Appendix

Importing data

```
> setwd("/Users/yangsong/Desktop/MyFolder/Academics/Courses/Spring_2018/MATH 352/project")
> mydata = read.csv("Benefits.csv")
> mydata$X <- NULL
```

Setting base levels for qualitative variables

```
mydata$joblost <- relevel(as.factor(trimws(mydata$joblost)), "other")
> mydata$nwhite <- relevel(as.factor(trimws(mydata$nwhite)), "no")
> mydata$school12 <- relevel(as.factor(trimws(mydata$school12)), "no")
> mydata$sex <- relevel(as.factor(trimws(mydata$sex)), "female")
> mydata$married <- relevel(as.factor(trimws(mydata$married)), "no")
> mydata$dkids <- relevel(as.factor(trimws(mydata$dkids)), "no")
> mydata$dykids <- relevel(as.factor(trimws(mydata$dykids)), "no")
> mydata$head <- relevel(as.factor(trimws(mydata$head)), "no")
> mydata$ui <- relevel(as.factor(trimws(mydata$ui)), "no")
```

Model 1：A linear model of almost all the variables

```
> model1 = lm(stateur ~ statemb + age + tenure + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + rr + head + ui,data = mydata)
>
> summary(model1)

Call:
lm(formula = stateur ~ statemb + age + tenure + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + rr +
    head + ui, data = mydata)

Residuals:
   Min     1Q  Median     3Q    Max
-5.4869 -1.4647 -0.0359  1.2467  9.0040

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    9.0630457  0.2317056  39.114  < 2e-16 ***
statemb        0.0080839  0.0009299   8.693  < 2e-16 ***
age           -0.0028225  0.0036320  -0.777 0.437133
tenure         0.0045395  0.0056327   0.806 0.420331
```

```
joblostposition_abolished  0.0419860  0.1152995   0.364 0.715764
joblostseasonal_job_ended -0.0788870  0.1661096  -0.475 0.634873
joblostslack_work          0.2513059  0.0663835   3.786 0.000155 ***
nwhiteyes                  0.1027083  0.0864991   1.187 0.235131
school12yes               -0.0246093  0.0777220  -0.317 0.751536
sexmale                    0.1711056  0.0826151   2.071 0.038400 *
marriedyes                -0.1583075  0.0748613  -2.115 0.034508 *
dkidsyes                   0.2039983  0.0777428   2.624 0.008717 **
dykidsyes                 -0.0224681  0.0908053  -0.247 0.804585
yrdispl                   -0.4825776  0.0124768 -38.678  < 2e-16 ***
rr                        -1.9523047  0.3187377  -6.125 9.78e-10 ***
headyes                   -0.1492884  0.0767043  -1.946 0.051678 .
uiyes                      0.4050584  0.0663045   6.109 1.08e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.091 on 4860 degrees of freedom
Multiple R-squared:  0.3029,      Adjusted R-squared:  0.3006
F-statistic:   132 on 16 and 4860 DF,  p-value: < 2.2e-16
```

Model 2：Deleted variables age and tenure since they are insignificant. Tried adding squares to statemb, rr, and yrdispl in modeltemp. Deleted rr^2 due to insignificance and arrived at model 2.

```
> modeltemp = lm(stateur ~ statemb + I(statemb^2)  + joblost + nwhite + school12 + sex + married +
dkids+ dykids + yrdispl + I(yrdispl^2) + rr + I(rr^2)  + head + ui,data = mydata)
>
> summary(modeltemp)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + I(rr^2) + head + ui, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-5.9954 -1.2742 -0.1477  1.2662  8.9125
```

```
Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                    1.322e+01  4.810e-01  27.479  < 2e-16 ***
statemb                       -2.533e-02  4.399e-03  -5.758 9.02e-09 ***
I(statemb^2)                   8.586e-05  1.159e-05   7.407 1.52e-13 ***
joblostposition_abolished     -1.881e-02  1.089e-01  -0.173  0.86285
joblostseasonal_job_ended     -1.012e-01  1.562e-01  -0.648  0.51694
joblostslack_work              1.388e-01  6.171e-02   2.250  0.02453 *
nwhiteyes                      8.252e-02  8.177e-02   1.009  0.31293
school12yes                   -7.474e-02  7.352e-02  -1.017  0.30939
sexmale                        1.088e-01  7.662e-02   1.420  0.15576
marriedyes                    -1.197e-01  6.729e-02  -1.778  0.07546 .
dkidsyes                       1.904e-01  7.283e-02   2.615  0.00895 **
dykidsyes                     -2.246e-02  8.233e-02  -0.273  0.78504
yrdispl                       -1.337e+00  4.399e-02 -30.389  < 2e-16 ***
I(yrdispl^2)                   7.781e-02  3.926e-03  19.817  < 2e-16 ***
rr                             9.579e-01  1.701e+00   0.563  0.57341
I(rr^2)                       -3.393e+00  2.126e+00  -1.596  0.11049
headyes                       -8.580e-02  7.014e-02  -1.223  0.22128
uiyes                          2.744e-01  6.245e-02   4.393 1.14e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.978 on 4859 degrees of freedom
Multiple R-squared:  0.3763,      Adjusted R-squared:  0.3741
F-statistic: 172.4 on 17 and 4859 DF,  p-value: < 2.2e-16
```

```
> model2 = lm(stateur ~ statemb + I(statemb^2)  + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr  + head + ui,data = mydata)
>
> summary(model2)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + head + ui, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-6.0219 -1.2802 -0.1378  1.2640  8.8990
```

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.362e+01	4.094e-01	33.270	< 2e-16	***
statemb	-2.446e-02	4.366e-03	-5.603	2.22e-08	***
I(statemb^2)	8.352e-05	1.150e-05	7.262	4.42e-13	***
joblostposition_abolished	-1.854e-02	1.089e-01	-0.170	0.86482	
joblostseasonal_job_ended	-1.016e-01	1.562e-01	-0.651	0.51527	
joblostslack_work	1.379e-01	6.171e-02	2.235	0.02547	*
nwhiteyes	8.633e-02	8.175e-02	1.056	0.29099	
school12yes	-8.083e-02	7.344e-02	-1.101	0.27109	
sexmale	1.112e-01	7.662e-02	1.452	0.14670	
marriedyes	-1.185e-01	6.730e-02	-1.761	0.07829	.
dkidsyes	1.910e-01	7.284e-02	2.623	0.00875	**
dykidsyes	-2.021e-02	8.233e-02	-0.245	0.80612	
yrdispl	-1.339e+00	4.398e-02	-30.443	< 2e-16	***
I(yrdispl^2)	7.805e-02	3.924e-03	19.888	< 2e-16	***
rr	-1.716e+00	2.982e-01	-5.754	9.28e-09	***
headyes	-8.687e-02	7.015e-02	-1.238	0.21565	
uiyes	2.780e-01	6.242e-02	4.454	8.63e-06	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.979 on 4860 degrees of freedom
Multiple R-squared: 0.3759, Adjusted R-squared: 0.3739
F-statistic: 183 on 16 and 4860 DF, p-value: < 2.2e-16

Model 3： Added interactions of statemb and other qualitative variables, observe the output.

```
> model3 = lm(stateur ~ statemb + I(statemb^2) + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr + head + ui + statemb*rr + I(statemb^2)*rr +statemb*joblost +
I(statemb^2)*joblost +statemb*nwhite + I(statemb^2)*nwhite +statemb*school12 +
I(statemb^2)*school12 +statemb*sex + I(statemb^2)*sex +statemb*married + I(statemb^2)*married
+statemb*dkids + I(statemb^2)*dkids +statemb*dykids + I(statemb^2)*dykids +statemb*yrdispl +
I(statemb^2)*yrdispl +statemb*head + I(statemb^2)*head +statemb*ui + I(statemb^2)*ui ,data =
mydata)
>
> summary(model3)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
```

```
    rr + head + ui + statemb * rr + I(statemb^2) * rr + statemb *
    joblost + I(statemb^2) * joblost + statemb * nwhite + I(statemb^2) *
    nwhite + statemb * school12 + I(statemb^2) * school12 + statemb *
    sex + I(statemb^2) * sex + statemb * married + I(statemb^2) *
    married + statemb * dkids + I(statemb^2) * dkids + statemb *
    dykids + I(statemb^2) * dykids + statemb * yrdispl + I(statemb^2) *
    yrdispl + statemb * head + I(statemb^2) * head + statemb *
    ui + I(statemb^2) * ui, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2557 -1.2357 -0.1198  1.2712  7.6172

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                      2.511e+01  2.067e+00  12.150  < 2e-16 ***
statemb                         -1.838e-01  2.336e-02  -7.870 4.34e-15 ***
I(statemb^2)                     6.363e-04  6.616e-05   9.618  < 2e-16 ***
joblostposition_abolished       -1.542e+00  1.501e+00  -1.027 0.304289
joblostseasonal_job_ended        1.566e+00  2.359e+00   0.664 0.506699
joblostslack_work                4.688e-01  8.192e-01   0.572 0.567201
nwhiteyes                        3.524e+00  9.863e-01   3.573 0.000357 ***
school12yes                     -1.315e+00  1.057e+00  -1.244 0.213664
sexmale                         -2.146e+00  1.001e+00  -2.144 0.032059 *
marriedyes                       1.269e+00  8.806e-01   1.441 0.149515
dkidsyes                         1.069e-01  9.891e-01   0.108 0.913914
dykidsyes                       -1.664e+00  1.111e+00  -1.498 0.134076
yrdispl                         -2.009e+00  1.580e-01 -12.722  < 2e-16 ***
I(yrdispl^2)                     1.121e-01  4.595e-03  24.396  < 2e-16 ***
rr                              -6.825e+00  3.587e+00  -1.902 0.057165 .
headyes                         -2.532e-01  9.426e-01  -0.269 0.788262
uiyes                            6.820e-01  8.858e-01   0.770 0.441412
statemb:rr                       6.024e-02  4.056e-02   1.485 0.137538
I(statemb^2):rr                 -1.756e-04  1.120e-04  -1.567 0.117192
statemb:joblostposition_abolished  1.743e-02  1.623e-02   1.074 0.282872
statemb:joblostseasonal_job_ended -1.710e-02  2.512e-02  -0.681 0.495978
statemb:joblostslack_work        -5.878e-03  8.887e-03  -0.661 0.508356
I(statemb^2):joblostposition_abolished -4.618e-05  4.243e-05  -1.089 0.276404
I(statemb^2):joblostseasonal_job_ended  4.404e-05  6.499e-05   0.678 0.498009
I(statemb^2):joblostslack_work   2.146e-05  2.345e-05   0.915 0.360076
statemb:nwhiteyes               -3.271e-02  1.079e-02  -3.031 0.002448 **
I(statemb^2):nwhiteyes           7.159e-05  2.858e-05   2.505 0.012264 *
statemb:school12yes              1.098e-02  1.121e-02   0.979 0.327439
I(statemb^2):school12yes        -2.169e-05  2.894e-05  -0.749 0.453637
```

```
statemb:sexmale                2.337e-02  1.086e-02   2.153 0.031406 *
I(statemb^2):sexmale          -5.772e-05  2.860e-05  -2.018 0.043630 *
statemb:marriedyes            -1.435e-02  9.526e-03  -1.507 0.131935
I(statemb^2):marriedyes        3.480e-05  2.502e-05   1.390 0.164448
statemb:dkidsyes              -2.123e-03  1.079e-02  -0.197 0.843984
I(statemb^2):dkidsyes          1.398e-05  2.859e-05   0.489 0.624783
statemb:dykidsyes              1.737e-02  1.214e-02   1.431 0.152558
I(statemb^2):dykidsyes        -4.350e-05  3.230e-05  -1.347 0.178059
statemb:yrdispl                9.192e-03  1.692e-03   5.433 5.80e-08 ***
I(statemb^2):yrdispl          -4.153e-05  4.748e-06  -8.747  < 2e-16 ***
statemb:headyes                5.317e-04  1.010e-02   0.053 0.958012
I(statemb^2):headyes           1.404e-06  2.627e-05   0.053 0.957400
statemb:uiyes                 -3.992e-03  9.509e-03  -0.420 0.674622
I(statemb^2):uiyes             8.055e-06  2.485e-05   0.324 0.745795
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.924 on 4834 degrees of freedom
Multiple R-squared:  0.4128,      Adjusted R-squared:  0.4077
F-statistic: 80.91 on 42 and 4834 DF,  p-value: < 2.2e-16
```

Model 4： Kept significant interaction terms (nwhite,sex,yrdispl). Also kept rr whose p is relatively small to give a good r-squared.

```
> model4 = lm(stateur ~ statemb + I(statemb^2)  + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr  + head + ui + statemb*rr + I(statemb^2)*rr     +statemb*nwhite +
I(statemb^2)*nwhite     +statemb*sex + I(statemb^2)*sex        +statemb*yrdispl +
I(statemb^2)*yrdispl  ,data = mydata)
> summary(model4)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + head + ui + statemb * rr + I(statemb^2) * rr + statemb *
    nwhite + I(statemb^2) * nwhite + statemb * sex + I(statemb^2) *
    sex + statemb * yrdispl + I(statemb^2) * yrdispl, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2469 -1.2292 -0.1239  1.2780  7.5012
```

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	2.555e+01	1.661e+00	15.385	< 2e-16	***
statemb	-1.903e-01	1.923e-02	-9.894	< 2e-16	***
I(statemb^2)	6.566e-04	5.614e-05	11.694	< 2e-16	***
joblostposition_abolished	4.194e-03	1.061e-01	0.040	0.968457	
joblostseasonal_job_ended	-2.412e-02	1.521e-01	-0.159	0.874001	
joblostslack_work	1.419e-01	6.009e-02	2.361	0.018267	*
nwhiteyes	3.391e+00	9.677e-01	3.505	0.000462	***
school12yes	-6.490e-02	7.149e-02	-0.908	0.364068	
sexmale	-2.043e+00	9.014e-01	-2.266	0.023490	*
marriedyes	-1.210e-01	6.555e-02	-1.846	0.064927	.
dkidsyes	1.961e-01	7.093e-02	2.765	0.005713	**
dykidsyes	-2.407e-02	8.013e-02	-0.300	0.763912	
yrdispl	-2.035e+00	1.527e-01	-13.322	< 2e-16	***
I(yrdispl^2)	1.119e-01	4.583e-03	24.407	< 2e-16	***
rr	-6.669e+00	3.421e+00	-1.950	0.051285	.
headyes	-1.050e-01	6.830e-02	-1.537	0.124270	
uiyes	2.365e-01	6.087e-02	3.885	0.000104	***
statemb:rr	5.844e-02	3.860e-02	1.514	0.130057	
I(statemb^2):rr	-1.697e-04	1.063e-04	-1.596	0.110606	
statemb:nwhiteyes	-3.162e-02	1.061e-02	-2.979	0.002902	**
I(statemb^2):nwhiteyes	6.978e-05	2.815e-05	2.479	0.013216	*
statemb:sexmale	2.167e-02	9.793e-03	2.213	0.026964	*
I(statemb^2):sexmale	-5.185e-05	2.584e-05	-2.006	0.044868	*
statemb:yrdispl	9.418e-03	1.639e-03	5.747	9.66e-09	***
I(statemb^2):yrdispl	-4.188e-05	4.629e-06	-9.049	< 2e-16	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.925 on 4852 degrees of freedom
Multiple R-squared: 0.4102, Adjusted R-squared: 0.4072
F-statistic: 140.6 on 24 and 4852 DF, p-value: < 2.2e-16

Model 5： Added interactions of rr and other qualitative variables, observe the output.

```
> model5 = lm(stateur ~ statemb + I(statemb^2) + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr + head + ui + statemb*rr + I(statemb^2)*rr +statemb*nwhite +
I(statemb^2)*nwhite +statemb*sex + I(statemb^2)*sex +statemb*yrdispl + I(statemb^2)*yrdispl
+rr*joblost + I(rr^2)*joblost +rr*nwhite + I(rr^2)*nwhite +rr*school12 + I(rr^2)*school12 +rr*sex
```

```
+ I(rr^2)*sex   +rr*married + I(rr^2)*married    +rr*dkids + I(rr^2)*dkids    +rr*dykids + I(rr^2)*dykids
+rr*yrdispl + I(rr^2)*yrdispl  +rr*head + I(rr^2)*head   +rr*ui + I(rr^2)*ui ,data = mydata)
> summary(model5)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + head + ui + statemb * rr + I(statemb^2) * rr + statemb *
    nwhite + I(statemb^2) * nwhite + statemb * sex + I(statemb^2) *
    sex + statemb * yrdispl + I(statemb^2) * yrdispl + rr * joblost +
    I(rr^2) * joblost + rr * nwhite + I(rr^2) * nwhite + rr *
    school12 + I(rr^2) * school12 + rr * sex + I(rr^2) * sex +
    rr * married + I(rr^2) * married + rr * dkids + I(rr^2) *
    dkids + rr * dykids + I(rr^2) * dykids + rr * yrdispl + I(rr^2) *
    yrdispl + rr * head + I(rr^2) * head + rr * ui + I(rr^2) *
    ui, data = mydata)

Residuals:
   Min      1Q  Median      3Q     Max
-6.2926 -1.2114 -0.1363  1.2731  7.5635

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2.368e+01  2.116e+00  11.193  < 2e-16 ***
statemb                     -1.910e-01  1.993e-02  -9.585  < 2e-16 ***
I(statemb^2)                 6.718e-04  5.842e-05  11.501  < 2e-16 ***
joblostposition_abolished   -5.038e-01  1.123e+00  -0.449  0.65380
joblostseasonal_job_ended   -1.367e+00  1.511e+00  -0.905  0.36567
joblostslack_work           -3.755e-01  7.124e-01  -0.527  0.59821
nwhiteyes                    3.589e+00  1.329e+00   2.701  0.00694 **
school12yes                  7.434e-01  7.064e-01   1.052  0.29265
sexmale                      5.368e-01  1.545e+00   0.347  0.72831
marriedyes                   5.501e-01  7.756e-01   0.709  0.47820
dkidsyes                    -6.437e-01  8.016e-01  -0.803  0.42201
dykidsyes                   -1.837e-02  9.603e-01  -0.019  0.98474
yrdispl                     -2.154e+00  1.836e-01 -11.729  < 2e-16 ***
I(yrdispl^2)                 1.127e-01  4.614e-03  24.418  < 2e-16 ***
rr                           4.026e+00  7.700e+00   0.523  0.60105
headyes                     -3.422e-01  9.432e-01  -0.363  0.71676
uiyes                       -1.439e-01  7.046e-01  -0.204  0.83823
I(rr^2)                     -1.311e+01  8.461e+00  -1.550  0.12123
statemb:rr                   5.265e-02  4.022e-02   1.309  0.19062
I(statemb^2):rr             -1.735e-04  1.111e-04  -1.561  0.11849
statemb:nwhiteyes           -3.382e-02  1.090e-02  -3.103  0.00192 **
```

```
I(statemb^2):nwhiteyes            7.654e-05  2.888e-05   2.650  0.00807 **
statemb:sexmale                   2.302e-02  9.930e-03   2.318  0.02050 *
I(statemb^2):sexmale             -5.635e-05  2.610e-05  -2.159  0.03091 *
statemb:yrdispl                   9.154e-03  1.652e-03   5.542 3.14e-08 ***
I(statemb^2):yrdispl             -4.187e-05  4.645e-06  -9.014  < 2e-16 ***
joblostposition_abolished:rr       1.945e+00  5.959e+00   0.326  0.74406
joblostseasonal_job_ended:rr       6.167e+00  8.325e+00   0.741  0.45885
joblostslack_work:rr               1.212e+00  3.675e+00   0.330  0.74161
joblostposition_abolished:I(rr^2) -1.702e+00  7.554e+00  -0.225  0.82176
joblostseasonal_job_ended:I(rr^2) -6.700e+00  1.093e+01  -0.613  0.53974
joblostslack_work:I(rr^2)         -9.398e-02  4.571e+00  -0.021  0.98360
nwhiteyes:rr                      2.596e-01  5.574e+00   0.047  0.96286
nwhiteyes:I(rr^2)                -7.561e-01  6.823e+00  -0.111  0.91176
school12yes:rr                   -4.609e+00  3.850e+00  -1.197  0.23129
school12yes:I(rr^2)               5.992e+00  4.980e+00   1.203  0.22894
sexmale:rr                       -1.348e+01  6.357e+00  -2.120  0.03404 *
sexmale:I(rr^2)                   1.607e+01  7.213e+00   2.227  0.02598 *
marriedyes:rr                    -3.770e+00  3.968e+00  -0.950  0.34204
marriedyes:I(rr^2)                4.865e+00  4.924e+00   0.988  0.32315
dkidsyes:rr                       3.729e+00  4.185e+00   0.891  0.37294
dkidsyes:I(rr^2)                 -3.933e+00  5.251e+00  -0.749  0.45386
dykidsyes:rr                      4.491e-01  5.042e+00   0.089  0.92903
dykidsyes:I(rr^2)                -1.001e+00  6.338e+00  -0.158  0.87446
yrdispl:rr                        6.043e-01  5.834e-01   1.036  0.30039
yrdispl:I(rr^2)                  -5.378e-01  7.256e-01  -0.741  0.45859
rr:headyes                        2.967e-01  4.651e+00   0.064  0.94915
headyes:I(rr^2)                   4.801e-01  5.607e+00   0.086  0.93177
rr:uiyes                          3.324e+00  3.638e+00   0.914  0.36089
uiyes:I(rr^2)                    -5.275e+00  4.528e+00  -1.165  0.24410
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.925 on 4827 degrees of freedom
Multiple R-squared:  0.4134,      Adjusted R-squared:  0.4074
F-statistic: 69.42 on 49 and 4827 DF,  p-value: < 2.2e-16
```

Model 6： Kept r*sex

```
> model6 = lm(stateur ~ statemb + I(statemb^2)  + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr  + head + ui + statemb*rr + I(statemb^2)*rr     +statemb*nwhite +
I(statemb^2)*nwhite     +statemb*sex + I(statemb^2)*sex        +statemb*yrdispl + I(statemb^2)*yrdispl
+rr*sex + I(rr^2)*sex   ,data = mydata)
> summary(model6)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + head + ui + statemb * rr + I(statemb^2) * rr + statemb *
    nwhite + I(statemb^2) * nwhite + statemb * sex + I(statemb^2) *
    sex + statemb * yrdispl + I(statemb^2) * yrdispl + rr * sex +
    I(rr^2) * sex, data = mydata)

Residuals:
   Min     1Q Median     3Q    Max
-6.207 -1.222 -0.125  1.278  7.478

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2.318e+01  1.931e+00  12.000  < 2e-16 ***
statemb                     -1.943e-01  1.939e-02 -10.020  < 2e-16 ***
I(statemb^2)                 6.665e-04  5.638e-05  11.820  < 2e-16 ***
joblostposition_abolished    3.354e-03  1.060e-01   0.032 0.974747
joblostseasonal_job_ended   -1.944e-02  1.520e-01  -0.128 0.898222
joblostslack_work            1.436e-01  6.003e-02   2.391 0.016831 *
nwhiteyes                    3.536e+00  9.683e-01   3.652 0.000263 ***
school12yes                 -5.427e-02  7.154e-02  -0.759 0.448151
sexmale                      3.229e-01  1.447e+00   0.223 0.823457
marriedyes                  -1.197e-01  6.552e-02  -1.827 0.067717 .
dkidsyes                     1.934e-01  7.089e-02   2.728 0.006402 **
dykidsyes                   -2.915e-02  8.008e-02  -0.364 0.715861
yrdispl                     -2.028e+00  1.526e-01 -13.288  < 2e-16 ***
I(yrdispl^2)                 1.118e-01  4.581e-03  24.401  < 2e-16 ***
rr                           7.086e+00  6.220e+00   1.139 0.254656
headyes                     -9.786e-02  6.831e-02  -1.433 0.152021
uiyes                        2.341e-01  6.089e-02   3.844 0.000123 ***
I(rr^2)                     -1.752e+01  6.087e+00  -2.879 0.004012 **
statemb:rr                   6.413e-02  3.873e-02   1.656 0.097846 .
I(statemb^2):rr             -1.814e-04  1.065e-04  -1.703 0.088560 .
statemb:nwhiteyes           -3.349e-02  1.062e-02  -3.153 0.001626 **
I(statemb^2):nwhiteyes       7.505e-05  2.818e-05   2.664 0.007757 **
```

```
statemb:sexmale            2.303e-02  9.905e-03   2.325 0.020097 *
I(statemb^2):sexmale      -5.607e-05  2.604e-05  -2.154 0.031326 *
statemb:yrdispl            9.374e-03  1.637e-03   5.725  1.1e-08 ***
I(statemb^2):yrdispl      -4.184e-05  4.625e-06  -9.049  < 2e-16 ***
sexmale:rr                -1.307e+01  5.777e+00  -2.262 0.023721 *
sexmale:I(rr^2)            1.611e+01  6.496e+00   2.480 0.013154 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.923 on 4849 degrees of freedom
Multiple R-squared:  0.4116,      Adjusted R-squared:  0.4083
F-statistic: 125.6 on 27 and 4849 DF,  p-value: < 2.2e-16
```

Model 7： Did a few interactions that we believed to have an impact.

```
> model7 = lm(stateur ~ statemb + I(statemb^2)  + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr  + head + ui + statemb*rr + I(statemb^2)*rr     +statemb*nwhite +
I(statemb^2)*nwhite     +statemb*sex + I(statemb^2)*sex        +statemb*yrdispl + I(statemb^2)*yrdispl
+rr*sex + I(rr^2)*sex     +joblost*yrdispl    +married*dkids  + married*dykids     + married*dkids*dykids
+ head*married   +head*dkids    + head*dykids ,data = mydata)
> summary(model7)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + head + ui + statemb * rr + I(statemb^2) * rr + statemb *
    nwhite + I(statemb^2) * nwhite + statemb * sex + I(statemb^2) *
    sex + statemb * yrdispl + I(statemb^2) * yrdispl + rr * sex +
    I(rr^2) * sex + joblost * yrdispl + married * dkids + married *
    dykids + married * dkids * dykids + head * married + head *
    dkids + head * dykids, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-6.1567 -1.2168 -0.1133  1.2881  7.6229

Coefficients: (2 not defined because of singularities)
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    2.367e+01  1.936e+00  12.224  < 2e-16 ***
statemb                       -1.976e-01  1.943e-02 -10.172  < 2e-16 ***
I(statemb^2)                   6.788e-04  5.658e-05  11.998  < 2e-16 ***
joblostposition_abolished      2.155e-01  2.137e-01   1.009 0.313240
joblostseasonal_job_ended      1.704e-01  3.050e-01   0.559 0.576387
joblostslack_work             -5.499e-02  1.163e-01  -0.473 0.636485
```

```
nwhiteyes                       3.464e+00  9.686e-01   3.577 0.000351 ***
school12yes                    -5.631e-02  7.159e-02  -0.787 0.431545
sexmale                         1.956e-01  1.449e+00   0.135 0.892580
marriedyes                     -1.694e-01  1.289e-01  -1.314 0.188957
dkidsyes                        1.091e-01  1.584e-01   0.689 0.491066
dykidsyes                      -1.928e-01  2.573e-01  -0.749 0.453805
yrdispl                        -2.047e+00  1.529e-01 -13.383  < 2e-16 ***
I(yrdispl^2)                    1.120e-01  4.580e-03  24.465  < 2e-16 ***
rr                              6.189e+00  6.234e+00   0.993 0.320859
headyes                        -1.684e-01  9.672e-02  -1.742 0.081635 .
uiyes                           2.311e-01  6.088e-02   3.796 0.000149 ***
I(rr^2)                        -1.694e+01  6.095e+00  -2.779 0.005477 **
statemb:rr                      6.903e-02  3.877e-02   1.780 0.075060 .
I(statemb^2):rr                -1.983e-04  1.067e-04  -1.858 0.063214 .
statemb:nwhiteyes              -3.265e-02  1.062e-02  -3.073 0.002131 **
I(statemb^2):nwhiteyes          7.279e-05  2.819e-05   2.582 0.009853 **
statemb:sexmale                 2.265e-02  9.913e-03   2.285 0.022381 *
I(statemb^2):sexmale           -5.542e-05  2.607e-05  -2.126 0.033585 *
statemb:yrdispl                 9.446e-03  1.637e-03   5.769 8.48e-09 ***
I(statemb^2):yrdispl           -4.220e-05  4.625e-06  -9.125  < 2e-16 ***
sexmale:rr                     -1.252e+01  5.787e+00  -2.163 0.030590 *
sexmale:I(rr^2)                 1.554e+01  6.506e+00   2.389 0.016935 *
joblostposition_abolished:yrdispl -3.838e-02  3.476e-02  -1.104 0.269560
joblostseasonal_job_ended:yrdispl -3.359e-02  4.943e-02  -0.680 0.496743
joblostslack_work:yrdispl       3.840e-02  1.932e-02   1.988 0.046923 *
marriedyes:dkidsyes             1.223e-01  1.618e-01   0.756 0.449908
marriedyes:dykidsyes           -1.381e-01  2.327e-01  -0.593 0.552994
dkidsyes:dykidsyes                  NA         NA      NA       NA
marriedyes:headyes              6.615e-02  1.592e-01   0.415 0.677846
dkidsyes:headyes               -9.275e-03  1.547e-01  -0.060 0.952199
dykidsyes:headyes               3.466e-01  1.977e-01   1.753 0.079634 .
marriedyes:dkidsyes:dykidsyes        NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.922 on 4841 degrees of freedom
Multiple R-squared:  0.4133,      Adjusted R-squared:  0.4091
F-statistic: 97.44 on 35 and 4841 DF,  p-value: < 2.2e-16
```

Model 8： Kept joblost*yrdispl

```
> model8 = lm(stateur ~ statemb + I(statemb^2)  + joblost + nwhite + school12 + sex + married + dkids+
dykids + yrdispl + I(yrdispl^2) + rr  + head + ui + statemb*rr + I(statemb^2)*rr     +statemb*nwhite +
I(statemb^2)*nwhite     +statemb*sex + I(statemb^2)*sex        +statemb*yrdispl + I(statemb^2)*yrdispl
+rr*sex + I(rr^2)*sex     +joblost*yrdispl,data = mydata)
> summary(model8)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + joblost + nwhite +
    school12 + sex + married + dkids + dykids + yrdispl + I(yrdispl^2) +
    rr + head + ui + statemb * rr + I(statemb^2) * rr + statemb *
    nwhite + I(statemb^2) * nwhite + statemb * sex + I(statemb^2) *
    sex + statemb * yrdispl + I(statemb^2) * yrdispl + rr * sex +
    I(rr^2) * sex + joblost * yrdispl, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-6.1108 -1.2194 -0.1181  1.2929  7.4185

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2.359e+01  1.935e+00  12.193  < 2e-16 ***
statemb                     -1.970e-01  1.941e-02 -10.147  < 2e-16 ***
I(statemb^2)                 6.769e-04  5.651e-05  11.979  < 2e-16 ***
joblostposition_abolished    2.186e-01  2.136e-01   1.023 0.306333
joblostseasonal_job_ended    1.699e-01  3.048e-01   0.557 0.577274
joblostslack_work           -6.129e-02  1.162e-01  -0.527 0.597908
nwhiteyes                    3.469e+00  9.681e-01   3.584 0.000342 ***
school12yes                 -5.454e-02  7.150e-02  -0.763 0.445657
sexmale                      1.818e-01  1.448e+00   0.126 0.900107
marriedyes                  -1.130e-01  6.552e-02  -1.725 0.084520 .
dkidsyes                     1.927e-01  7.085e-02   2.720 0.006549 **
dykidsyes                   -2.831e-02  8.004e-02  -0.354 0.723562
yrdispl                     -2.045e+00  1.529e-01 -13.377  < 2e-16 ***
I(yrdispl^2)                 1.120e-01  4.579e-03  24.465  < 2e-16 ***
rr                           5.947e+00  6.228e+00   0.955 0.339686
headyes                     -9.681e-02  6.827e-02  -1.418 0.156235
uiyes                        2.327e-01  6.087e-02   3.824 0.000133 ***
I(rr^2)                     -1.664e+01  6.091e+00  -2.731 0.006332 **
statemb:rr                   6.874e-02  3.875e-02   1.774 0.076147 .
I(statemb^2):rr             -1.974e-04  1.066e-04  -1.851 0.064176 .
statemb:nwhiteyes           -3.261e-02  1.062e-02  -3.070 0.002149 **
I(statemb^2):nwhiteyes       7.234e-05  2.818e-05   2.567 0.010276 *
statemb:sexmale              2.260e-02  9.901e-03   2.282 0.022520 *
```

```
I(statemb^2):sexmale             -5.540e-05  2.602e-05  -2.129 0.033340 *
statemb:yrdispl                   9.406e-03  1.637e-03   5.747 9.62e-09 ***
I(statemb^2):yrdispl             -4.207e-05  4.622e-06  -9.101  < 2e-16 ***
sexmale:rr                       -1.220e+01  5.782e+00  -2.110 0.034914 *
sexmale:I(rr^2)                   1.519e+01  6.502e+00   2.337 0.019504 *
joblostposition_abolished:yrdispl -3.951e-02  3.475e-02  -1.137 0.255572
joblostseasonal_job_ended:yrdispl -3.447e-02  4.940e-02  -0.698 0.485355
joblostslack_work:yrdispl         3.961e-02  1.930e-02   2.053 0.040148 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.922 on 4846 degrees of freedom
Multiple R-squared:  0.4127,      Adjusted R-squared:  0.4091
F-statistic: 113.5 on 30 and 4846 DF,  p-value: < 2.2e-16
```

An intermediate step:

```
> mycor = mydata
> mycor$state <- NULL
> mycor$joblost <- NULL
> mycor$nwhite <- NULL
> mycor$school12 <- NULL
> mycor$sex <- NULL
> mycor$bluecol <- NULL
> mycor$smsa <- NULL
> mycor$married <- NULL
> mycor$dkids <- NULL
> mycor$dykids <- NULL
> mycor$head <- NULL
> mycor$ui <- NULL

> cor(mycor)
            stateur     statemb         age      tenure     yrdispl          rr
stateur  1.00000000 -0.242385521 -0.03542044  0.022494002 -0.52492326 -0.08656481
statemb -0.24238552  1.000000000  0.02108761  0.008032231  0.60780839  0.27002287
age     -0.03542044  0.021087610  1.00000000  0.486452638  0.04724166 -0.12542447
tenure   0.02249400  0.008032231  0.48645264  1.000000000 -0.02305965 -0.14510825
yrdispl -0.52492326  0.607808386  0.04724166 -0.023059646  1.00000000  0.07230758
rr      -0.08656481  0.270022866 -0.12542447 -0.145108248  0.07230758  1.00000000
```

Model 9： Added age*tenure, which turned out to be good.

```
> model9 = lm(stateur ~ statemb + I(statemb^2) +age + tenure + age*tenure+ joblost + nwhite + school12
+ sex + married + dkids+ dykids + yrdispl + I(yrdispl^2) + rr  + head + ui + statemb*rr + I(statemb^2)*rr
+statemb*nwhite + I(statemb^2)*nwhite     +statemb*sex + I(statemb^2)*sex        +statemb*yrdispl +
I(statemb^2)*yrdispl     +rr*sex + I(rr^2)*sex     +joblost*yrdispl,data = mydata)
> summary(model9)

Call:
lm(formula = stateur ~ statemb + I(statemb^2) + +age + tenure +
    age * tenure + joblost + nwhite + school12 + sex + married +
    dkids + dykids + yrdispl + I(yrdispl^2) + rr + head + ui +
    statemb * rr + I(statemb^2) * rr + statemb * nwhite + I(statemb^2) *
    nwhite + statemb * sex + I(statemb^2) * sex + statemb * yrdispl +
    I(statemb^2) * yrdispl + rr * sex + I(rr^2) * sex + joblost *
    yrdispl, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max
-6.0542 -1.2125 -0.1148  1.2757  7.4479

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2.345e+01  1.941e+00  12.083  < 2e-16 ***
statemb                     -1.965e-01  1.941e-02 -10.120  < 2e-16 ***
I(statemb^2)                 6.749e-04  5.652e-05  11.940  < 2e-16 ***
age                         -3.271e-04  3.969e-03  -0.082 0.934331
tenure                       5.422e-02  2.261e-02   2.398 0.016501 *
joblostposition_abolished    2.384e-01  2.139e-01   1.115 0.264987
joblostseasonal_job_ended    1.894e-01  3.056e-01   0.620 0.535450
joblostslack_work           -4.869e-02  1.170e-01  -0.416 0.677213
nwhiteyes                    3.451e+00  9.686e-01   3.562 0.000371 ***
school12yes                 -5.110e-02  7.170e-02  -0.713 0.476083
sexmale                      2.533e-01  1.448e+00   0.175 0.861199
marriedyes                  -9.860e-02  6.899e-02  -1.429 0.153037
dkidsyes                     1.637e-01  7.241e-02   2.260 0.023856 *
dykidsyes                   -3.644e-02  8.377e-02  -0.435 0.663570
yrdispl                     -2.055e+00  1.529e-01 -13.438  < 2e-16 ***
I(yrdispl^2)                 1.124e-01  4.581e-03  24.545  < 2e-16 ***
rr                           6.198e+00  6.228e+00   0.995 0.319672
headyes                     -9.399e-02  7.085e-02  -1.326 0.184744
uiyes                        2.229e-01  6.153e-02   3.622 0.000295 ***
I(rr^2)                     -1.668e+01  6.090e+00  -2.738 0.006195 **
age:tenure                  -9.644e-04  4.579e-04  -2.106 0.035220 *
```

```
statemb:rr                      6.778e-02  3.875e-02   1.749 0.080310 .
I(statemb^2):rr                -1.948e-04  1.066e-04  -1.827 0.067813 .
statemb:nwhiteyes              -3.257e-02  1.063e-02  -3.065 0.002189 **
I(statemb^2):nwhiteyes          7.257e-05  2.820e-05   2.573 0.010099 *
statemb:sexmale                 2.244e-02  9.902e-03   2.266 0.023496 *
I(statemb^2):sexmale           -5.497e-05  2.603e-05  -2.111 0.034783 *
statemb:yrdispl                 9.458e-03  1.636e-03   5.780 7.95e-09 ***
I(statemb^2):yrdispl           -4.216e-05  4.621e-06  -9.123  < 2e-16 ***
sexmale:rr                     -1.248e+01  5.782e+00  -2.158 0.030975 *
sexmale:I(rr^2)                 1.548e+01  6.501e+00   2.381 0.017319 *
joblostposition_abolished:yrdispl -4.130e-02  3.475e-02  -1.188 0.234762
joblostseasonal_job_ended:yrdispl -3.317e-02  4.940e-02  -0.671 0.501999
joblostslack_work:yrdispl        4.079e-02  1.930e-02   2.113 0.034620 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.921 on 4843 degrees of freedom
Multiple R-squared:  0.4136,      Adjusted R-squared:  0.4096
F-statistic: 103.5 on 33 and 4843 DF,  p-value: < 2.2e-16
```

Compare the complete and reduced model using anova.

```
> anova(model1,model9)
Analysis of Variance Table

Model 1: stateur ~ statemb + age + tenure + joblost + nwhite + school12 +
    sex + married + dkids + dykids + yrdispl + rr + head + ui
Model 2: stateur ~ statemb + I(statemb^2) + +age + tenure + age * tenure +
    joblost + nwhite + school12 + sex + married + dkids + dykids +
    yrdispl + I(yrdispl^2) + rr + head + ui + statemb * rr +
    I(statemb^2) * rr + statemb * nwhite + I(statemb^2) * nwhite +
    statemb * sex + I(statemb^2) * sex + statemb * yrdispl +
    I(statemb^2) * yrdispl + rr * sex + I(rr^2) * sex + joblost *
    yrdispl
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   4860 21254
2   4843 17879 17      3375 53.779 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```