

Statistical Models for Time Use Data:  
An Application to Housework and Childcare  
Activities Using the Austrian Time Use Surveys  
from 1992 and 2008

Master Thesis, University of Vienna

Bernhard Hammer

November 8, 2012

# Objectives

1. Analyzing changes in the time use of men and women for housework and childcare between 1992 and 2008
2. Evaluating statistical models for time use data:
  - ▶ Linear model
  - ▶ Tobit model for censored data
  - ▶ GLM with a negative binomial distribution
  - ▶ GLM with a Poisson-gamma distribution

# Time use for household work: What should we expect?

Most of the research based on the Multinational Time Use Survey containing time use surveys from 1960-

## **For women:**

- ▶ Housework: Strong decline
- ▶ Paid work: Increase
- ▶ Childcare: Increase

## **For men:**

- ▶ Housework: Constant/slightly increasing
- ▶ Paid work: Constant/decline
- ▶ Childcare: Increase (from a very low level)

# Why expecting changes between 1992 and 2008?

- ▶ Increase in the labour force participation of women
  - ▶ age-group 25-54: 67.1% in 1992, 81.5% in 2008
- ▶ Increase in public child care enrolment:
  - ▶ 3 year old: 29% in 1992, 53% in 2008;
  - ▶ 4-5 year old: 75% in 1992, 90% in 2008;
- ▶ Introduction of parental leave for fathers (1990). Share of men on parental leave: 1990: 0.18%, 2001: 2%, 2012: <5%
- ▶ Attitudes towards gender-roles became more egalitarian
- ▶ *Ganze Männer machen Halbe-Halbe* (Full men share half-half)  
Campaign for an obligation spouses to an equal division of household work by law (1996)

# Austrian Time Use Surveys

- ▶ 1992: 25,233 individuals in 12,169 households
- ▶ 2008: 8,234 individuals in 4,757 households
- ▶ Time diaries from all household members older than 10
- ▶ 15 minutes times-slots from 5:00 to 23:00
- ▶ 30 minute time-slots from 23:00 to 5:00

# The Subsample

The analysis is restricted to a subsample, focus is on working age adults:

- ▶  $15 < \text{Age} < 55$
- ▶ children of household head or her/his partner are excluded
- ▶ Sample size: 11,733 in 1992 and 4,155 in 2008

# Household Structure

Table: Household Structure

	1992	2008
	Share in %	Share in %
Single households	14.5	27.8
Couples with children	49.1	43.2
Couples without children	13.5	15.4
Single parents	7.0	6.4
Other	15.9	7.2
<b>Total</b>	<b>100.0</b>	<b>100.0</b>
Average Household Size	3.20	2.65

# Time Use for Housework and Childcare

Table: Time Use for Housework in Minutes per Day

	Men		Women	
	1992	2008	1992	2008
Single households	92	102	161	132
Couples with children	93	102	347	250
Couples without children	99	107	230	177
Single parents	160		262	202
Other households	87	83	299	193
<b>Total</b>	<b>93</b>	<b>102</b>	<b>299</b>	<b>209</b>

Table: Time Use for Childcare in Minutes per Day

	Men		Women	
	1992	2008	1992	2008
Couple households	38	70	114	137
Single parents			96	116



# Statistical Models - Hypotheses

- ▶ Relationship between factors other than sex and household type and time use for housework and childcare.
- ▶ Men in 2008 devote significantly more time to housework and childcare than in 1992. This holds after controlling for other characteristics such as education, age, or the number of children.

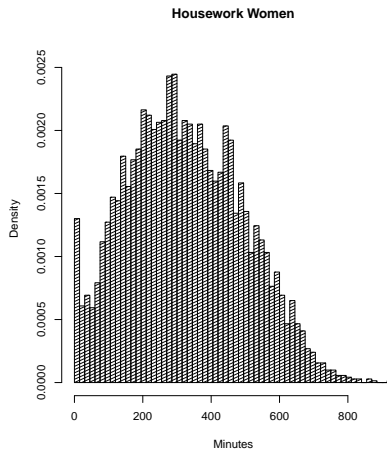
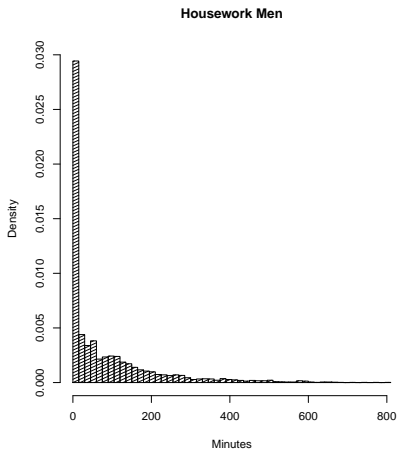
# Characteristics of Time Use Data

- ▶ Non-negativity
- ▶ Frequent observation of zeros
- ▶ Skewness

**The data generating process depends on the activity and the subpopulation under investigation.**

# Empirical Density of Time Devoted Housework

Figure: Time Devoted to Housework by Couples with Children: Empirical Density



# A Linear Model for the Housework

$$\begin{aligned} \text{Minutes used for housework on the survey day} = & \\ & = \text{const.} + \beta_1 * \text{year 2008} + \\ & + \beta_2 * \text{upper secondary} + \beta_3 * \text{tertiary} + \\ & + \beta_4 * \text{age} + \beta_5 * \text{year 2008} * \text{age} + \\ & + \beta_6 * \text{household size} + \beta_7 * \text{no. of children below 6} + \\ & + \beta_8 * \text{living in city} + \beta_9 * \text{partner is working} + \epsilon \quad (1) \end{aligned}$$

# Results: Linear Model for Housework

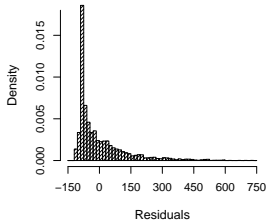
Table: Housework: Results from the Linear Model, Couples with Children

	Housework Men	Housework Women
Year 2008	-11.57	-96.57***
Upper Secondary (Matura)	-8.249	-45.36***
Tertiary Education	-1.509	-72.41***
Age	0.431	2.801***
Age*Year 2008	0.686	0.157
Household Size	0.483	32.92***
Number of Children below 6	-3.550	10.20**
Living in City	-1.027	-27.96***
Partner is Employed	-0.0914	16.52**
Constant	65.71***	112.2***
Observations	4,068	4,715
R-squared	0.006	0.139

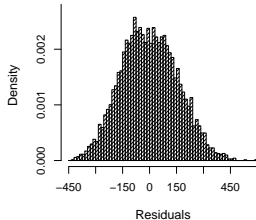
Standard errors in parentheses; \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# Residual Analysis of the Linear Model for Housework

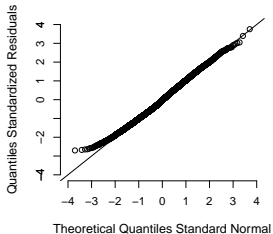
Residuals Men



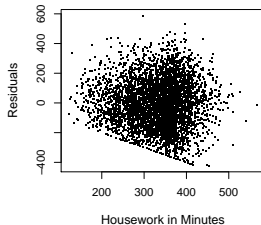
Residuals Women



Q-Q Plot Women



Women; Predicted vs. Residuals



# Generalized Linear Models

Basic idea: Model the expectation  $E[Y|X_i] = \mu_i$  and assume a (conditional) distribution for  $Y$  given  $\mu_i$

Advantage: We can choose a distribution which is non-negative, right skewed, has a point-mass at zero and where the variance depends on the mean.

Components of GLM:

- ▶ The distribution of  $Y$  given  $\mu_i$ , which is called the *random component*
- ▶ The second component is the linear predictor  $\eta_i = \sum_{j=1}^k x_{ij}\beta_j$
- ▶ The third component is the link function  $g(\mu_i) = \eta_i$  e.g.  $\log(\mu_i) = \eta_i$

We use a Poisson-gamma distribution with a log-link.

# GLM with a Poisson-Gamma Distribution

Judith Brown and Peter Dunn (2011). *Comparisons of Tobit, Linear, and Poisson-Gamma Regression Models: An Application of Time Use Data*

Structure of the data:

- ▶ Several episodes of different length
- ▶ Model for the number of episodes  $N$ : Poisson process
- ▶ Model for the duration : Gamma distribution

$Y = \sum_{i=1}^N Z_i$ . The sum  $Y$  represents the total time spent in this activity on a certain day and follows a Poisson-gamma distribution.



# Exponential Dispersion Models: Tweedie Distributions

The density of exponential dispersion models:

$$f(y; \theta, \phi) = c(y, \phi) \exp \left\{ \frac{y\theta - \kappa(\theta)}{\phi} \right\}. \quad (2)$$

These models have a unique variance function  $V(\mu)$  for fixed  $\phi$ .

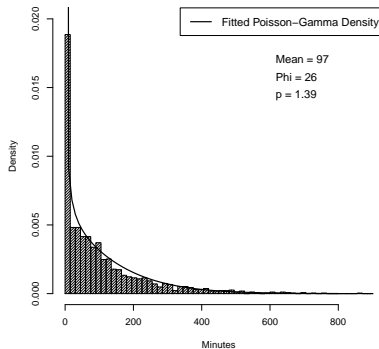
**Tweedie distributions:**  $V(\mu) = \mu^p$ ,  $p \notin (0, 1)$ ;

Examples: Normal distribution ( $p = 0$ ), the Poisson distribution ( $p = 1$ ), gamma distribution ( $p = 2$ ).

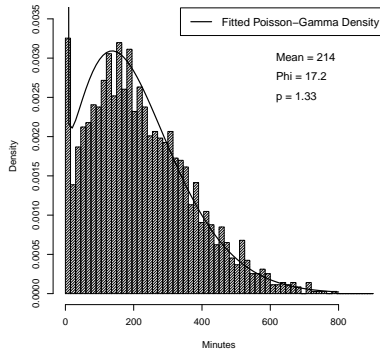
Tweedie distributions with  $p \in (1, 2)$  correspond to Poisson-gamma distributions

# Fitted Poisson-Gamma Distribution

Housework Men 2008



Housework Women 2008



# Estimation

GLM with a Tweedie distribution:

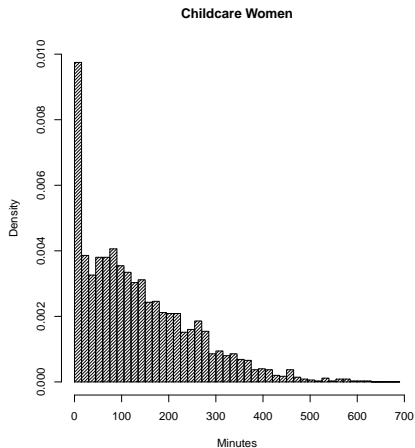
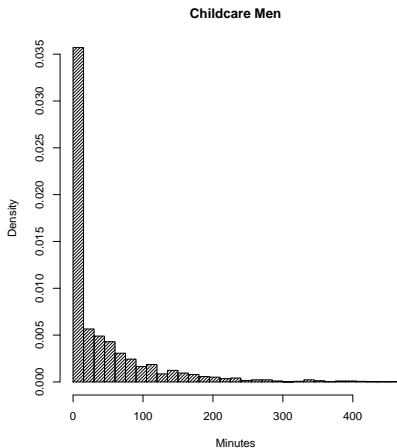
R-package *statmod*; function `glm(,family=tweedie())`

You have to specify the power  $p$  of the Poisson-gamma (resp. Tweedie-) variance function!

Estimation of  $p$ : Function `tweedie.profile()` from the R-package *tweedie*.

# Empirical Density: Childcare

Figure: Density Childcare



# Linear Predictor for the Childcare

$$\begin{aligned} \log(\mu) = & \text{const.} + \beta_1 * \text{year 2008} + \\ & + \beta_2 * \text{upper secondary} + \beta_3 * \text{tertiary} + \\ & + \beta_4 * \text{age} + \beta_5 * \text{year 2008} * \text{age} + \\ & + \beta_6 * \text{no. children below 10} + \beta_7 * \text{no. of children aged 0-2} + \\ & + \beta_8 * \text{no. of children aged 3-5} + \\ & + \beta_9 * \text{living in city} + \beta_{10} * \text{partner is working} \quad (3) \end{aligned}$$

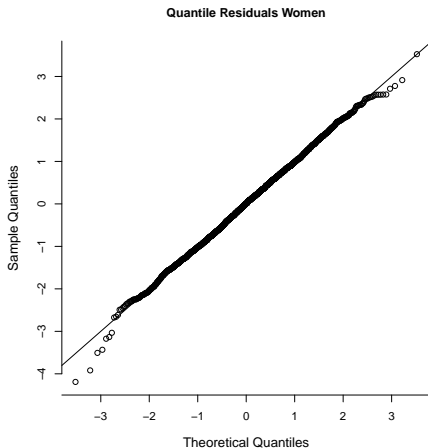
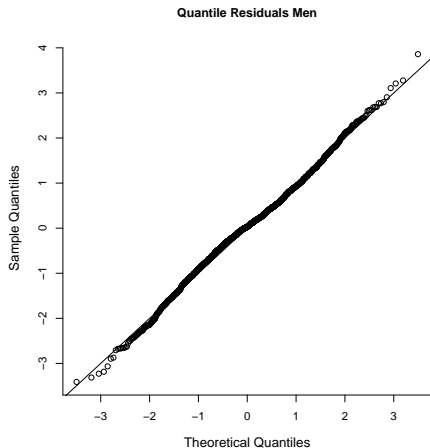
# Results: Poisson-Gamma for Childcare

Table: Childcare: Results from the Poisson-Gamma Model

	(Childcare Men)	(Childcare Women)
(Intercept)	1.389***	2.216***
Year 2008	0.013	-0.273
Upper Secondary (Matura)	0.018	0.094*
Tertiary	0.297**	0.169***
Age	-0.031***	-0.021***
Age*Year 2008	0.018	0.013*
No. Children below 10	-0.028	0.054
No. Children aged 0-2	0.950***	0.775***
No. Children aged 3-5	0.323***	0.213***
Living in City	0.368***	0.064
Partner is Working	0.193**	0.169**
N	2112	2332

# Model Evaluation

Figure: Childcare: Quantile Plots for the Poisson-Gamma Model



# Conclusion

Changes in time use for housework:

- ▶ For women: Strong decline
- ▶ For men: No significant change; amount of housework independent of other individual characteristics

Changes in time use for childcare:

- ▶ Increase in averages; not significant after controlling for other characteristics

Models for time use data:

- ▶ Dependent on the activity and the subpopulation under investigation
- ▶ GLM with a Poisson-gamma random component is a useful tool for time use data