Use Patterns Have Changed From 2003 to 2017

Group July

06 December 2018

Introduction

Introduction

The ATUS dataset [1] used is based on research carried out as part of the American Time Use Survey [2] from 2003 to 2017, containing time use data for 431 different activities grouped into 17 over-arching categories.

Aim: Hypothesise, validate and present long-term trends based on the data

Approach:

- EDA
  - Exploring data from even months as training data (July excluded and used for validation);
- Validation using left out "unseen" data
  - Formal hypothesis tests on the initial beliefs from the EDA using the odd months (and July) as the validation data;
- Summary plots of findings

Tidying and Structuring the Data

- Missing value were confined to 9 columns of factor variables so these were removed
- The 17 categories were built from the activity data
- As the aim is to present long-term trends, weighted means for **each year** were calculated
  - This also allows for comparisons and ensures that each group is correctly represented in the population

The weighted means are then calculated using

$$\overline{T_j} = \frac{\sum_i weight_i T_{ij}}{\sum_i weight_i}$$

where i corresponds to the individual, and j corresponds to the type of individual

Observation 1:

Participation in Caring for & Helping Non-HH Members

Exploratory Data Analysis

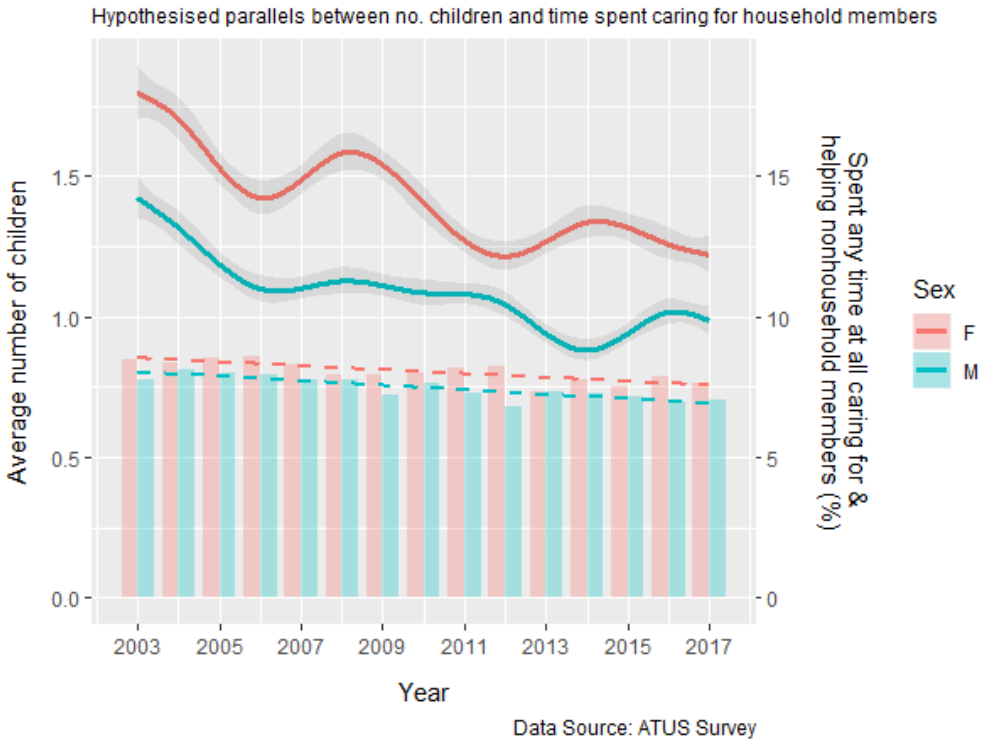
- Activities with % change larger than 10% and variance greater than 0.5

Table 1: Change in Participation of Activities

Measure	tu04	tu08	tu13	tu14	tu16
Variance	2.54	0.89	1.44	0.60	1.93
% Change	-32.11	-26.46	10.17	12.32	-24.08

- Fitting a linear model and performing best subsets regression with `Year` forced in as an explanatory variable gives `Sex` and `Number of Household Children`
- Fit a more complex model: glm with log link and multiplicative errors
  7
  - Use natural cubic splines to show fluctuations
  - $tu04participation \sim -1 + Sex + Sex : ns(Year, knots = 2003, 2005, ..., 2017)$
- Performing an F-test on the model shows this is a significant improvement on
  $tu04participation \sim -1 + ns(Year, knots = 2003, 2005, ..., 2017)$

Observation 1:

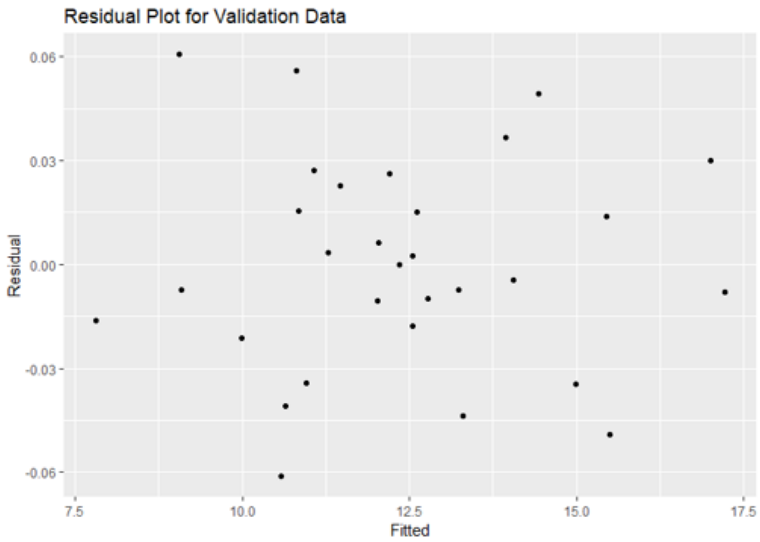


- The plot indicates that over the period, the participation in 'tu04' has decreased for both men and women
- Changes in the 'average no. of household children' seem to follow the trend in participation, however the link is weak (correlations of 0.65 for Men and 0.49 for Women)

A one-sided t-test on a linear model simplification of the generalised linear model without splines - $(tu04participation \sim Year + Sex)$ - gives a p-value of **5.7e-09** $<<$ 0.05

Validation

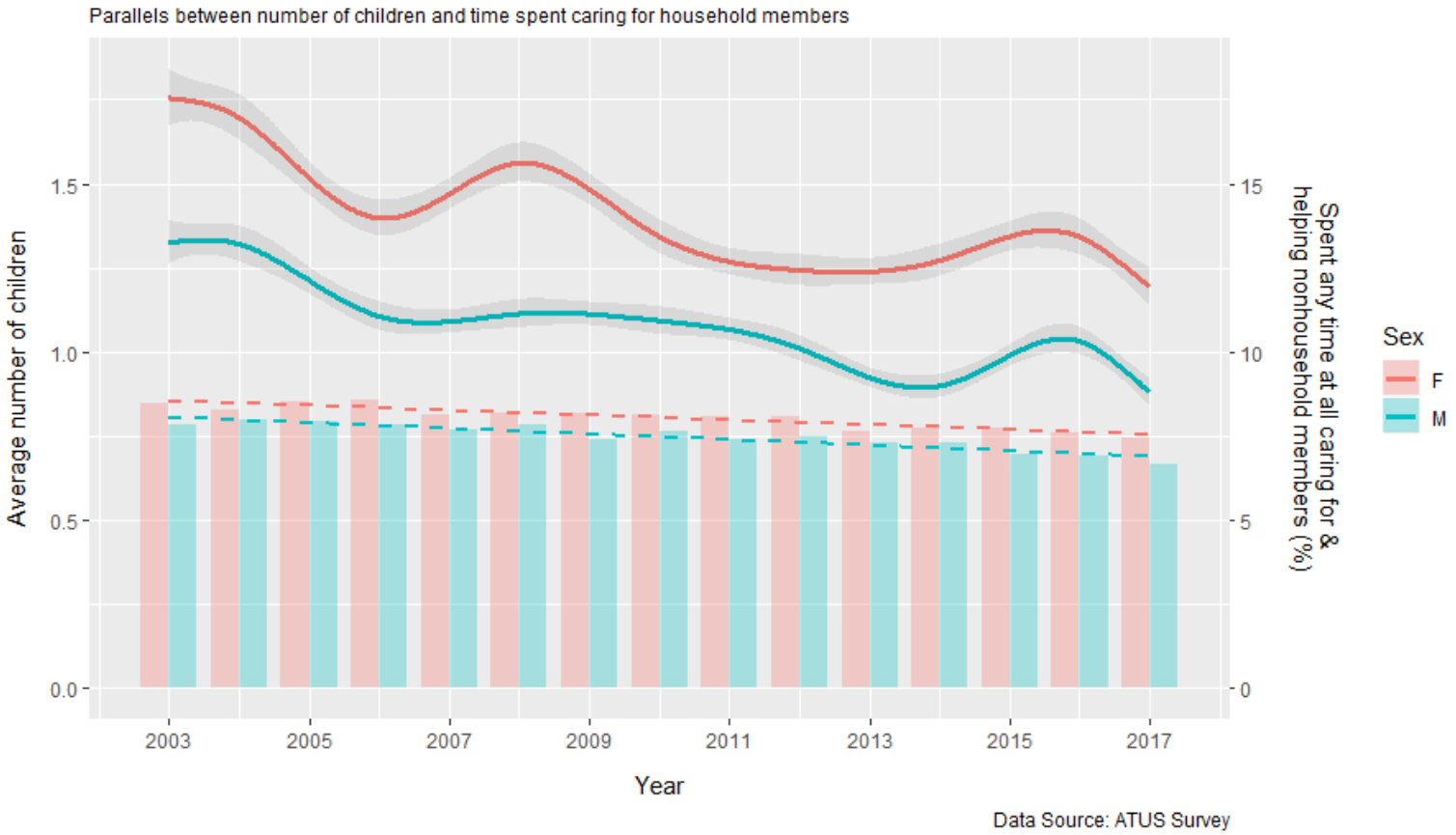
Formal one-sided t-test on linear model build on *validation data* gives a p-value of **1.4e-07** $<<$ 0.05



- To test the suitability of the model on the validation dataset a residual plot was created
- The errors:
  - Are uncorrelated
  - Have mostly equal variance
  - Seem to have mean 0

Performing a formal t-test when `average number of household children` is added confirms that this has a significant effect on participation in `caring for & helping non-household children`

Final Plot Built On All Data Excluding July



Observation 2:

How Time Spent on Traditionally Gendered Activities has Changed as Gender Roles have Broken Down?

Table 2: Traditionalist Gender Actitivies

Male Activites	Female Activities
Working	Housework
House Maintenance	Cooking
Vehicle Maintenance	Childcare

Table 3: Generations

Generation	Birth Years
Silent Generation	1928 - 1945
Baby Boomers	1946 - 1964
Generation X	1965 - 1980
Millennials	1981 - 1996

- Table 2 is a simplified version of Talcott Parsons' [3] study on gender roles
- Table 3 shows how to break down the respondents into different generation groups [4]

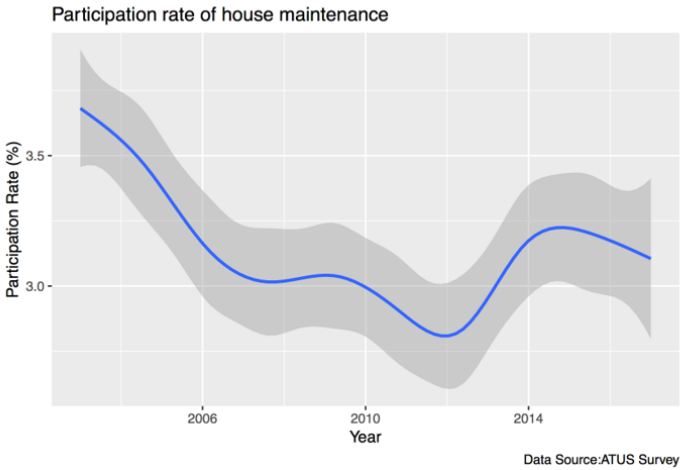
- Division between genders in terms of societal roles is consistently featured in the news; evident in the "#MeToo" movement amongst others
- The report aimed to investigate how this division changed over the given period through investigating long-term trends in each of the traditional gender activities

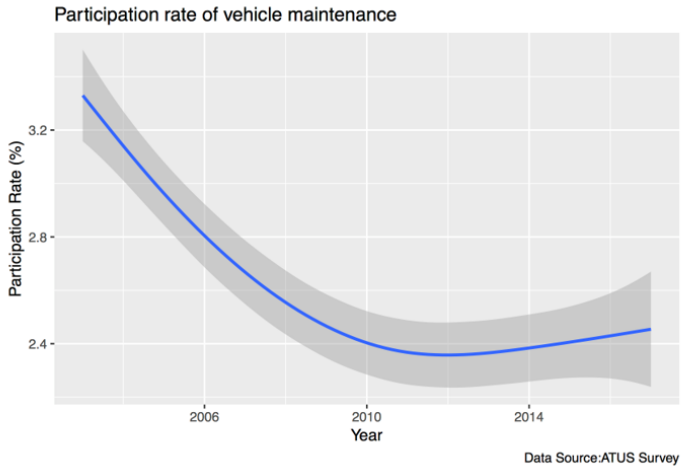
Exploratory Data Analysis

- The first stage of the analysis looked at participation rates at a total population level for the different activities to check they were popular enough for comparison
- Following this initial check, general linear models were developed for all suitable activities and different parameters were checked including:
  - Sex
  - Year
  - Generation
  - Region
- After performing formal F-tests, comparing simpler models, the following model was settled upon for all activities $Activity \sim -1 + ns(Year) + Sex + Sex : ns(Year, knots = 2003, 2005, ..., 2017) + Generation + Generation : ns(Year, knots = 2003, 2005, ..., 2017)$
- The simpler models were variants of this model, but without sex or generational information respectively.

House Maintenance and Vehicle Maintenance (Males)

- The participation rate was too low to warrant deeper analysis
- Whilst the findings represented that there existed a separation in gender, the participation rates of around 3% for both reflected that these were more uncommon activities
- It was decided that there was not enough data to reflect the time spent on these activities in a suitable linear model



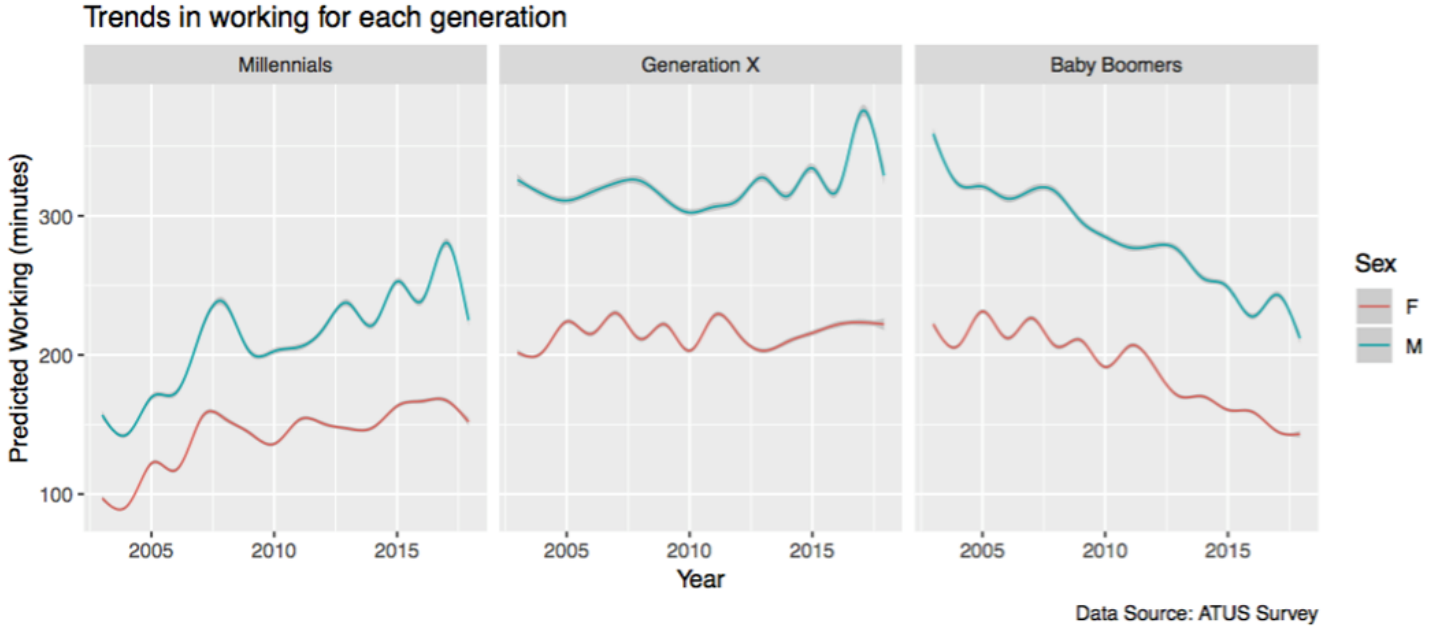


Validation

- **Males:** Working, House Maintenance and Vehicle Maintenance;
- **Females:** Housework, Cooking and Childcare
- The models shown on the following slides showcase the results of the analysis, plotting all of the data with the exception of July as required
- Despite using 11 months of the data here, it is critical to reiterate that all of the EDA and validation was carried out on entirely separate 6 month subsets of each year to ensure validity of the conclusions and testing
- Formal one-sided t-test were performed on the simple linear model below for each activity, using weighted yearly averages for the data

$$\textit{Gender time difference} \sim \textit{Year}$$

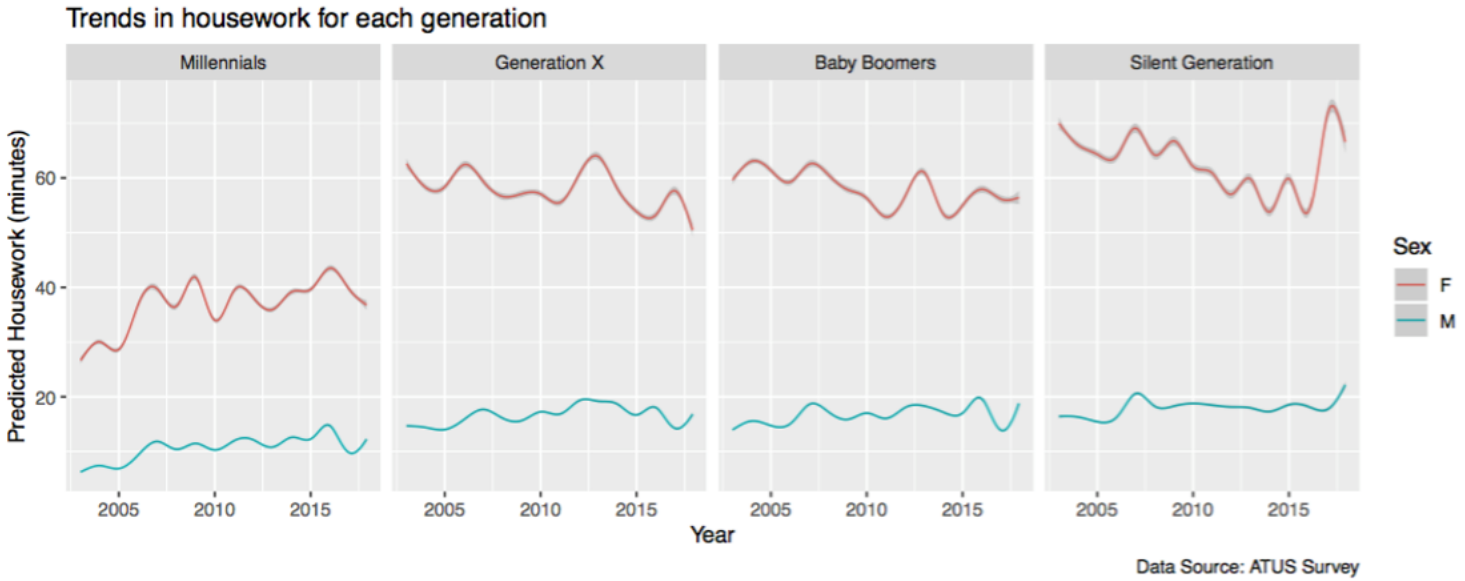
Validation - Working



- The plots show the changes in working patterns between 3 generations. There's an increase in working time for *Millenials* as this generation begin to work. Likewise, there's a decrease in time for *Baby Boomers* as many of this generation begin to retire.
- *Generation X* shows that there is an increase in women working time and for, the first 13 years, a decrease in time working for men. There is a slight uptick for men working in 2016, however, this begins to return to previous lower levels in 2017.
- The t-test on a population level for this gave a p-value of $0.047 < 0.05$
- The F-test is shown below

Model	Residual DF	Residual Deviance	DF drop	Deviance drop	F	p-value
1	98098	56241.17				
2	98068	52250.68	30	3990.49	733.51	$<$ 2.2e-16
3	98020	23983.07	48	28267.62	3247.51	$<$ 2.2e-16

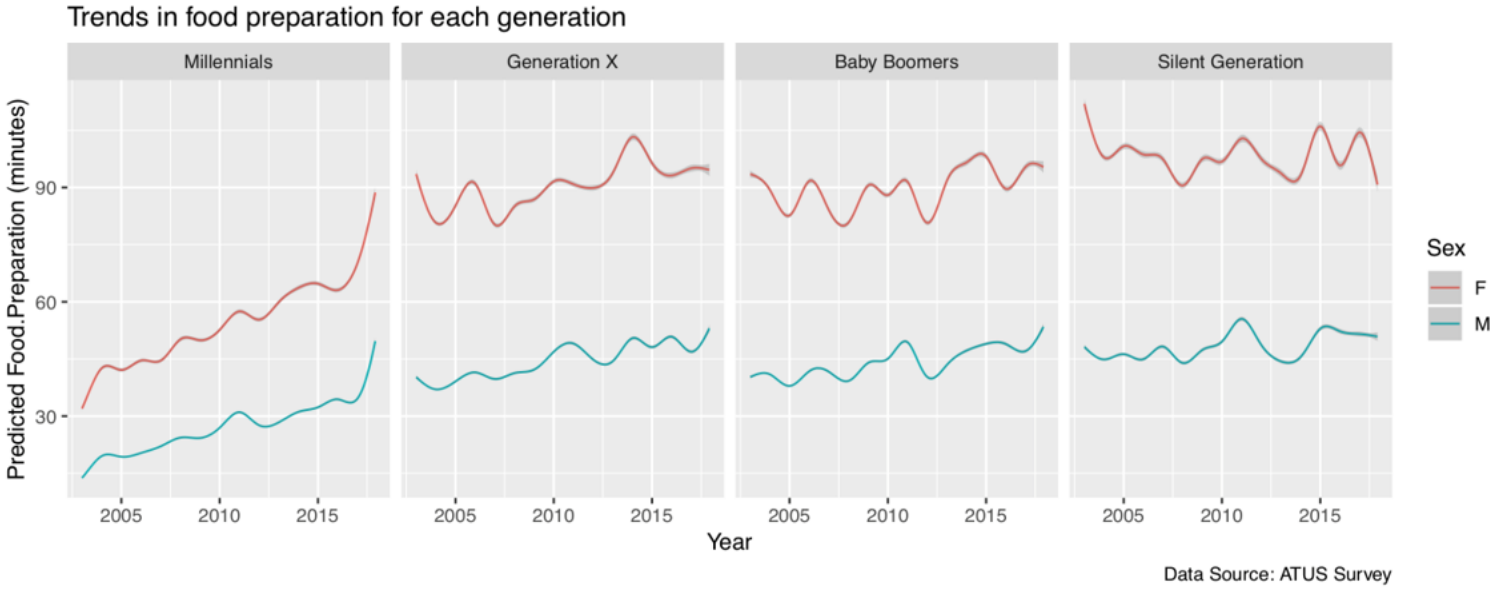
Validation - Housework



- Except for *Millennials*, all generations have the decrease in time spent on housework. But the decrease for women is sharper than any *increase* for men, which is clear through observation of *Generation X*.
- On the other hand, the gap seems to have increased slightly for *Millennials* - both sexes are increasing the amount of time spent on housework, confounding effect of increased time due to age / moving out.
- Notably, *Millennials* also spend less time doing housework than the others.
- The t-test on a population level for this gave p-value of 6.8e-05 $<<$ 0.05
- The F-test is shown below

Model	Residual DF	Residual Deviance	DF drop	Deviance drop	F	p-value
1	98098	66513.61				
2	98068	30837.95	30	35675.66	4231.51	$<$ 2.2e-16
3	98020	27722.58	48	3115.37	230.95	$<$ 2.2e-16

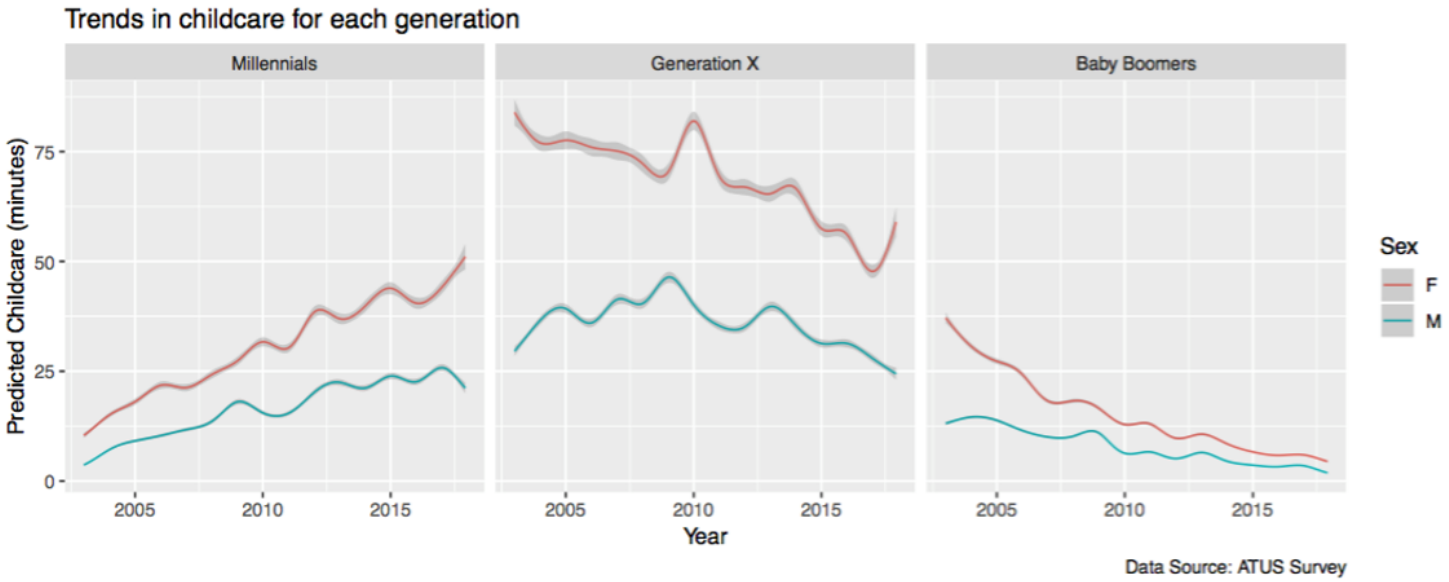
Validation - Cooking



- Both genders from all the generations other than the *Silent Generation* are actually spending more time than previously on food preparation
- For men, there is a sharper increase than in time spent by women which is evidence of erosion in this particular gender stereotype
- Notably, *Millennials* spend more and more time on cooking and the nearly same increasing rate of both gender leads to the small gap
- The t-test on a population level for this gave p-value of 0.0054 $<<$ 0.05
- The F-test is shown below

Model	Residual DF	Residual Deviance	DF drop	Deviance drop	F	p-value
1	98098	30146.10				
2	98068	17578.80	30	12567.29	2889.4	$<$ 2.2e-16
3	98020	13478.53	48	4100.28	589.2	$<$ 2.2e-16

Validation - Childcare



- Note that the drop off for women is sharper than it is for men, leading to a convergence in the weighted means for both.
- The t-test on a population level for this gave a p-value of $0.037 < 0.05$
- The F-test is shown below

Model	Residual DF	Residual Deviance	DF drop	Deviance drop	F	p-value
1	98098	176671.00				
2	98068	162588.37	30	14082.63	298.72	$<$ 2.2e-16
3	98020	59257.72	48	103330.65	1369.92	$<$ 2.2e-16

Limitations of the Data

Limitations of the Data

- 15 years is a relatively short period within which to observe long term trends
- Sporadic subset sizes due to the filtering and sub-setting required
- The reliance on people to remember the way in which they spend their time (i.e. They could forget smaller tasks and focus on more memorable or time-consuming ones)

References

References

[1] "ATUS datasets." https://www.bls.gov/tus/datafiles_0317.htm.

[2] Bureau of Labor Statistics, "The american time use survey." https://www.bls.gov/tus/, 2017.

[3] T. Parsons, "Age and sex in the social structure of the united states," *American Sociological Review*, vol. 7, no. 5, pp. 604–616, 1942 [Online]. Available: http://www.jstor.org/stable/2085686

[4] "Millennials projected to overtake baby boomers as america's largest generation." http://www.pewresearch.org/fact-tank/2018/03/01/millennials-overtake-baby-boomers/.