

Analyses

Philip & Philipp

14 Apr 2016

```
if(!file.exists("data_final.rda")) {  
  source('data_combine.R')  
}  
  
# Opens data_final.rda if not already loaded as an object  
if(!exists("z.df" )) {  
  load("data_final.rda")  
}
```

Motivation

- Populate from google docs

Descriptive results

Distribution of overall and life satisfaction

!The distribution of the well-being variables should probably be presented in a table.!

```
happy_dist <- table(z$happy)/sum(table(z$happy))*100  
satjob_dist <- table(z$satjob)/sum(table(z$happy))*100  
  
par(mfrow=c(1,2))  
  
barplot(happy_dist,  
        col=c("navyblue"),  
        xlab = "Overall life satisfaction categories",  
        legend = c("Life satisfaction"),  
        ylim = c(0, 80))  
  
barplot(satjob_dist,  
        col=c("navyblue"),  
        xlab = "Job satisfaction categories",  
        legend = c("Job satisfaction"),  
        ylim = c(0, 80))
```

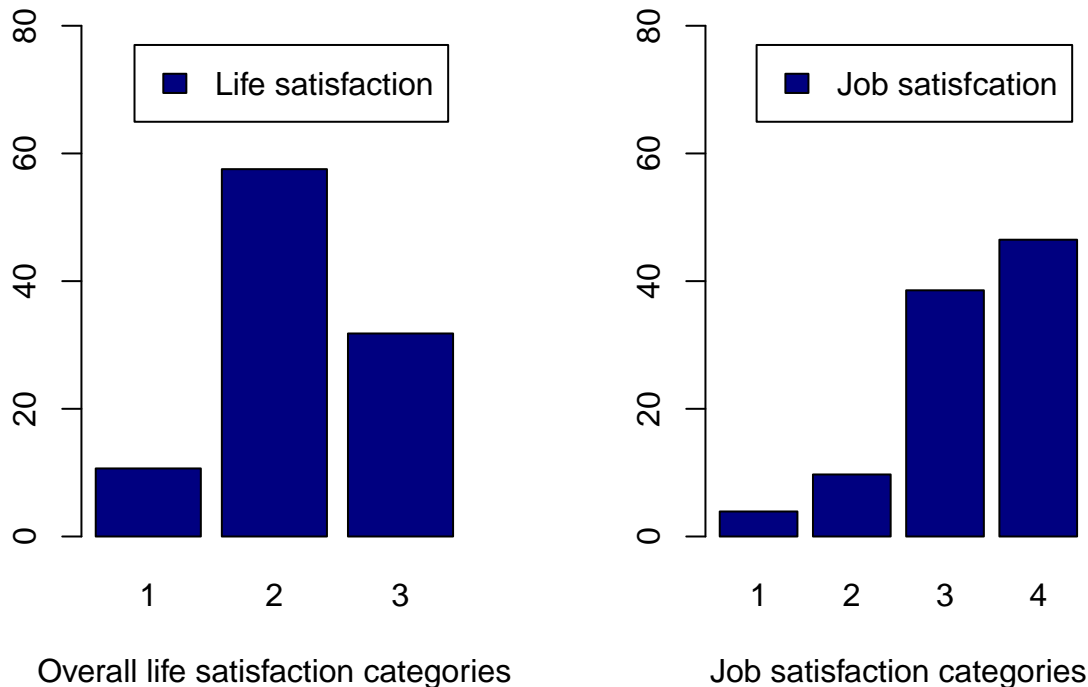


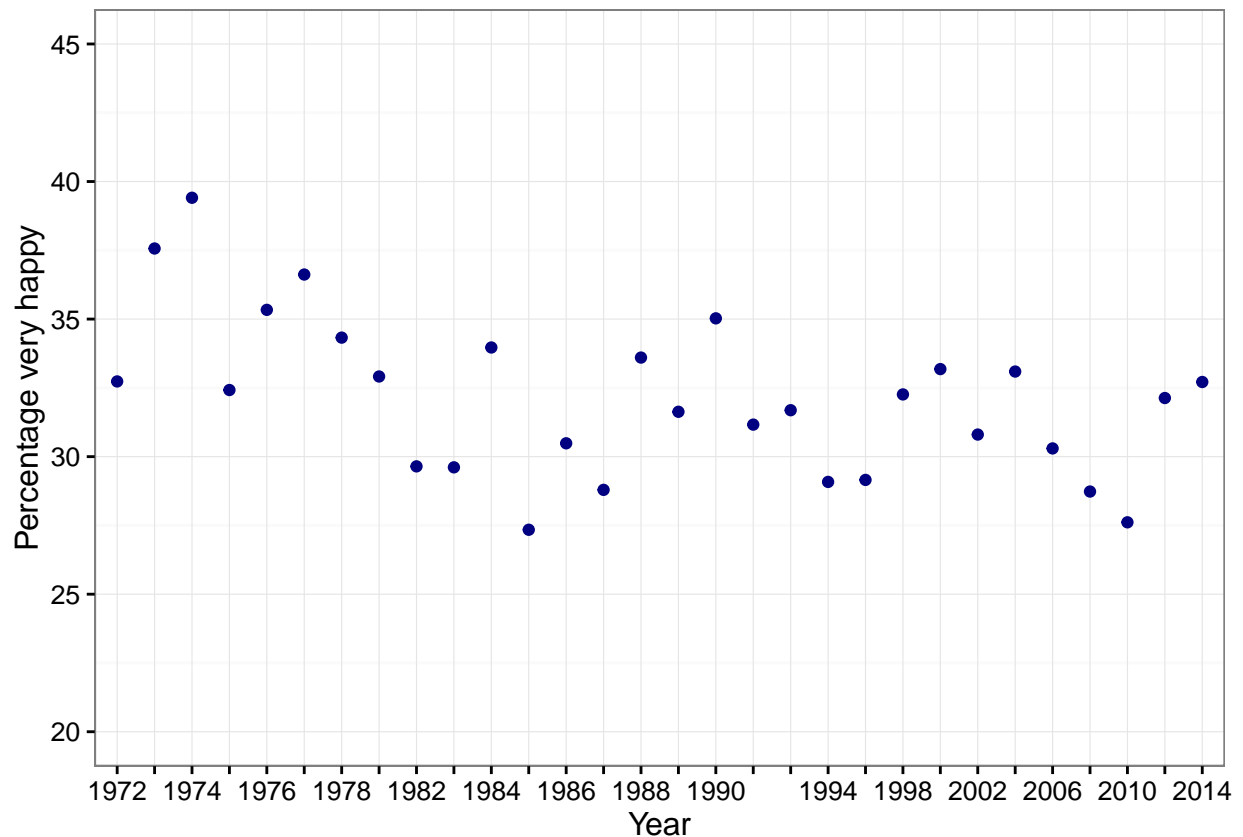
Figure x shows the distribution of survey respondents' answer to the two questions: "Taken all together, how would you say things are these days?" and "On the whole, how satisfied are you with the work you do?". The former is measured on a three point scale, of which around 60 % of the sample respond a mid-level happiness level. On the job-satisfaction variable, a majority of the sample responds either moderate or high satisfaction (score 3 and 4).

Happiness in different survey years

The GSS is conducted between 1972 and 2014. Due to year specific events, unintended differences in how the survey was implemented or trends in overall happiness, there could be year-specific differences. Figure x shows the average share of the population who report high-happiness.

```
t <- z
t$vhappy <- t$vhappy*100
ggplot(t, aes(x=factor(year), y=vhappy)) +
  stat_summary(fun.y="mean", geom="point", col="Navyblue") +
  scale_x_discrete("Year", labels = c("1973" = "", "1975" = "", "1977" = "", "1980" = "", "1983" = "",
  scale_y_continuous("Percentage very happy") +
  expand_limits(y=c(20,45)) +
  theme_bw()
```

```
## Warning: Removed 2461 rows containing non-finite values (stat_summary).
```



The figure shows that there is considerable variation between years, and a weakly negative trend particularly between 1972 and 1983. It is not directly possible to disentangle what is caused by random noise and what is caused by structural differences, however, it signifies that it is pragmatic to control for survey year in a regression design to control for survey-year fixed effects.

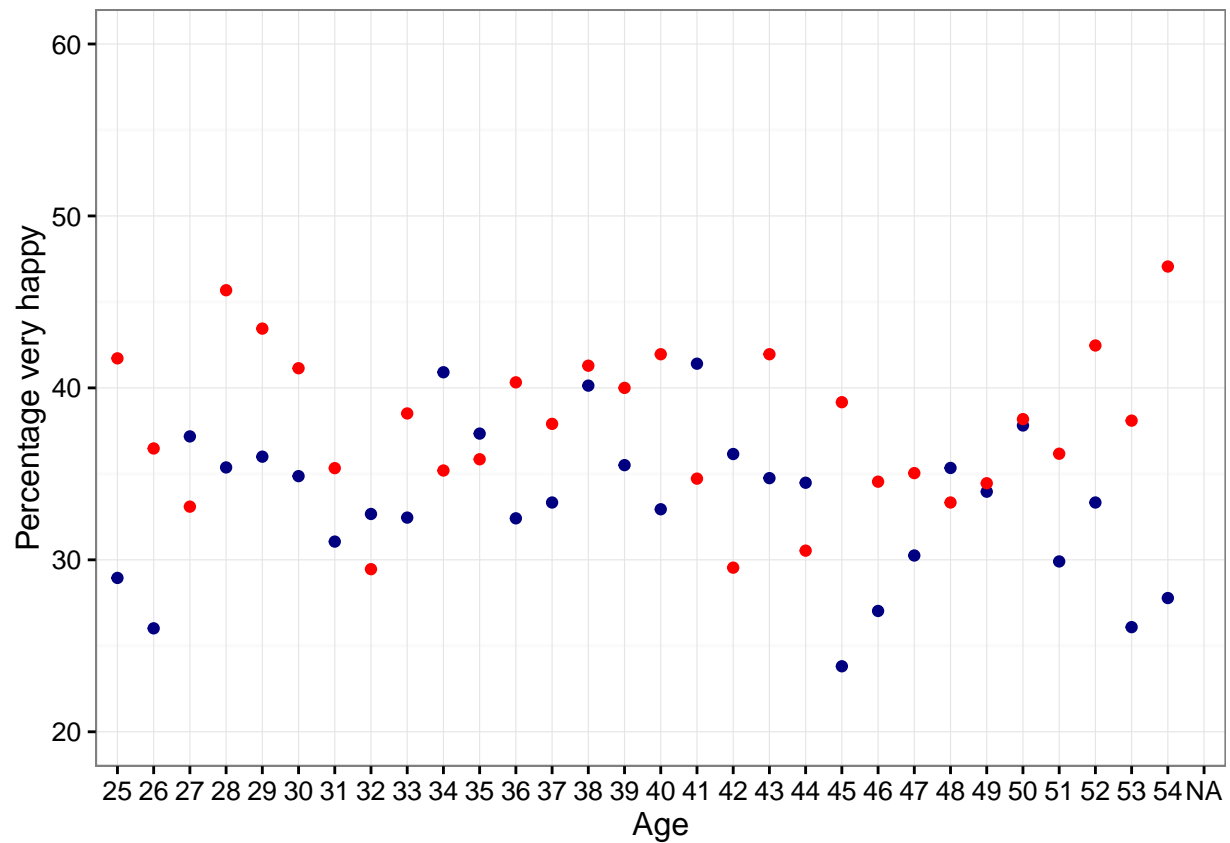
Happiness across age for college educated men and women

Previous studies have found that there is a correlation between life satisfaction and age. However, this is not apparent. Likewise, there could be gender specific differences.

```
ggplot() +
  stat_summary(data = t[t$sex == 1 & t$educat == 4,], aes(x=factor(age), y=vhappy),
    fun.y="mean", geom="point", col="Navyblue") +
  stat_summary(data = t[t$sex == 2 & t$educat == 4,], aes(x=factor(age), y=vhappy),
    fun.y="mean", geom="point", col="Red") +
  scale_x_discrete("Age") +
  scale_y_continuous("Percentage very happy") +
  expand_limits(y=c(20,60)) +
  theme_bw()
```

```
## Warning: Removed 374 rows containing non-finite values (stat_summary).
```

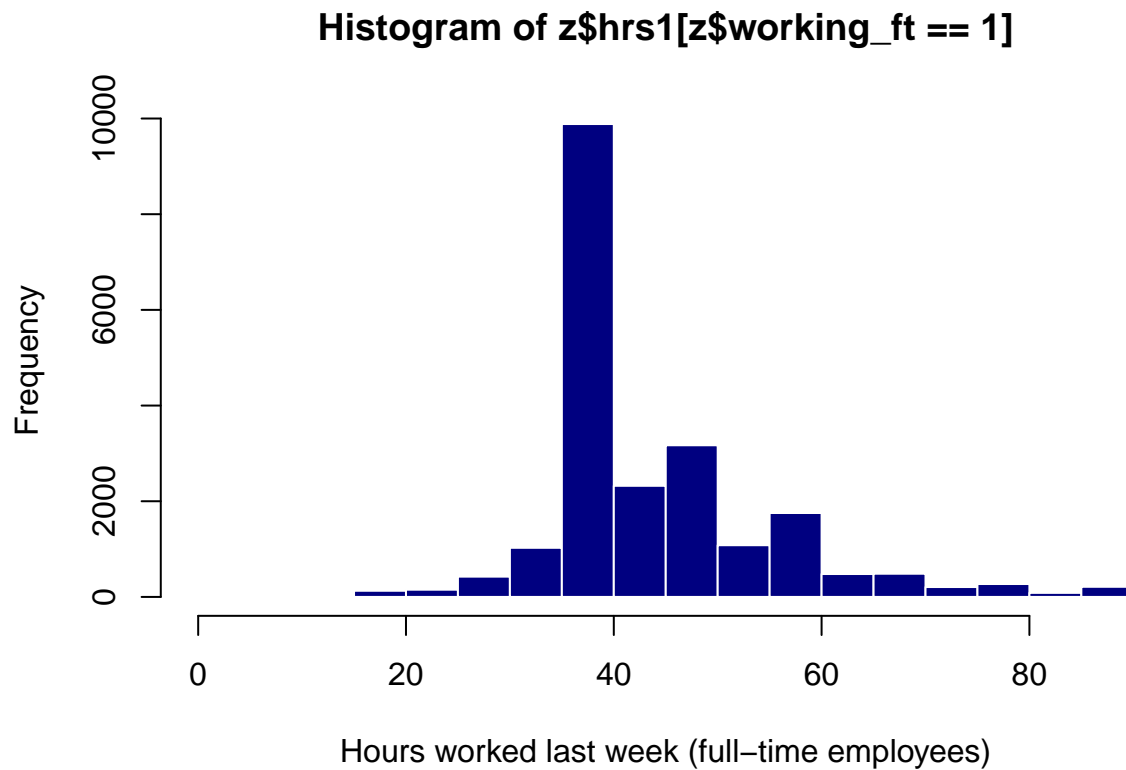
```
## Warning: Removed 439 rows containing non-finite values (stat_summary).
```



College educated women has a slightly higher reported happiness overall relative to men (38% vs. 34%).

Distribution of work-hours for full-time workers

```
hist(z$hrs1[z$working_ft==1],
     col=c("navyblue"),
     border = "White",
     xlab = "Hours worked last week (full-time employees)"
)
```



Analysis

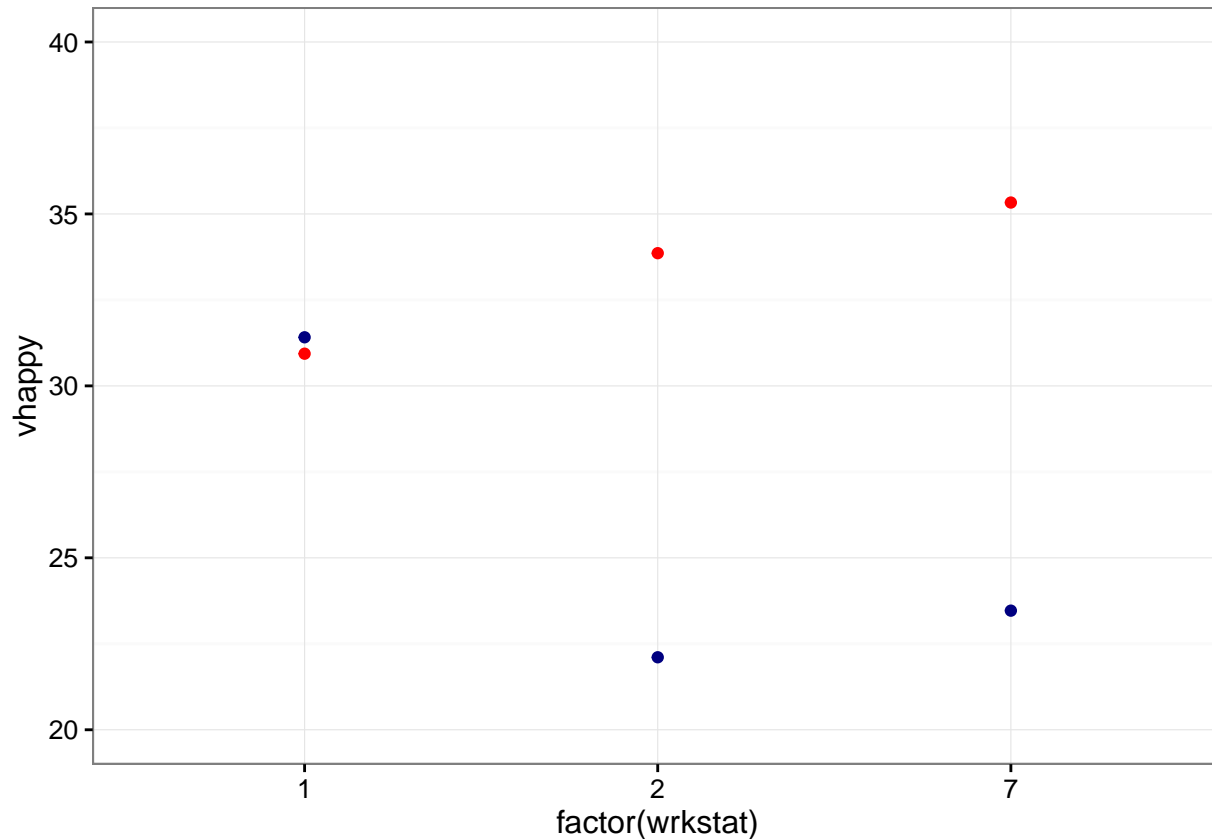
Average reported happiness and workstatus for men and women

This graph is probably redundant. Perhaps the text can be used.

```
ggplot() +
  stat_summary(data = t[t$sex == 2 ,], aes(x=factor(wrkstat), y=vhappy),
    fun.y="mean", geom="point", col="Red") +
  stat_summary(data = t[t$sex == 1 ,], aes(x=factor(wrkstat), y=vhappy),
    fun.y="mean", geom="point", col="Navyblue") +
  expand_limits(y=c(20,40)) +
  theme_bw()
```

```
## Warning: Removed 1366 rows containing non-finite values (stat_summary).
```

```
## Warning: Removed 1095 rows containing non-finite values (stat_summary).
```



The figure shows that women in full-time positions are less likely to report high life-satisfaction than women in both part-time jobs and women who do housework. The opposite is apparent for men, where full-time employed shows the highest levels of life happiness. Explanations?: Income is a very likely omitted variable, insofar men primarily are the bread-winner, then part-time jobs could proxy low household income.

Gender

In contemporary society the strive for gender equality often clashes with labour market realities and the persistence of traditional gender norms. Previous studies found that the reconcilability of having a family and pursuing a successful career is still difficult to achieve for women while it is rather the norm for men (e.g. Hipp reference). Thus, women who clash with traditional customs might face a “life happiness penalty” as shown in Bertrand (2013).

Figure 3 shows that gender only plays a role when people do not have a high income. For women (red) the probability to be very happy does not change with having a low income. Men (blue) on the other hand face a large happiness reduction when having a low income [the sample is limited to people with college education]. Figure 4 further differentiates in four possible combinations of having a family (married and kids) and having a high income job. Both, men and women, are happier when having a family. When not having a family, higher income improves life satisfaction for both genders although the increase is slightly larger for men. Differences become more pronounced when having a family. With a family but without a career women are the happiest, while men are considerably less happy, compared to having a career and a family. Thus, men and women interestingly have different career-family configurations under which they appear most happy.

Figure 3: Happiness for women and men depending whether they have a career

Figure 4: Happiness for women and men depending on combination of having a career and a family

```
z$meanhap <- NA
z$meanhap[z$family==0 & z$career==0] <- "No career, no family"
```

```

z$meanhap[z$family==0 & z$career==1] <- "Career, no family"
z$meanhap[z$family==1 & z$career==0] <- "No career, family"
z$meanhap[z$family==1 & z$career==1] <- "Career, family"

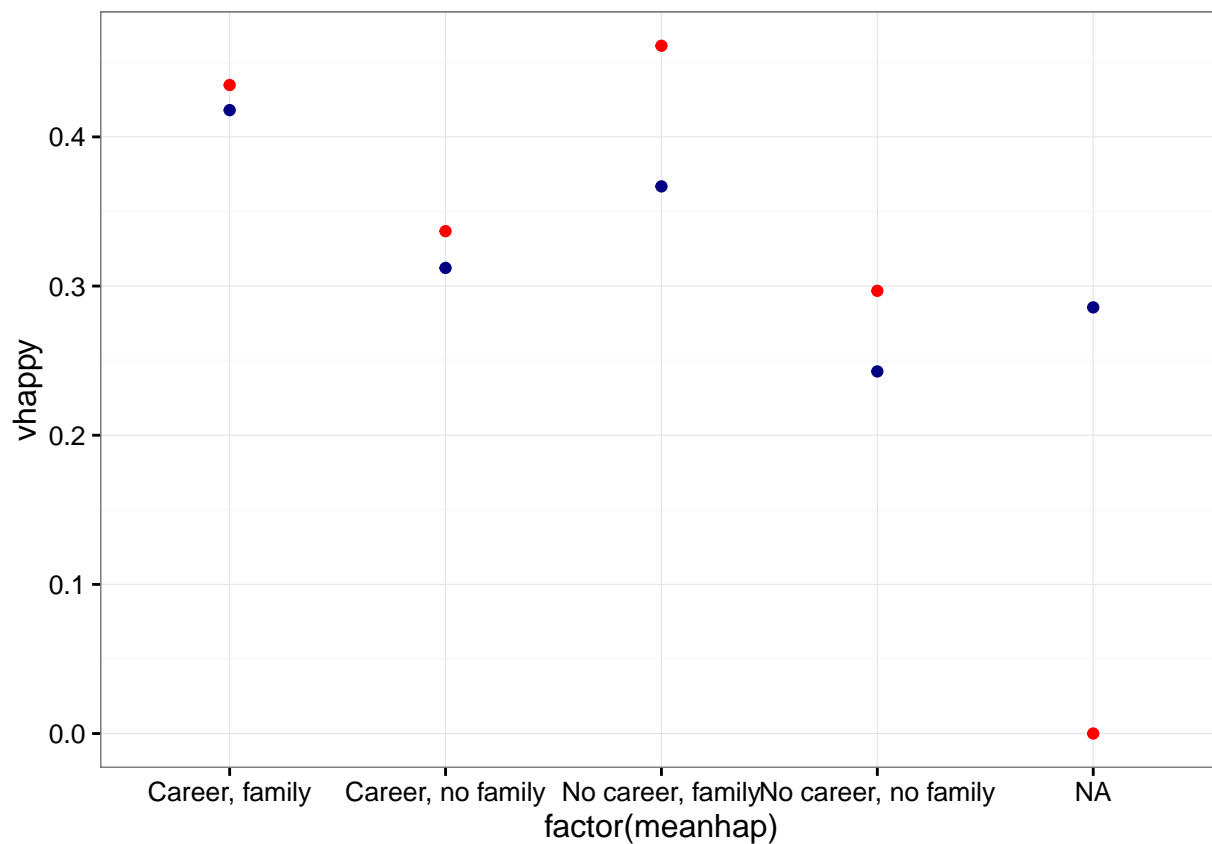
t <- subset(z, z$educat==4)

ggplot() +
  stat_summary(data = t[t$sex == 1,], aes(x=factor(meanhap), y=vhappy),
    fun.y="mean", geom="point", col="Navyblue") +
  stat_summary(data = t[t$sex == 2,], aes(x=factor(meanhap), y=vhappy),
    fun.y="mean", geom="point", col="Red") +
  theme_bw()

```

```
## Warning: Removed 346 rows containing non-finite values (stat_summary).
```

```
## Warning: Removed 400 rows containing non-finite values (stat_summary).
```



Interaction effects of married*career for working men and women

In a second step we replicate a linear regression model by Bertrand (2013) which estimates the interaction effect of being married and having a high paid job (career) on the binary variable being very happy. While Bertrand (2013) limits her analysis on college-educated women who are working we compare these findings to the respective male group. The model controls for age, age-squared, the survey year, race and decade of birth.

Figure 5 shows the effect of the interaction term on the probability of being very happy. On the left-hand side the effect of marriage on life satisfaction is stronger for women who do not occupy a high paying job. For women who have a high paying job, however, the effect of marriage on life satisfaction is much weaker. Although this difference is not significant at the 5% level, the career effect is much stronger than for men where having a career or not hardly influences the effect of marriage on happiness.

Figure 5: Interaction effects of being married and high income job on life satisfaction

```
z$working_ft <- as.numeric(z$working_ft)
z$working_pt <- as.numeric(z$working_pt)

M1a <- lm(vhappy ~ career*married + age + agesq + as.factor(year) + as.factor(race) + as.factor(bdec),
         data = subset(z, sex==2 & educat == 4))
M2a <- lm(vhappy ~ career*married + age + agesq + as.factor(year) + as.factor(race) + as.factor(bdec),
         data = subset(z, sex==1 & educat == 4))

# compare men and women

require("interplot")
require("gridExtra")

interM1 <- interplot(M1a, var1 = "married", var2 = "career") +
  ggtitle("Working women") +
  xlab("career") +
  ylab("effect of marriage on life satisfaction") +
  expand_limits(y=c(0.05,0.3)) +
  theme_bw()

interM2 <- interplot(M2a, var1 = "married", var2 = "career") +
  ggtitle("Working men") +
  xlab("career") +
  expand_limits(y=c(0.05,0.3)) +
  theme_bw()

grid.arrange(interM1, interM2, ncol = 2)
```

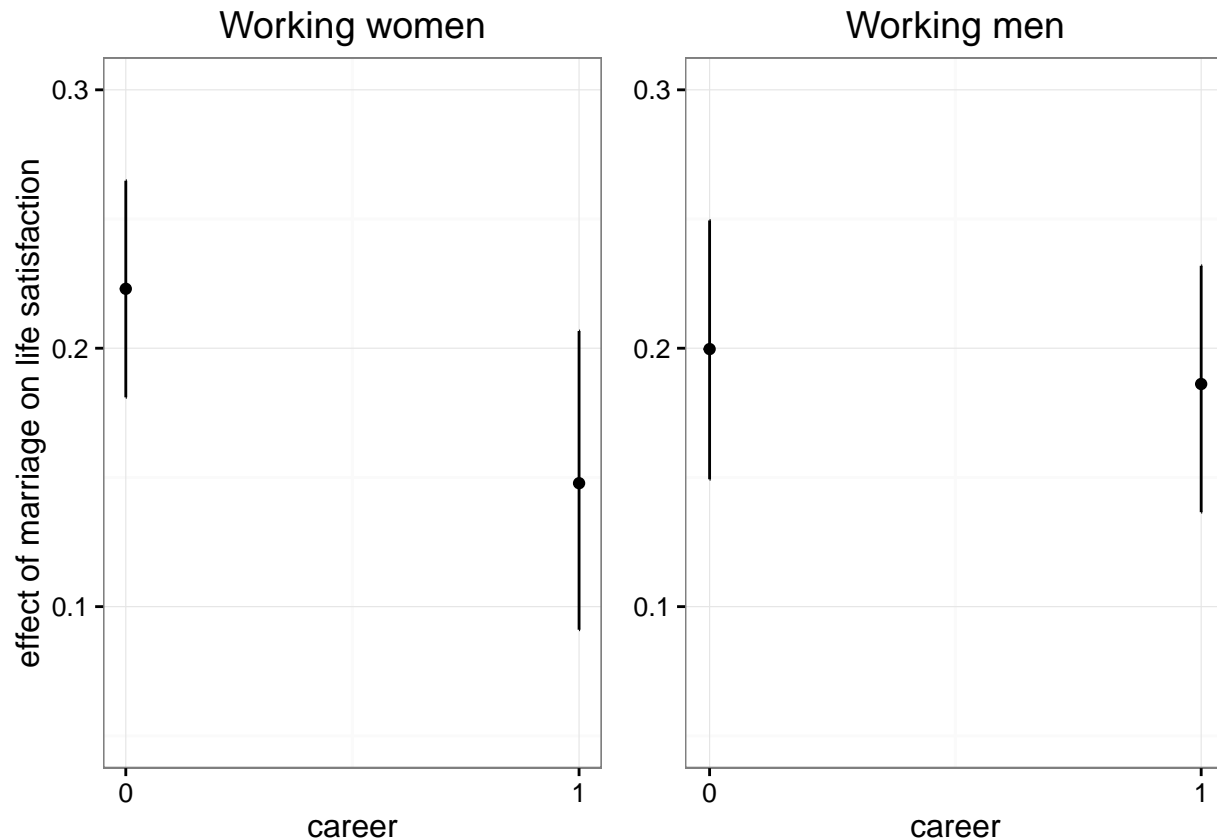



Figure 6: Interaction effects of family-career for working men and women

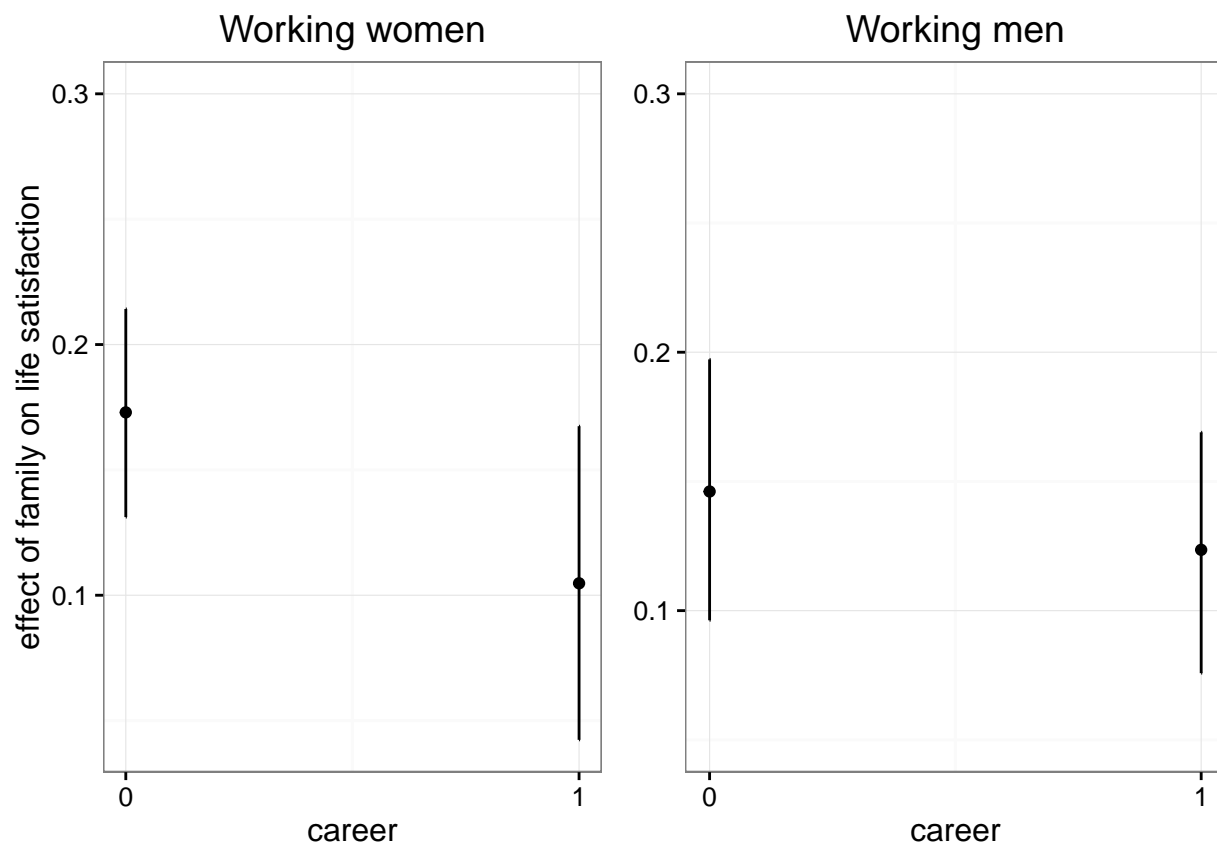
```
M3 <- lm(vhappyy ~ career*family + age + agesq + as.factor(year) + as.factor(race) + as.factor(bdec),
        data = subset(z, sex==2 & educat == 4))
M4 <- lm(vhappyy ~ career*family + age + agesq + as.factor(year) + as.factor(race) + as.factor(bdec),
        data = subset(z, sex==1 & educat == 4))

# Compare men and women (Career-Family interaction)

interM3 <- interplot(M3, var1 = "family", var2 = "career") +
  ggtitle("Working women") +
  xlab("career") +
  ylab("effect of family on life satisfaction") +
  expand_limits(y=c(0.05,0.3)) +
  theme_bw()

interM4 <- interplot(M4, var1 = "family", var2 = "career") +
  ggtitle("Working men") +
  xlab("career") +
  ylab("effect of family on life satisfaction") +
  expand_limits(y=c(0.05,0.3)) +
  theme_bw()

grid.arrange(interM3, interM4, ncol = 2)
```



Software and packages used for the analysis

The analysis is done in R [CiteR] with the use of the following packages: “ggplot2” [R-ggplot2], “repmis” [R-repmis], “plyr” [R-plyr], “dplyr” [R-dplyr], “MASS” [R-MASS], “Hmisc” [R-Hmisc], “interplot” [R-interplot] and “gridExtra” [R-gridExtra]

References