```
***************************************************************************************
***************************************************************************************
* The Elasticity of Substitution between Skilled and Unskilled Labor: A Meta-Analysis *
***************************************************************************************
***************************************************************************************
* July 01, 2020
log using skill.log, replace
import excel skill.xlsx, sheet("data") firstrow
set more off
***************************************************************************************
* Summary statistics
***************************************************************************************
winsor2 elasticity, suffix(_w25) cuts(2.5 97.5)
winsor2 se, suffix(_w25) cuts(2.5 97.5)
winsor2 elasticity, suffix(_w30) cuts(3 97)
winsor2 se, suffix(_w30) cuts(3 97)
winsor2 elasticity, suffix(_w35) cuts(3.5 96.5)
winsor2 se, suffix(_w35) cuts(3.5 96.5)
winsor2 elasticity, suffix(_w40) cuts(4 96)
winsor2 se, suffix(_w40) cuts(4 96)
winsor2 elasticity, suffix(_w45) cuts(4.5 95.5)
winsor2 se, suffix(_w45) cuts(4.5 95.5)
winsor2 elasticity, suffix(_w50) cuts(5 95)
winsor2 se, suffix(_w50) cuts(5 95)
gen elasticity_w = elasticity_w40
gen se_w = se_w40
gen t_statistics_w = elasticity_w/se_w
gen precision_w = 1/se_w

univar elasticity elasticity_w se se_w
sum elasticity_w se_w hicks_elasticity other_elasticity inverted_estimate direct_estimate
skilled_by_college skilled_by_high_school skilled_by_occupation higher_frequency annual_frequency
lower_frequency micro_data sectoral_data aggregated_data cross_section data_midyear data_length
data_size united_states developing_country male_workers manufacturing_sector onelevel_ces_function
multilevel_ces_function other_function time_control location_control education_control macro_control
population_control sector_control age_control ethnicity_control capital_control dynamic_model
unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method ml_method
impact_factor citations published_study
sum elasticity_w se_w hicks_elasticity other_elasticity inverted_estimate direct_estimate
skilled_by_college skilled_by_high_school skilled_by_occupation higher_frequency annual_frequency
lower_frequency micro_data sectoral_data aggregated_data cross_section data_midyear data_length
data_size united_states developing_country male_workers manufacturing_sector onelevel_ces_function
multilevel_ces_function other_function time_control location_control education_control macro_control
population_control sector_control age_control ethnicity_control capital_control dynamic_model
unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method ml_method
impact_factor citations published_study [aweight=weight]
correlate se_w hicks_elasticity other_elasticity inverted_estimate skilled_by_college
skilled_by_high_school skilled_by_occupation higher_frequency annual_frequency lower_frequency
micro_data sectoral_data aggregated_data cross_section data_midyear data_length data_size
united_states developing_country male_workers manufacturing_sector onelevel_ces_function
multilevel_ces_function other_function time_control location_control education_control macro_control
population_control sector_control age_control ethnicity_control capital_control dynamic_model
unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method ml_method
impact_factor citations published_study
collin se_w hicks_elasticity inverted_estimate skilled_by_college skilled_by_occupation
higher_frequency lower_frequency micro_data sectoral_data cross_section data_midyear data_length
united_states developing_country male_workers manufacturing_sector onelevel_ces_function
multilevel_ces_function time_control location_control education_control macro_control
population_control sector_control age_control ethnicity_control capital_control dynamic_model
unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method impact_factor
citations published_study

mean elasticity_w
mean elasticity_w if elasticity_w>1
```

```
mean elasticity_w if 0<elasticity_w & elasticity_w<1
mean elasticity_w if 1<elasticity_w & elasticity_w<2
mean elasticity_w if longrun_effect==0
mean elasticity_w if longrun_effect==1
mean elasticity_w if hicks_elasticity==1
mean elasticity_w if other_elasticity==1
mean elasticity_w if inverted_estimate==1
mean elasticity_w if inverted_estimate==0
mean elasticity_w if skilled_by_college==1
mean elasticity_w if skilled_by_high_school==1
mean elasticity_w if skilled_by_occupation==1
mean elasticity_w if higher_frequency==1
mean elasticity_w if annual_frequency==1
mean elasticity_w if lower_frequency==1
mean elasticity_w if micro_data==1
mean elasticity_w if sectoral_data==1
mean elasticity_w if aggregated_data==1
mean elasticity_w if cross_section==1
mean elasticity_w if cross_section==0
mean elasticity_w if united_states==1
mean elasticity_w if united_states==0
mean elasticity_w if developing_country==1
mean elasticity_w if developing_country==0
mean elasticity_w if male_workers==1
mean elasticity_w if male_workers==0
mean elasticity_w if manufacturing_sector==1
mean elasticity_w if manufacturing_sector==0
mean elasticity_w if onelevel_ces_function==1
mean elasticity_w if multilevel_ces_function==1
mean elasticity_w if other_function==1
mean elasticity_w if dynamic_model==1
mean elasticity_w if unit_fixed_effects==1
mean elasticity_w if time_fixed_effects==1
mean elasticity_w if ols_method==1
mean elasticity_w if iv_method==1
mean elasticity_w if sur_method==1
mean elasticity_w if ml_method==1
mean elasticity_w if published_study==0
mean elasticity_w if published_study==1
mean elasticity_w if top_journal==1

mean elasticity_w [aweight=weight]
mean elasticity_w [aweight=weight] if longrun_effect==0
mean elasticity_w [aweight=weight] if longrun_effect==1
mean elasticity_w [aweight=weight] if hicks_elasticity==1
mean elasticity_w [aweight=weight] if other_elasticity==1
mean elasticity_w [aweight=weight] if inverted_estimate==1
mean elasticity_w [aweight=weight] if inverted_estimate==0
mean elasticity_w [aweight=weight] if skilled_by_college==1
mean elasticity_w [aweight=weight] if skilled_by_high_school==1
mean elasticity_w [aweight=weight] if skilled_by_occupation==1
mean elasticity_w [aweight=weight] if higher_frequency==1
mean elasticity_w [aweight=weight] if annual_frequency==1
mean elasticity_w [aweight=weight] if lower_frequency==1
mean elasticity_w [aweight=weight] if micro_data==1
mean elasticity_w [aweight=weight] if sectoral_data==1
mean elasticity_w [aweight=weight] if aggregated_data==1
mean elasticity_w [aweight=weight] if cross_section==1
mean elasticity_w [aweight=weight] if cross_section==0
mean elasticity_w [aweight=weight] if united_states==1
mean elasticity_w [aweight=weight] if united_states==0
mean elasticity_w [aweight=weight] if developing_country==1
mean elasticity_w [aweight=weight] if developing_country==0
mean elasticity_w [aweight=weight] if male_workers==1
mean elasticity_w [aweight=weight] if male_workers==0
```

```
mean elasticity_w [aweight=weight] if manufacturing_sector==1
mean elasticity_w [aweight=weight] if manufacturing_sector==0
mean elasticity_w [aweight=weight] if onelevel_ces_function==1
mean elasticity_w [aweight=weight] if multilevel_ces_function==1
mean elasticity_w [aweight=weight] if other_function==1
mean elasticity_w [aweight=weight] if dynamic_model==1
mean elasticity_w [aweight=weight] if unit_fixed_effects==1
mean elasticity_w [aweight=weight] if time_fixed_effects==1
mean elasticity_w [aweight=weight] if ols_method==1
mean elasticity_w [aweight=weight] if iv_method==1
mean elasticity_w [aweight=weight] if sur_method==1
mean elasticity_w [aweight=weight] if ml_method==1
mean elasticity_w [aweight=weight] if published_study==0
mean elasticity_w [aweight=weight] if published_study==1
mean elasticity_w [aweight=weight] if top_journal==1

histogram elasticity if elasticity >-10 & elasticity<15, bin(90) fcolor(gs14) lstyle(thin) frequency
xtitle("Estimate of the elasticity of substitution between skilled and unskilled labor") xline(1 2,
lcolor(red)) xlabel(-5 0 1 2 5 10 15) ylabel( ,glcolor(ltbluishgray))
graphregion(color(ltbluishgray)) saving(histogram, replace)
bysort idstudy: egen elasticity_med = median(elasticity)
bysort idstudy: egen midyear_med = median(midyear)
generate elasticity_med_w=elasticity_med
graph twoway (scatter elasticity_med_w midyear_med if elasticity_med_w<7 & elasticity_med_w>0,
msize(*1) msymbol(Oh) yline(1, lpattern(dott)) ylabel( ,glcolor(ltbluishgray))
graphregion(color(ltbluishgray))) (lfit elasticity_med_w midyear_med, lcolor(black) lpattern(dash)),
xtitle("Median year of data") ytitle("Median estimate of the elasticity of substitution") legend(off)
saving(trend, replace)

twoway(hist elasticity if skilled_by_occupation==1 & elasticity >-2 & elasticity<8, bin(80) freq
fcolor(navy) lcolor(navy) legend(label(1 "Skilled by occupation"))) (hist elasticity if
skilled_by_college==1 & elasticity >-2 & elasticity<8, bin(80) freq fcolor(cranberry)
lcolor(cranberry) legend(label(2 "Skilled by college"))) (hist elasticity if
skilled_by_high_school==1 & elasticity >-2 & elasticity<8, bin(80) gap(20) freq  lcolor(gs12)
fcolor(gs12) legend(label(3 "Skilled by high school"))), legend(ring(0) position(2) bmargin(medium)
rows(3) region(lstyle(none))) xtitle("Estimate of the elasticity of substitution between skilled and
unskilled labor")  saving(pattern1, replace)
twoway(hist elasticity if cross_section==0 & elasticity >-2 & elasticity<8, bin(80) freq fcolor(navy)
lcolor(navy) legend(label(1 "Time series or panel"))) (hist elasticity if cross_section==1 &
elasticity >-2 & elasticity<8, bin(80) freq fcolor(cranberry) lcolor(cranberry) legend(label(2
"Cross-sectional data"))), legend(ring(0) position(2) bmargin(medium) rows(2) region(lstyle(none)))
xtitle("Estimate of the elasticity of substitution between skilled and unskilled labor")
saving(pattern2, replace)
twoway(hist elasticity if sectoral_data==1 & elasticity >-2 & elasticity<8, bin(80) freq
fcolor(cranberry) lcolor(cranberry) legend(label(1 "Sectoral data"))) (hist elasticity if
aggregated_data==1 & elasticity >-2 & elasticity<8, bin(80) gap(20) freq  lcolor(gs12) fcolor(gs12)
legend(label(2 "Aggregated data")))(hist elasticity if micro_data==1 & elasticity >-2 & elasticity<8,
bin(80) freq fcolor(navy) lcolor(navy) legend(label(3 "Micro data"))), legend(ring(0) position(2)
bmargin(medium) rows(3) region(lstyle(none)))  xtitle("Estimate of the elasticity of substitution
between skilled and unskilled labor")  saving(pattern3, replace)
twoway(hist elasticity if developing_country==0 & elasticity >-2 & elasticity<8, bin(80) freq
fcolor(cranberry) lcolor(cranberry) legend(label(1 "Developed country")))(hist elasticity if
developing_country==1 & elasticity >-2 & elasticity<8, bin(80) gap(20) freq fcolor(navy) lcolor(navy)
legend(label(2 "Developing country"))) , legend(ring(0) position(2) bmargin(medium) rows(2)
region(lstyle(none))) xtitle("Estimate of the elasticity of substitution between skilled and
unskilled labor")  saving(pattern4, replace)

graph hbox elasticity if elasticity >-15 & elasticity<15, over(author,label(grid)) xsize(2.5)
ysize(4) scale(0.55) yline(1, lcolor (red))  box( 1,lcolor(black) fcolor(none))
marker(1,msymbol(circle_hollow)  mcolor(gs12)) ytitle("Estimated elasticity of substitution between
skilled and unskilled labor") ylabel(, nogrid) saving(studies, replace)

graph hbox elasticity if elasticity >-3 & elasticity<8, over(author,label(grid)) xsize(2.5) ysize(4)
scale(0.55) yline(1, lcolor (red))  box( 1,lcolor(black) fcolor(none))
marker(1,msymbol(circle_hollow)  mcolor(gs12)) ytitle("Estimated elasticity of substitution between
```

```
skilled and unskilled labor") ylabel(, nogrid) saving(studies, replace)
graph hbox elasticity if elasticity >-3 & elasticity<8 & (idcountry!=6 & idcountry!=7 & idcountry!=15
& idcountry!=20 & idcountry!=21 & idcountry!=23 & idcountry!=28), over(country,label(grid)) xsize(6)
ysize(5) scale(0.8) yline(1, lcolor (red))  box( 1,lcolor(black) fcolor(none))
marker(1,msymbol(circle_hollow)  mcolor(gs12)) ytitle("Estimated elasticity of substitution between
skilled and unskilled labor") ylabel(, nogrid) saving(countries, replace)
*************************************************************************************
* PUBLICATION BIAS - Funnel plot (Egger et al., 1997)
*************************************************************************************
twoway scatter precision elasticity if precision<40 & elasticity>-10 & elasticity<10,
ytitle("Precision of the estimate (1/SE)") ylabel( ,glcolor(ltbluishgray)) xtitle("Estimate of the
elasticity of substitution between skilled and unskilled labor") xline(2, lpattern(dott) lcolor
(black)) xline(1.4, lpattern(dash) lcolor (black)) msymbol(smcircle_hollow)
graphregion(color(ltbluishgray)) saving(funnel, replace)
graph twoway (scatter precision elasticity if longrun_effect==1, msize(*1) msymbol(Oh) mcolor(black)
legend(label(1 "long-run"))) (scatter precision elasticity if longrun_effect==0, msize(*1)
msymbol(Oh) legend(label(2 "short-run"))) if precision<100 & elasticity>-15 & elasticity<15,
ytitle("Precision of the estimate (1/SE)") ylabel( ,glcolor(ltbluishgray)) xtitle("Estimate of the
elasticity of substitution") xline(2, lpattern(dott) lcolor (black))
graphregion(color(ltbluishgray)) saving(funnel_both, replace)
*************************************************************************************
*************************************************************************************
* PUBLICATION BIAS - FAT-PET (Stanley, 2005)
*************************************************************************************
*************************************************************************************
xtset idstudy
eststo: ivreg2 elasticity_w se_w, cluster(idstudy)
boottest se_w
boottest _cons

ivreg2 elasticity_w se_w if published_study==1, cluster(idstudy)
ivreg2 elasticity_w se_w if published_study==0, cluster(idstudy)

eststo: xtreg elasticity_w se_w, fe vce(cluster idstudy)
eststo: xtreg elasticity_w se_w, be
eststo: ivreg2 elasticity_w se_w [pweight=weight*weight], cluster(idstudy)
boottest se_w
boottest _cons
eststo: ivreg2 elasticity_w se_w [pweight=precision_w*precision_w], cluster(idstudy)
boottest se_w
boottest _cons
eststo: ivreg2 elasticity_w se_w if top_journal==1, cluster(idstudy)
boottest se_w
boottest _cons
*gen instrument = ln(nobs)
*eststo: ivreg2 elasticity_w (se_w=instrument), cluster(idstudy)
*weak instrument, tried  1/sqrt(nobs), 1/nobs, 1/(nobs*nobs)
esttab using table_bias.tex, se booktabs replace compress title(FAT-PET all\label{tab:fatpet})
star(\sym{*} 0.10 \sym{**} 0.05 \sym{***} 0.01)
eststo clear

*LONG-RUN only
eststo: ivreg2 elasticity_w se_w if longrun_effect==1, cluster(idstudy)
boottest se_w
boottest _cons
eststo: xtreg elasticity_w se_w if longrun_effect==1, fe vce(cluster idstudy)
eststo: ivreg2 elasticity_w se_w [pweight=weight*weight] if longrun_effect==1, cluster(idstudy)
boottest se_w
boottest _cons
eststo: ivreg2 elasticity_w se_w [pweight=precision_w*precision_w] if longrun_effect==1,
cluster(idstudy)
boottest se_w
boottest _cons
eststo: ivreg2 elasticity_w se_w if (top_journal==1 & longrun_effect==1), cluster(idstudy)
boottest se_w
```

```
boottest _cons
esttab using table_bias_long.tex, se booktabs replace compress title(FAT-PET all\label{tab:fatpet})
star(\sym{*} 0.10 \sym{**} 0.05 \sym{***} 0.01)
eststo clear


*HICKS only
eststo: ivreg2 elasticity_w se_w if hicks_elasticity==1, cluster(idstudy)
boottest se_w
boottest _cons
eststo: xtreg elasticity_w se_w if hicks_elasticity==1, fe vce(cluster idstudy)
eststo: ivreg2 elasticity_w se_w [pweight=weight*weight] if hicks_elasticity==1, cluster(idstudy)
boottest se_w
boottest _cons
eststo: ivreg2 elasticity_w se_w [pweight=precision_w*precision_w] if hicks_elasticity==1,
cluster(idstudy)
boottest se_w
boottest _cons
eststo: ivreg2 elasticity_w se_w if (top_journal==1 & hicks_elasticity==1), cluster(idstudy)
boottest se_w
boottest _cons
esttab using table_bias_hicks.tex, se booktabs replace compress title(FAT-PET all\label{tab:fatpet})
star(\sym{*} 0.10 \sym{**} 0.05 \sym{***} 0.01)
eststo clear
*************************************************************************************
*************************************************************************************
* PUBLICATION BIAS - Caliper test (Gerber & Malhotra, 2008)
*************************************************************************************
*************************************************************************************
histogram t_statistics_w if t_statistics_w<20, bin(60) fcolor(gs14) lstyle(thin) frequency normal
xtitle("t-statistics of the elasticity of substitution between skilled and unskilled labor") xlabel(0
1.96 4 6 8 10 12 14 16 18 20) xline(0 1.96 , lcolor(red)) ylabel( ,glcolor(ltbluishgray))
graphregion(color(ltbluishgray)) saving(caliper, replace)

generate significant_w = 0
replace significant_w = 1 if t_statistics_w > 1.96
reg significant_w if t_statistics_w > 1.56 & t_statistics_w < 2.36
 lincom _cons - 0.5
reg significant_w if t_statistics_w > 1.46 & t_statistics_w < 2.46
 lincom _cons - 0.5
reg significant_w if t_statistics_w > 1.36 & t_statistics_w < 2.56
 lincom _cons - 0.5

reg significant_w if t_statistics_w > 1.56 & t_statistics_w < 2.36 & hicks_elasticity==1
 lincom _cons - 0.5
reg significant_w if t_statistics_w > 1.46 & t_statistics_w < 2.46 & hicks_elasticity==1
 lincom _cons - 0.5
reg significant_w if t_statistics_w > 1.36 & t_statistics_w < 2.56 & hicks_elasticity==1
 lincom _cons - 0.5


replace significant_w = 0
replace significant_w = 1 if t_statistics_w > 0
reg significant_w if t_statistics_w > -0.6 & t_statistics_w < 0.6
 lincom _cons - 0.5
reg significant_w if t_statistics_w > -0.7 & t_statistics_w < 0.7
 lincom _cons - 0.5
reg significant_w if t_statistics_w > -0.8 & t_statistics_w < 0.8
 lincom _cons - 0.5

reg significant_w if t_statistics_w > -0.6 & t_statistics_w < 0.6 & hicks_elasticity==1
 lincom _cons - 0.5
reg significant_w if t_statistics_w > -0.7 & t_statistics_w < 0.7 & hicks_elasticity==1
 lincom _cons - 0.5
reg significant_w if t_statistics_w > -0.8 & t_statistics_w < 0.8 & hicks_elasticity==1
```

```
 lincom _cons - 0.5
*************************************************************************************
*************************************************************************************
* PUBLICATION BIAS - FAT-PET (hierarchical) in R
*************************************************************************************
*************************************************************************************
/*
library(bayesm)
datalabor = read.table("clipboard-512", sep="\t", header=TRUE)
str(datalabor)
study <- levels(datalabor$author)
nreg <- length(study); nreg
regdata <- NULL
for (i in 1:nreg) {
        filter <- datalabor$author==study[i]
        y <- datalabor$elasticity_w[filter]
        X <- cbind(1,
        datalabor$se_w[filter])
        regdata[[i]] <- list(y=y, X=X)
}
Data <- list(regdata=regdata)
Mcmc <- list(R=6000)
 out <- bayesm::rhierLinearModel(
         Data=Data,
         Mcmc=Mcmc)
cat("Summary of Delta Draws", fill=TRUE)
summary(out$Deltadraw)
*/
*************************************************************************************
*************************************************************************************
* PUBLICATION BIAS - Stem-based method in R (Furukawa, 2019)
*************************************************************************************
*************************************************************************************
/*
source("stem_method.R") #github.com/Chishio318/stem-based_method
datalabor = read.table("clipboard-512", sep="\t", header=TRUE)
stem_results = stem(datalabor$elasticity, datalabor$se, param)
view(stem_results$estimates)
*/
*************************************************************************************
*************************************************************************************
* PUBLICATION BIAS - TOP10 method (Stanley et al., 2010)
*************************************************************************************
*************************************************************************************
summarize precision_w, detail
gen top10bound = r(p90)
summarize elasticity_w precision_w if precision_w > top10bound

summarize precision_w if longrun_effect==1, detail
gen top10bound_long = r(p90)
summarize elasticity_w precision_w if (longrun_effect==1 & precision_w > top10bound_long)
*************************************************************************************
*************************************************************************************
* PUBLICATION BIAS - WAAP (Ioannidis et al., 2017)
*************************************************************************************
*************************************************************************************
summarize elasticity_w [aweight=1/(se_w*se_w)]
gen waapbound = abs(r(mean))/2.8
summarize elasticity_w if se < waapbound

summarize elasticity_w if longrun_effect==1 [aweight=1/(se_w*se_w)]
gen waapbound_long = abs(r(mean))/2.8
summarize elasticity_w if se < waapbound_long
*************************************************************************************
*************************************************************************************
```

```
* HETEROGENEITY - OLS
*************************************************************************************
*************************************************************************************
ivreg2 elasticity_w se_w hicks_elasticity inverted_estimate skilled_by_college skilled_by_occupation
higher_frequency lower_frequency micro_data sectoral_data cross_section data_midyear data_length
united_states developing_country male_workers manufacturing_sector onelevel_ces_function
multilevel_ces_function time_control location_control education_control macro_control
population_control sector_control age_control ethnicity_control capital_control dynamic_model
unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method impact_factor
citations published_study, cluster(idstudy)
stepwise, pr(.05): regress elasticity_w se_w hicks_elasticity inverted_estimate skilled_by_college
skilled_by_occupation higher_frequency lower_frequency micro_data sectoral_data cross_section
data_midyear data_length united_states developing_country male_workers manufacturing_sector
onelevel_ces_function multilevel_ces_function time_control location_control education_control
macro_control population_control sector_control age_control ethnicity_control capital_control
dynamic_model unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method
impact_factor citations published_study, cluster(idstudy)
ivreg2 elasticity_w se_w inverted_estimate higher_frequency cross_section developing_country
onelevel_ces_function multilevel_ces_function time_control sector_control unit_fixed_effects
iv_method, cluster(idstudy)

ivreg2 elasticity_w se_w hicks_elasticity inverted_estimate skilled_by_college skilled_by_occupation
higher_frequency lower_frequency micro_data sectoral_data cross_section data_midyear data_length
united_states developing_country male_workers manufacturing_sector onelevel_ces_function
multilevel_ces_function time_control location_control education_control macro_control
population_control sector_control age_control ethnicity_control capital_control dynamic_model
unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method impact_factor
citations published_study [aweight=weight*weight], cluster(idstudy)
stepwise, pr(.05): regress elasticity_w se_w hicks_elasticity inverted_estimate skilled_by_college
skilled_by_occupation higher_frequency lower_frequency micro_data sectoral_data cross_section
data_midyear data_length united_states developing_country male_workers manufacturing_sector
onelevel_ces_function multilevel_ces_function time_control location_control education_control
macro_control population_control sector_control age_control ethnicity_control capital_control
dynamic_model unit_fixed_effects time_fixed_effects longrun_effect ols_method iv_method sur_method
impact_factor citations published_study [aweight=weight*weight], cluster(idstudy)
ivreg2 elasticity_w se_w hicks_elasticity skilled_by_college sectoral_data data_midyear
onelevel_ces_function multilevel_ces_function time_control education_control macro_control
population_control ols_method iv_method impact_factor citations published_study
[aweight=weight*weight], cluster(idstudy)

*************************************************************************************
*************************************************************************************
* HETEROGENEITY - BAYEASIAN MODEL AVERAGING ////CODE for R////
*************************************************************************************
*************************************************************************************
/*
library(BMS)
    *ctrl+C from labor.xlsx, sheet("data.r")
datalabor = read.table("clipboard-512", sep="\t", header=TRUE)
labor = bms(datalabor, burn=1e5,iter=3e5, g="UIP", mprior="uniform", nmodel=50000, mcmc="bd",
user.int=FALSE)
labor1 = bms(datalabor, burn=1e5,iter=3e5, g="UIP", mprior="dilut", nmodel=50000, mcmc="bd",
user.int=FALSE)
labor2 = bms(datalabor, burn=1e5,iter=3e5, g="BRIC", mprior="random", nmodel=50000, mcmc="bd",
user.int=FALSE)
labor3 = bms(datalabor, burn=1e5,iter=3e5, g="hyper=BRIC", mprior="random", nmodel=50000, mcmc="bd",
user.int=FALSE)
coef(labor, order.by.pip = F, exact=T, include.constant=T)
image(labor, yprop2pip=FALSE, order.by.pip=TRUE, do.par=TRUE, do.grid=TRUE, do.axis=TRUE, cex.axis =
0.7)
summary(labor)
plot(labor)
print(labor$topmod[1])

library(corrplot)
```

```
datalabor = read.table("clipboard-512", sep="\t", header=TRUE)
col<- colorRampPalette(c("red", "white", "blue"))
M <- cor(datalabor)
corrplot.mixed(M, lower = "number", upper = "circle", lower.col=col(200), upper.col=col(200), tl.pos
= c("lt"), diag = c("u"), tl.col="black", tl.srt=45, tl.cex=0.85, number.cex = 0.5, cl.cex=0.8,
cl.ratio=0.1)
*/
*************************************************************************************
*************************************************************************************
* HETEROGENEITY - FREQUENTIST MODEL AVERAGING (MALLOWS) ////CODE for R////
*************************************************************************************
*************************************************************************************
/*
library(foreign)
library(xtable)
library(LowRankQP)
datalabor=read.table("clipboard-512", sep="\t", header=TRUE)
datalabor <-na.omit(datalabor)
x.data <- datalabor[,-1]
const_<-c(1)
x.data <-cbind(const_,x.data)

x <- sapply(1:ncol(x.data),function(i){x.data[,i]/max(x.data[,i])})
scale.vector <- as.matrix(sapply(1:ncol(x.data),function(i){max(x.data[,i])}))
Y <- as.matrix(datalabor[,1])
output.colnames <- colnames(x.data)
full.fit <- lm(Y~x-1)
beta.full <- as.matrix(coef(full.fit))
M <- k <- ncol(x)
n <- nrow(x)
beta <- matrix(0,k,M)
e <- matrix(0,n,M)
K_vector <- matrix(c(1:M))
var.matrix <- matrix(0,k,M)
bias.sq <- matrix(0,k,M)

for(i in 1:M)
{
  X <- as.matrix(x[,1:i])
  ortho <- eigen(t(X)%*%X)
  Q <- ortho$vectors ; lambda <- ortho$values
  x.tilda <- X%*%Q%*%(diag(lambda^-0.5,i,i))
  beta.star <- t(x.tilda)%*%Y
  beta.hat <- Q%*%diag(lambda^-0.5,i,i)%*%beta.star
  beta[1:i,i] <- beta.hat
  e[,i] <- Y-x.tilda%*%as.matrix(beta.star)
  bias.sq[,i] <- (beta[,i]-beta.full)^2
  var.matrix.star <- diag(as.numeric(((t(e[,i])%*%e[,i])/(n-i))),i,i)
  var.matrix.hat <- var.matrix.star%*%(Q%*%diag(lambda^-1,i,i)%*%t(Q))
  var.matrix[1:i,i] <- diag(var.matrix.hat)
  var.matrix[,i] <- var.matrix[,i]+ bias.sq[,i]
}

e_k <- e[,M]
sigma_hat <- as.numeric((t(e_k)%*%e_k)/(n-M))
G <- t(e)%*%e
a <- ((sigma_hat)^2)*K_vector
A <- matrix(1,1,M)
b <- matrix(1,1,1)
u <- matrix(1,M,1)
optim <- LowRankQP(Vmat=G,dvec=a,Amat=A,bvec=b,uvec=u,method="LU",verbose=FALSE)
weights <- as.matrix(optim$alpha)
beta.scaled <- beta%*%weights
final.beta <- beta.scaled/scale.vector
std.scaled <- sqrt(var.matrix)%*%weights
```

```
final.std <- std.scaled/scale.vector
results.reduced <- as.matrix(cbind(final.beta,final.std))
rownames(results.reduced) <- output.colnames; colnames(results.reduced) <- c("Coefficient", "Sd.
Err")
MMA.fls <- round(results.reduced,4)
MMA.fls <- data.frame(MMA.fls)
t <- as.data.frame(MMA.fls$Coefficient/MMA.fls$Sd..Err)
MMA.fls$pv <-round( (1-apply(as.data.frame(apply(t,1,abs)), 1, pnorm))*2,3)
MMA.fls$names <- rownames(MMA.fls)
names <- c(colnames(datalabor))
names <- c(names,"const_")
MMA.fls <- MMA.fls[match(names, MMA.fls$names),]
MMA.fls$names <- NULL
MMA.fls
*/
*************************************************************************************
*************************************************************************************
* BEST-PRACTICE (SE calculation to BMA estimate)
*************************************************************************************
*************************************************************************************
summarize data_midyear, detail
summarize data_length, detail
summarize impact_factor, detail
summarize citations, detail


ivreg2 elasticity_w se_w inverted_estimate higher_frequency cross_section developing_country
onelevel_ces_function multilevel_ces_function time_control sector_control unit_fixed_effects
iv_method, cluster(idstudy)

* Our best-practice

*Overall
lincom
_cons+se_w*0+inverted_estimate*0.534+higher_frequency*0+cross_section*0+developing_country*0.311+onel
evel_ces_function*0+multilevel_ces_function*1+time_control*1+sector_control*0.044+unit_fixed_effects*
1+iv_method*1
*Homogenous data
lincom
_cons+se_w*0+inverted_estimate*0.534+higher_frequency*0+cross_section*0+developing_country*0.311+onel
evel_ces_function*0+multilevel_ces_function*1+time_control*1+sector_control*0.044+unit_fixed_effects*
1+iv_method*1
*Developing
lincom
_cons+se_w*0+inverted_estimate*0.534+higher_frequency*0+cross_section*0+developing_country*1+onelevel
_ces_function*0+multilevel_ces_function*1+time_control*1+sector_control*0.044+unit_fixed_effects*1+iv
_method*1
*USA
lincom
_cons+se_w*0+inverted_estimate*0.534+higher_frequency*0+cross_section*0+developing_country*0+onelevel
_ces_function*0+multilevel_ces_function*1+time_control*1+sector_control*0.044+unit_fixed_effects*1+iv
_method*1


* Autor (2014)


*Overall
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*0.00+developing_country*0
.311+0.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*1.00+sector_control*0.
00+unit_fixed_effects*0.00+iv_method*0.00
*Homogenous data
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*0.00+developing_country*0
```

```
.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*1.00+sector_control*0.00+uni
t_fixed_effects*0.00+iv_method*0.00
*Developing
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*0.00+developing_country*1
.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*1.00+sector_control*0.00+uni
t_fixed_effects*0.00+iv_method*0.00
*USA
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*0.00+developing_country*0
.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*1.00+sector_control*0.00+uni
t_fixed_effects*0.00+iv_method*0.00


* Card (2009)


*Overall
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*1.00+developing_country*0
.311+0.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*0.00+sector_control*0.
00+unit_fixed_effects*0.00+iv_method*1.00
*Homogenous data
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*1.00+developing_country*0
.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*0.00+sector_control*0.00+uni
t_fixed_effects*0.00+iv_method*1.00
*Developing
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*1.00+developing_country*1
.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*0.00+sector_control*0.00+uni
t_fixed_effects*0.00+iv_method*1.00
*USA
lincom
_cons+se_w*0.00+inverted_estimate*0.534+higher_frequency*0.00+cross_section*1.00+developing_country*0
.00+onelevel_ces_function*1.00+multilevel_ces_function*0.00+time_control*0.00+sector_control*0.00+uni
t_fixed_effects*0.00+iv_method*1.00

clear
```