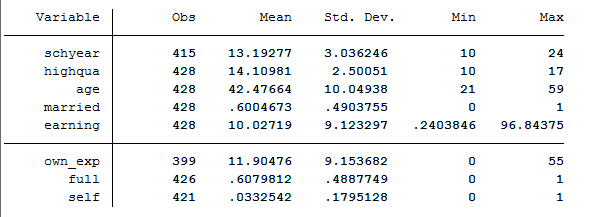
\*统计描述（表1-第六列数据）

sum schyear highqua age married earning own\_exp full self



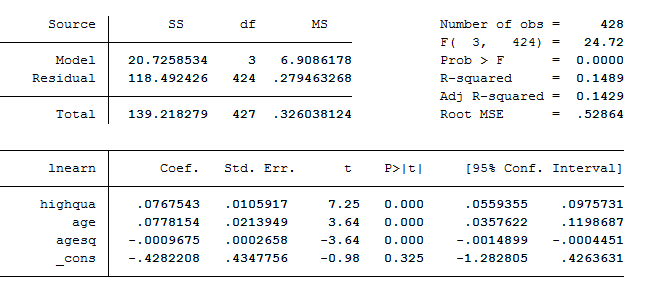
\*定义每小时工资对数和年龄平方

gen lnearn=ln(earning)

gen agesq=age\*age

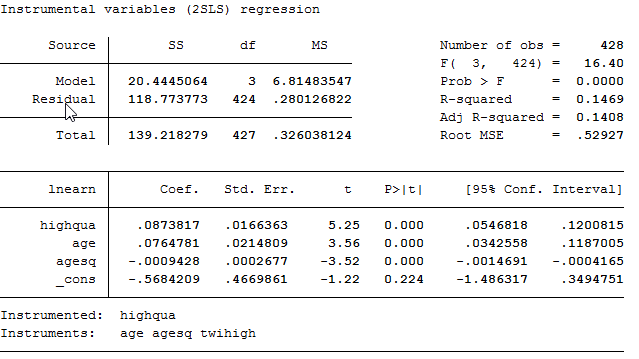
\*做线性回归 工资对数关于教育年限、年龄的二元二次项回归（表2-第二列数据）

reg lnearn highqua age agesq



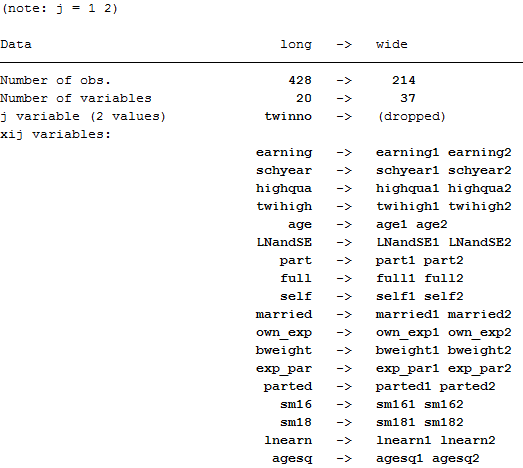
\*工具变量回归（原来的教育年限被twihigh取代）（表2—第三列数据）

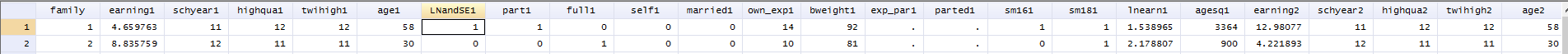
ivreg lnearn (highqua=twihigh) age agesq



\*把数据对应同一家庭的两行双胞胎数据合成一行

reshape wide earning-agesq, i(family) j(twinno)

\



\*定义差分

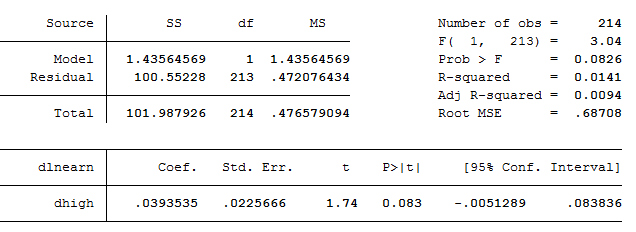
gen dlnearn = lnearn1-lnearn2

gen dhigh = highqua1-highqua2

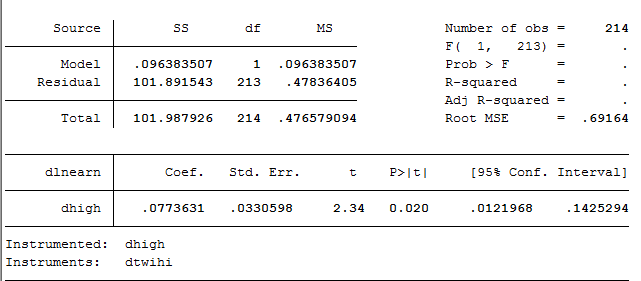
gen dtwihi=twihigh1-twihigh2

\*做线性回归 工资对数的差分关于教育年限的差分（无常数项）（表2-第四列）

reg dlnearn dhigh, noc



\*工具变量线性回归 (同上把dhigh用dtwini取代)（无常数项）（表2-第五列）



\*定义其他变量的差分

gen dpart = part1-part2

gen dmarried = married1-married2

gen dLNandSE = LNandSE1-LNandSE2

gen down\_exp = own\_exp1-own\_exp2

\*删去差分数据都为0的样本，即双胞胎数据在变量上来看完全一样的删去（删去27个）

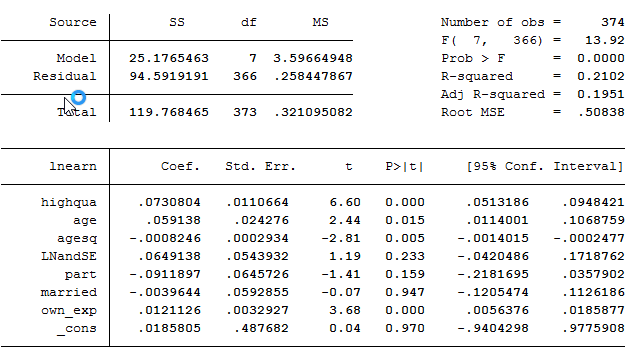
drop if dlnearn==.|dhigh==.|dtwihi==.|dpart==.|dmarried==.|dln==.|down\_exp==

\*把一行数据拆分成两行，即恢复成原来的初始样子（样本数从428变为374）

reshape long

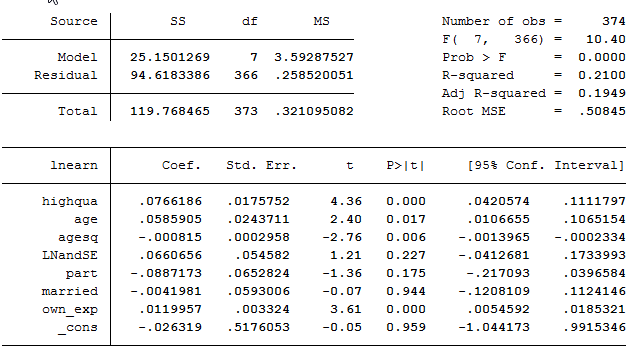
\*做线性回归 工资对数的多元线性回归（表2-第六列）

reg lnearn highqua age agesq LNandSE part married own\_exp



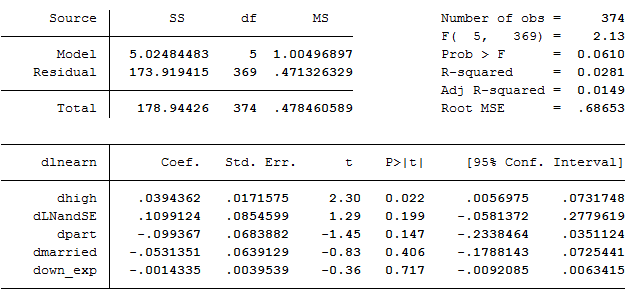
\*工具变量线性回归 （同上把highqua用twihigh取代，对教育年限进行修正）（表2-第七列）

ivreg lnearn (highqua=twihigh) age agesq LNandSE part married own\_exp



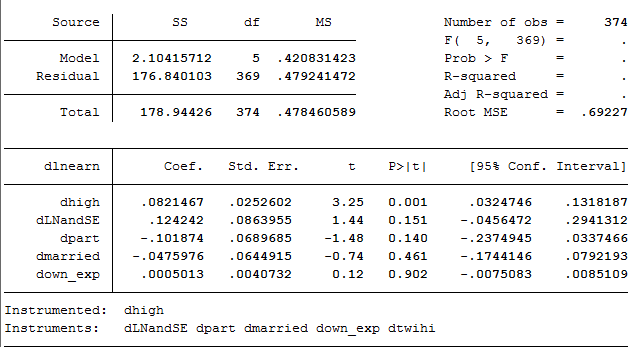
\*做关于工资对数差分的多元线性回归（无常数项）（表2-第八列）

reg dlnearn dhigh dLNandSE dpart dmarried down\_exp, noc



\*做工具变量线性回归（同上把dhigh用dtwihi取代，修正教育年限差分数据）（无常数项）（表2-第九列）

ivreg dlnearn (dhigh=dtwihi) dLNandSE dpart dmarried down\_exp, noc



\*清理数据

Clear

\*定义每小时工资对数和年龄平方

gen lnearn=ln(earning)

gen agesq=age\*age

\*把数据对应同一家庭的两行双胞胎数据合成一行

reshape wide earning-agesq, i(family) j(twinno)

\*定义差分

gen dweight = bweight1 - bweight2

gen aveweigh = (bweight1 + bweight2)/2

gen dmarry = married1 - married2

gen avemarry = (married1 + married2)/2

gen dself = self1 - self2

gen aveself = (self1+self2)/2

gen dhigh = highqua1 - highqua2

gen avehigh = (highqua1+highqua2)/2

gen dpart = part1 - part2

gen avepart = (part1+part2)/2

gen dpartexp = exp\_par1 - exp\_par2

gen avepexp = (exp\_par1 + exp\_par2)/2

gen aveped = (parted1+parted2)/2

gen dped = parted1 - parted2

gen avesm16 = (sm161+sm162)/2

gen dsm16 = sm161-sm162

gen avesm18 = (sm181+sm182)/2

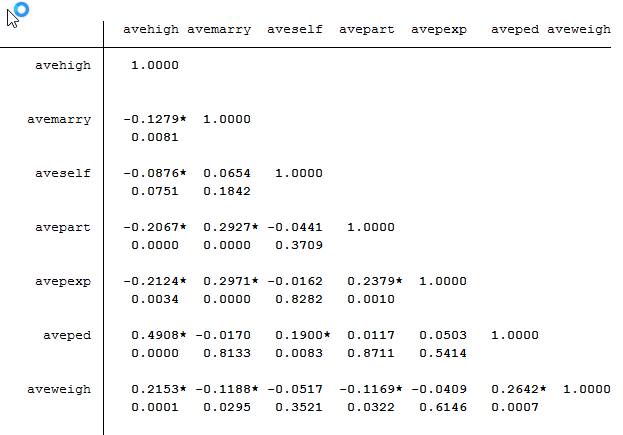
gen dsm18 = sm181-sm182

\*把一行数据拆分成两行，即恢复成原来的初始样子（样本数依旧428，没有删去）

reshape long

\*多变量协方差矩阵（10%显著性）（表3-1-6行的第一列数据）

pwcorr avehigh avemarry aveself avepart avepexp aveped aveweigh, sig st(10)



\*同上（5%显著性）

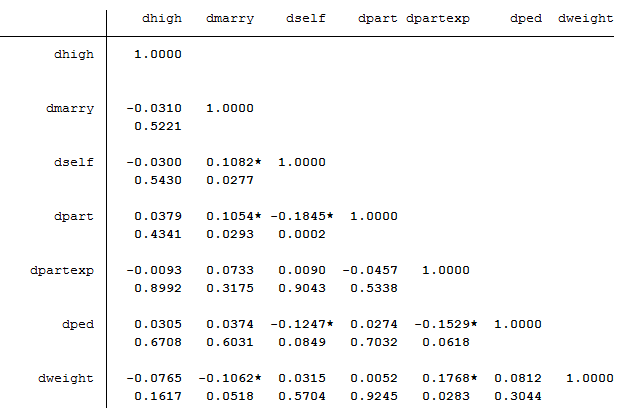
pwcorr avehigh avemarry aveself avepart avepexp aveped aveweigh, sig st(5)

\* 同上（1%显著性）

pwcorr avehigh avemarry aveself avepart avepexp aveped aveweigh, sig st(1)

\*多变量协方差矩阵（差分）（10%显著性）（表3-1-6行的第二列数据）

pwcorr dhigh dmarry dself dpart dpartexp dped dweight, sig st(10)



\*同上（5%显著性）

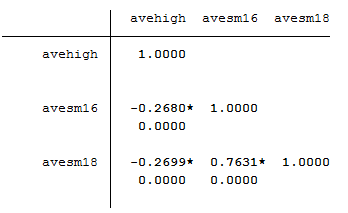
pwcorr dhigh dmarry dself dpart dpartexp dped dweight, sig st(5)

\* 同上（1%显著性）

pwcorr dhigh dmarry dself dpart dpartexp dped dweight, sig st(1)

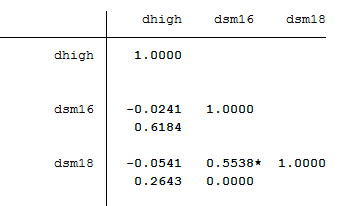
\*多变量协方差矩阵（1%显著性）（表4-第一列数据）

pwcorr avehigh avesm16 avesm18, sig st(1)



\*多变量协方差矩阵（差分）（10%显著性）（表4-第二列数据）

pwcorr dhigh dsm16 dsm18, sig st(10)



\*同上（5%显著性）

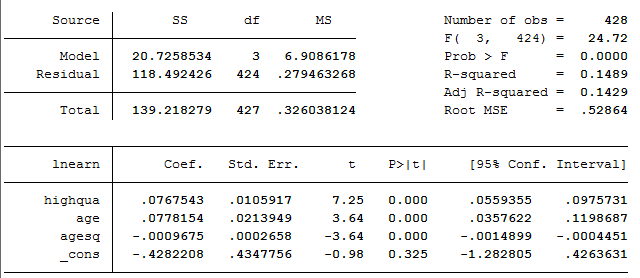
pwcorr dhigh dsm16 dsm18, sig st(5)

\* 同上（1%显著性）

pwcorr dhigh dsm16 dsm18, sig st(1)

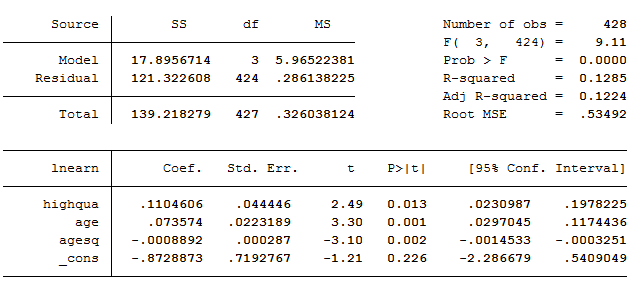
\*做线性回归 工资对数关于教育年限、年龄的二元二次项回归（表五-第一列数据）

reg lnearn highqua age agesq



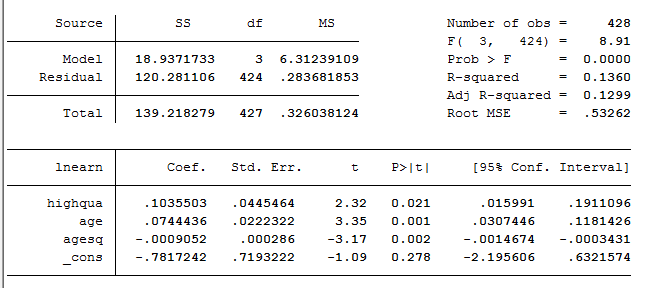
\*同上，工具变量线性回归（把highqua用sm16取代）（表五-第二列数据）

ivreg lnearn (highqua=sm16) age agesq



\*同上，工具变量线性回归（把highqua用sm18取代）（表五-第三列数据）

ivreg lnearn (highqua=sm18) age agesq



log close

clear