Project

Kevin Choi

2022-11-13

Data Cleaning

```r
data <- read_sav(here("data_files", "pj_data.sav"))


data<- data %>% #don't need household data
   filter(RECTYPE != 1)

data_filled<- data %>%
  group_by(CASEID) %>%
  mutate(FULLPART = FULLPART[1],
         DAY=DAY[1],
         AGE=AGE[1],
         SEX=SEX[1],
         MARST=MARST[1],
         UHRSWORKT=UHRSWORKT[1],
         EARNWEEK=EARNWEEK[1],
         PAIDHOUR=PAIDHOUR[1],
         HH_NUMOWNKIDS=HH_NUMOWNKIDS[1],
         WB_RESP=WB_RESP[1],
         Leisure_All=Leisure_All[1])%>%
  subset(select = -c(PERNUM, LINENO, WT06, IND2_CPS8, YEAR, AWBWT)) %>%
  filter(RECTYPE!=2) #after pasting values, remove rows with only level 2
head(data_filled)
```

```
## # A tibble: 6 x 20
## # Groups:   CASEID [1]
##          RECTYPE CASEID      DAY   AGE     SEX    MARST FULLPART UHRSWORKT EARNWEEK
##       <dbl+lbl>   <dbl> <dbl+l> <dbl> <dbl+l> <dbl+l> <dbl+lb>     <dbl>    <dbl>
## 1 3 [Activity] 2.02e13 5 [Thu~    60 2 [Fem~ 4 [Div~ 1 [Full~        40      516
## 2 3 [Activity] 2.02e13 5 [Thu~    60 2 [Fem~ 4 [Div~ 1 [Full~        40      516
## 3 3 [Activity] 2.02e13 5 [Thu~    60 2 [Fem~ 4 [Div~ 1 [Full~        40      516
## 4 3 [Activity] 2.02e13 5 [Thu~    60 2 [Fem~ 4 [Div~ 1 [Full~        40      516
## 5 3 [Activity] 2.02e13 5 [Thu~    60 2 [Fem~ 4 [Div~ 1 [Full~        40      516
## 6 3 [Activity] 2.02e13 5 [Thu~    60 2 [Fem~ 4 [Div~ 1 [Full~        40      516
## # ... with 11 more variables: PAIDHOUR <dbl+lbl>, HH_NUMOWNKIDS <dbl+lbl>,
## #   WB_RESP <dbl+lbl>, Leisure_All <dbl>, ACTLINE <dbl>, ACTIVITY <dbl+lbl>,
## #   DURATION <dbl>, SCHAPPY <dbl+lbl>, SCSTRESS <dbl+lbl>, MEANING <dbl+lbl>,
## #   WBELIG <dbl+lbl>
```

```r
#Recoding Hourly Status Variable (1=Not Paid Hourly, 2 = Paid Hourly)
data_filled$PAIDHOUR <- as.numeric(data_filled$PAIDHOUR)
data_filled <- data_filled %>%
  mutate(PAIDHOUR=recode(PAIDHOUR,
                     '1' = 2,
                     '2' = 1))


#Recoding Day of Week (1=Weekday, 2=Weekend)
data_filled$DAY <- as.numeric(data_filled$DAY)
data_filled <- data_filled %>%
  mutate(DAY=recode(DAY,
                     '1' = 2,
                     '2' = 1,
                     '3' = 1,
                     '4' = 1,
                     '5' = 1,
                     '6' = 1,
                     '7' = 2))


#Recoding Marital Stauts (1=Not Married, 2= Married)
data_filled$MARST <- as.numeric(data_filled$MARST)
data_filled <- data_filled %>%
  mutate(MARST=recode(MARST,
                     '1' = 2,
                     '2' = 2,
                     .default = 1))


#Recoding Well-Being Measures
data_filled$MEANING <- as.numeric(data_filled$MEANING)
data_filled <- data_filled %>%
  mutate(MEANING=recode(MEANING,
                     '0' = 1,
                     '1' = 2,
                     '2' = 3,
                     '3' = 4,
                     '4' = 5,
                     '5' = 6,
                     '6' = 7))

data_filled$SCHAPPY <- as.numeric(data_filled$SCHAPPY)
data_filled <- data_filled %>%
  mutate(SCHAPPY=recode(SCHAPPY,
                     '0' = 1,
                     '1' = 2,
                     '2' = 3,
                     '3' = 4,
                     '4' = 5,
                     '5' = 6,
                     '6' = 7))


data_filled$SCSTRESS <- as.numeric(data_filled$SCSTRESS)
data_filled <- data_filled %>%
```

```
mutate(SCSTRESS=recode(SCSTRESS,
                        '0' = 1,
                        '1' = 2,
                        '2' = 3,
                        '3' = 4,
                        '4' = 5,
                        '5' = 6,
                        '6' = 7))
```

Exclusions

```
#include participants who answered the well-being module
data_short <- data_filled %>%
  filter(WBELIG==1) %>%
  subset(select = -c(WBELIG, RECTYPE))

#randomly selected to report well-being during leisurely activites (ATUS activity coding category 12 an
data_short <- data_short %>%
  filter(ACTIVITY < 140000) %>%
  filter(ACTIVITY >=120000) %>%
  filter(SCHAPPY %in% (1:7)) %>%
  filter(SCSTRESS %in% (1:7)) %>%
  filter(MEANING %in% (1:7)) %>%
  filter(PAIDHOUR != 99) %>% #don't have information on hourly payment status
  filter(UHRSWORKT != 9999) %>% #don't have number of hours worked
  filter(UHRSWORKT != 9995) %>%#responded as hours varied
  filter(EARNWEEK != 0)

data_short <- data_short %>%
  group_by(CASEID) %>%
  filter(n()>=2) %>%
  mutate (CASEID=cur_group_id())%>%
  ungroup()
```

Variable Description

- CASEID: unique identifying number for participant
- DAY: whether the day of the week of the diary day was a weekday (DAY = 1) or a weekend (DAY = 2)
- AGE: person's age in years as of his/her last birthday
- SEX: whether the individual is male (SEX = 1) or female (SEX = 2)
- MARST: person's marital status, either not married (MARST = 1) or married (MARST = 2)
- FULLPART: indicates whether the individual usually works full time (FULLPART = 1) or part time (FULLPART = 2).
- UHRSWORKT: total number of hours the respondent usually works per week
- EARNWEEK: income per week
- PAIDHOUR: whether the respondent is paid on an hourly basis (PAIDHOUD = 2) or not (PAIDHOUR = 1)
- HH_NUMOWNKIDS: the number of the respondent's own children under the age of 18 who live in the household

3

- DURATION: length of the activity in minutes
- SCHAPPY: seven point scale that indicates how much happiness the respondent felt during (1 - Not At all, 7 - Very) the activity.
- SCSTRESS: seven point scale that indicates how much stress the respondent felt during the activity (1 - Not At all, 7 - Very)
- MEANING: seven point scale that indicates how meaningful the activity was to the respondent (1 - Not At all, 7 - Very)

Model Equation

Lv-1:

$$\text{SCHAPPY}_{ij} = \beta_{0j} + \beta_{1j}\text{DURATION}_{ij} + e_{ij}$$

Lv-2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{PAIDHOUR}_j + \gamma_{02}\log(\text{INCOME})_j + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Descriptive Statistics

```
summary(data_short)
```

```
##      CASEID           DAY             AGE             SEX
##  Min.   :  1.0   Min.   :1.000   Min.   :15.00   Min.   :1.000
##  1st Qu.:109.0   1st Qu.:1.000   1st Qu.:33.00   1st Qu.:1.000
##  Median :214.5   Median :2.000   Median :43.00   Median :1.000
##  Mean   :215.4   Mean   :1.616   Mean   :44.14   Mean   :1.435
##  3rd Qu.:322.2   3rd Qu.:2.000   3rd Qu.:55.00   3rd Qu.:2.000
##  Max.   :429.0   Max.   :2.000   Max.   :85.00   Max.   :2.000
##      MARST          FULLPART        UHRSWORKT        EARNWEEK
##  Min.   :1.000   Min.   :1.000   Min.   : 4.00   Min.   :  35.0
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:40.00   1st Qu.: 593.2
##  Median :1.000   Median :1.000   Median :40.00   Median : 990.4
##  Mean   :1.456   Mean   :1.184   Mean   :39.24   Mean   :1190.7
##  3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:42.25   3rd Qu.:1634.2
##  Max.   :2.000   Max.   :2.000   Max.   :70.00   Max.   :2884.6
##     PAIDHOUR      HH_NUMOWNKIDS      WB_RESP   Leisure_All        ACTLINE
##  Min.   :1.000   Min.   :0.0000   Min.   :1   Min.   :   0.0   Min.   : 1.000
##  1st Qu.:1.000   1st Qu.:0.0000   1st Qu.:1   1st Qu.: 198.0   1st Qu.: 5.000
##  Median :2.000   Median :0.0000   Median :1   Median : 322.0   Median : 9.000
##  Mean   :1.536   Mean   :0.4766   Mean   :1   Mean   : 359.6   Mean   : 9.901
##  3rd Qu.:2.000   3rd Qu.:1.0000   3rd Qu.:1   3rd Qu.: 515.0   3rd Qu.:13.000
##  Max.   :2.000   Max.   :4.0000   Max.   :1   Max.   :1145.0   Max.   :35.000
##     ACTIVITY         DURATION        SCHAPPY          SCSTRESS        MEANING
##  Min.   :120101   Min.   :   5    Min.   :1.000   Min.   :1.00    Min.   :1.000
##  1st Qu.:120303   1st Qu.:  45    1st Qu.:4.000   1st Qu.:1.00    1st Qu.:3.000
##  Median :120303   Median :  90    Median :6.000   Median :1.00    Median :5.000
##  Mean   :121337   Mean   : 120    Mean   :5.295   Mean   :1.98    Mean   :4.831
##  3rd Qu.:120307   3rd Qu.: 160    3rd Qu.:7.000   3rd Qu.:3.00    3rd Qu.:7.000
##  Max.   :130299   Max.   :1105    Max.   :7.000   Max.   :7.00    Max.   :7.000
```

```
data_short %>%
    select(PAIDHOUR, EARNWEEK, UHRSWORKT, DURATION) %>%
    psych::pairs.panels(ellipses = FALSE, cex = 0.2, cex.cor = 1)
```
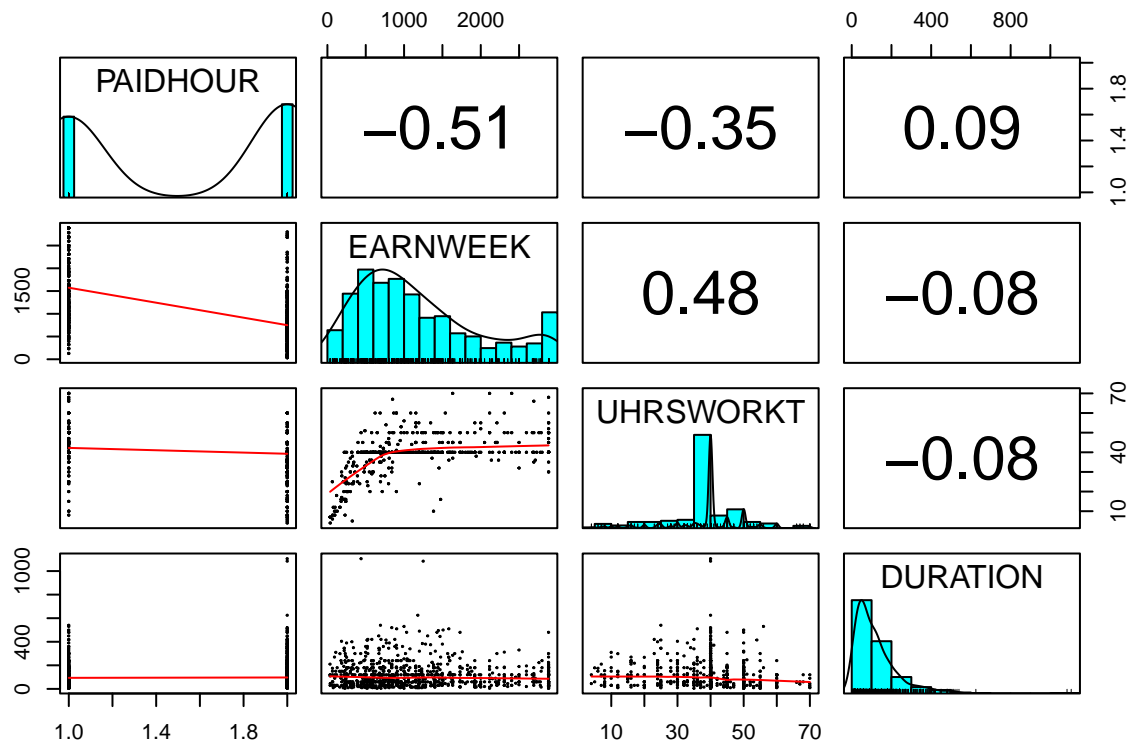


```
data_short %>%
    select(SCHAPPY, MEANING, SCSTRESS) %>%
    psych::pairs.panels(ellipses = FALSE, cex = 0.2, cex.cor = 1)
```

ICC
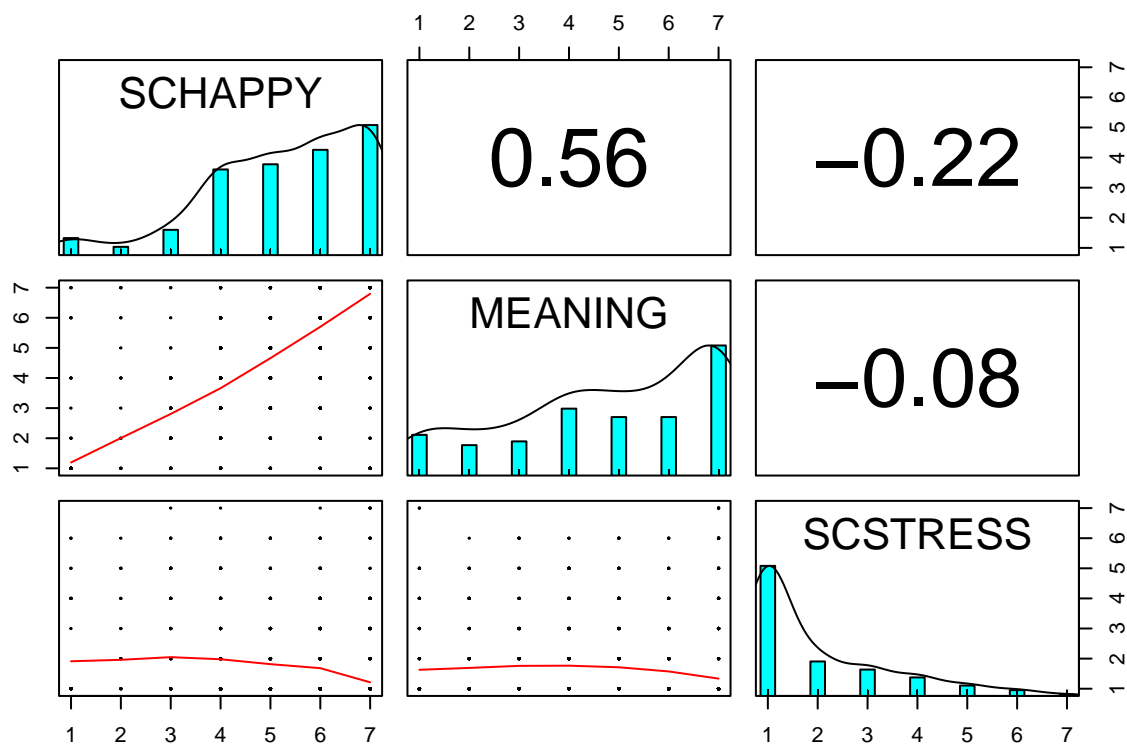
```r
m0_hap <- lmer(SCHAPPY ~ (1 | CASEID), data = data_short)
performance::icc(m0_hap)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.621
##   Unadjusted ICC: 0.621
```

Multilevel Analysis

```r
m1 <- lmer(SCHAPPY ~ PAIDHOUR * DURATION * log(EARNWEEK) + (1 | CASEID),
          data = data_short)
```

```
## Warning: Some predictor variables are on very different scales: consider
## rescaling
```

```
## Warning: Some predictor variables are on very different scales: consider
## rescaling
```

```
summary(m1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: SCHAPPY ~ PAIDHOUR * DURATION * log(EARNWEEK) + (1 | CASEID)
##    Data: data_short
##
## REML criterion at convergence: 3146.5
##
## Scaled residuals:
##     Min     1Q  Median     3Q     Max
## -3.7136 -0.4293  0.1439  0.4167  3.8653
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  CASEID   (Intercept) 1.5191   1.2325
##  Residual             0.8997   0.9485
## Number of obs: 896, groups:  CASEID, 429
##
## Fixed effects:
##                                Estimate Std. Error         df t value
## (Intercept)                    1.971710   3.148857 723.401164   0.626
## PAIDHOUR                       1.612238   1.758803 720.539556   0.917
## DURATION                       0.016549   0.017104 642.149268   0.968
## log(EARNWEEK)                  0.412585   0.438300 724.487191   0.941
## PAIDHOUR:DURATION             -0.005422   0.009415 657.569120  -0.576
## PAIDHOUR:log(EARNWEEK)        -0.198693   0.249987 719.618843  -0.795
## DURATION:log(EARNWEEK)        -0.002087   0.002409 645.577581  -0.866
## PAIDHOUR:DURATION:log(EARNWEEK) 0.000658  0.001344 662.059919   0.490
##                              Pr(>|t|)
## (Intercept)                     0.531
## PAIDHOUR                        0.360
## DURATION                        0.334
## log(EARNWEEK)                   0.347
## PAIDHOUR:DURATION               0.565
## PAIDHOUR:log(EARNWEEK)          0.427
## DURATION:log(EARNWEEK)          0.387
## PAIDHOUR:DURATION:log(EARNWEEK) 0.625
##
## Correlation of Fixed Effects:
##                 (Intr) PAIDHOUR DURATION l(EARN PAIDHOUR:DURATION PAIDHOUR:(
## PAIDHOUR        -0.969
## DURATION        -0.582  0.562
## l(EARNWEEK)     -0.996  0.972    0.584
## PAIDHOUR:DURATION 0.570 -0.585   -0.973   -0.576
## PAIDHOUR:(E      0.952 -0.994   -0.556   -0.965  0.584
## DURATION:(E      0.577 -0.561   -0.996   -0.584  0.974            0.559
## PAIDHOUR:DURATION: -0.561 0.581  0.961    0.571 -0.996           -0.586
##                 DURATION:
## PAIDHOUR
## DURATION
## l(EARNWEEK)
## PAIDHOUR:DURATION
```

```
## PAIDHOUR:(E
## DURATION:(E
## PAIDHOUR:DURATION: -0.969
## fit warnings:
## Some predictor variables are on very different scales: consider rescaling
```

```r
#Covariate
m1_c <- lmer(SCHAPPY ~ DURATION * PAIDHOUR * log(EARNWEEK) + AGE + MARST + DAY + UHRSWORKT + HH_NUMOWNK
           data = data_short)
```

```
## Warning: Some predictor variables are on very different scales: consider
## rescaling
```

```
## Warning: Some predictor variables are on very different scales: consider
## rescaling
```

```r
summary(m1_c)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: SCHAPPY ~ DURATION * PAIDHOUR * log(EARNWEEK) + AGE + MARST +
##     DAY + UHRSWORKT + HH_NUMOWNKIDS + (1 | CASEID)
##    Data: data_short
##
## REML criterion at convergence: 3165.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7147 -0.4083  0.1173  0.4246  3.8355
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  CASEID   (Intercept) 1.5199   1.2328
##  Residual             0.9002   0.9488
## Number of obs: 896, groups:  CASEID, 429
##
## Fixed effects:
##                                Estimate Std. Error         df t value
## (Intercept)                   1.936e+00  3.198e+00  7.075e+02   0.605
## DURATION                      1.739e-02  1.713e-02  6.398e+02   1.015
## PAIDHOUR                      1.658e+00  1.779e+00  7.067e+02   0.932
## log(EARNWEEK)                 3.501e-01  4.419e-01  7.137e+02   0.792
## AGE                           4.091e-03  4.875e-03  4.225e+02   0.839
## MARST                         7.145e-02  1.618e-01  4.189e+02   0.442
## DAY                           8.577e-03  1.415e-01  4.342e+02   0.061
## UHRSWORKT                     2.526e-03  8.622e-03  4.205e+02   0.293
## HH_NUMOWNKIDS                 1.318e-01  8.795e-02  4.228e+02   1.498
## DURATION:PAIDHOUR            -6.302e-03  9.432e-03  6.543e+02  -0.668
## DURATION:log(EARNWEEK)       -2.201e-03  2.412e-03  6.432e+02  -0.912
## PAIDHOUR:log(EARNWEEK)       -2.054e-01  2.527e-01  7.061e+02  -0.813
## DURATION:PAIDHOUR:log(EARNWEEK)  7.826e-04  1.347e-03  6.589e+02   0.581
##                               Pr(>|t|)
## (Intercept)                      0.545
```

```
## DURATION                              0.310
## PAIDHOUR                              0.352
## log(EARNWEEK)                         0.428
## AGE                                   0.402
## MARST                                 0.659
## DAY                                   0.952
## UHRSWORKT                             0.770
## HH_NUMOWNKIDS                         0.135
## DURATION:PAIDHOUR                     0.504
## DURATION:log(EARNWEEK)               0.362
## PAIDHOUR:log(EARNWEEK)               0.416
## DURATION:PAIDHOUR:log(EARNWEEK)      0.561


##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)   or
##      vcov(x)          if you need it


## fit warnings:
## Some predictor variables are on very different scales: consider rescaling
```

```r
msummary(list(
    "M1" = m1,
    "M1-Covarites" = m1_c
))
```

```
## Warning: 'modelsummary' uses the 'performance' package to extract goodness-of-fit statistics from mo
## This warning is displayed once per session.
```

Results

```r
msummary(m1,
         estimate = c("{estimate} [{conf.low}, {conf.high}]"),
         statistic = NULL,  # suppress the extra rows for SEs
         shape = effect + term ~ model,
         title = "Table 1: Model coefficients")
```

If hourly workers have a greater propensity to make economic evaluation of time, they should display greater sensitivity to the opportunity cost of time and as a consequence, derive less happiness from leisurely activities. To explore people's subjective experience of leisurely activities, I used the happiness measure from the ATUS well-being module. To worker's hourly work status was dummy coded "1" for non-hourly workers and "2" for hourly workers. To account for the worker's opportunity cost of time, we used the duration spent on leisurely activities and their income as an indicator of the foregone earning from spending (wasting) time on leisure. To control for potential differences between hourly and non hourly workers, we controlled for various variables: age, marital status, number of hours worked, number of children under 18 and the day of diary. To test the hypothesis that happiness derived from engaging in leisurely activities is influenced by both hourly payment status and the opportunity cost of time, I fitted a multilevel model with hourly status, duration, and logged income on happiness to explore the three-way interaction among the three variables.

	M1	M1-Covarites
(Intercept)	1.972	1.936
	(3.149)	(3.198)
PAIDHOUR	1.612	1.658
	(1.759)	(1.779)
DURATION	0.017	0.017
	(0.017)	(0.017)
log(EARNWEEK)	0.413	0.350
	(0.438)	(0.442)
PAIDHOUR × DURATION	−0.005	
	(0.009)	
PAIDHOUR × log(EARNWEEK)	−0.199	−0.205
	(0.250)	(0.253)
DURATION × log(EARNWEEK)	−0.002	−0.002
	(0.002)	(0.002)
PAIDHOUR × DURATION × log(EARNWEEK)	0.001	
	(0.001)	
SD (Intercept)	1.233	1.233
SD (Observations)	0.949	0.949
AGE		0.004
		(0.005)
MARST		0.071
		(0.162)
DAY		0.009
		(0.141)
UHRSWORKT		0.003
		(0.009)
HH_NUMOWNKIDS		0.132
		(0.088)
DURATION × PAIDHOUR		−0.006
		(0.009)
DURATION × PAIDHOUR × log(EARNWEEK)		0.001
		(0.001)
Num.Obs.	896	896
RMSE	0.75	0.75

Table 1: Table 1: Model coefficients

	Model 1
(Intercept)	1.972 [−4.208, 8.152]
PAIDHOUR	1.612 [−1.840, 5.064]
DURATION	0.017 [−0.017, 0.050]
log(EARNWEEK)	0.413 [−0.448, 1.273]
PAIDHOUR × DURATION	−0.005 [−0.024, 0.013]
PAIDHOUR × log(EARNWEEK)	−0.199 [−0.689, 0.292]
DURATION × log(EARNWEEK)	−0.002 [−0.007, 0.003]
PAIDHOUR × DURATION × log(EARNWEEK)	0.001 [−0.002, 0.003]
SD (Intercept)	1.233
SD (Observations)	0.949
Num.Obs.	896
RMSE	0.75