The Stata Journal (2015)
15, Number 2, pp. 554–573

Estimating almost-ideal demand systems with endogenous regressors

Sébastien Lecocq
Food and Social Science Research Unit (ALISS)
The French National Institute for Agricultural Research (INRA)
Paris, France
sebastien.lecocq@ivry.inra.fr

Jean-Marc Robin
Sciences Po
Paris, France
and University College London
London, UK
jeanmarc.robin@sciencespo.fr

Abstract. In this article, we present the new **aidsills** command for estimating almost-ideal demand systems and their quadratic extensions. In contrast with Poi's (2012, *Stata Journal* 12: 433–446) **quaids** command, which is based on the nonlinear **nlsur** command, **aidsills** uses the computationally attractive iterated linear least-squares estimator developed by Blundell and Robin (1999, *Journal of Applied Econometrics* 14: 209–232). The new command further allows one to account for endogenous prices and total expenditure by using instrumental-variable techniques. Elasticities and their standard errors can be obtained using the **aidsills_elas** postestimation command.

Keywords: st0393, aidsills, aidsills_pred, aidsills_elas, almost ideal, demand system, quadratic, endogeneity, iterated, linear, least squares

1 Introduction

In a series of three articles, Poi shows how to fit Deaton and Muellbauer's (1980) original almost-ideal demand system (AIDS) and Banks, Blundell, and Lewbel's (1997) quadratic almost-ideal demand system (QUAIDS) in Stata. The first two articles are based on Stata's `ml` (Poi 2002) and `nlsur` (Poi 2008) commands, respectively, which require the user to write or modify one or more program files. The third introduces the `quaids` command (Poi 2012), which obviates the need for any programming and further allows for the inclusion of demographic variables as well as the computation of expenditure and price elasticities through postestimation tools. Although the `quaids` command shows undeniable qualities, it has two main shortcomings.

The first, and most important, is that it does not allow one to handle endogeneity. This point is crucial because endogeneity is a typical ingredient of demand systems. Several right-hand-side variables—such as prices and total expenditure—may indeed be correlated with the equation errors. This may be because of the simultaneous equation nature of demand models (total expenditure is the sum of expenditures on individual commodities and, because these expenditures are assumed to be endogenous, it might be expected to be jointly endogenous); because of some unobserved (or uncontrolled for) features of the commodities (goods may differ in quality from one household to another, and their prices may reflect these differences in quality and, therefore, depend on tastes); or because of measurement errors. If expenditure or prices are correlated with the equation errors, resulting estimators will be both biased and inconsistent.

The other shortcoming of the `quaids` command is that it requires the use of nonlinear techniques (`nlsur`) that can be computationally demanding, especially for large and disaggregated demand systems. A fundamental problem in demand estimation is the number of parameters to estimate. The `nlsur` estimates are obtained by iterating a series of linear regressions on a first-order linear expansion of the model. If there are 20 commodities and 20 equations, the linear expansion is a system of 20 equations with a number of the order of 400 parameters in each equation. Yet the number of products in most markets is often higher than 20, and even when using a multistage budgeting approach (Hausman, Leonard, and Zona 1994), the number of goods in each segment may still be large.

In this article, we present an alternative to the `quaids` command, called `aidsills`, where the potential endogeneity of prices and total expenditure can be tested and controlled for and where the estimation is performed using linear techniques. `aidsills` is based on Gauss's `aids.src` program written by J.-M. Robin to estimate the AIDS and QUAIDS using Blundell and Robin's (1999) iterated linear least-squares (ILLS) estimator. Although nonlinear, almost-ideal (AI) demand models, as most popular parametric demand systems, share a common property: they are conditionally linear. That is, they are linear in all the parameters conditional on a set of functions of explanatory variables and parameters. Browning and Meghir (1991) exploited this conditional linearity to construct a simple ILLS estimator for the AI demand model, and Blundell and Robin (1999) generalized it and derived the conditions for its consistency and asymptotic normality. Blundell and Robin (1999) also showed how to account for the endogeneity of total expenditure by using the instrumental-variable (IV) and augmented regression techniques of Hausman (1978) and Holly and Sargan (1982).

ILLS is a preferred alternative to nonlinear seemingly unrelated regressions (SUR) and nonlinear three-stage least squares for large demand systems. By exploiting the conditional linearity property, the ILLS estimator requires only a series of linear SUR when using the `reg3` command. This makes the estimation much faster, which may be helpful for models with many equations and to those estimated on large datasets. In the above example of 20 commodities and 20 equations, the `aidsills` estimates are obtained by iterating least squares applied to a system of 20 equations, each of approximately 20 (instead of 400) parameters.

Other advantages of the new command compared with Poi's are that it allows the user to fit unconstrained, homogeneity constrained, or homogeneity and symmetry constrained models, and it allows the user to test whether these theoretical restrictions hold. On the other hand, `aidsills` does not provide the `vce()` option available with `quaids`. Eventually, expenditure and price (compensated or not) elasticities can be obtained using the postestimation command `aidsills_elas`. This postestimation command does not compute elasticities for individual observations, as can be done with Poi's postestimation commands, but it computes elasticities at the mean point of a user-defined sample with their standard errors.

This article is organized as follows. In section 2, we present the model and the estimation procedure. In section 3, we provide the `aidsills` syntax and options. In section 4, we give some examples.

2 AI demand systems

In this section, we briefly describe AI demand models and elasticities, we show how endogenous regressors can be dealt with, and we present the principle of the ILLS estimator.

2.1 Overview

Let's consider the quadratic extension of Deaton and Muellbauer's (1980) AIDS. In the QUAIDS, introduced by Banks, Blundell, and Lewbel (1997), the budget share w_i^h on good $i = 1, \ldots, N$ for household $h = 1, \ldots, H$ with log total-expenditure x^h and the log price N-vector $\mathbf{p}^h$ is given by

$$w_i^h = \alpha_i + \boldsymbol{\gamma}_i'\mathbf{p}^h + \beta_i \left\{x^h - a\left(\mathbf{p}^h, \boldsymbol{\theta}\right)\right\} + \lambda_i \frac{\left\{x^h - a\left(\mathbf{p}^h, \boldsymbol{\theta}\right)\right\}^2}{b(\mathbf{p}^h, \boldsymbol{\theta})} + u_i^h \tag{1}$$

with the nonlinear price aggregators

$$\begin{aligned} a\left(\mathbf{p}^h, \boldsymbol{\theta}\right) &= \alpha_0 + \boldsymbol{\alpha}'\mathbf{p}^h + \frac{1}{2}\mathbf{p}^{h\prime}\boldsymbol{\Gamma}\boldsymbol{p}^h \\ b\left(\mathbf{p}^h, \boldsymbol{\theta}\right) &= \exp\left(\boldsymbol{\beta}'\mathbf{p}^h\right) \end{aligned}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)'$, $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_N)'$, $\boldsymbol{\theta}$ is the set of all parameters, and u_i^h is an error term.[1] These parameters must satisfy three sets of theoretical restrictions: all must sum to zero over all equations except the constant term, which must sum to one (additivity); log price-parameters must sum to zero within each equation (homogeneity); and the effect of log price i on budget share j must equal the effect of log price j on budget share i (symmetry). The additivity constraint is mechanically satisfied in AI-type demand models. This is not the case for homogeneity and symmetry constraints, which can be tested and imposed.

There are different ways to introduce demographic variables in a demand system. Here households' heterogeneity enters the demand system through the α's, which are

1. Parameter α_0 in the first price aggregator is unidentified and can be set to 0 or to any other fixed value.

modeled as linear combinations of a set of sociodemographic variables ($\mathbf{s}^h$) observed in the data

$$\boldsymbol{\alpha}^h = \mathbf{A}\mathbf{s}^h, \quad \mathbf{A} = (\boldsymbol{\alpha}_i')$$

This is called the translating approach (Pollak and Wales 1981), and it allows the level of demand to depend upon demographic variables. Note that modeling heterogeneity like this does not mean that heterogeneity enters the model linearly. As seen in (1), heterogeneity appears not only linearly in the intercepts but also nonlinearly in all expenditure terms through the first price aggregator. Although this approach is more restrictive than the scaling approach (Ray 1983) used by Poi (2012), which allows the level and slope (of total expenditure terms) to depend upon demographic variables, it preserves the conditional linearity of the model.

2.2 Elasticities

One of the main motivations for estimating demand systems is to derive expenditure and price elasticities. Omitting h superscripts, differentiating (1) with respect to x and p_j yields, respectively,

$$\begin{aligned} \mu_i &= \beta_i + 2\tau_i \frac{\{x - a(\mathbf{p}, \boldsymbol{\theta})\}}{b(\mathbf{p}, \boldsymbol{\theta})} \\ \mu_{ij} &= \gamma_{ij} - \mu_i(\boldsymbol{\alpha}_j + \boldsymbol{\gamma}_j \mathbf{p}) - \lambda_i \beta_j \frac{\{x - a(\mathbf{p}, \boldsymbol{\theta})\}^2}{b(\mathbf{p}, \boldsymbol{\theta})} \end{aligned}$$

Expenditure elasticities are then given by $e_i = \mu_i / w_i + 1$; uncompensated price elasticities by $e_{ij}^u = \mu_{ij}/w_i - \delta_{ij}$, where δ_{ij} is the Kronecker delta; and compensated price elasticities by $e_{ij}^c = e_{ij}^u + e_i w_j$.

2.3 Handling endogeneity

Ordinary least squares (OLS) or SUR (linear or nonlinear) generally do not provide consistent estimators for (1) because of the potential endogeneity of some right-hand-side variables. In each share equation, the error term u_i^h may be correlated with the log total-budget variable x^h (common shocks determine both taste and total expenditure changes). It may also be correlated with log prices $\mathbf{p}^h$. In most empirical studies, prices are unit values that are computed for each good as the ratio of expenditures and physical quantities. These unit values mostly depend on actual market prices. Because a given good may differ in quality by household, its calculated unit values may reflect these differences in quality and, therefore, may depend on tastes (see Deaton [1988]). These correlations, which are sources of potential biases, can be accounted for with IV and augmented regression techniques (Hausman 1978; Holly and Sargan 1982).

Assuming that there exists a set of IVs (for budget alone or for prices alone, or for both), the specification in (1) can be augmented with the error vector $\widehat{\mathbf{v}}^h$ predicted from estimating reduced forms for x^h and $\mathbf{p}^h$. The error u_i^h is written as the orthogonal decomposition

$$u_i^h = \boldsymbol{\rho}_i \widehat{\mathbf{v}}^h + \varepsilon_i^h$$

and $E(\varepsilon_i^h | x^h, \mathbf{p}^h) = 0$ is assumed for all i and h. The SUR estimator of the demand parameters in this augmented regression framework is identical to the three-stage least-squares estimator, and testing for the significance of the coefficients of $\widehat{\mathbf{v}}^h$ is a direct test of the exogeneity of x^h and each element of $\mathbf{p}^h$. Independent variables in the reduced-form equations are all the variables included in $\mathbf{s}^h$ (sociodemographic variables, but it could also include time dummies), the log prices or log total-expenditure if exogenous, and the proper identifying instruments.

2.4 ILLS estimator

As mentioned previously, an attractive feature of (1) is that it is conditionally linear in price aggregators: all equations are linear in all parameters conditional on price aggregators. Estimation using the iterated moment estimator developed in Blundell and Robin (1999) is, therefore, straightforward. This estimator consists of the following series of iterations: for given values of price aggregators, we estimate the parameters using a linear moment estimator, we use these estimates to update price aggregators, and then we continue the iteration until numerical convergence occurs. If numerical convergence occurs, this procedure yields a consistent and asymptotically normal estimator of $\boldsymbol{\theta}$.

Specifically, unbiased estimates of parameters in (1) are obtained by iterating a series of SUR or OLS regressions of w_i^h on $\mathbf{s}^h$, $\mathbf{p}^h$, $\{x^h - a(\cdot)\}$, $\{x^h - a(\cdot)\}^2/b(\cdot)$, and $\widehat{\mathbf{v}}^h$. [2] The Stone price index, which can be written as $\overline{\mathbf{w}}'\mathbf{p}^h$ (where $\overline{\mathbf{w}}$ is the N-vector containing the sample average budget shares), and the unit vector are used as initial values for $a(\cdot)$ and $b(\cdot)$, respectively. Within each iteration, the estimation is performed by SUR, whether constraining for homogeneity or not. Additivity is automatically satisfied and homogeneity can easily be imposed by considering $N-1$ relative prices instead of N absolute prices in each equation (see Deaton and Muellbauer [1980, 318]). Convergence occurs when the relative-difference criterion $\max |(\boldsymbol{\theta}_{n+1} - \boldsymbol{\theta}_n) \oslash (\boldsymbol{\theta}_n + 1)| \leq \text{tol}$, where $\oslash$ is the element-by-element division operator and tol is a predefined tolerance level (10^{-5}, for instance), is satisfied. Once convergence has occurred, a last estimation can be performed imposing the symmetry constraint.[3] Standard errors of all parameters in all equations are then simultaneously calculated using the asymptotic variance–covariance matrix given in Blundell and Robin (1999), which takes into account the predicted regressors $\widehat{\mathbf{v}}^h$ introduced in each equation as well as the correlation of the error terms ε_i^h across equations.

2. OLS and SUR give strictly identical parameter estimates because the same set of variables appears in the right-hand side of each equation.
3. We could have imposed symmetry within the iteration process, but we did not do this because it increases the number of estimations where convergence fails while also giving almost identical results.

3 The aidsills command

In this section, we present the syntax of the `aidsills` command, its available options, and its two postestimation commands, `aidsills_pred` and `aidsills_elas`.

3.1 Syntax

The syntax for `aidsills` is as follows:

`aidsills` *varlist*$_{\text{shares}}$ [*if*] [*in*], `prices(`*varlist*$_{\text{prices}}$`)` `expenditure(`*varname*`)` [`intercept(`*varlist*`)` `ivprices(`*varlist*`)` `ivexpenditure(`*varlist*`)` `quadratic` `homogeneity` `symmetry` `nofirst` `tolerance(`#`)` `iteration(`#`)` `alpha_0(`#`)` `level(`#`)`]

where *varlist*$_{\text{shares}}$ is a list of N variables for budget shares, the last being used as the reference. They must sum to one for each observation.

Note that, within the program, the first $N-1$ prices listed in `prices()` are introduced as relative prices, using the Nth price as the reference. Formally, omitting the unnecessary terms, the ith equation of the unconstrained model (1) can be written as

$$w_i^h = \cdots + \gamma_{i1}\left(p_1^h - p_N^h\right) + \cdots + \gamma_{iN-1}\left(p_{N-1}^h - p_N^h\right) + \gamma_{iN} p_N^h + \cdots$$

where p_j^h is the logarithm of the jth price listed in `prices()`. Under the null hypothesis of homogeneity and for real total expenditure $\{x^h - a(\cdot)\}$ held constant, only relative prices matter. Therefore, the impact of the reference (absolute) price on each budget share must be zero ($\gamma_{iN} = 0$ for all i). A direct test of homogeneity then consists in fitting the unconstrained model, keeping the Nth absolute price as a regressor, and testing whether it has a jointly significant impact. Of course, a simple way to constrain for homogeneity is to remove this Nth price from the list of regressors. Further note that absolute price effects are easily recovered from relative price effects and that the output displays the absolute (not the relative) price effects.

The postestimation command `aidsills_pred` can be used following `aidsills` to obtain the linear prediction (`xb`, the default) or residuals using the estimates of the equation specified in `equation()`. Predictions are available both in and out of sample; type `aidsills_pred` ... `if e(sample)` ... if predictions are wanted only for the estimation sample. The syntax is as follows:

`aidsills_pred` *newvar* [*if*] [*in*], `equation(`*varname*$_{\text{share}}$`)` [`residuals`]

The postestimation command `aidsills_elas` can be used following `aidsills` to obtain budget and uncompensated and compensated price elasticities computed at the mean point of the sample defined by `if` and `in` with their standard errors. Results are presented using Jann's (2005) `estout` command, which can be downloaded from within Stata by typing `search estout`.[4] The syntax is as follows:

`aidsills_elas` [*if*] [*in*]

3.2 Options for aidsills

`prices(`*varlist*$_{\text{prices}}$`)` specifies a list of N variables for prices, in level (not logarithm). Prices must appear in the same order as shares. `prices()` is required.

`expenditure(`*varname*`)` specifies the total expenditure variable, in level (not logarithm). *varname* must represent the total amount of money spent on the N goods of the system for each observation. `expenditure()` is required.

`intercept(`*varlist*`)` specifies the variables used as sociodemographic shifters; a constant term is added by default, whether the `intercept()` option is specified or not.

`ivprices(`*varlist*`)` specifies that the potentially endogenous prices (or unit values) are to be instrumented by all exogenous variables listed in *varlist* of `intercept()`, the log of *varname* in `expenditure()` if expenditure is exogenous, and identifying IVs listed in *varlist* of `ivprices()`—the number of variables in `ivprices()` must be at least equal to the number of prices—and `ivexpenditure()`.

`ivexpenditure(`*varlist*`)` specifies that the potentially endogenous total expenditure is to be instrumented by all exogenous variables listed in *varlist* of `intercept()`, the log of variables listed in *varlist* of `prices()` if prices are exogenous, and identifying IVs listed in *varlist* of `ivprices()` and `ivexpenditure()`.

Note: Variables in *varlist* of `ivprices()` cannot enter *varlist* of `ivexpenditure()`, and vice versa.

`quadratic` indicates that the quadratic version of the AIDS must be considered.

`homogeneity` indicates that the log price-parameters must satisfy the homogeneity constraint; a homogeneity chi-squared test is provided when the unconstrained model is fit.

`symmetry` indicates that the log price-parameters must satisfy the homogeneity and symmetry constraints; a symmetry chi-squared test is provided when the homogeneity constrained model is fit.

`nofirst` indicates that the output from the first-stage instrumental regressions must be omitted.

4. See http://www.ats.ucla.edu/stat/stata/faq/estout.htm for more information.

`tolerance(#)` specifies the criterion used to declare convergence of the ILLS estimator. The default is `tolerance(1e-5)`.

`iteration(#)` specifies the maximum number of iterations; `iteration(0)` estimates the linearized version of the model, where $a(\cdot)$ is replaced by the Stone price index and $b(\cdot) = 1$. The default is `iteration(50)`.

`alpha_0(#)` specifies the value of α_0 in the price index $a(\cdot)$. The default is `alpha_0(0)`.

`level(#)` specifies the confidence level, as a percentage, for confidence intervals. The default is `level(95)`.

3.3 Options of aidsills_pred

`equation(`*varname*$_{\text{share}}$`)` specifies the variable for which predictions are calculated. `equation()` is required.

`residuals` calculates the residuals rather than the linear prediction (`xb`, the default) for the specified equation.

3.4 Stored results

`aidsills` stores the following in `e()`:

Scalars	
`e(N)`	number of observations
`e(alpha_0)`	value of α_0
`e(iteration)`	maximum number of iterations
Macros	
`e(cmd)`	`aidsills`
`e(model)`	name of the model
`e(const)`	constraint label used in the output header
`e(shares)`	budget share variables
`e(prices)`	price variables
`e(expenditure)`	expenditure variable
`e(ivprices)`	IVs for price variables
`e(ivexpenditure)`	IVs for expenditure variable
`e(intercept)`	demographic variables
Matrices	
`e(b)`	coefficient vector
`e(V)`	variance–covariance matrix of the estimators
`e(alpha)`	estimated $\boldsymbol{\alpha}$ vector
`e(gamma)`	estimated $\boldsymbol{\Gamma}$ matrix
`e(beta)`	estimated $\boldsymbol{\beta}$ vector
`e(lambda)`	estimated $\boldsymbol{\lambda}$ vector
`e(rho)`	estimated $\boldsymbol{\rho}$ vector
Functions	
`e(sample)`	marks estimation sample

4 Examples

For comparison, we apply `aidsills` to the dataset used by Poi (2012) to illustrate his `quaids` command. We drop variables `lnp1` to `lnp4` and `lnexp` because they are not used in estimation (but they could be kept, it does not matter). We randomly generate two demographic variables.

```
. version 13
. webuse food
. drop lnp1 lnp2 lnp3 lnp4 lnexp
. set seed 1
. generate nkids = int(runiform()*4)
. generate rural = (runiform() > 0.7)
```

AI demand models are singular (the dependent variables sum to one), and one equation must be eliminated when estimating. In the `aidsills` command, it is the last equation (here the fourth), which does not matter because singular systems are invariant to this choice. Estimated parameters for the dropped equation and corresponding elements in the variance–covariance matrix are recovered from additivity.

In the base output, the header gives information on which model is estimated; a first table gives some overall statistics for each equation; and a second table gives (equation by equation) the detailed parameter estimates, their asymptotic standard errors, and the usual related statistics. Independent variables include the four logs of (absolute) prices, $\texttt{lnp1} = \ln(\texttt{p1})$ to $\texttt{lnp4} = \ln(\texttt{p4})$, and the log of expenditure terms: $\texttt{lnx} = \ln(\texttt{expfd}) - a(\cdot)$ and $\texttt{lnx2} = (\texttt{lnx}^2)/b(\cdot)$, sociodemographic variables, and a constant. The name of the corresponding parameter vector or matrix (alpha, gamma, beta, and [if applicable] lambda and rho) is added in front of each independent variable.

An issue raised by using `food.dta` above is that `lnx` and `lnx2` are so strongly correlated that we cannot get satisfying estimates for any quadratic demand system (many parameters are unusually insignificant).[5] Indeed, estimating any equation of any quadratic model by using `regress` and typing `estat vif` shows a very strong collinearity between expenditure terms, with the highest variance inflation factors ranging from around 60 when the `iteration(0)` option is specified to around 200 when it is not. Models presented below are thus nonquadratic AIDS.

4.1 Comparisons

We consider a model without demographic variables and that assumes exogeneity for all regressors, given that symmetry constrained estimates and elasticities obtained from `aidsills` and `quaids` commands must be almost identical. Indeed, except small differences due to rounding precision, the estimates below are equivalent to those obtained using Poi's `quaids` command: `quaids w1-w4, prices(p1-p4) expenditure(expfd) noquadratic anot(10)`.

5. Because this issue does not occur when we use other much larger datasets, it seems to be specific to these data.

```
. aidsills w1-w4, prices(p1-p4) expenditure(expfd) symmetry alpha_0(10)
Iteration = 1       Criterion = .17980229
Iteration = 2       Criterion = .00611116
Iteration = 3       Criterion = .00012707
Iteration = 4       Criterion = 1.563e-06
AIDS - PROPER ESTIMATION WITH FIXED ALPHA_0 = 10
HOMOGENEITY AND SYMMETRY CONSTRAINED ESTIMATES

Equation          Obs  Parms        RMSE    "R-sq"    F(  5,  4042)   Prob > F

w1               4048      5    .1326944    0.1233         142.18     0.0000
w2               4048      5    .1024636    0.0753          82.34     0.0000
w3               4048      5    .0537182    0.1442         170.28     0.0000
w4               4048      5    .1064814    0.0515          54.93     0.0000


             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

w1           |
  gamma_lnp1 |   .1230886   .0059509    20.68   0.000      .111425    .1347521
  gamma_lnp2 |  -.0546438    .006042    -9.04   0.000    -.0664859   -.0428017
  gamma_lnp3 |  -.0352279   .0045005    -7.83   0.000    -.0440487   -.0264071
  gamma_lnp4 |  -.0332169   .0046497    -7.14   0.000    -.0423302   -.0241036
    beta_lnx |   .0157531   .0036641     4.30   0.000     .0085717    .0229345
  alpha_cons |   .3947989   .0238983    16.52   0.000     .3479591    .4416387

w2           |
  gamma_lnp1 |  -.0546438   .0045337   -12.05   0.000    -.0635296    -.045758
  gamma_lnp2 |   .0680193   .0045612    14.91   0.000     .0590794    .0769591
  gamma_lnp3 |  -.0012362   .0034042    -0.36   0.717    -.0079083    .0054359
  gamma_lnp4 |  -.0121393   .0036722    -3.31   0.001    -.0193366   -.0049419
    beta_lnx |  -.0260689    .002678    -9.73   0.000    -.0313176   -.0208201
  alpha_cons |   .1408526   .0174654     8.06   0.000     .1066211    .1750841

w3           |
  gamma_lnp1 |  -.0352279   .0023811   -14.79   0.000    -.0398948    -.030561
  gamma_lnp2 |  -.0012362   .0023942    -0.52   0.606    -.0059287    .0034563
  gamma_lnp3 |   .0425736   .0017881    23.81   0.000     .0390689    .0460782
  gamma_lnp4 |  -.0061095   .0018934    -3.23   0.001    -.0098205   -.0023985
    beta_lnx |   .0013848   .0018103     0.76   0.444    -.0021634     .004933
  alpha_cons |   .1109648   .0118041     9.40   0.000     .0878292    .1341004

w4           |
  gamma_lnp1 |  -.0332169   .0046833    -7.09   0.000     -.042396   -.0240378
  gamma_lnp2 |  -.0121393   .0047169    -2.57   0.010    -.0213843   -.0028943
  gamma_lnp3 |  -.0061095   .0035087    -1.74   0.082    -.0129864    .0007674
  gamma_lnp4 |   .0514657   .0036728    14.01   0.000     .0442671    .0586642
    beta_lnx |    .008931   .0037124     2.41   0.016     .0016548    .0162072
  alpha_cons |   .3533837   .0242125    14.60   0.000      .305928    .4008394
```

Comparing elasticities at the sample mean point shows that values calculated after the `aidsills` and `quaids` commands are very close. After running `aidsills`, the postestimation command `aidsills_elas` gives the predicted shares, budget, and price elasticities with their standard errors.

```
. aidsills_elas
PREDICTED SHARES, BUDGET AND (UN)COMPENSATED OWN-PRICE ELASTICITIES
---------------------------------------------------------------
                  shares        budget       u_price       c_price
                    b/se          b/se          b/se          b/se
---------------------------------------------------------------
w1              0.401***      1.039***     -0.713***     -0.296***
               (0.002)       (0.009)       (0.015)       (0.015)
w2              0.239***      0.891***     -0.708***     -0.494***
               (0.001)       (0.011)       (0.019)       (0.019)
w3              0.102***      1.014***     -0.585***     -0.481***
               (0.001)       (0.018)       (0.018)       (0.018)
w4              0.257***      1.035***     -0.811***     -0.545***
               (0.002)       (0.014)       (0.015)       (0.014)
---------------------------------------------------------------
* p<0.05, ** p<0.01, *** p<0.001
UNCOMPENSATED CROSS-PRICE ELASTICITIES
---------------------------------------------------------------
                      p1            p2            p3            p4
                    b/se          b/se          b/se          b/se
---------------------------------------------------------------
w1             -0.713***     -0.139***     -0.092***     -0.095***
               (0.015)       (0.015)       (0.011)       (0.012)
w2             -0.174***     -0.708***      0.007        -0.017
               (0.020)       (0.019)       (0.014)       (0.014)
w3             -0.351***     -0.013        -0.585***     -0.064***
               (0.025)       (0.024)       (0.018)       (0.019)
w4             -0.147***     -0.050**      -0.028*       -0.811***
               (0.019)       (0.019)       (0.014)       (0.015)
---------------------------------------------------------------
* p<0.05, ** p<0.01, *** p<0.001
COMPENSATED CROSS-PRICE ELASTICITIES
---------------------------------------------------------------
                      p1            p2            p3            p4
                    b/se          b/se          b/se          b/se
---------------------------------------------------------------
w1             -0.296***      0.110***      0.014         0.172***
               (0.015)       (0.015)       (0.011)       (0.012)
w2              0.184***     -0.494***      0.098***      0.212***
               (0.019)       (0.019)       (0.014)       (0.014)
w3              0.055*        0.230***     -0.481***      0.197***
               (0.023)       (0.023)       (0.018)       (0.018)
w4              0.269***      0.198***      0.078***     -0.545***
               (0.018)       (0.019)       (0.014)       (0.014)
---------------------------------------------------------------
* p<0.05, ** p<0.01, *** p<0.001
```

After fitting the same model with Poi's `quaids` command—that is, `quaids w1-w4, prices(p1-p4) expenditure(expfd) noquadratic anot(10)`—elasticities and their standard errors can be obtained as follows:

```
. estat expenditure, atmeans stderrs
. matrix list r(expelas)
r(expelas)[1,4]
           c1         c2         c3         c4
r1  1.0401845  .89145306  1.0141862  1.0331519
. matrix list r(sd)
r(sd)[1,4]
           c1         c2         c3         c4
r1  .00851918  .01091724  .01347971  .01066943
. estat uncompensated, atmeans stderrs
. matrix list r(uncompelas)
r(uncompelas)[4,4]
             c1          c2          c3          c4
r1  -.71352724  -.13889921  -.09181671  -.09594133
r2  -.17218979  -.70896104   .00636933  -.01667156
r3  -.34971894  -.01450851  -.58618298   -.0637758
r4  -.14692494  -.04962317  -.02731562  -.80928815
. matrix list r(sd)
r(sd)[4,4]
           c1         c2         c3         c4
r1  .01480274  .00999946  .00549943   .0092104
r2  .01704632   .0187432  .00830413  .01257617
r3  .02214005  .01961331  .01727322  .01622405
r4  .01489985  .01187526  .00645927  .01430764
. estat compensated, atmeans stderrs
. matrix list r(compelas)
r(compelas)[4,4]
             c1          c2          c3          c4
r1  -.29659618   .11092802    .0144787   .17118946
r2   .18512612  -.49485551   .09746602   .21226337
r3   .05679139   .22907457  -.48254431   .19667835
r4   .26718728     .198515   .07826113  -.54396341
. matrix list r(sd)
r(sd)[4,4]
           c1         c2         c3         c4
r1  .01427661  .00980769  .00543718  .00904256
r2  .01636742   .0185529  .00822818  .01236247
r3  .02132646  .01933838  .01721192  .01599929
r4   .0141134  .01156203  .00636639  .01417522
```

The main difference between the two commands is computing time. `aidsills` takes 4 iterations and lasts less than 1 second, which is very fast, while Poi's `quaids` command is slower, taking 14 iterations and lasting 3 seconds—see table 1 below, part 1. After adding two demographic variables, estimation still requires 4 iterations and lasts 1 second using `aidsills` and requires 41 iterations and lasts 16 seconds using `quaids`. Note that the number of iterations does not matter as long as the process converges.

But time matters, especially when the dataset and the number of equations become large—see table 1, part 2. For the same model, but with almost twice as many equations and estimated on data more than 6 times larger than in part 1, `aidsills` needs only 7 seconds versus more than 1 minute for `quaids`. This difference becomes much bigger with the two demographic variables added because of the way they are incorporated in the model: computation time does not change much for the `aidsills` command (only 3 seconds longer), but it really explodes for the `quaids` command (more than 5 minutes). Given that demand systems can be expected to cover more than 20 commodities and several thousands of individual observations characterized by tens of demographic variables, this is a huge difference.

Table 1. Convergence speed of Stata commands (Stata/MP 13.1)

	`aidsills`		`quaids`	
Data and model	iterations	time	iterations	time
(1) Obs. = 4048, eq. = 4,				
demo. = 0:	4	<1s	14	3s
demo. = 2:	4	1s	41	16s
(2) Obs. = 25776, eq. = 7,				
demo. = 0:	5	7s	19	1min31s
demo. = 2:	6	10s	68	5min09s

Note: HP Z400 Workstation, Intel Xeon CPU, 3.20 GHz, 7.98 Go RAM.

4.2 Endogenous regressors

Beyond the computation speed, the main interest of the `aidsills` command is to allow for endogenous regressors. Indeed, most articles dealing with demand-system estimations at least consider that the total budget is endogenous.

Let's consider that total expenditure might be endogenous and generate a variable—say, `lninc`—for the logarithm of income, partly random and partly correlated to total expenditure. This variable is used as the identifying IV (of course, using more than one IV is possible). It is indicated by the addition of the `ivexpenditure(lninc)` term in the `aidsills` command. Let's also incorporate our two demographic variables (`nkids`, `rural`) in the model.

Except when the `nofirst` option is specified, the output now contains an additional table presenting the reduced-form estimates. The dependent variable is the log total-expenditure, and the independent variables are all exogenous variables entering the model (here four log prices and two demographic variables) and the identifying IV. From these estimates, residuals `vexpfd` (that is, `v` + *varname* in `expenditure(`*varname*`)`) are computed and introduced in the demand model as additional variables.

When prices are not exogenous, each log price is regressed on exogenous variables (that is, total expenditure if exogenous, demographic variables) and the identifying IVs, which there are at least four of in `ivprices()`. Again results from these first-stage regressions are reported in a series of tables preceding the demand-system estimates. Residuals `vp1` to `vp4` (that is, `v` + *varlist* in `prices(`*varlist*`)`) are then predicted and added to the set of regressors in each demand equation.

```
. generate lninc = ln(100+(runiform()*1000)+10*expfd)

. aidsills w1-w4, prices(p1-p4) expenditure(expfd) intercept(nkids rural)
> ivexpenditure(lninc) symmetry alpha_0(10)
INSTRUMENTAL REGRESSION(S)

      Source |       SS           df       MS      Number of obs   =     4,048
-------------+----------------------------------   F(7, 4040)      =    401.62
       Model |  552.014066         7  78.8591523   Prob > F        =    0.0000
    Residual |  793.261605     4,040  .196351883   R-squared       =    0.4103
-------------+----------------------------------   Adj R-squared   =    0.4093
       Total |  1345.27567     4,047  .332413064   Root MSE        =    .44312

------------------------------------------------------------------------------
     lnexpfd |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        lnp1 |   .1439361   .0209383     6.87   0.000     .1028978    .1849745
        lnp2 |  -.0616234   .0224917    -2.74   0.006    -.1057064   -.0175405
        lnp3 |  -.0139445   .0151455    -0.92   0.357    -.0436292    .0157401
        lnp4 |   .0035425   .0165306     0.21   0.830     -.028857    .0359419
       lninc |   .8736148   .0168881    51.73   0.000     .8405147    .9067149
       nkids |    .008288    .006216     1.33   0.182    -.0038951    .0204712
       rural |   .0021161   .0151447     0.14   0.889    -.0275671    .0317992
       _cons |  -2.429511   .1187146   -20.47   0.000    -2.662188   -2.196835
------------------------------------------------------------------------------

Iteration = 1       Criterion = .2153067
Iteration = 2       Criterion = .00907554
Iteration = 3       Criterion = .00024835
Iteration = 4       Criterion = 5.910e-06
AIDS - PROPER ESTIMATION WITH FIXED ALPHA_0 = 10
HOMOGENEITY AND SYMMETRY CONSTRAINED ESTIMATES
------------------------------------------------------------------------------
Equation           Obs  Parms        RMSE    "R-sq"    F(  8,  4039)   Prob > F
------------------------------------------------------------------------------
w1                4048      8    .1327288    0.1235          81.33     0.0000
w2                4048      8    .1024318    0.0766          47.87     0.0000
w3                4048      8    .0536455    0.1471          99.57     0.0000
w4                4048      8    .1064861    0.0522          31.76     0.0000
------------------------------------------------------------------------------
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
w1						
gamma_lnp1	.1233584	.0059918	20.59	0.000	.1116147	.1351022
gamma_lnp2	-.055359	.0061231	-9.04	0.000	-.0673601	-.043358
gamma_lnp3	-.0347832	.0045051	-7.72	0.000	-.043613	-.0259533
gamma_lnp4	-.0332162	.0046255	-7.18	0.000	-.042282	-.0241504
beta_lnx	.0169582	.0059297	2.86	0.004	.0053361	.0285802
rho_vexpfd	-.0013014	.0076341	-0.17	0.865	-.0162639	.0136612
alpha_nkids	.001139	.0019357	0.59	0.556	-.0026549	.0049328
alpha_rural	.0028354	.0047168	0.60	0.548	-.0064095	.0120802
alpha_cons	.3997159	.0378613	10.56	0.000	.3255091	.4739226
w2						
gamma_lnp1	-.055359	.004543	-12.19	0.000	-.0642632	-.0464549
gamma_lnp2	.0702557	.0047201	14.88	0.000	.0610045	.079507
gamma_lnp3	-.0023602	.0034068	-0.69	0.488	-.0090374	.0043169
gamma_lnp4	-.0125365	.0037755	-3.32	0.001	-.0199364	-.0051365
beta_lnx	-.031956	.0040512	-7.89	0.000	-.0398962	-.0240158
rho_vexpfd	.011132	.0054667	2.04	0.042	.0004175	.0218465
alpha_nkids	-.0007932	.0014412	-0.55	0.582	-.0036179	.0020314
alpha_rural	.0032516	.0035104	0.93	0.354	-.0036287	.010132
alpha_cons	.104769	.0259033	4.04	0.000	.0539994	.1555386
w3						
gamma_lnp1	-.0347832	.0023909	-14.55	0.000	-.0394692	-.0300971
gamma_lnp2	-.0023602	.0024528	-0.96	0.336	-.0071676	.0024471
gamma_lnp3	.042702	.0017913	23.84	0.000	.039191	.0462129
gamma_lnp4	-.0055586	.0019351	-2.87	0.004	-.0093514	-.0017658
beta_lnx	.0068281	.0035393	1.93	0.054	-.0001088	.013765
rho_vexpfd	-.0105338	.0040452	-2.60	0.009	-.0184623	-.0026053
alpha_nkids	.000036	.000759	0.05	0.962	-.0014516	.0015236
alpha_rural	-.0005303	.0018491	-0.29	0.774	-.0041545	.0030939
alpha_cons	.14483	.0225554	6.42	0.000	.1006223	.1890377
w4						
gamma_lnp1	-.0332162	.0046813	-7.10	0.000	-.0423915	-.024041
gamma_lnp2	-.0125365	.0048012	-2.61	0.009	-.0219466	-.0031263
gamma_lnp3	-.0055586	.0035068	-1.59	0.113	-.0124318	.0013146
gamma_lnp4	.0513113	.0037496	13.68	0.000	.0439623	.0586603
beta_lnx	.0081698	.0074473	1.10	0.273	-.0064266	.0227662
rho_vexpfd	.0007032	.0083846	0.08	0.933	-.0157304	.0171367
alpha_nkids	-.0003817	.0014938	-0.26	0.798	-.0033096	.0025461
alpha_rural	-.0055566	.0036396	-1.53	0.127	-.0126901	.0015769
alpha_cons	.3506851	.0474526	7.39	0.000	.2576797	.4436905

Postestimation commands `test` and `testnl` can be used to perform tests on parameters. For instance, testing whether `vexpfd` is significant in each equation separately provides direct (and independent) tests of total expenditure exogeneity in each equation. A joint test can be obtained as follows:

```
. test rho_vexpfd
 ( 1)  [w1]rho_vexpfd = 0
 ( 2)  [w2]rho_vexpfd = 0
 ( 3)  [w3]rho_vexpfd = 0
 ( 4)  [w4]rho_vexpfd = 0
       Constraint 2 dropped
           chi2(  3) =   11.05
         Prob > chi2 =    0.0114
```

Additivity implies that Stata must drop one constraint (here the second, but it could have been any other): if the coefficient of `vexpfd` is zero in three equations, then it must also be zero in the fourth. Jointly, the null hypothesis of exogeneity can be rejected at the 5% level. It can be rejected at the 5% level in the second equation and at the 1% level in the third equation, separately.

When the `ivprices()` option is used—with at least as many arguments as there are prices—in addition to (or instead of) `ivexpenditure()`, four additional independent variables, `vp1-vp4`, appear in each equation, preceding (or replacing) `vexpfd`. Testing the exogeneity of a given price or all prices in the whole demand system, or in each equation separately, can then be done exactly the same way as above. This option can be used, for example, to fit Deaton's (1988) AI model with endogenous unit values, or it can be used to fit the lower level of Hausman, Leonard, and Zona's (1994) multistage budgeting model.

4.3 Other features

The `aidsills` command allows the user to fit proper AI models as well as linearized, constrained, and unconstrained models. It also allows one to test the validity of these constraints. For example, assuming the exogeneity of all regressors and removing the `symmetry` option while specifying the `iteration(0)` option, we fit the unconstrained linearized AIDS.

Note that when the estimation is unconstrained, a joint test of homogeneity is reported between the two tables. This is also the case when fitting the proper version of the unconstrained AIDS model (that is, without the `iteration(0)` option). Here homogeneity is jointly rejected at the 1% level. Testing for homogeneity can also be done after the estimation for each equation separately or jointly.

```
. aidsills w1-w4, prices(p1-p4) expenditure(expfd) intercept(nkids rural)
> iteration(0) alpha_0(10)
AIDS - LINEARIZED WITH STONE PRICE INDEX
UNCONSTRAINED ESTIMATES

Equation          Obs  Parms        RMSE    "R-sq"    F(  7,  4040)   Prob > F

w1               4048      7    .1325483    0.1259          83.13       0.0000
w2               4048      7    .1023763    0.0776          48.55       0.0000
w3               4048      7    .0530703    0.1653         114.31       0.0000
w4               4048      7    .1064415    0.0530          32.27       0.0000

HOMOGENEITY TEST:   Chi2(  3) =   121.84    Prob > chi2 =   0.0000

                    Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

w1
  gamma_lnp1     .1372222    .006269    21.89   0.000     .1249352    .1495093
  gamma_lnp2    -.0333321   .0068252    -4.88   0.000    -.0467092    -.019955
  gamma_lnp3    -.0588387   .0045422   -12.95   0.000    -.0677412   -.0499362
  gamma_lnp4    -.0314309   .0052674    -5.97   0.000    -.0417548    -.021107
    beta_lnx     .0210047   .0036465     5.76   0.000     .0138578    .0281517
 alpha_nkids     .0010471   .0018576     0.56   0.573    -.0025937    .0046879
 alpha_rural     .0031973   .0045257     0.71   0.480     -.005673    .0120676
  alpha_cons     .2376424   .0151096    15.73   0.000      .208028    .2672568

w2
  gamma_lnp1    -.0467974    .004842    -9.66   0.000    -.0562876   -.0373073
  gamma_lnp2     .0709456   .0052716    13.46   0.000     .0606136    .0812777
  gamma_lnp3     .0044428   .0035082     1.27   0.205    -.0024332    .0113188
  gamma_lnp4     .0095794   .0040684     2.35   0.019     .0016056    .0175533
    beta_lnx    -.0233401   .0028164    -8.29   0.000    -.0288602     -.01782
 alpha_nkids    -.0008194   .0014347    -0.57   0.568    -.0036314    .0019927
 alpha_rural     .0033278   .0034955     0.95   0.341    -.0035233    .0101789
  alpha_cons     .3928153   .0116702    33.66   0.000     .3699421    .4156885

w3
  gamma_lnp1    -.0389351     .00251   -15.51   0.000    -.0438546   -.0340155
  gamma_lnp2    -.0168701   .0027327    -6.17   0.000    -.0222261   -.0115141
  gamma_lnp3     .0399693   .0018186    21.98   0.000     .0364049    .0435337
  gamma_lnp4    -.0172411    .002109    -8.18   0.000    -.0213747   -.0131076
    beta_lnx    -.0038375     .00146    -2.63   0.009     -.006699   -.0009759
 alpha_nkids     .0001842   .0007437     0.25   0.804    -.0012735     .001642
 alpha_rural    -.0006798    .001812    -0.38   0.708    -.0042314    .0028717
  alpha_cons     .1008339   .0060497    16.67   0.000     .0889767     .112691

w4
  gamma_lnp1    -.0514897   .0050343   -10.23   0.000    -.0613567   -.0416227
  gamma_lnp2    -.0207434   .0054809    -3.78   0.000    -.0314858   -.0100011
  gamma_lnp3     .0144266   .0036476     3.96   0.000     .0072775    .0215757
  gamma_lnp4     .0390926   .0042299     9.24   0.000      .030802    .0473831
    beta_lnx     .0061728   .0029283     2.11   0.035     .0004335    .0119121
 alpha_nkids     -.000412   .0014917    -0.28   0.782    -.0033357    .0025117
 alpha_rural    -.0058453   .0036343    -1.61   0.108    -.0129685    .0012779
  alpha_cons     .2687084   .0121336    22.15   0.000     .2449269    .2924899
```

When homogeneity constrained models are fit, a test for symmetric price effects is provided, as can be seen in the proper AIDS below.

```
. aidsills w1-w4, prices(p1-p4) expenditure(expfd) intercept(nkids rural)
> homogeneity alpha_0(10)
Iteration = 1       Criterion = .179615
Iteration = 2       Criterion = .0061031
Iteration = 3       Criterion = .0001267
Iteration = 4       Criterion = 1.567e-06
AIDS - PROPER ESTIMATION WITH FIXED ALPHA_0 = 10
HOMOGENEITY CONSTRAINED ESTIMATES
```

Equation	Obs	Parms	RMSE	"R-sq"	F(7, 4040)	Prob > F
w1	4048	7	.1327138	0.1235	94.90	0.0000
w2	4048	7	.1024759	0.0756	55.05	0.0000
w3	4048	7	.0537307	0.1442	113.49	0.0000
w4	4048	7	.1064734	0.0522	37.06	0.0000

```
SYMMETRY TEST:      Chi2(  3) =    62.07    Prob > chi2 =   0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
w1						
gamma_lnp1	.1317659	.0058962	22.35	0.000	.1202096	.1433221
gamma_lnp2	-.0474768	.0059874	-7.93	0.000	-.059212	-.0357416
gamma_lnp3	-.0628936	.0044601	-14.10	0.000	-.0716353	-.054152
gamma_lnp4	-.0213955	.0046054	-4.65	0.000	-.0304219	-.012369
beta_lnx	.0166045	.0036438	4.56	0.000	.0094629	.0237461
alpha_nkids	.0011289	.0019139	0.59	0.555	-.0026222	.00488
alpha_rural	.0031723	.0046638	0.68	0.496	-.0059685	.0123131
alpha_cons	.4131347	.0240166	17.20	0.000	.366063	.4602065
w2						
gamma_lnp1	-.055147	.0045336	-12.16	0.000	-.0640327	-.0462613
gamma_lnp2	.0670054	.0045601	14.69	0.000	.0580678	.075943
gamma_lnp3	.0021435	.0034037	0.63	0.529	-.0045276	.0088146
gamma_lnp4	-.0140018	.0036739	-3.81	0.000	-.0212026	-.006801
beta_lnx	-.0261695	.00268	-9.76	0.000	-.0314223	-.0209168
alpha_nkids	-.0008039	.0014395	-0.56	0.577	-.0036253	.0020176
alpha_rural	.0031937	.0035067	0.91	0.362	-.0036794	.0100667
alpha_cons	.1379505	.0176701	7.81	0.000	.1033177	.1725833
w3						
gamma_lnp1	-.0296808	.0023732	-12.51	0.000	-.0343323	-.0250293
gamma_lnp2	-.0029737	.0023855	-1.25	0.213	-.0076491	.0017017
gamma_lnp3	.0446189	.0017819	25.04	0.000	.0411264	.0481114
gamma_lnp4	-.0119644	.001887	-6.34	0.000	-.0156629	-.0082659
beta_lnx	.0012039	.0018237	0.66	0.509	-.0023705	.0047782
alpha_nkids	.0001089	.0007499	0.15	0.885	-.001361	.0015787
alpha_rural	-.000582	.0018271	-0.32	0.750	-.0041631	.0029991
alpha_cons	.1019298	.0119737	8.51	0.000	.0784617	.125398

```
w4           |
  gamma_lnp1 |  -.0469381   .0047305    -9.92   0.000    -.0562097   -.0376664
  gamma_lnp2 |  -.0165549   .0047642    -3.47   0.001    -.0258926   -.0072172
  gamma_lnp3 |   .0161312   .0035437     4.55   0.000     .0091858    .0230767
  gamma_lnp4 |   .0473617   .0037069    12.78   0.000     .0400964     .054627
    beta_lnx |   .0083612   .0037706     2.22   0.027      .000971    .0157513
 alpha_nkids |  -.0004339   .0015107    -0.29   0.774    -.0033947     .002527
 alpha_rural |   -.005784   .0036811    -1.57   0.116    -.0129989    .0014309
  alpha_cons |   .3469849   .0247559    14.02   0.000     .2984641    .3955057
```

Again, testing for symmetry (or for any other hypotheses on parameters) can be done after the estimation.

```
. quietly test [w1]gamma_lnp2=[w2]gamma_lnp1, notest
. quietly test [w1]gamma_lnp3=[w3]gamma_lnp1, notest accumulate
. test [w2]gamma_lnp3=[w3]gamma_lnp2, accumulate
 ( 1)  [w1]gamma_lnp2 - [w2]gamma_lnp1 = 0
 ( 2)  [w1]gamma_lnp3 - [w3]gamma_lnp1 = 0
 ( 3)  [w2]gamma_lnp3 - [w3]gamma_lnp2 = 0
           chi2(  3) =   62.07
         Prob > chi2 =    0.0000
```

These values correspond to those reported in the output. Symmetry is rejected at any conventional level.

5 Acknowledgment

We thank an anonymous referee for helpful comments and suggestions on how to improve the code and the manuscript.

6 References

Banks, J., R. Blundell, and A. Lewbel. 1997. Quadratic Engel curves and consumer demand. *Review of Economics and Statistics* 79: 527–539.

Blundell, R., and J.-M. Robin. 1999. Estimation in large and disaggregated demand systems: An estimator for conditionally linear systems. *Journal of Applied Econometrics* 14: 209–232.

Browning, M., and C. Meghir. 1991. The effects of male and female labor supply on commodity demands. *Econometrica* 59: 925–951.

Deaton, A. S. 1988. Quality, quantity, and spatial variation of price. *American Economic Review* 78: 418–430.

Deaton, A. S., and J. Muellbauer. 1980. An almost ideal demand system. *American Economic Review* 70: 312–326.

Hausman, J., G. Leonard, and J. D. Zona. 1994. Competitive analysis with differenciated products. *Annales d'Économie et de Statistique* 34: 159–180.

Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251–1271.

Holly, A., and D. Sargan. 1982. Testing for exogeneity in a limited information framework. Cahiers de Recherches Economiques 8204, Université de Lausanne.

Jann, B. 2005. Making regression tables from stored estimates. *Stata Journal* 5: 288–308.

Poi, B. P. 2002. From the help desk: Demand system estimation. *Stata Journal* 2: 403–410.

———. 2008. Demand-system estimation: Update. *Stata Journal* 8: 554–556.

———. 2012. Easy demand-system estimation with quaids. *Stata Journal* 12: 433–446.

Pollak, R. A., and T. J. Wales. 1981. Demographic variables in demand analysis. *Econometrica* 49: 1533–1551.

Ray, R. 1983. Measuring the costs of children: An alternative approach. *Journal of Public Economics* 22: 89–102.

About the authors

Sébastien Lecocq is a research fellow in economics at the French National Institute for Agricultural Research (INRA).

Jean-Marc Robin is a professor of economics at the Department of Economics of Sciences Po, Paris, and at University College London.