Do-File:

clear all

use "/Users/mariahhastings/Documents/Stata/Empirical Paper/cps_00013.dta"
*data extract is from IPUMS CPS with samples from 2010 and 2014 and variables incwelfr
(income from welfare assistance), cpi99 (cpi value that incwelfr must be multiplied by to control
for inflation), health (health status), region, age, sex, marst (marital status), famsize (number of
own family members living in household), educ (education level), and mthwelfr (number of
months received welfare)*

log using "/Users/mariahhastings/Documents/Stata/Empirical Paper/Multiple Regression
Log.smcl", replace

to get rid of other variables given automatically in data set
drop year serial month cpsid asecflag asecwth pernum cpsidp asecwt hflag
br incwelfr health region sex age marst famsize educ mthwelfr

cleaning data
replace incwelfr = . if incwelfr == 999999
drop if incwelfr == .
drop if incwelfr == 0

cap gen adj_welfr = cpi99*incwelfr
label variable adj_welfr "adjusted welfare income"

tab health
tab health, nolabel
cap gen goodhealth = 0
replace goodhealth = 1 if health == 1 | health == 2 | health == 3
tab goodhealth
label variable goodhealth "good health status"

tab region, nolabel
tab region
cap gen midwest = 0
replace midwest = 1 if region >= 21 & region <= 23
cap gen south = 0
replace south = 1 if region >= 31 & region <= 34
cap gen west = 0

```
replace west = 1 if region >= 41 & region <= 43
tab midwest
tab south
tab west
label variable midwest "living in midwest"
label variable south "living in south"
label variable west "living in west"

tab sex
tab sex, nolabel
cap gen sex_rev = 0
replace sex_rev = 1 if sex == 2
label variable sex_rev "sex: 1 = female, 0 = male"

tab age

tab marst
tab marst, nolabel
cap gen married = 0
replace married = 1 if marst >= 1 & marst <= 2
replace married = . if marst == .
tab married
label variable married "marital status: married or unmarried"

tab famsize

tab educ
tab educ, nolabel
replace educ = . if educ == 1
drop if educ == .
cap gen educ_rev = educ
replace educ_rev = 0 if educ == 2
replace educ_rev = 4 if educ == 10
replace educ_rev = 6 if educ == 20
replace educ_rev = 8 if educ == 30
replace educ_rev = 9 if educ == 40
replace educ_rev = 10 if educ == 50
replace educ_rev = 11 if educ == 60 | educ == 71
replace educ_rev = 12 if educ == 73
replace educ_rev = 13 if educ == 81
```

```
replace educ_rev = 14 if educ == 91 | educ == 92
replace educ_rev = 16 if educ == 111
replace educ_rev = 18 if educ == 123
replace educ_rev = 21 if educ == 124 | educ == 125
label variable educ_rev "revised education level"

tab mthwelfr

*interaction terms*
cap gen adj_welfr_famsize = adj_welfr*famsize
label variable adj_welfr_famsize "interaction term of adjusted welfare income and number of
family members in home"

cap gen adj_welfr_educ_rev = adj_welfr*educ_rev
label variable adj_welfr_educ_rev "interaction term of adjusted welfare income and education
level"

*linear simple regression*
reg goodhealth adj_welfr
outreg2 using myreg.doc, replace ctitle(Model 1)

*quadratic simple regression*
cap gen adj_welfr_sq = adj_welfr*adj_welfr
label variable adj_welfr_sq "adjusted welfare income squared"

reg goodhealth adj_welfr_sq
outreg2 using myreg.doc, append ctitle(Model 2)

*cubic simple regression*
cap gen adj_welfr_cu = adj_welfr*adj_welfr*adj_welfr
label variable adj_welfr_cu "adjusted welfare income cubed"

reg goodhealth adj_welfr_cu
outreg2 using myreg.doc, append ctitle(Model 3)

*lin-log simple regression*
cap gen adj_welfr_log = ln(adj_welfr)
label variable adj_welfr_log "natural log of adjusted welfare income"
hist adj_welfr
hist adj_welfr_log
```

```
reg goodhealth adj_welfr_log
outreg2 using myreg.doc, append ctitle(Model 4)

*linear multiple regression*
reg goodhealth adj_welfr educ_rev married west south midwest mthwelfr famsize sex_rev age
adj_welfr_famsize adj_welfr_educ_rev
outreg2 using myreg2.doc, replace ctitle(Model 5)

*quadratic multiple regression*
cap gen age_sq = age*age
label variable age_sq "age squared"
cap gen famsize_sq = famsize*famsize
label variable famsize_sq "number of family members in household squared"
cap gen educ_rev_sq = educ_rev*educ_rev
label variable educ_rev_sq "education level squared"
cap gen mthwelfr_sq = mthwelfr*mthwelfr
label variable mthwelfr_sq "number of months received welfare income squared"
cap gen adj_welfr_famsize_sq = adj_welfr_famsize*adj_welfr_famsize
label variable adj_welfr_famsize_sq "interaction term of adjusted welfare income and number of
family members in home squared"
cap gen adj_welfr_educ_rev_sq = adj_welfr_educ_rev*adj_welfr_educ_rev
label variable adj_welfr_educ_rev_sq "interaction term of adjusted welfare income and
education level squared"

reg goodhealth adj_welfr_sq educ_rev_sq married west south midwest mthwelfr_sq famsize_sq
sex_rev age_sq adj_welfr_famsize_sq adj_welfr_educ_rev_sq
outreg2 using myreg2.doc, append ctitle(Model 6)

*cubic multiple regression*
cap gen age_cu = age*age*age
label variable age_cu "age cubed"
cap gen famsize_cu = famsize*famsize*famsize
label variable famsize_cu "number of family members in household cubed"
cap gen educ_rev_cu = educ_rev*educ_rev*educ_rev
label variable educ_rev_cu "education level cubed"
cap gen mthwelfr_cu = mthwelfr*mthwelfr*mthwelfr
label variable mthwelfr_cu "number of months received welfare income cubed"
cap gen adj_welfr_famsize_cu = adj_welfr_famsize*adj_welfr_famsize*adj_welfr_famsize
```

label variable adj_welfr_famsize_cu "interaction term of adjusted welfare income and number of family members in home cubed"
cap gen adj_welfr_educ_rev_cu = adj_welfr_educ_rev*adj_welfr_educ_rev*adj_welfr_educ_rev
label variable adj_welfr_educ_rev_cu "interaction term of adjusted welfare income and education level cubed"

reg goodhealth adj_welfr_cu educ_rev_cu married west south midwest mthwelfr_cu famsize_cu sex_rev age_cu adj_welfr_famsize_cu adj_welfr_educ_rev_cu
outreg2 using myreg2.doc, append ctitle(Model 7)

lin-log multiple regression
gen age_log = ln(age)
label variable age_log "natural log of age"
gen famsize_log = ln(famsize)
label variable famsize_log "natural log of number of family members in household"
gen educ_rev_log = ln(educ_rev)
label variable educ_rev_log "natural log of education level"
gen mthwelfr_log = ln(mthwelfr)
label variable mthwelfr_log "natural log of months received welfare income"
cap gen adj_welfr_famsize_log = ln(adj_welfr)*ln(famsize)
label variable adj_welfr_famsize_log "natural log of interaction term of adjusted welfare income and number of family members in home"
cap gen adj_welfr_educ_rev_log = ln(adj_welfr)*ln(educ_rev)
label variable adj_welfr_educ_rev_log "natural log of interaction term of adjusted welfare income and education level"

reg goodhealth adj_welfr_log educ_rev_log married west south midwest mthwelfr_log famsize_log sex_rev age_log adj_welfr_famsize_log adj_welfr_educ_rev_log
outreg2 using myreg2.doc, append ctitle(Model 8)

descriptive statistics
sum incwelfr health goodhealth adj_welfr adj_welfr_sq adj_welfr_cu adj_welfr_log educ_rev educ_rev_sq educ_rev_cu educ_rev_log married west south midwest mthwelfr mthwelfr_sq mthwelfr_cu mthwelfr_log famsize famsize_sq famsize_cu famsize_log sex_rev age age_sq age_cu age_log adj_welfr_famsize adj_welfr_famsize_sq adj_welfr_famsize_cu adj_welfr_famsize_log adj_welfr_educ_rev adj_welfr_educ_rev_sq adj_welfr_educ_rev_cu adj_welfr_educ_rev_log

outreg2 using summary.doc, replace sum(log) keep(incwelfr health goodhealth adj_welfr adj_welfr_sq adj_welfr_cu adj_welfr_log educ_rev educ_rev_sq educ_rev_cu educ_rev_log

married west south midwest mthwelfr mthwelfr_sq mthwelfr_cu mthwelfr_log famsize
famsize_sq famsize_cu famsize_log sex_rev age age_sq age_cu age_log adj_welfr_famsize
adj_welfr_famsize_sq adj_welfr_famsize_cu adj_welfr_famsize_log adj_welfr_educ_rev
adj_welfr_educ_rev_sq adj_welfr_educ_rev_cu adj_welfr_educ_rev_log)

table 3
cap gen fiftypercent_adj_welfr = 0
replace fiftypercent_adj_welfr = 1 if adj_welfr >= 2486.58
label variable fiftypercent_adj_welfr "top fifty percent of adjusted welfare income"

bys fiftypercent_adj_welfr: sum incwelfr health goodhealth adj_welfr adj_welfr_sq adj_welfr_cu
adj_welfr_log educ_rev educ_rev_sq educ_rev_cu educ_rev_log married west south midwest
mthwelfr mthwelfr_sq mthwelfr_cu mthwelfr_log famsize famsize_sq famsize_cu famsize_log
sex_rev age age_sq age_cu age_log adj_welfr_famsize adj_welfr_famsize_sq
adj_welfr_famsize_cu adj_welfr_famsize_log adj_welfr_educ_rev adj_welfr_educ_rev_sq
adj_welfr_educ_rev_cu adj_welfr_educ_rev_log

bys fiftypercent_adj_welfr: outreg2 using descriptivestats.doc, replace sum(log) eqkeep(N mean)
keep(incwelfr health goodhealth adj_welfr adj_welfr_sq adj_welfr_cu adj_welfr_log educ_rev
educ_rev_sq educ_rev_cu educ_rev_log married west south midwest mthwelfr mthwelfr_sq
mthwelfr_cu mthwelfr_log famsize famsize_sq famsize_cu famsize_log sex_rev age age_sq
age_cu age_log adj_welfr_famsize adj_welfr_famsize_sq adj_welfr_famsize_cu
adj_welfr_famsize_log adj_welfr_educ_rev adj_welfr_educ_rev_sq adj_welfr_educ_rev_cu
adj_welfr_educ_rev_log)

save "/Users/mariahhastings/Documents/Stata/Empirical Paper/Multiple Regression Cleaned
Data.dta", replace
log close