Econométrie appliquée

Licence Economie-Gestion. Parcours Analyse Economique

Professeur Georges Bresson

Session Juin 2019

Exercice 1 (5 pts.)
On s'intéresse au choix d'achat de Ketchup dans des magasins de Springfield (Montana) et de Sioux Falls (Dakota du Sud) en 1986 et 1988.[1] L'échantillon concernent 93744 achats de Ketchup de 4 marques différentes: Heinz (`Heinz`), Hunt's (`Hunts`), Del Monte (`DelMonte`) et le produit distributeur (`Store`) dans les magasins de Springfield (`Springfield_market`) et de Sioux Falls (`Sioux_falls_market`).
Les 3 premiers produits sont classés dans la catégorie marque commerciale (`Trademark`) et le dernier dans la catégorie marque distributeur (`BrandStore`). Pour chaque achat, on dispose d'informations relatives au produit, au magasin, au consommateur et à sa famille.
On souhaite estimer un modèle logit emboîté de choix du produit (`choice = 1`) en fonction du prix du produit en \$ (`price`), d'un indice de disponibilité du coupon de réduction (`coupon_avail`), de la valeur du coupon de réduction en \$ (`coupon_val`), de la moyenne de la valeur des coupons de réduction par marque (`coupon_av`), de la présence (ou non) du produit sur le présentoir (`display`), du packaging (ou non) du produit sur le flyer (`featured`), du magasin (`Springfield_market` ou `Sioux_falls_market`), de la classe de revenu du ménage en \$ (`income_1` à `income_5`) et du nombre de membres de la famille (`members`). Commentez les Tables 1 à 4.

Exercice 2 (5 pts.)
On s'intéresse aux dépenses de consommation de cigarettes journalières (`cig_expend`) en US cents par paquet relatives à 807 individus sondés aux USA.[2] On souhaite expliquer le logarithme de ces dépenses à l'aide d'un certain nombre de régresseurs : l'âge (`age`), le niveau d'éducation (`educ`), le logarithme du prix du paquet de cigarettes (`cig_price`) en US cents et le logarithme du revenu annuel de la personne sondée (`income`) en dollars. On estime les modèles volontairement sans constante.

1. On estime d'abord un modèle Tobit et on calcule les effets marginaux. Commentez les Tables 5 et 6. Interprétez les résultats au risque de 7% (et non de 5% comme d'habitude).

2. On estime ensuite un modèle de sélection à la Heckman dans lequel l'équation de sélection dépend uniquement du logarithme du revenu annuel. On calcule les effets marginaux, les probabilités et valeurs estimées. Commentez les Tables 7 et 8. Interprétez les résultats au risque de 7% (et non de 5% comme d'habitude).

[1] Les données proviennent de Chung, A., Erdem, T. and M. Keane, (2009), The price consideration model of brand choice, *Journal of Applied Econometrics*, 24, 393-420.

[2] Les données proviennent de Mullahy, J., (1997), Instrumental-variable estimation of count data models: Applications to models of cigarette smoking behavior, *Review of Economics and Statistics*, 79, 596-593.

Exercice 3 (5 pts.)
On s'intéresse maintenant au nombre de cigarettes fumées. On estime un modèle binomial négatif à excès de zéros (ZINB) du nombre de cigarettes fumées (`cigs`) en fonction des prédicteurs: `cig_price`, `age`, `educ` et `income`. Commentez les Tables 9 à 11. Interprétez les résultats au risque de 7% (et non de 5% comme d'habitude).

Exercice 4 (5 pts.)
On s'intéresse au modèle de croissance inclusive (*inclusive growth*).[3] Soit $\overline{y}_p$, le revenu moyen cumulé associé au percentile p de la population. L'indice de mobilité social (SMI (*social mobility index*)) est défini comme la surface sous la courbe: $y^* = \int_0^{100} \overline{y}_p dp$. Si le revenu est le même pour tous, alors $\overline{y} = y^*$, et il n'y a pas d'inégalité dans la distribution des revenus. Plus SMI augmente, plus "l'ascenseur social" fonctionne et plus les inégalités se réduisent.
On souhaite estimer le logarithme du SMI (`smi`) en fonction d'un certain nombre de prédicteurs pour 63 pays et 24 années (1990-2013).[4]

1. On estime d'abord un modèle statique doublement logarithmique dans lequel les prédicteurs sont: le PIB par tête (`pib`), le taux d'investissement en pourcentage du PIB (`invest`), le ratio des années d'étude de la population par rapport au niveau le plus faible d'études (`school`) (Burkina Faso, 1990, school = 1), les dépenses de R&D par tête (`RD`), les crédits bancaires par tête (`credit`), les dépots de marques commerciales (`TM`) et les termes de l'échange (prix des exports / prix des imports) (en base 100 en 2005 pour USA) (`TT`). Toutes les valeurs monétaires sont en US dollars 2005 corrigés de la PPA. On suspecte l'endogénéité du PIB et on estime des modèles statiques à effets fixes et à erreurs composées. Commentez les Tables 12 et 13.

2. On estime ensuite un modèle dynamique. Commentez la Table 14. Calculez les effets de court terme et de long terme du PIB et des dépenses de R&D. Interprétez.

Aucun document autorisé
Calculatrice et tables statistiques autorisées

Table 1.

```
. des choice price coupon_avail coupon_val coupon_av display featured Springfield_market
.      Sioux_falls_market members income_1- income_5

variable name                        variable label
-------------------------------------------------------------------------------------------
choice                          chosen brand
price                           price of brand
coupon_avail                    coupon availability for brand
coupon_val                      total value of coupon used (in US$)
coupon_av                       average of coupon value by brand
display                         none (or least one brand) (0/1) is on display
featured                        none (or least one brand) (0/1) is featured in the flyer
Springfield_market
Sioux_falls_market
members                         number of members in a household
income_1                        household income < 20,000 $
income_2                        household income [20,000; 40,000[ $
```

[3] La croissance inclusive est une croissance économique qui est répartie équitablement dans la société et crée des opportunités pour tous.

[4] Les données proviennent de Bresson, G., Etienne, J-M. et P. Mohnen, (2015), Inclusive growth and innovation: a dynamic simultaneous equations model on a panel of countries, *STI Policy Review*, 6, 1, 1-23.

```
income_3                          household income [40,000; 60,000[ $
income_4                          household income [60,000; 100,000[ $
income_5                          household income > 100,000 $


. sum choice price coupon_avail coupon_val coupon_av display featured Springfield_market
.      Sioux_falls_market members income_1- income_5

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      choice |     93,744         .25     .433015          0          1
       price |     93,744    1.054951    .1918314        .59       1.98
coupon_avail |     93,744    .0678016    .0867932          0    .586667
  coupon_val |     93,744    .1068422    .2403362          0          5
   coupon_av |     93,744    .1091367    .0987126          0    .586667
     display |     93,744    .0464243    .2104033          0          1
    featured |     93,744    .0920806    .2891412          0          1
Springfiel~t |     93,744    .3343574    .4717678          0          1
Sioux_fall~t |     93,744    .6656426    .4717678          0          1
     members |     93,744    3.353516    1.379027          1          8
    income_1 |     93,744    .2652756    .4414822          0          1
    income_2 |     93,744    .5333248    .4988909          0          1
    income_3 |     93,744    .1636371     .369948          0          1
    income_4 |     93,744    .0286738     .166889          0          1
    income_5 |     93,744    .0090886    .0949003          0          1
```

Table 2.

```
. label define brand 1 "Heinz" 2 "Hunts" 3 "DelMonte" 4 "Store"
. label variable brand " Type of Brand"
. nlogitgen type = brand(Trademark: 1|2|3, BrandStore: 4)

tree structure specified for the nested logit model

 type         N       brand      N     k
-------------------------------------------
 Trademark  70308 --- Heinz    23436 14951
                  |- Hunts    23436  3874
                  +- DelMonte 23436  1793
 BrandStore 23436 --- Store    23436  2818
-------------------------------------------
                        total  93744 23436

k = number of times alternative is chosen
N = number of observations at each level

. estat alternatives

   Alternatives summary for type
  +-------------------------------------------------------------------------+
  |          Alternative                 |    Cases   Frequency    Percent |
  |  index         value            label |  present   selected   selected |
  |--------------------------------------+----------------------------------|
  |      1            1        Trademark |    70308      20618      87.98 |
  |      2            2       BrandStore |    23436       2818      12.02 |
  +-------------------------------------------------------------------------+

   Alternatives summary for brand
  +-------------------------------------------------------------------------+
  |          Alternative                 |    Cases   Frequency    Percent |
  |  index         value            label |  present   selected   selected |
  |--------------------------------------+----------------------------------|
  |      1            1            Heinz |    23436      14951      63.80 |
  |      2            2            Hunts |    23436       3874      16.53 |
  |      3            3         DelMonte |    23436       1793       7.65 |
  |      4            4            Store |    23436       2818      12.02 |
  +-------------------------------------------------------------------------+
```

Table 3.

```
. constraint 1 [BrandStore_tau]_cons=1
. nlogit choice price coupon_avail display featured  || type: coupon_av coupon_val  ///
               income_1 income_3 income_4 income_5, base(BrandStore) ///
               || brand: Springfield_market members, base(Store) noconst constraint(1)
```

```
note: branch 2 of level 1 is degenerate and the associated dissimilarity parameter
       ([BrandStore_tau]_cons) is not defined.

RUM-consistent nested logit regression        Number of obs      =     93,744
Case variable: number_Obs                     Number of cases    =      23436

Alternative variable: brand                   Alts per case: min =          4
                                                             avg =        4.0
                                                             max =          4

                                                 Wald chi2(16)   =    2929.34
Log likelihood = -18810.563                      Prob > chi2     =     0.0000

 ( 1)  [BrandStore_tau]_cons = 1
------------------------------------------------------------------------------------
            choice |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
brand              |
             price |  -.3428895   .0502513    -6.82   0.000    -.4413802   -.2443989
      coupon_avail |   .7501788   .0933402     8.04   0.000     .5672354    .9331223
           display |    1.00521   .0556408    18.07   0.000     .8961555    1.114264
          featured |   1.204041   .0630302    19.10   0.000     1.080504    1.327578
------------------------------------------------------------------------------------
type equations
------------------------------------------------------------------------------------
Trademark          |
         coupon_av |   46.41574   1.264055    36.72   0.000     43.93824    48.89325
        coupon_val |   .3434516   .3326287     1.03   0.302    -.3084886    .9953917
          income_1 |  -.3835035   .0537836    -7.13   0.000    -.4889174   -.2780896
          income_3 |   .2170897    .076787     2.83   0.005     .0665899    .3675895
          income_4 |   .6498739   .1974841     3.29   0.001     .2628123    1.036936
          income_5 |  -.4256298    .264914    -1.61   0.108    -.9448517    .0935921
-------------------+----------------------------------------------------------------
BrandStore         |
         coupon_av |          0  (base)
        coupon_val |          0  (base)
          income_1 |          0  (base)
          income_3 |          0  (base)
          income_4 |          0  (base)
          income_5 |          0  (base)
------------------------------------------------------------------------------------
brand equations
------------------------------------------------------------------------------------
Heinz              |
Springfield_market |  -.6820936   .0489233   -13.94   0.000    -.7779815   -.5862057
           members |   .0923363   .0106554     8.67   0.000      .071452    .1132206
-------------------+----------------------------------------------------------------
Hunts              |
Springfield_market |  -.5750404   .0519559   -11.07   0.000     -.676872   -.4732087
           members |  -.1270398   .0172201    -7.38   0.000    -.1607906   -.0932891
-------------------+----------------------------------------------------------------
DelMonte           |
Springfield_market |   -.990929   .0592808   -16.72   0.000    -1.107117   -.8747408
           members |  -.2419293   .0228026   -10.61   0.000    -.2866215    -.197237
-------------------+----------------------------------------------------------------
Store              |
Springfield_market |          0  (base)
           members |          0  (base)
------------------------------------------------------------------------------------
dissimilarity parameters
------------------------------------------------------------------------------------
type               |
    /Trademark_tau |   .6684414   .0350825                       .599681    .7372019
   /BrandStore_tau |          1  (constrained)
------------------------------------------------------------------------------------
LR test for IIA (tau=1): chi2(1) = 75.39                  Prob > chi2 = 0.0000
```

Table 4.

```
. predict pred_lev1 pred_lev2, pr

. tabulate type, summarize(pred_lev1)

            |  Summary of Pr(type alternatives)
```

```
        type |        Mean   Std. Dev.       Freq.
------------+------------------------------------
   Trademark |   .87803244   .19884922      70,308
   BrandStor |   .12196756   .19885205      23,436
------------+------------------------------------
       Total |   .68901622   .38304492      93,744

. tabulate brand, summarize(pred_lev2)

     Type of |  Summary of Pr(brand alternatives)
       Brand |        Mean   Std. Dev.       Freq.
------------+------------------------------------
       Heinz |   .60021025   .25019018      23,436
       Hunts |   .18641603   .15782197      23,436
    DelMonte |   .09140616   .10555373      23,436
       Store |   .12196756   .19885205      23,436
------------+------------------------------------
       Total |         .25   .27677539      93,744

. sum pred_lev1 pred_lev2 if choice==1 & brand==Heinz

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   pred_lev1 |    14,951    .9709601     .076892   .0943968          1
   pred_lev2 |    14,951    .7075434    .2054447   .0335811   .9967773

. sum pred_lev1 pred_lev2 if choice==1 & brand==Hunts

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   pred_lev1 |     3,874    .8548264    .1870214   .0815174          1
   pred_lev2 |     3,874    .2988125    .2172267   .0046167   .9418151

. sum pred_lev1 pred_lev2 if choice==1 & brand==DeMonte

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   pred_lev1 |     1,793    .7101472    .1784531   .0544408          1
   pred_lev2 |     1,793    .1757824    .1839247   .0011811   .9389681

. sum pred_lev1 pred_lev2 if choice==1 & brand==Store

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   pred_lev1 |     2,818    .4762763     .197838   .0000168   .9441794
   pred_lev2 |     2,818    .4762763     .197838   .0000168   .9441794
```

Table 5.

```
. des smoker cigs cig_expend cig_expend_smoker log_cig_expend cig_price age educ income

variable name      variable label
-------------------------------------------------------------------------------------------
smoker                 smoker : (1/0) if cigarettes expenditures per day
cigs                   cigarettes smoked per day
cig_expend             cigarettes expenditures per day in cents/pack
cig_expend_smoker      cigarettes expenditures per day in cents/pack for a smoker
log_cig_expend         log of cigarettes expenditures per day in cents/pack
cig_price              state cigarettes price, cents/pack
age                    in years
educ                   years of schooling
income                 annual income, $

. sum smoker cigs cig_expend cig_expend_smoker log_cig_expend cig_price age educ income

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      smoker |       807    .3841388    .4866926          0          1
        cigs |       807    8.686493    13.72152          0         80
  cig_expend |       807    523.0598    832.4359          0    4869.84
cig_expend~r |       310    1361.643    813.6054     57.352    4869.84
log_cig_ex~d |       310    6.984103    .8004979   4.049208   8.490816
   cig_price |       807    60.30041    4.738469     44.004     70.129
         age |       807    41.23792    17.02729         17         88
        educ |       807    12.47088    3.057161          6         18
      income |       807    19304.83    9142.958        500      30000
```

Table 6.

```
. * min cig_expend =  57.352 -> log(min) =  4.0492077
. global xlist_log log_cig_price age educ log_income
. tobit log_cig_expend $xlist_log, ll(4.04) noconstant

Tobit regression                                  Number of obs     =        310
                                                  LR chi2(4)        =          .
Log likelihood = -361.60377                       Prob > chi2       =          .

------------------------------------------------------------------------------
log_cig_exp~d |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
log_cig_price |     1.2551   .1531764     8.19   0.000     .9536879    1.556513
          age |   .0058071   .0030112     1.93   0.055    -.0001183    .0117324
         educ |   .0350262   .0176206     1.99   0.048     .0003534     .069699
   log_income |   .1237152   .0650917     1.90   0.058    -.0043687    .2517991
-------------+----------------------------------------------------------------
       /sigma |   .7768758      .0312                      .715482    .8382696
------------------------------------------------------------------------------
            0  left-censored observations
          310     uncensored observations
            0 right-censored observations

. mfx compute, predict(e(4.04,.))

Marginal effects after tobit
      y  = E(log_cig_expend|log_cig_expend>4.04) (predict, e(4.04,.))
         =  6.9841609
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
log_ci~e |   1.253655      .15302    8.19   0.000   .953739  1.55357   4.09311
     age |   .0058004      .00301    1.93   0.054  -.000095  .011695   39.1129
    educ |   .0349858       .0176    1.99   0.047    .00049  .069482   11.9968
log_in~e |   .1235727      .06502    1.90   0.057  -.003858  .251004   9.69426
------------------------------------------------------------------------------
```

Table 7.

```
. heckman log_cig_expend log_cig_price educ age  , select(log_income, noconstant) noconstant

Heckman selection model                           Number of obs     =        807
(regression model with sample selection)          Censored obs      =        497
                                                  Uncensored obs    =        310

                                                  Wald chi2(3)      =    9406.11
Log likelihood = -863.3086                        Prob > chi2       =     0.0000

-------------------------------------------------------------------------------
log_cig_expend |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
log_cig_expend |
 log_cig_price |   1.825727   .0542324    33.66   0.000     1.719434    1.932021
          educ |   .0403156   .0137511     2.93   0.003     .0133638    .0672673
           age |   .0055566   .0025411     2.19   0.029     .0005761    .0105371
--------------+----------------------------------------------------------------
select        |
    log_income |  -.0304111    .004547    -6.69   0.000    -.0393231    -.021499
--------------+----------------------------------------------------------------
       /athrho |   -2.09929   .1551362   -13.53   0.000    -2.403351   -1.795228
      /lnsigma |   .2320019   .0558348     4.16   0.000     .1225677     .341436
--------------+----------------------------------------------------------------
           rho |  -.9704106    .009045                     -.983783   -.9463096
         sigma |   1.261122   .0704144                     1.130396    1.406966
        lambda |  -1.223806   .0760472                    -1.372856   -1.074756
-------------------------------------------------------------------------------
LR test of indep. eqns. (rho = 0):   chi2(1) =    75.63   Prob > chi2 = 0.0000

. margins, dydx(*) predict(psel)
Average marginal effects                          Number of obs     =        807

Expression   : Pr(select), predict(psel)
dy/dx w.r.t. : log_income
```

```
------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  log_income |  -.0116145   .0015852    -7.33   0.000    -.0147214   -.0085076
------------------------------------------------------------------------------

. margins, dydx(*) predict(ycond)
Average marginal effects                        Number of obs     =        807

Expression   : E(log_cig_expend|Zg>0), predict(ycond)
dy/dx w.r.t. : log_cig_price educ age log_income

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
log_cig_price |   1.825727   .0542324    33.66   0.000     1.719434    1.932021
        educ |   .0403156   .0137511     2.93   0.003     .0133638    .0672673
         age |   .0055566   .0025411     2.19   0.029     .0005761    .0105371
  log_income |  -.0258982   .0047352    -5.47   0.000     -.035179   -.0166173
------------------------------------------------------------------------------
```

Table 8.

```
. predict pred_prob, psel  (prediction of smoking)
. predict pred_y, ycond    (prediction of log. cigarettes expenditures)
. gen resid = log_cig_expend -  pred_y

. sum smoker pred_prob

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      smoker |        807    .3841388    .4866926          0          1
   pred_prob |        807    .3841745    .0083204   .3769481   .4250492

. sum log_cig_expend pred_y resid if log_cig_expend != .

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
log_cig_ex~d |        310    6.984103    .8004979   4.049208   8.490816
      pred_y |        310    6.956709    .2016302   6.189157   7.411317
       resid |        310    .0273941    .7859795  -3.000527   1.482906
```

Table 9.

```
. generate cigs_range = cigs
. recode cigs_range (11/19 = 19) (21/39 = 39) (40/80 = 80)
. tabulate cigs_range

 cigs_range |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        497       61.59       61.59
          1 |          7        0.87       62.45
          2 |          5        0.62       63.07
          3 |          5        0.62       63.69
          4 |          2        0.25       63.94
          5 |          7        0.87       64.81
          6 |          3        0.37       65.18
          7 |          2        0.25       65.43
          8 |          3        0.37       65.80
          9 |          2        0.25       66.05
         10 |         28        3.47       69.52
         19 |         37        4.58       74.10
         20 |        101       12.52       86.62
         39 |         55        6.82       93.43
         80 |         53        6.57      100.00
------------+-----------------------------------
      Total |        807      100.00
```

Table 10.

```
. global xlist cig_price age income
. global xlist_inf educ
. zinb cigs $xlist, inf($xlist_inf) vuong

Zero-inflated negative binomial regression      Number of obs     =        807
                                                Nonzero obs       =        310
                                                Zero obs          =        497

Inflation model = logit                         LR chi2(3)        =      10.67
Log likelihood  = -1746.372                     Prob > chi2       =     0.0137

------------------------------------------------------------------------------
        cigs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
cigs         |
   cig_price |   .0011261   .0068825     0.16   0.870    -.0123634    .0146155
         age |   .0046967   .0024699     1.90   0.057    -.0001443    .0095376
      income |   .0000101   3.95e-06     2.56   0.011     2.35e-06    .0000178
       _cons |   2.664666   .4237615     6.29   0.000     1.834109    3.495223
-------------+----------------------------------------------------------------
inflate      |
        educ |   .0837303   .0242069     3.46   0.001     .0362858    .1311749
       _cons |  -.5673557   .3065086    -1.85   0.064    -1.168101      .03339
-------------+----------------------------------------------------------------
    /lnalpha |  -1.110532   .0951144   -11.68   0.000    -1.296952   -.9241108
-------------+----------------------------------------------------------------
       alpha |   .3293838   .0313291                      .2733637    .3968842
------------------------------------------------------------------------------
Vuong test of zinb vs. standard negative binomial: z =    10.45  Pr>z = 0.0000

. margeff

Average marginal effects on Prob(cigs) after zinb

------------------------------------------------------------------------------
        cigs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   cig_price |   .0098657   .0603124     0.16   0.870    -.1083444    .1280759
         age |   .0411488    .022141     1.86   0.063    -.0022467    .0845442
      income |   .0000884   .0000359     2.46   0.014      .000018    .0001588
        educ |  -.4451001   .1258867    -3.54   0.000    -.6918335   -.1983668
------------------------------------------------------------------------------
```

Table 11.

```
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
ZINB        0.023         10      0.006

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.616       0.616      0.000     0.000
1        0.009       0.002      0.007    24.090
2        0.006       0.003      0.003     2.955
3        0.006       0.004      0.002     0.669
4        0.002       0.006      0.003     1.471
5        0.009       0.007      0.002     0.305
6        0.004       0.008      0.005     2.030
7        0.002       0.009      0.007     4.104
8        0.004       0.010      0.007     3.425
9        0.002       0.011      0.009     5.437
10       0.035       0.012      0.023    35.953
------------------------------------------------
Sum      0.695       0.688      0.067    80.438
```

```
. prvalue, x(cig_price=58.141 age=28 educ=10 income=12500) max(10) brief

zinb: Predictions for cigs

  Expected y:         8.5897
  Pr(Always0|z):      0.5671
  Pr(y=0|x,z):        0.5680
  Pr(y=1|x):          0.0025
  Pr(y=2|x):          0.0043
  Pr(y=3|x):          0.0063
  Pr(y=4|x):          0.0083
  Pr(y=5|x):          0.0101
  Pr(y=6|x):          0.0117
  Pr(y=7|x):          0.0131
  Pr(y=8|x):          0.0143
  Pr(y=9|x):          0.0152
  Pr(y=10|x):         0.0158

. prvalue, x(cig_price=63.179 age=54 educ=13.5 income=30000) max(10) brief

zinb: Predictions for cigs

  Expected y:         9.7608
  Pr(Always0|z):      0.6371
  Pr(y=0|x,z):        0.6375
  Pr(y=1|x):          0.0010
  Pr(y=2|x):          0.0017
  Pr(y=3|x):          0.0026
  Pr(y=4|x):          0.0035
  Pr(y=5|x):          0.0045
  Pr(y=6|x):          0.0054
  Pr(y=7|x):          0.0062
  Pr(y=8|x):          0.0070
  Pr(y=9|x):          0.0077
  Pr(y=10|x):         0.0084
```

Table 12.

```
. des smi pib invest school RD credit TM TT

variable      variable label
---------------------------------------------------------------------------------------------
smi       social mobility index
pib       GDP per capita, (PPP 2005 international $)
invest    Investment rate in %GDP, (PPP 2005 international $)
school    Ratio of years of schooling above age 15
RD        R&D per capita, (PPP 2005 international $)
credit    Domestic credit to private sector by banks per capita, (PPP 2005 international $)
TM        Trademark applications per capita, total
TT        Terms of trade, exports prices/import prices, USA GDP in 2005=1

. xtsum smi pib invest school RD credit TM TT

Variable         |      Mean   Std. Dev.       Min        Max |    Observations
-----------------+--------------------------------------------+----------------
smi      overall |  7992.817   7104.976   298.3306   32900.35 |     N =    1512
         between |             6956.238   445.5028   28583.53 |     n =      63
         within  |             1681.667  -193.5752   14852.11 |     T =      24
                 |                                            |
pib      overall |  13774.67   11081.74   579.8603   49101.77 |     N =    1512
         between |             10815.02   867.2201   42906.58 |     n =      63
         within  |             2760.586   721.0392   23657.89 |     T =      24
                 |                                            |
invest   overall |  .2269372   .0827923   .0287776   .6838483 |     N =    1512
         between |             .0600315   .0904753   .4251315 |     n =      63
         within  |             .0574947   .0526941   .6041038 |     T =      24
                 |                                            |
school   overall |  7.257627   1.950286          1    10.9507 |     N =    1512
         between |             1.904434   1.719045   10.63521 |     n =      63
         within  |             .4816152   5.822669   8.611562 |     T =      24
                 |                                            |
RD       overall |  18981.03    29309.8   2.948833   129030.4 |     N =    1512
         between |             28204.81   129.0843   104299.7 |     n =      63
         within  |             8698.382  -33874.73    60737.4 |     T =      24
                 |                                            |
```

```
credit   overall |  10418.87   14169.97   20.80623   85600.23 |     N =    1512
         between |             12374.91   82.26069   43137.08 |     n =      63
         within  |             7069.702  -23970.59   53808.21 |     T =      24
                 |                                            |
TM       overall |  .0010802   .0009679   7.21e-08   .0079859 |     N =    1512
         between |             .0008466   2.32e-06   .0030876 |     n =      63
         within  |             .0004807  -.0018981   .0060855 |     T =      24
                 |                                            |
TT       overall |  1.046661   .1241951    .421012   2.165286 |     N =    1512
         between |             .0810081   .7845962   1.289462 |     n =      63
         within  |             .0946678   .5296467   1.965391 |     T =      24

. correlate log_pib log_smi L.log_pib L.log_smi L(7).log_pib L(7).log_smi


             |                         L.       L.      L7.      L7.
             |  log_pib  log_smi  log_pib  log_smi  log_pib  log_smi
-------------+------------------------------------------------------
     log_pib |   1.0000
     log_smi |   0.9840   1.0000
   L.log_pib |   0.9993   0.9827   1.0000
   L.log_smi |   0.9840   0.9990   0.9840   1.0000
  L7.log_pib |   0.9834   0.9631   0.9857   0.9658   1.0000
  L7.log_smi |   0.9738   0.9824   0.9754   0.9850   0.9825   1.0000

. correlate D.log_pib D.log_smi

             |       D.       D.
             |  log_pib  log_smi
-------------+------------------
   D.log_pib |  1.0000
   D.log_smi |  0.8494   1.0000
```

Table 13.

```
. xtivreg log_smi log_invest log_school log_RD log_credit log_TM log_TT (log_pib = L(1/7).log_pib) , fe
. estimates store fe2sls

Fixed-effects (within) IV regression           Number of obs     =      1,071
Group variable: ident_N                        Number of groups  =         63

R-sq:                                          Obs per group:
     within  = 0.9255                                        min =         17
     between = 0.9698                                        avg =       17.0
     overall = 0.9683                                        max =         17

                                               Wald chi2(7)      =   2.89e+07
corr(u_i, Xb)  = -0.1350                       Prob > chi2       =     0.0000

------------------------------------------------------------------------------
     log_smi |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     log_pib |   1.099057   .0214647    51.20   0.000     1.056987    1.141127
  log_invest |   .0153058   .0094711     1.62   0.106    -.0032572    .0338688
  log_school |   .1785745   .0471941     3.78   0.000     .0860759    .2710732
      log_RD |  -.0071584   .0032876    -2.18   0.029    -.0136019   -.0007149
  log_credit |   -.016398    .005689    -2.88   0.004    -.0275482   -.0052478
      log_TM |  -.0042753   .0043713    -0.98   0.328    -.0128429    .0042923
      log_TT |  -.0288911   .0229497    -1.26   0.208    -.0738717    .0160894
       _cons |  -1.698002   .1393727   -12.18   0.000    -1.971168   -1.424837
-------------+----------------------------------------------------------------
     sigma_u |  .17604726
     sigma_e |  .05252013
         rho |  .91827311   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F  test that all u_i=0:     F(62,1001) =   140.34         Prob > F    = 0.0000
------------------------------------------------------------------------------
Instrumented:   log_pib
Instruments:    log_invest log_school log_RD log_credit log_TM log_TT
                L.log_pib L2.log_pib L3.log_pib L4.log_pib L5.log_pib
                L6.log_pib L7.log_pib
------------------------------------------------------------------------------

. xtivreg log_smi log_invest log_school log_RD log_credit log_TM log_TT (log_pib = L(1/7).log_pib), ec2sls
. estimates store ec2sls
```

```
EC2SLS random-effects IV regression            Number of obs     =      1,071
Group variable: ident_N                        Number of groups  =         63

R-sq:                                          Obs per group:
     within  = 0.9254                                        min =         17
     between = 0.9708                                        avg =       17.0
     overall = 0.9693                                        max =         17

                                               Wald chi2(7)      =   14454.63
corr(u_i, X)       = 0 (assumed)               Prob > chi2       =     0.0000

------------------------------------------------------------------------------
     log_smi |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     log_pib |   1.093547   .0199155    54.91   0.000     1.054514    1.132581
  log_invest |    .017434   .0093849     1.86   0.063    -.0009602    .0358282
  log_school |   .1659505   .0426024     3.90   0.000     .0824513    .2494497
      log_RD |  -.0047136   .0032683    -1.44   0.149    -.0111195    .0016922
  log_credit |   -.016026   .0056024    -2.86   0.004    -.0270064   -.0050456
      log_TM |  -.0082663   .0042338    -1.95   0.051    -.0165643    .0000317
      log_TT |  -.0258808    .022469    -1.15   0.249    -.0699192    .0181576
       _cons |  -1.672291   .1313399   -12.73   0.000    -1.929713    -1.41487
-------------+----------------------------------------------------------------
     sigma_u |  .15215886
     sigma_e |  .05252013
         rho |  .89354337   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:   log_pib
Instruments:    log_invest log_school log_RD log_credit log_TM log_TT
                L.log_pib L2.log_pib L3.log_pib L4.log_pib L5.log_pib
                L6.log_pib L7.log_pib
------------------------------------------------------------------------------

. hausman fe2sls ec2sls

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fe2sls       ec2sls       Difference          S.E.
-------------+----------------------------------------------------------------
     log_pib |    1.099057     1.093547        .0055099        .0080067
  log_invest |    .0153058      .017434       -.0021282        .0012744
  log_school |    .1785745     .1659505         .012624        .0203055
      log_RD |   -.0071584    -.0047136       -.0024448         .000355
  log_credit |    -.016398     -.016026       -.0003719         .000989
      log_TM |   -.0042753    -.0082663         .003991        .0010879
      log_TT |   -.0288911    -.0258808       -.0030103        .0046726
------------------------------------------------------------------------------
                         b = consistent under Ho and Ha; obtained from xtivreg
          B = inconsistent under Ha, efficient under Ho; obtained from xtivreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        0.49
                Prob>chi2 =      0.9995
                (V_b-V_B is not positive definite)
```

Table 14.

```
. xtabond log_smi log_invest log_school log_RD log_credit log_TM log_TT, ///
         endogenous(log_pib)  twostep noconstant

Arellano-Bond dynamic panel-data estimation     Number of obs     =      1,386
Group variable: ident_N                         Number of groups  =         63
Time variable: year
                                                Obs per group:
                                                              min =         22
                                                              avg =         22
                                                              max =         22

Number of instruments =    512                  Wald chi2(8)      =   80987.00
                                                Prob > chi2       =     0.0000
Two-step results
------------------------------------------------------------------------------
     log_smi |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     log_smi |
         L1. |   .5296909   .0179635    29.49   0.000      .494483    .5648988
             |
     log_pib |   .4944891   .0186348    26.54   0.000     .4579656    .5310127
  log_invest |   .0445557   .0029747    14.98   0.000     .0387254     .050386
  log_school |    .200413   .0162962    12.30   0.000     .1684731    .2323529
      log_RD |  -.0151801   .0011879   -12.78   0.000    -.0175083    -.012852
  log_credit |  -.0114279   .0012793    -8.93   0.000    -.0139352   -.0089206
      log_TM |   .0020701   .0005288     3.91   0.000     .0010338    .0031065
      log_TT |   .0679877   .0046631    14.58   0.000     .0588483    .0771272
------------------------------------------------------------------------------
Instruments for differenced equation
        GMM-type: L(2/.).log_smi L(2/.).log_pib
        Standard: D.log_invest D.log_school D.log_RD D.log_credit D.log_TM
                  D.log_TT

. predict pred_logsmi, xb
. predict resid_logsmi, e

. estat abon
Arellano-Bond test for zero autocorrelation in first-differenced errors
  +-----------------------+
  |Order |  z     Prob > z|
  |------+----------------|
  |   1  |-4.0864  0.0000 |
  |   2  |-.09898  0.9212 |
  +-----------------------+
   H0: no autocorrelation

. estat sargan
Sargan test of overidentifying restrictions
        H0: overidentifying restrictions are valid

        chi2(504)    =  56.29296
        Prob > chi2  =    1.0000

. sum log_smi pred_logsmi resid_logsmi

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     log_smi |      1,512     8.53664    1.023534   5.698203   10.40124
 pred_logsmi |      1,449     9.13546    1.010958    6.17464   10.94133
resid_logsmi |      1,449   -.5898971    .1068451  -.9474829  -.2053567
```