** person.do
** gets person characteristics for poverty mapping
** JKG 18 Dec 2016

clear
scalar drop _all
capture log close
cd "C:\Current\Poverty\SolomonIslands\Mapping"
log using person.log, text replace
set more off
set mem 600m

macro define data "C:\Current\Poverty\SolomonIslands\revised_HIES"
macro define output "C:\Current\Poverty\SolomonIslands\Mapping"

****************************************************************************
***
get household head characteristics
****************************************************************************

use $data\PERSON.dta
drop if p1000!=1
gen head_fem=cond(p1002==2,1,0)
gen head_age=p1003
gen head_married=cond(p1106==2|p1106==3|p1106==4,1,0)
gen head_birth_prov=p1105
gen head_nonmelanesian=cond(p1107!=1,1,0)
gen head_school=0
replace head_school=1 if p3102==2 & p3103==6
replace head_school=2 if p3102==3
replace head_school=3 if p3102==4
replace head_school=4 if p3102>4 & p3102!=. 
label define head_school 0 "Nil/incomplete primary" 1 "Grade 6" 2 "Junior secondary" 3 "Senior secondary" 4 "Tertiary"
label values head_school head_school

gen head_i

***************************************************************************
***
get household averages or totals of personal characteristics
***************************************************************************

use $data\PERSON.dta

* remove list 2 people who used to live in the household (not needed for HH Head, who are all list 1)
drop if p_status==2

* remove list 1 people who are absent during the survey (to match census list of people staying in the HH

drop if p1111==3

* remove anyone with missing age data

drop if p1003==.

* household size and age/sex demographic ratios
gen hhsize=1
gen m06=cond(p1003<7 & p1002==1,1,0)
gen f02=cond(p1003<7 & p1002==2,1,0)
gen m714=cond(p1003>6 & p1003<15 & p1002==1,1,0)
gen f714=cond(p1003>6 & p1003<15 & p1002==2,1,0)
gen m1550=cond(p1003>14 & p1003<51 & p1002==1,1,0)
gen f1550=cond(p1003>14 & p1003<51 & p1002==2,1,0)
gen m51=cond(p1003>=50 & p1002==1,1,0)
gen f51=cond(p1003>=50 & p1002==2,1,0)

* education for ages 12+
gen age12plus=cond(p1003==11,1,0)
gen noschool=cond(p1003==1 & age12plus==1,1,0)
** Jnr Sec is completing Form 3, Snr Sec is completing Form 5, 6 or 7, everything above is tertiary

- those who have completed school
  - gen jnrprimary=cond(p3102==3 & age12plus==1,1,0)
  - gen snrprimary=cond(p3102==4 & age12plus==1,1,0)
  - gen tertiary=cond(p3102>4 & p3102!=. & age12plus==1,1,0)

- those currently in school
  - replace jnrprimary=1 if p3105==3
  - replace snrprimary=1 if p3105==4
  - replace tertiary=1 if p3105>4 & p3105!=.

** Work activity for ages 12+

- Use fewer categories than for household-heads, combine Volunteer with inactive
  - gen empl_inactive=cond(p1200>9,1,0)
  - gen empl_employer=cond(p1200==1,1,0)
  - gen empl_govt=cond(p1200==2,1,0)
  - gen empl_private=cond(p1200==3|p1200==4|p1200==5,1,0)
  - gen empl_self_emp=cond(p1200==6,1,0)
  - gen empl_ownact=cond(p1200==7,1,0)
  - gen empl_unpaid=cond(p1200==8|p1200==9,1,0)
  - gen empl_self=empl_self_emp+empl_ownact+empl_unpaid

for var noschool preprimary someprimary fullprimary jnrsecondary snrsecondary tertiary empl_* : replace X=.
if age12plus==0
collapse (sum) hhsize - empl_self, by(seq_num)
save $output\demogs.dta, replace

** combine with hh_head and get the EA level variables
merge seq_num using $output\head.dta, sort uniquising

tab _merge
keep if _merge==3
drop _merge

** combine with the reweighting file to calculate national level means and SD for comparing with
census
merge seq_num using $output\rweight.dta, sort uniquising keep(rpweight rhhwght)
tab _merge
keep if _merge==3
drop _merge

- gen head_migrant=cond(province!=head_birth_prov,1,0)
- gen head_gr6=cond(head_school==6,1,0)
- gen head_gr7=cond(head_school==7,1,0)
- gen head_tertiary=cond(head_school==8,1,0)
- collapse (sd) head_* [aw=rhhwght]
drop head_*

- collapse (sd) hhsize-empl_self [aw=rhhwght]
- stop
- log close
** c_person.do
** gets person characteristics from CENSUS files
** JKG 26 Dec 2016

clear
scalar drop _all
capture log close
cd "C:\Current\Poverty\SolomonIslands\Mapping"
log using c_person.log, text replace
set more off
set mem 600m

macro define data "C:\Current\Poverty\SolomonIslands\Mapping\Census\R_Working"
macro define output "C:\Current\Poverty\SolomonIslands\Mapping"

****************************************************************************
*** get household head characteristics
****************************************************************************
import delimited $data CensusPersonsHH.csv

```stata
gen head_migrant=cond(province!=head_birth_prov,1,0)
gen head_subgr6=cond(head_school==0,1,0)
gen head_gr6=cond(head_school==1,1,0)
gen head_jnr=cond(head_school==2,1,0)
gen head_snr=cond(head_school==3,1,0)
gen head_tertiary=cond(head_school==4,1,0)
rename head_self head_self_emp
gen head_self = head_self_emp+head_ownact+head_unpaid
sum head_self
```

order province ward ea hhno head_fem head_age head_married head_birth_prov head_nonmelanesian head_school head_inactive head_employer head_govt head_private head_self_emp head_self ownact head_unpaid:

```stata
save $output\c.hhead.dta, replace
```

******************************************************************************
*** get household averages or totals of personal characteristics
******************************************************************************
clear
import delimited "$data HH Characteristics.csv"
egen myhhid=group(province ward ea hhno)

codebook myhhid

" sum the three categories of self-employment and own-account
```
gen empl_self=empl_self_emp+empl_ownact+empl_unpaid
```
order empl_self, after(empl_unpaid)
collapse (mean) province ward ea hhno (sum) hhsize-empl_self, by(myhhid)
summ hhsize-empl_self, sep(0)
```
save $output\c_demogs.dta, replace
```
log close
** house.do

** gets dwelling and hh characteristics to match Section H of census

** JKG 18 Dec 2016

clear

scalar drop _all

capture log close
cd "C:\Current\Poverty\SolomonIslands\Mapping"

log using house.log, text replace

set more off

set mem 600m

macro define data "C:\Current\Poverty\SolomonIslands\revised_HIES"
macro define output "C:\Current\Poverty\SolomonIslands\Mapping"

* restrict to the reweighted data
use $data
merge seq_num using $output

keep if _merge==3

drop _merge

rename c_province province

rename c_ward ward

rename c_ea ea

** best match to census is h16_1 + h16_3 (car/bus/truck)
** also has_motorboat matches with h16_6 (has outboard motor)
** and has_fridge matches with h16_7

gen has_car=cond(h2402_111>0 & h2402_111 !=.,1,0)
gen has_ute=cond(h2402_112>0 & h2402_112 !=.,1,0)
gen has_truck=cond(h2402_113>0 & h2402_113 !=.,1,0)
gen has_motoboat=cond(h2402_114>0 & h2402_114 !=.,1,0)
gen has_canoe=cond(h2402_115>0 & h2402_115 !=.,1,0)
gen has_carbustruck=cond(has_car==1|has_ute==1|has_truck==1,1,0)

** note change in format - while vehicles are counts, these are 1/2 for yes/no

gen has_fridge=cond(h2302_313==1,1,0)
gen has_tv =cond(h2302_611 ==1,1,0)
gen has_computer=cond(h2302_711==1,1,0)

summ has_* [aw=rhhwght]

** dwelling tenure - does not match well to census

gen tenure_rent=cond(h1410==1,1,0)
gen tenure_own=cond(h1410==2,1,0)
gen tenure_mortgage=cond(h1410==3,1,0)
gen tenure_free=cond(h1410==4,1,0)

summ tenure_* [aw=rhhwght]

** dwelling characteristics
* tin roof, makeshift walls and modern floor look like best matches
* number of rooms is a good match

gen dwell_tinroof=cond(h1420_r==2,1,0)
gen dwell_modfloor=cond(h1420_f==3,1,0)
gen dwell_tempwall=cond(h1420_w==5,1,0)
gen dwell_detached=cond(h1420==1,1,0)
gen dwell_nrooms=h1422

summ dwell_* [aw=rhhwght]

** main source of lighting is a bad match

gen light_elec=cond(h1430==1,1,0)
gen light_solar=cond(h1430==4,1,0)
gen light_kero=cond(h1430==5,1,0)

summ light_* [aw=rhhwght]

** main fuel for cooking is better match

gen cook_wood=cond(h1432==1,1,0)
gen cook_gas=cond(h1432==3,1,0)

summ cook_* [aw=rhhwght]

** main source of drinking water also a good match

gen water_meter=cond(h1440==1,1,0)
gen water_standpipe=cond(h1440==2,1,0)
gen water_hhtank=cond(h1440==3,1,0)
gen water_comtank=cond(h1440==4,1,0)
gen water_river=cond(h1440==7,1,0)

** washing water not a good match, but adding river+sea might work [census h7 options 8&9]
gen wash_river_lake_sea=cond(h1447==6|h1447==7,1,0)
** private flush toilet and private pit latrine (census option 1 and 5) seem an ok match

gen toilet_ownflush=cond(h1448==1,1,0)
gen toilet_ownpit=cond(h1448==6,1,0)

summ water_* wash_* toilet_* [aw=rhhwght]

keep has_* tenure_* dwell_* water_* wash_* toilet_* cook_* light_* province ward ea seq_num rhhwght
summ

save $output\dwellings.dta, replace

* collapse (sd) has_* tenure_* dwell_* water_* wash_* toilet_* cook_* light_* [aw=rhhwght]
stop
log close
** c_house.do
** gets dwelling and hh characteristics from Section H of census
** JKG 26 Dec 2016

clear
scalar drop _all
capture log close
cd "C:\Current\Poverty\SolomonIslands\Mapping"
log using c_house.log, text replace
set more off
set mem 600m
macro define output "C:\Current\Poverty\SolomonIslands\Mapping"
use C:\Current\Poverty\SolomonIslands\Mapping\Census\housing_gps.dta
** best match to census is h16_1 + h16_3 (car/bus/truck)
** also has_motorboat matches with h16_6 (has outboard motor)
** and has_fridge matches with h16_7

gen has_car=cond(h16_1>0,1,0)
gen has_ute=cond(h16_3>0,1,0)
gen has_truck=cond(h16_3>0,1,0)
gen has_motobike=cond(h16_2>0,1,0)
gen has_motorboat=cond(h16_6==0,1,0)
gen has_canoe=cond(h16_4>0,1,0)
gen has_carbusstruck=cond(h16_1>0|h16_3>0,1,0)
gen has_fridge=cond(h16_7>0,1,0)
gen has_tv =cond(h16_8>0,1,0)
gen has_computer=cond(h16_10>0,1,0)

summ has_*
** dwelling tenure - based on h12
gen tenure_rent=cond(h12==2,1,0)
gen tenure_own=cond(h12==1,1,0)
gen tenure_mortgage=cond(h12==1,1,0)
gen tenure_free=cond(h12==4,1,0)

summ tenure_*
** dwelling characteristics, based on H2, H3, H4, H5
gen dwell_tinroof=cond(h4==2,1,0)
gen dwell_modfloor=cond(h3==3,1,0)
gen dwell_tempwall=cond(h2==5,1,0)
gen dwell_detached=cond(h1==1,1,0)
gen dwell_nrooms=h5

summ dwell_*
** main source of lighting based on H10
gen light_elec=cond(h10==1,1,0)
gen light_solar=cond(h10==3,1,0)
gen light_kero=cond(h10==5,1,0)

summ light_*
** main fuel for cooking based on H11 is better match
gen cook_wood=cond(h11==3,1,0)
gen cook_gas=cond(h11==5,1,0)

summ cook_*
** main source of drinking water based on H6
gen water_meter=cond(h6==2,1,0)
gen water_standpipe=cond(h6==2,1,0)
gen water_hhtank=cond(h6==3,1,0)
gen water_comtank=cond(h6==4,1,0)
gen water_river=cond(h6==7,1,0)

** washing water not a good match, but adding river+sea might work {census h7 options 849}
gen wash_river_lake_sea=cond(h7==8|h7==9,1,0)

** private flush toilet and private pit latrine {census option 1 and 5} seem an ok match
gen toilet_ownflush=cond(h8==1,1,0)
gen toilet_ownpit=cond(h8==5,1,0)

summ water_* wash_* toilet_*
** keep hh-level of village has church, as a cluster variable
gen Echurch=cond(church==1,1,0)

keep has_* tenure_* dwell_* water_* wash_* toilet_* cook_* light_* Echurch province ward ea hhno
summ

save $output\c_dwelling.dta, replace

collapse (sd) has_* tenure_* dwell_* water_* wash_* toilet_* cook_* light_*

log close
Appendix B: Files Used for Database Construction and Initial Estimation of Beta and Alpha Models in Stata

** c_merge.do
** merges constructed census files and links to survey—census—link
** JRG 26 Dec 2016

clear
capture log close
cd "C:\Current\Poverty\SolomonIslands\Mapping"
set more off
set mem 600m
use c_dwelling.dta, replace
merge province ward ea hhno using c_head.dta, sort uniquising tab _merge
drop _merge
merge province ward ea hhno using c_demogs.dta, sort uniquising tab _merge
drop _merge

** drop unneeded variables and ones that don’t merge
drop v1 excludee myhhid
drop head_fem head_birth_prov head_school head_inactive head_govt head_private head_self_emp head_ownact head_unpaid head_gr6 head_jnr head_snr head_tertiary
drop empl_self_emp empl_owact empl_unpaid noschool preprimary fullprimary jnrsecondary snrsecondary tertiary

* turn demographics into ratios rather than counts
for var someprimary - empl_self : replace X=X/age12plus
for var m06 - f51: replace X=X/hhsize
drop age12plus

drop has_car has_ute has_truck has_motobike has_canoe has_tv has_computer tenure_own tenure_mortgage light_elec light_solar light_kero

** merge with HIES Sample Design file to get urban indicator for each PSU
merge province ward ea using HIES_sample_design, sort keep(urb_rural_code) uniquising
gen urban=cond(urb_rural_code==1,1,0)
replace urban=1 if _merge==1
drop _merge urb_rural_code
sum, sep(0)
for var has_motorboat: egen mean_X=mean(X), by(province ward ea)
merge province using povline.dta, sort uniquising tab _merge
drop _merge

gen hhsize2=hhsize^2
gen dwell_nroom2=dwell_nrooms^2
gen double HHID = (province * 10000000) + (ward * 100000) + (ea * 1000) + hhno

gen double clusterID = (province * 10000) + (ward * 100) + ea

order HHID clusterID province ward ea hhno urban
saveold c_databank.dta, replace
** merge.do  
** merges constructed HIES files and uses survey-census-link to bring in EA means  
** also brings in the consumption and poverty line data  
** JKG 26 Dec 2016  

clear  
scalar drop _all  
cd "C:\Current\Poverty\SolomonIslands\Mapping"  
tempfile survey  
set more off  
set mem 600m  
** use dwelling.dta for the merges because it already has the rhhwght and correct sample size (n=4364)  
use dwelling.dta, replace  
drop has_car has_ute has_truck has_motobike has_canoe has_tv tenure_own tenure_mortgage light_elec light_solar light_kero  
** hh head file needs to have same education categories as census file  
merge seq_num using hhead.dta, sort uniqusing  
tag_merge keep if _merge==3  
gen head_subgr6=cond(head_school==0,1,0)  
gen head_gr6=cond(head_school==1,1,0)  
gen head_jnr=cond(head_school==2,1,0)  
gen head_snr=cond(head_school==3,1,0)  
gen head_tertiary=cond(head_school==4,1,0)  
gen head_migrant=cond(head_birth_prov!=province,1,0)  
drop _merge head_fem head_birth_prov head_school head_inactive head_govt head_private head_self_emp  
head_ownact head_unpaid head_gr6 head_jnr head_snr head_tertiary p1200 p3100 p3102 p3103  
** demogs file needs to turn counts into ratios  
merge seq_num using demogs.dta, sort uniqusing  
tag_merge keep if _merge==3  
drop _merge empl_self_emp empl_ownact empl_unpaid empl_govt empl_private noschool preprimary fullprimary jnsecondary srnsecondary tertiary  
* turn demographics into ratios rather than counts  
for var someprimary = empl_self : replace X=X/age12plus  
for var m06 - f51: replace X=X/hhsize  
drop age12plus  
** merge with HIES Sample Design file to get urban indicator for each PSU  
** and to get the bridge to the EA means  
merge province ward ea using census_survey_link, sort uniqusing keep(urban c_province c_ward c_ea)  
tag_merge  
drop _merge  
save 'survey', replace  
use c_databank.dta, clear  
collapse (p50) mean *, by(province ward ea)  
rename province c_province  
rename ward c_ward  
rename ea c_ea  
merge c_province c_ward c_ea using 'survey', sort uniqmaster  
tag_merge keep if _merge==3  
drop _merge  
order seq_num province ward ea urban rhhwght has_motorboat has_car has_truck has_motobike has_fridge has_computer  
tenure_rent tenure_free dwell_tinroof dwell_modfloor dwell_temppwall dwell_dettached dwell_nrooms cook_gas cook_water_meter water_standpipe water_comtank water_river lake sea toilet_sumpflush toilet_nompit head_age head_married head_nonmelanesian head_employer head_volun head_self_subgr6 head_migrant hhsize m06 f06 m714 f714 m1550 f1550 m51 f51 someprimary empl_inactive empl_employer empl_self  
** now bring in the expenditure data and poverty lines  
merge seq_num using C:\Current\Poverty\SolomonIslands\PovAnalysis_2014\Analysis\welfare_data.dta, sort uniqusing keep(texpae zu)  
tag_merge  
drop _merge  
gen ltexpae=ln(texpae)  
gen hhsize2=hhsize^2  
gen dwell_nrooms2=dwell_nrooms^2  
** generate hierarchical ID (base on c_* to match census)
gen hhno=((seq_num/1000)-(int(seq_num/1000)))*1000
gen double HHID = (c_province * 10000000) + (c_ward * 100000) + (c_ea * 1000) + hhno
gen double clusterID = (c_province * 10000) + (c_ward * 100) + c_ea
order HHID clusterID seq_num province ward ea hhno urban rhhwght texpae ltxpae zu
saveold databank.dta, replace
drop mean_f51 meanempl_govt mean_urban
preserve
** beta model
stepwise, pr(.1): regress ltexpae has_motorboat has_carbustruck has_fridge tenure_rent tenure_free
dwell_tinroof dwell_modfloor dwell_tempwall dwell_detached dwell_nrooms dwell_nroom2 cook wood cook gas
water_meter water_standpipe water_hhtank water_comtank water_river wash_river_lake_sea toilet_owndish
toilet_owndish head_age head_married head_nonmelanesian head_employer head_volun head_self head_subgr6
head_migrant hhsize hhsize2 m06 f06 m714 f714 m1550 f1550 m51 someprimary empl_inactive empl_employer
empl_self mean.* [pw=rhhwght]
*outreg using betatable, replace 10pct
display_result(8)
indeplist
** alpha model
predict yhat
gen yhat2=yhat^2
predict uhat, resid
egen muhat=mean(uhat), by(clusterID)
gen epsilon2=(uhat-muhat)^2
egen a=max(epsilon2)
gen alpha=ln(epsilon2/(a-epsilon2))
local myx "has_motorboat has_carbustruck has_fridge mean_has_carbustruck mean_water_hhtank dwell_tinroof
mean_empl_self mean_tenure_free mean_hhsize dwell_nrooms mean_m06 cook_wood mean_head_age water_meter
water_standpipe water_comtank water_river wash_river_lake_sea toilet_owndish mean_water_river
mean_has_fridge mean_water_standpipe head_nonmelanesian mean_empl_inactive mean_m1550 mean_water_meter
mean_dwell_tempwall mean_dwell_nroom2 hhsize hhsize2 m06 f06 mean_head_self mean_dwell_nrooms m1550 f1550
mean_EAchurch mean_dwell_tinroof empl_inactive mean_head_employer empl_self mean_has_motorboat"
foreach x of local myx {
gen y`x' = yhat*`x'
gen y2`x' = yhat2*`x'
}
* stepwise, pr(.05): regress alpha 'myx' y* [pw=rhhwght]
display_result(8)
indeplist
** on separate domains
** Honiara
preserve
keep if province==10
stepwise, pr(.1): regress ltexpae has_motorboat has_carbustruck has_fridge tenure_rent tenure_free
dwell_tinroof dwell_modfloor dwell_tempwall dwell_detached dwell_nrooms dwell_nroom2 cook wood cook gas
water_meter water_standpipe water_hhtank water_comtank water_river wash_river_lake_sea toilet_owndish
toilet_owndish head_age head_married head_nonmelanesian head_employer head_volun head_self head_subgr6
head_migrant hhsize hhsize2 m06 f06 m714 f714 m1550 f1550 m51 someprimary empl_inactive empl_employer
empl_self mean.* [pw=rhhwght]
*outreg using betatable, append 10pct
display_result(8)
indeplist
** alpha model
predict yhat
gen yhat2=yhat^2
predict uhat, resid
egen muhat=mean(uhat), by(clusterID)
gen epsilon2=(uhat-muhat)^2
egen a=max(epsilon2)
gen alpha=ln(epsilon2/(a-epsilon2))
local myx "mean_has_fridge has_carbustruck has_fridge mean_toilet_owndish tenure_free dwell_tinroof
mean_hhsize dwell_tinroof dwell_detached mean_EAchurch dwell_nroom2 mean_water_river cook_wood
mean_toilet_owndish mean_wash_river_lake_sea mean_head_self mean_cook_wood empl_inactive
wash_river_lake_sea toilet_owndish mean_dwell_tinroof head_married head_nonmelanesian head_employer
mean_head_nonmelanesian mean_water_standpipe mean_tenure_free hhsize hhsize2 mean_dwell_tempwall
someprimary"
foreach x of local myx {
gen y`x' = yhat*`x'
gen y2`x' = yhat2*`x'
}
* stepwise, pr(.05): regress alpha 'myx' y* [pw=rhhwght]
display_result(8)
indeplist
restore
** rural sector
preserve
drop if province==10
keep if urban==0
stepwise, pr(.1): regress ltxpae has_motorboat has_carbuscruck has_fridge tenure_rent tenure_free
dwell_tinroof dwell_modfloor dwell_tempwall dwell_detached dwell_nrrooms dwell_nrroom2 cook_wood cook_gas
water_meter water_standpipe water_hhtank water_comtank water_river wash_river_lake_sea toilet_ownflush
toilet_ownpit head_age head_married head_nonmelanesian head_employer head_volun head_self head_subgr6
head_migrant hhsize hhsize2 m06 f06 f714 f714 f1550 f1550 m51 someprimary empl_inactive empl_employer
empl_self mean_* [pw=rhhwght]
*outreg using betatable, append 10pct
display_result() indeplist

** alpha model
predict yhat
gen yhat2=yhat^2
predict yhat, resid
egen muhat=mean(yhat), by(clusterID)
gen epsilon2=(yhat-muhat)^2
egen a=max(epsilon2)
gen alpha=ln(epsilon2/(a-epsilon2))
local myx "has_motorboat mean_cook_gas mean_has_fridge mean_water_river dwell_tinroof mean_dwell_tempwall
mean_head_subgr6 mean_tenant paid mean_dwell_modfloor mean_dwell_tempwall mean_dwell_detached mean_dwell_nrrooms mean_dwell_nrroom2
mean_carbuscruck mean_f714 head_self head_subgr6 mean_water_hhtank hhsize hhsize2 m06 f06 mean_dwell_nrrooms mean_head_age
mean_m1550 f1550 mean_m06 mean_m714 empl_inactive mean_dwell_tinroof empl_self mean_has_motorboat"
foreach x of local myx {
gen y'x' = yhat'*x'
gen y2'x' = yhat2'*x'
}
stepwise, pr(.05): regress alpha 'myx' y* [pw=rhhwght]
display_result() indeplist restore

** urban sector
preserve
drop if province==10
keep if urban==1
drop mean_f06 mean_f714 mean_f1550 mean_m51 mean_empl_employer
stepwise, pr(.1): regress ltxpae has_motorboat has_carbuscruck has_fridge tenure_rent tenure_free
dwell_tinroof dwell_modfloor dwell_tempwall dwell_detached dwell_nrrooms dwell_nrroom2 cook_wood cook_gas
water_meter water_standpipe water_hhtank water_comtank water_river wash_river_lake_sea toilet_ownflush
toilet_ownpit head_age head_married head_nonmelanesian head_employer head_volun head_self head_subgr6
head_migrant hhsize hhsize2 m06 f06 f714 f714 f1550 f1550 m51 someprimary empl_inactive empl_employer
empl_self mean_* [pw=rhhwght]
*outreg using betatable, append 10pct
display_result() indeplist restore

** alpha model
predict yhat
gen yhat2=yhat^2
predict yhat, resid
egen muhat=mean(yhat), by(clusterID)
gen epsilon2=(yhat-muhat)^2
egen a=max(epsilon2)
gen alpha=ln(epsilon2/(a-epsilon2))
local myx "has_motorboat has_carbuscruck has_fridge mean_someprimary mean_toilet_ownpit dwell_modfloor
dwell_tempwall mean_head_married mean_head_paid mean_dwell_modfloor mean_dwell_tempwall mean_dwell_detached mean_dwell_nrrooms mean_dwell_nrroom2
mean_carbuscruck water_standpipe empl_self m1550 f1550 mean_water_hhtank toilet_ownflush toilet_ownpit mean_m1550
mean_has_motorboat head_nonmelanesian mean_cook_gas mean_cook_wood mean_has_fridge mean_dwell_modfloor mean_dwell_tempwall
mean_tenant paid mean_head_subgr6"
foreach x of local myx {
gen y'x' = yhat'*x'
gen y2'x' = yhat2'*x'
}
stepwise, pr(.05): regress alpha 'myx' y* [pw=rhhwght]
display_result() indeplist restore