caso_pp

May 26, 2021

1 Datos panel: una aplicación a temas de salud y bienestar

Alumna: Alinne Fuentes Trujillo

Fecha: 25 de Mayo de 2021

1.1 0. Reading data

[2]:
```
* Reads gbdcovar data
local path1 /Users/cesar/github/bananalinne/Examen
use "`path1'/gbdcovar", clear
```

[3]:
```
describe
```

```
Contains data from /Users/cesar/github/bananalinne/Examen/gbdcovar.dta
  obs:       528,080
 vars:            44                          12 May 2021 17:13
------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------------------------
location_id     int     %8.0g
location_name   str37   %37s
sex_id          byte    %8.0g
age_group_id    int     %8.0g
year_id         int     %ty
asfr            float   %9.0g                 Age specific fertility rate
educ            float   %9.0g                 Education years per capita
majdep          float   %9.0g                 Major depression
physact         float   %9.0g                 Physical activity met-min/week
prevobes        float   %9.0g                 Obesity prevalence
prevover        float   %9.0g                 Obesity and overweight prevalence
sev_act         float   %9.0g                 SEV for low physical activity
sev_transfat    float   %9.0g                 SEV high transfat diet
sev_kidney      float   %9.0g                 SEV impaired kidney function
sev_sickle      float   %9.0g                 SEV hemoglobinopathies prevalence
covariate_id    int     %8.0g
```

```
covariate_nam~t str17   %17s
adult_under     float   %9.0g                 Age-standardized proportion adult
                                                underweight
bmd             float   %9.0g                 Bone mineral density 60+
alc_binge       float   %9.0g                 Alcohol binge drinker proportion
birthwgt        float   %9.0g                  Mean birthweight
choles          float   %9.0g                 Mean cholesterol
bmi             float   %9.0g                 Mean BMI
tfa             float   %9.0g                  Diet high in trans fatty acids
undwgtall       float   %9.0g                 All age underweight
sev_sc_kidney   float   %9.0g                 SEV age st.  impaired kidney
                                                function
diabetes_prev   float   %9.0g                 Diabetes prevalence age st.
diabetes_015    float   %9.0g                 Diabetes death rate under 15
latitude        float   %9.0g                 Absolute value of average latitude
anc1            float   %9.0g                 Proportion receiving antenatal
                                                care from skilled provider
teen_birth      float   %9.0g                 Tasa de nacimientos de
                                                adolescentes, 10-19
edu_aaid        float   %9.0g                 Education Absolute Inequality
fiber           float   %9.0g                 Fiber g/p/d available
gallneg         float   %9.0g                 Gallup: Negative Experience Index
gdp             float   %9.0g                 PIB per capita base 2010
haqi            float   %9.0g                 Healthcare access and quality
                                                index
sugar           float   %9.0g                 Sugar g/p/d avail
tfr             float   %9.0g                 Total fertility rate
uhc             float   %9.0g                 Universal health coverage
sev_alcohol     float   %9.0g                 Age-standardized SEV for Impaired
                                                kidney function
sev_drugs       float   %9.0g                 Age-standardized SEV for illicit
                                                drugs
sev_chabuse     float   %9.0g                 Age-standardized SEV for child
                                                abuse
sev_ipabuse     float   %9.0g                 Age-standardized SEV for intimate
                                                partner abuse
age_yrs_med     float   %9.0g                 Median age of age group
------------------------------------------------------------------------------
Sorted by: uhc
```

[57]:
```
summarize
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
 location_id |    528,080    714.7909    1408.035          6       4674
location_n~e |          0
      sex_id |    528,080         1.5    .5000005          1          2
```

age_group_id	528,080	23.34783	45.89274	2	235
year_id	528,080	1999.5	11.54341	1980	2019
asfr	528,080	.0134368	.043781	0	.3806646
educ	528,080	7.325114	4.22949	.0010855	15.99808
majdep	528,080	.0246385	.0184985	0	.1302928
physact	528,080	2885.607	2591.049	0	12641.01
prevobes	528,080	.1327815	.1308562	0	.7343736
prevover	528,080	.3544757	.2501232	0	.9345977
sev_act	528,080	.0055198	.0044641	0	.0153319
sev_transfat	528,080	.1068829	.1973945	0	.9821594
sev_kidney	528,080	.0580895	.0644674	0	.2459489
sev_sickle	528,080	2.531625	2.34816	.3105117	10.07606
covariate_id	528,080	1100	0	1100	1100
covariate_~t	0				
adult_under	528,080	.0528687	.0663087	.0003889	.4543971
bmd	528,080	.7717294	.0734517	.59158	.9379107
alc_binge	528,080	.2628549	.1642318	.0008222	.848402
birthwgt	528,080	1.448149	.6690489	.3112738	4.062741
choles	528,080	2.997545	.4806424	1.428009	4.055423
bmi	528,080	25.64284	2.352385	19.81361	32.74129
tfa	528,080	.0079072	.0065415	.0013538	.0271751
undwgtall	528,080	.0544941	.0683714	.0003273	.4589998
sev_sc_kid~y	528,080	.3557911	.1577356	.0223919	.6329808
diabetes_p~v	528,080	.0617895	.0352453	.0172169	.2729402
diabetes_015	528,080	.3004895	.3540529	.0120945	4.020514
latitude	528,080	27.67945	15.98249	.4079866	74.72861
anc1	528,080	.9035499	.148769	.1140436	.9993807
teen_birth	528,080	.0151456	.0101564	.0003696	.0567445
edu_aaid	528,080	3.398558	.8230227	.4799097	5.976593
fiber	528,080	25.93144	8.586826	3.212216	59.36975
gallneg	528,080	27.76113	4.755153	15.16364	50.92962
gdp	528,080	19587.75	21781.28	131.8428	314172.7
haqi	528,080	56.09664	23.70696	5.057151	96.77751
sugar	528,080	97.41193	49.15931	5.749934	219.96
tfr	528,080	3.092655	1.669233	.8659061	8.238059
uhc	528,080	61.4848	20.50467	7.386258	96.70595
sev_alcohol	528,080	.1225478	.0674823	.00003	.3383737
sev_drugs	528,080	.0074801	.0061761	.0016291	.0433674
sev_chabuse	528,080	.0736899	.0293877	.0185638	.1679942
sev_ipabuse	528,080	.0259368	.0232403	.0015454	.1382152

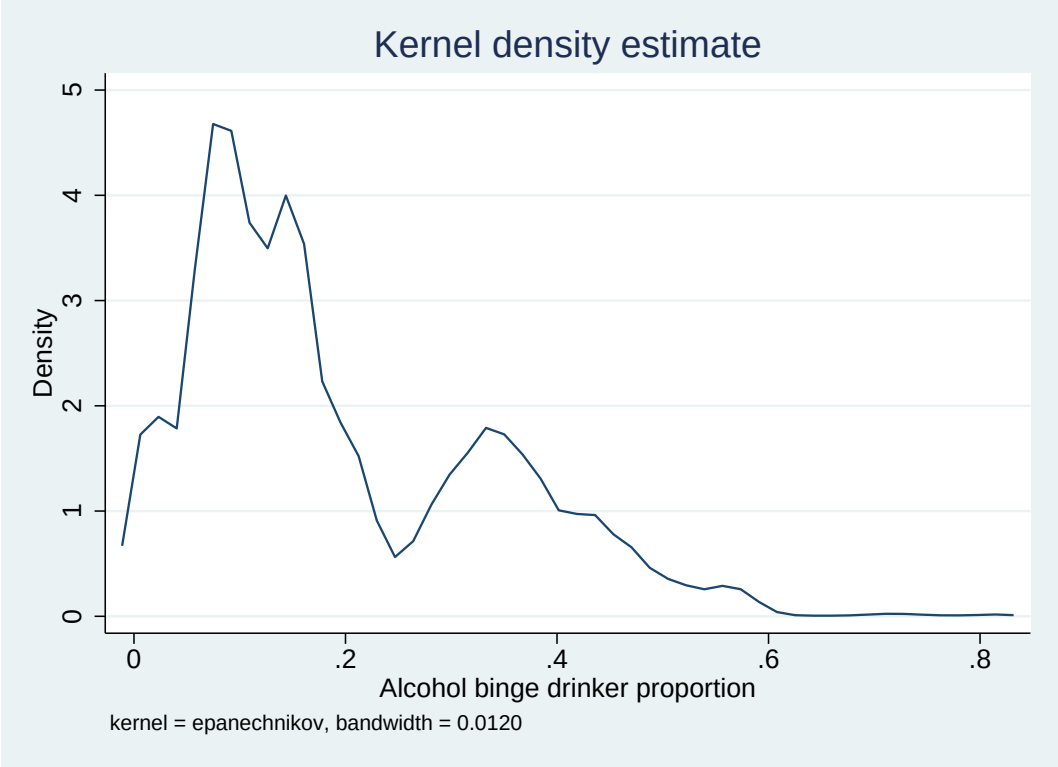
```
 age_yrs_med |     528,080     43.55156     31.62718    .0192308        97.5
```

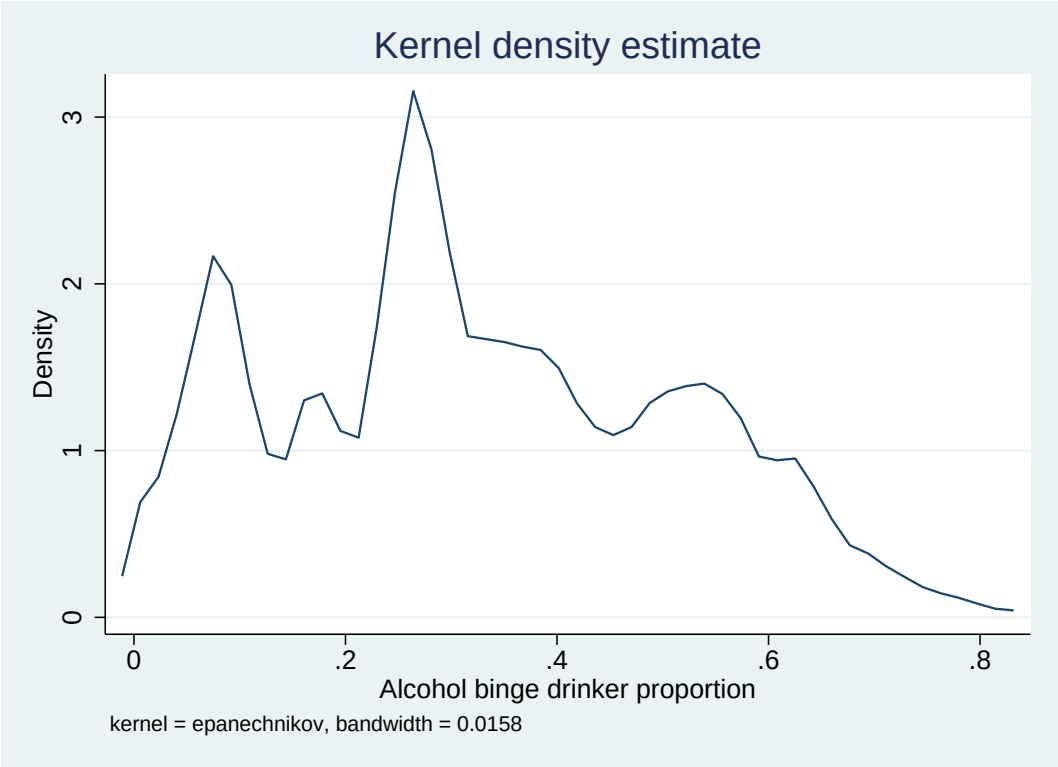
1.2 0.1. Filtering our dataset

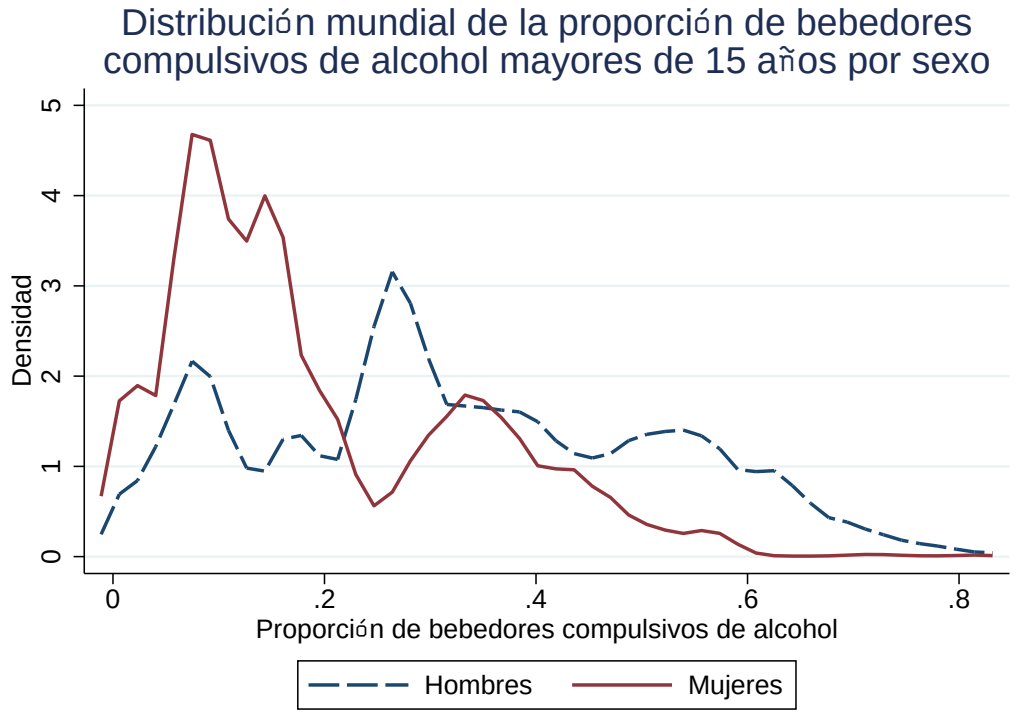
[4]:
```
* Selecciona registros que pertenecen solo a paises (no estados, o sub␣
 ↪regiones)
local filter_country location_id > 1 & location_id < 436

*Dependiendo de lo que se investigue, se pueden restringir los grupos de edad.
* age_group == 27 es el valor estandarizado para todas las edades,
* age_group_id == 22 es el valor de todas las edades sin estandarizar
* age_group_id > 7 - grupos desde 15 a mas de 95 anios

local filter_age (age_group_id != 22| age_group_id != 27) & age_group_id > 7
```

1.3 1. Some graphs Graphs

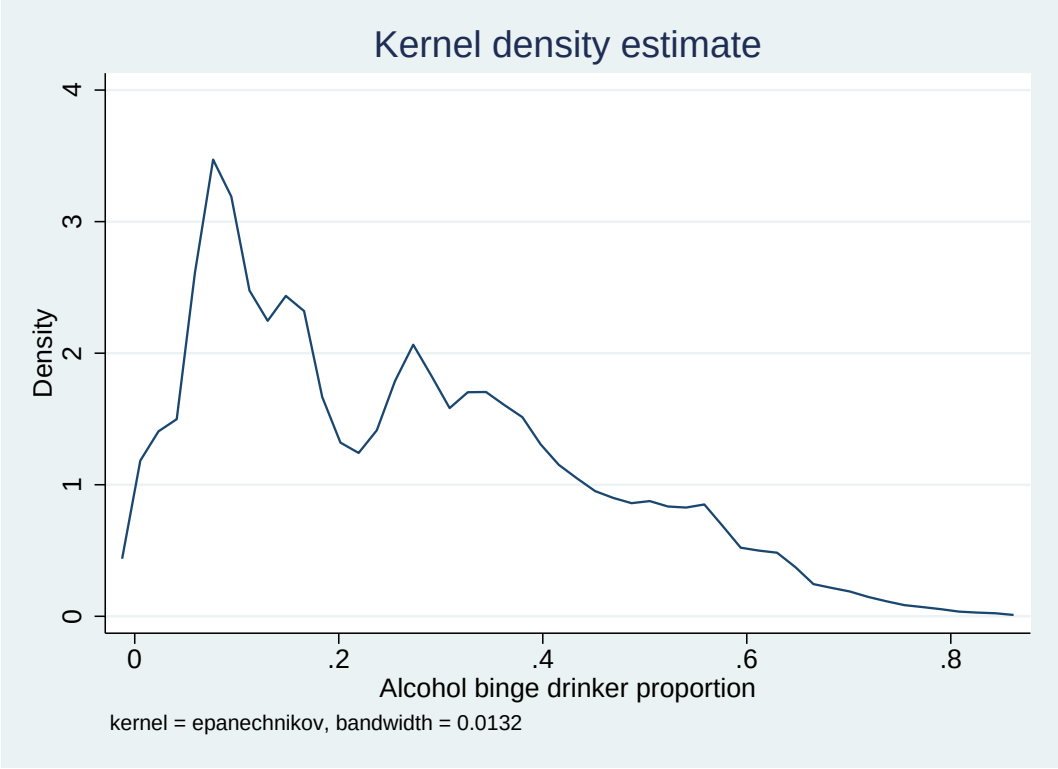
[5]:
```
*1.a
* Grafico de la distribucion de proporcion de bebedores de alcohol
kdensity alc_binge if sex_id == 2 & `filter_country' & `filter_age', generate(x␣
 ↪fx0)
kdensity alc_binge if sex_id == 1 & `filter_country' & `filter_age',␣
 ↪generate(fx1) at(x)
label var fx0 "Mujeres"
label var fx1 "Hombres"
line fx1 fx0 x, sort ytitle(Densidad) title("Distribución mundial de la␣
 ↪proporción de bebedores" "compulsivos de alcohol mayores de 15 años por␣
 ↪sexo") lpattern(longdash solid) lwidth(medthick medthick)␣
 ↪graphregion(color(white)) xtitle("Proporción de bebedores compulsivos de␣
 ↪alcohol")
drop x fx0 fx1
```

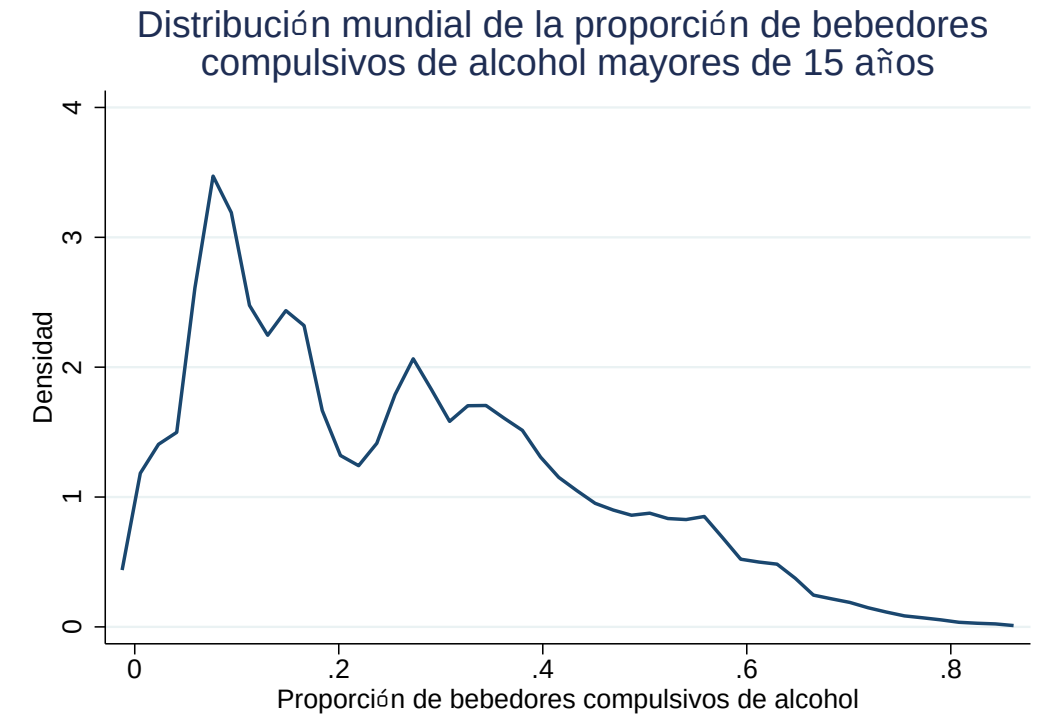

Kernel density estimate
Density
0
1
2
3
4
5
0
.2
.4
.6
.8
Alcohol binge drinker proportion
kernel = epanechnikov, bandwidth = 0.0120


Kernel density estimate
Density
0
1
2
3
0
.2
.4
.6
.8
Alcohol binge drinker proportion
kernel = epanechnikov, bandwidth = 0.0158

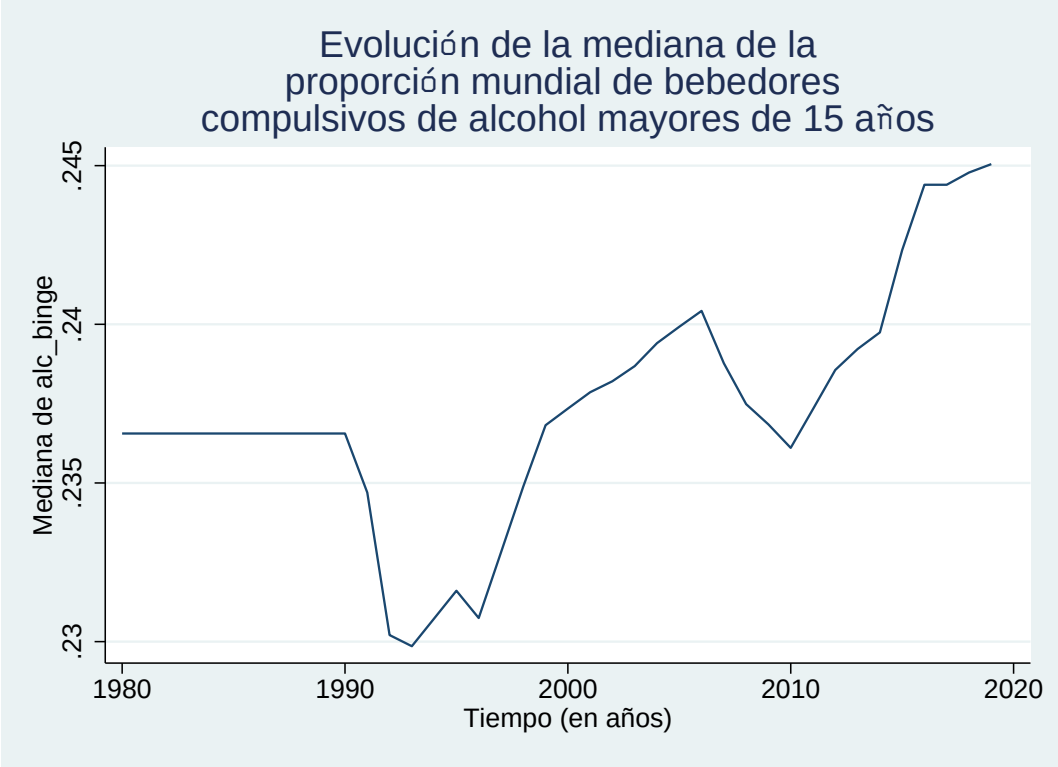


[6]:
```
kdensity alc_binge if `filter_country' & `filter_age', generate(x fx0)
line fx0 x, sort ytitle(Densidad) title("Distribución mundial de la proporción␣
 ↪de bebedores " "compulsivos de alcohol mayores de 15 años") lpattern(solid)␣
 ↪lwidth(medthick) graphregion(color(white)) xtitle("Proporción de bebedores␣
 ↪compulsivos de alcohol")
drop x fx0
```

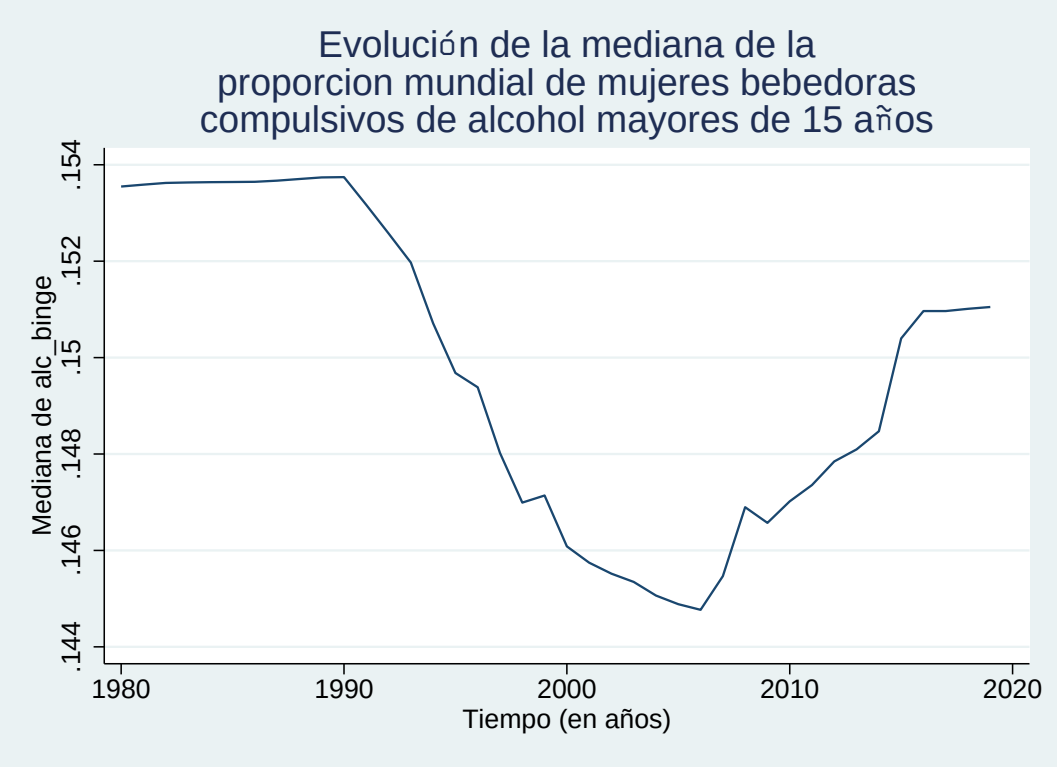

Kernel density estimate
Density
0
1
2
3
4
0
.2
.4
.6
.8
Alcohol binge drinker proportion
kernel = epanechnikov, bandwidth = 0.0132


Distribución mundial de la proporción de bebedores compulsivos de alcohol mayores de 15 aÑos
Densidad
0
1
2
3
4
0
.2
.4
.6
.8
Proporción de bebedores compulsivos de alcohol

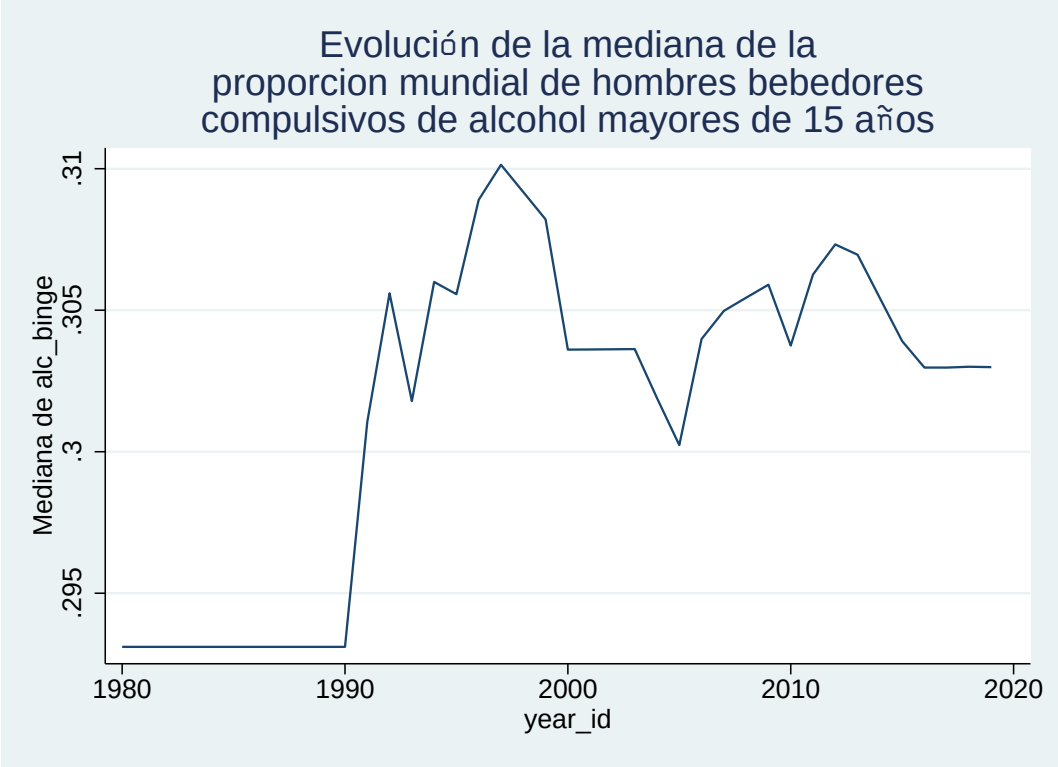
[7]:
```
twoway mband alc_binge year_id if `filter_country' & `filter_age',␣
↪title("Evolución de la mediana de la" "proporción mundial de bebedores "␣
↪"compulsivos de alcohol mayores de 15 años") ytitle("Mediana de alc_binge")␣
↪xtitle("Tiempo (en años)")
```



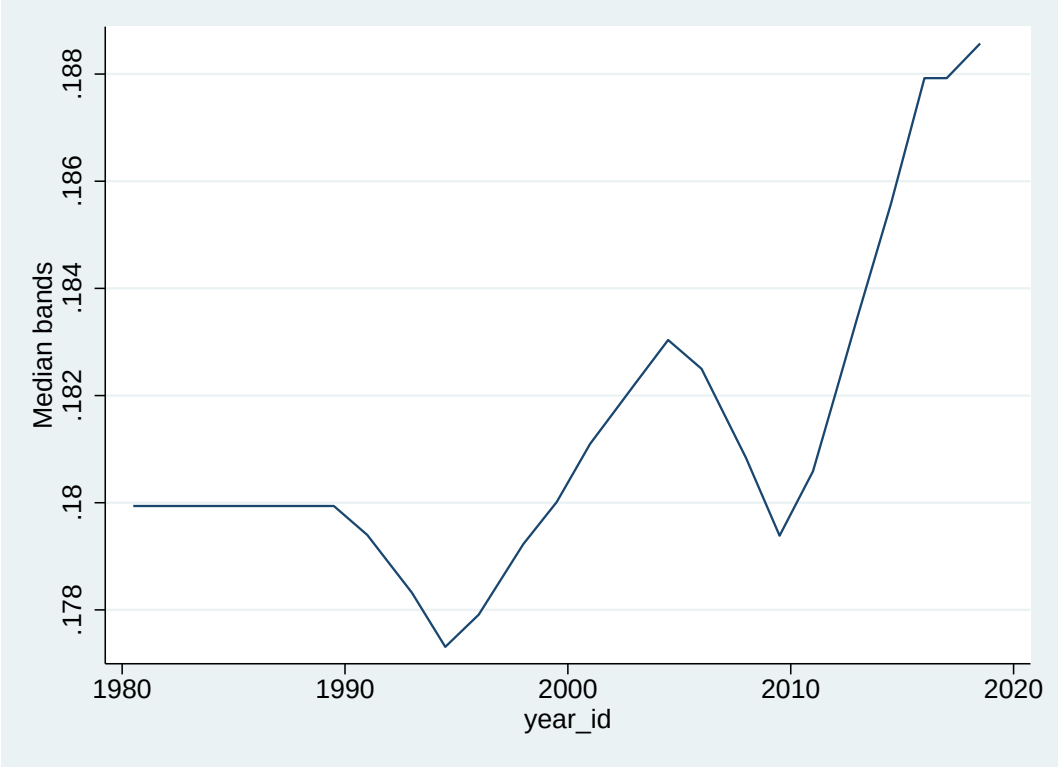
[8]:
```
twoway mband alc_binge year_id if `filter_country' & `filter_age' & sex_id ==␣
↪2, title("Evolución de la mediana de la" "proporcion mundial de mujeres␣
↪bebedoras" "compulsivos de alcohol mayores de 15 años") ytitle("Mediana de␣
↪alc_binge") xtitle("Tiempo (en años)")
```



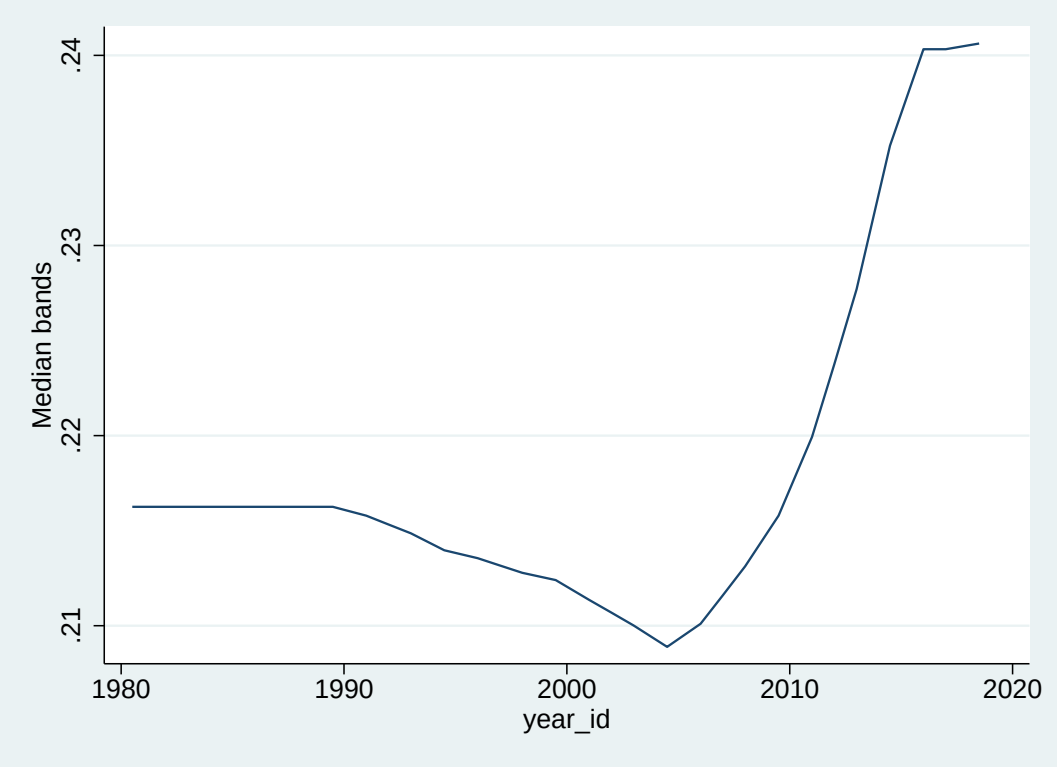
[9]:
```
twoway mband alc_binge year_id if `filter_country' & `filter_age' & sex_id ==
↪1, title("Evolución de la mediana de la" "proporcion mundial de hombres
↪bebedores" "compulsivos de alcohol mayores de 15 años") ytitle("Mediana de
↪alc_binge")
```



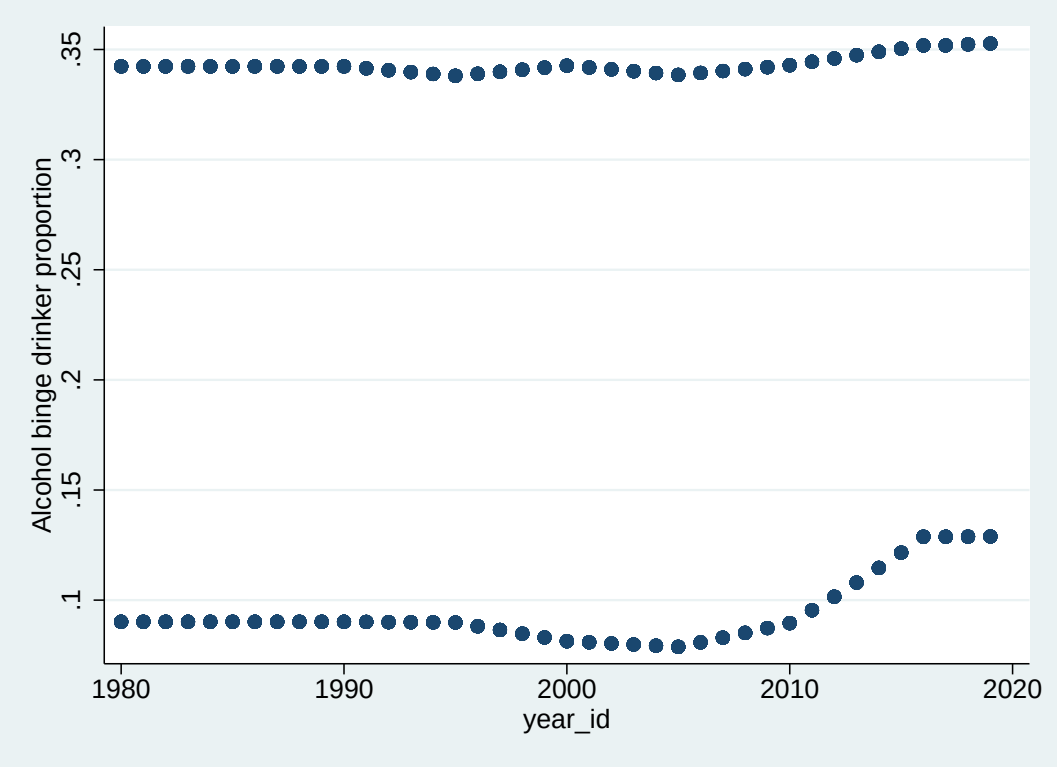
[10]:
```
twoway mband alc_binge year_id if `filter_country' & `filter_age' &␣
↪location_name == "China"
```



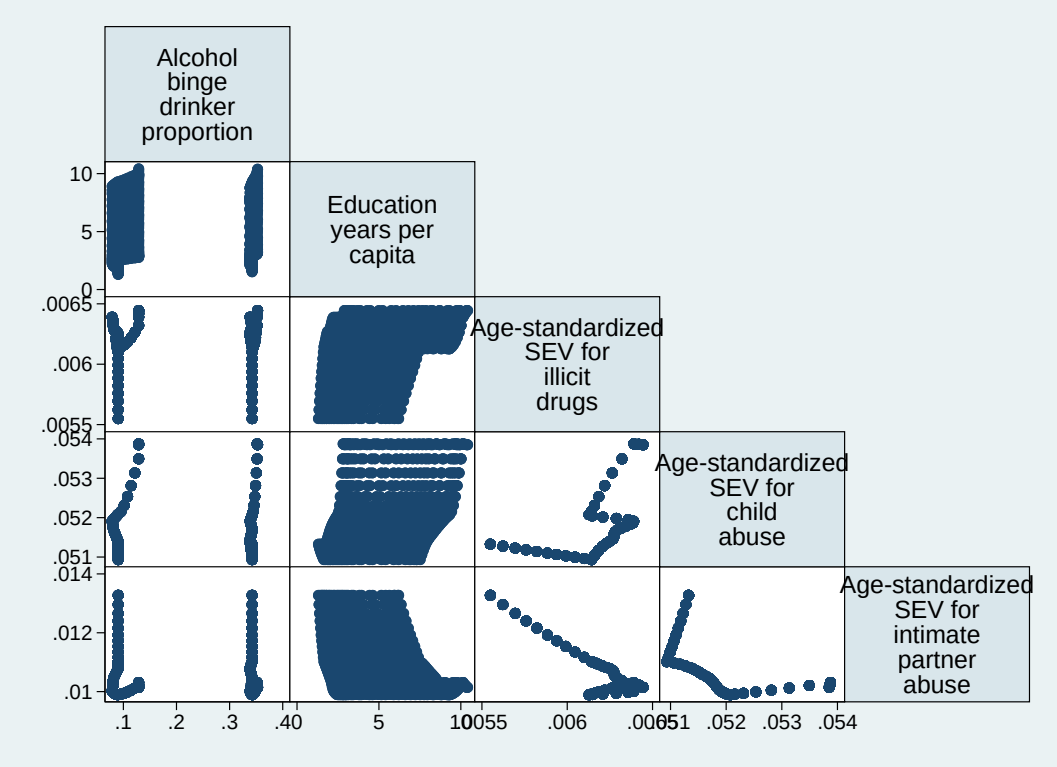
[11]:
```
twoway mband alc_binge year_id if `filter_country' & `filter_age' &␣
↪location_name == "Mexico"
```



[12]:
```
graph twoway scatter alc_binge year_id if `filter_country' & `filter_age' &␣
↪location_name == "Mexico"
```



[14]:
```
graph matrix alc_binge educ sev_drugs sev_chabuse sev_ipabuse if␣
↪`filter_country' & `filter_age' & location_name == "Mexico", half
```



2 2. Modelos

$$\int f(x)dx$$

[20]:
```
*2.
*2.a
regress alc_binge educ sev_drugs prevobes teen_birth bmi majdep year_id if␣
↪`filter_country' & `filter_age' , r
```

```
Linear regression                               Number of obs     =    277,440
                                                F(7, 277432)      =   23642.46
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3467
                                                Root MSE          =     .14495

------------------------------------------------------------------------------
             |               Robust
  alc_binge |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0222416   .0000917   242.56   0.000     .0220619    .0224213
   sev_drugs |  -10.73357   .0869456  -123.45   0.000    -10.90398   -10.56316
```

```
    prevobes |  -.2241608   .0036404   -61.58   0.000    -.2312958   -.2170257
  teen_birth |  -2.291745   .0289726   -79.10   0.000     -2.34853   -2.234959
         bmi |   .0055429   .0002203    25.16   0.000     .0051111    .0059748
      majdep |    .240966    .019002    12.68   0.000     .2037226    .2782094
     year_id |  -.0014518   .0000249   -58.39   0.000    -.0015005    -.001403
       _cons |   3.005712   .0498757    60.26   0.000     2.907957    3.103467
------------------------------------------------------------------------------
```

[ ]:

[ ]:

[ ]:

[41]:
```
regress alc_binge uhc sugar sex_id fiber choles tfa year_id if `filter_age' &␣
↪`filter_country', r
```

```
Linear regression                               Number of obs     =    277,440
                                                F(7, 277432)      =   30015.62
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4209
                                                Root MSE          =     .13648

------------------------------------------------------------------------------
             |               Robust
  alc_binge |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         uhc |   .0055982   .0000255   219.43   0.000     .0055482    .0056482
       sugar |  -.0005959   9.02e-06   -66.08   0.000    -.0006135   -.0005782
      sex_id |  -.1327018   .0005379  -246.72   0.000     -.133756   -.1316476
       fiber |  -.0019673   .0000286   -68.74   0.000    -.0020234   -.0019112
      choles |  -.0039567   .0009598    -4.12   0.000    -.0058378   -.0020756
         tfa |  -.1353735   .0818868    -1.65   0.098    -.2958695    .0251224
     year_id |  -.0027203   .0000246  -110.47   0.000    -.0027685    -.002672
       _cons |    5.69018   .0495358   114.87   0.000     5.593092    5.787269
------------------------------------------------------------------------------

command varlabels is unrecognized

r(199);
r(199);
```

[88]:
```
regress alc_binge sev_alcohol sev_drugs educ, r
```

```
Linear regression                               Number of obs     =    528,080
                                                F(3, 528076)      >   99999.00
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4177
                                                Root MSE          =     .12533

------------------------------------------------------------------------------
             |               Robust
  alc_binge  |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sev_alcohol |   1.179868   .0030616   385.38   0.000     1.173867    1.185868
   sev_drugs |  -6.673979   .0303331  -220.02   0.000    -6.733431   -6.614527
        educ |    .011366   .0000552   206.02   0.000     .0112579    .0114741
       _cons |   .0849292   .0003365   252.35   0.000     .0842696    .0855889
------------------------------------------------------------------------------
```

[89]:
```
regress alc_binge sev_alcohol sev_drugs sev_chabuse sev_ipabuse if `filter_age'␣
↪& `filter_country',r
```

```
Linear regression                               Number of obs     =    277,440
                                                F(4, 277435)      =   64700.31
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4574
                                                Root MSE          =     .13211

------------------------------------------------------------------------------
             |               Robust
  alc_binge  |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sev_alcohol |   1.303757   .0040401   322.70   0.000     1.295839    1.311676
   sev_drugs |  -6.855672   .0673823  -101.74   0.000     -6.98774   -6.723605
 sev_chabuse |   -.232463   .0076255   -30.49   0.000    -.2474087   -.2175172
 sev_ipabuse |  -1.347665   .0111331  -121.05   0.000    -1.369485   -1.325844
       _cons |   .2152978   .0011879   181.24   0.000     .2129695    .2176261
------------------------------------------------------------------------------
```

[51]:
```
sum alc_binge if `filter_age' & `filter_country'
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
   alc_binge |    277,440    .2605897    .1793433   .0008222    .848402
```

2.0.1 Modelos sin y con efectos fijos

[58]:
```
* 3.a
xtset location_id
xtreg alc_binge sev_alcohol sev_drugs sev_chabuse sev_ipabuse educ gdp if␣
 ↪`filter_age' & `filter_country', vce(cluster location_id)
```

```
       panel variable:  location_id (balanced)


Random-effects GLS regression                   Number of obs     =    277,440
Group variable: location_id                     Number of groups  =        204

R-sq:                                           Obs per group:
     within  = 0.0236                                         min =      1,360
     between = 0.3351                                         avg =    1,360.0
     overall = 0.2424                                         max =      1,360

                                                Wald chi2(6)      =     173.47
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

                         (Std. Err. adjusted for 204 clusters in location_id)
------------------------------------------------------------------------------
             |               Robust
   alc_binge |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sev_alcohol |   .1383107   .0549617     2.52   0.012     .0305877    .2460337
   sev_drugs |   1.036315   .7642984     1.36   0.175    -.4616829    2.534312
 sev_chabuse |  -.4079397    .477923    -0.85   0.393    -1.344652    .5287722
 sev_ipabuse |   .6546751   .1125036     5.82   0.000     .4341721    .8751781
        educ |   .0047616   .0004064    11.72   0.000     .0039649    .0055582
         gdp |   5.62e-07   1.69e-07     3.33   0.001     2.31e-07    8.92e-07
       _cons |   .2114304   .0361001     5.86   0.000     .1406755    .2821853
-------------+----------------------------------------------------------------
     sigma_u |  .09214547
     sigma_e |  .08370517
         rho |  .54788665   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

[60]:
```
* 3.b
xtset location_id
xtreg alc_binge sev_alcohol sev_drugs sev_chabuse sev_ipabuse educ gdp if␣
 ↪`filter_age' & `filter_country', fe vce(cluster location_id)
```

```
       panel variable:  location_id (balanced)
```

```
Fixed-effects (within) regression               Number of obs      =       277,440
Group variable: location_id                     Number of groups   =           204

R-sq:                                           Obs per group:
     within  = 0.0237                                         min =         1,360
     between = 0.3036                                         avg =       1,360.0
     overall = 0.2143                                         max =         1,360

                                                F(6,203)           =         27.82
corr(u_i, Xb)  = 0.3786                         Prob > F           =        0.0000

                          (Std. Err. adjusted for 204 clusters in location_id)
------------------------------------------------------------------------------
             |               Robust
   alc_binge |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sev_alcohol |   .1239753   .0559731     2.21   0.028      .013612    .2343386
   sev_drugs |   1.407192   .7853143     1.79   0.075    -.1412266    2.955611
 sev_chabuse |   -.253774   .6496555    -0.39   0.696    -1.534712    1.027164
 sev_ipabuse |   .7009588   .1156497     6.06   0.000     .4729301    .9289875
        educ |   .0047529   .0004072    11.67   0.000     .0039501    .0055557
         gdp |   5.66e-07   1.69e-07     3.35   0.001     2.33e-07    8.99e-07
       _cons |   .1982514   .0460011     4.31   0.000     .1075503    .2889526
-------------+----------------------------------------------------------------
     sigma_u |   .1464002
     sigma_e |  .08370517
         rho |  .75363359   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

[61]:
```
* 3.c
xtset location_id
xtreg alc_binge sev_alcohol sev_drugs sev_chabuse educ gdp if `filter_age' &␣
 ↪`filter_country', fe vce(cluster location_id)
```

```
       panel variable:  location_id (balanced)


Fixed-effects (within) regression               Number of obs      =       277,440
Group variable: location_id                     Number of groups   =           204

R-sq:                                           Obs per group:
     within  = 0.0224                                         min =         1,360
     between = 0.5156                                         avg =       1,360.0
     overall = 0.3626                                         max =         1,360

                                                F(5,203)           =         33.11
```

```
corr(u_i, Xb)  = 0.5597                         Prob > F          =     0.0000

                          (Std. Err. adjusted for 204 clusters in location_id)
------------------------------------------------------------------------------
             |               Robust
  alc_binge |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sev_alcohol |   .1081832   .0580915     1.86   0.064     -.006357    .2227234
   sev_drugs |  -.5874798   .6715231    -0.87   0.383    -1.911535     .736575
 sev_chabuse |   .0708656   .6383804     0.11   0.912    -1.187841    1.329572
        educ |   .0045146   .0003912    11.54   0.000     .0037432     .005286
         gdp |   5.13e-07   1.67e-07     3.07   0.002     1.84e-07    8.42e-07
       _cons |   .2119235   .0469379     4.51   0.000     .1193751    .3044719
-------------+----------------------------------------------------------------
     sigma_u |  .14049314
     sigma_e |  .08375983
         rho |  .73777001   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

[54]:
```
* 3.b
xtset location_id
xtreg alc_binge sev_alcohol sev_drugs educ majdep if `filter_age' &␣
 ↪`filter_country', fe vce(cluster location_id)
```

```
      panel variable:  location_id (balanced)


Random-effects GLS regression                   Number of obs     =    277,440
Group variable: location_id                     Number of groups  =        204

R-sq:                                           Obs per group:
     within  = 0.1903                                         min =      1,360
     between = 0.0995                                         avg =    1,360.0
     overall = 0.1137                                         max =      1,360

                                                Wald chi2(4)      =     167.70
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

                          (Std. Err. adjusted for 204 clusters in location_id)
------------------------------------------------------------------------------
             |               Robust
  alc_binge |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 sev_alcohol |   .2554981   .0555389     4.60   0.000     .1466438    .3643524
   sev_drugs |   4.081867   1.276085     3.20   0.001     1.580787    6.582947
        educ |  -.0031268   .0006997    -4.47   0.000    -.0044981   -.0017554
     majdep |  -3.345438   .3152704   -10.61   0.000    -3.963357    -2.72752
```

```
       _cons |   .3423668     .01778    19.26   0.000     .3075186     .377215
-------------+----------------------------------------------------------------
     sigma_u |  .09220068
     sigma_e |  .07622596
         rho |  .59400113   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

[50]:
```
drop educ2 majdep2
```

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]: