The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$ 98	1-year subscription	$138
2-year subscription	$165	2-year subscription	$245
3-year subscription	$225	3-year subscription	$345
1-year student subscription	$ 75	1-year student subscription	$ 99
1-year institutional subscription	$245	1-year institutional subscription	$285
2-year institutional subscription	$445	2-year institutional subscription	$525
3-year institutional subscription	$645	3-year institutional subscription	$765
Electronic only		**Electronic only**	
1-year subscription	$ 75	1-year subscription	$ 75
2-year subscription	$125	2-year subscription	$125
3-year subscription	$165	3-year subscription	$165
1-year student subscription	$ 45	1-year student subscription	$ 45

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Stata, Stata Press, Mata, Mata, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2014)
14, Number 2, pp. 259–279

Simulated multivariate random-effects probit models for unbalanced panels

Alexander Plum
Otto von Guericke University Magdeburg
Magdeburg, Germany
alexander.plum@ovgu.de

Abstract. This article develops a method for implementing a simulated multivariate random-effects probit model for unbalanced panels (with gaps) and illustrates the model by using artificial data. Halton draws generated by `mdraws` are used to simulate multivariate normal probabilities with the `mvnp() egen` function. The estimator can be easily adjusted, for example, to allow for autocorrelated errors. The advantages of this simulated estimation, when compared with existing commands such as `redpace`, are high accuracy and improved stability.

Keywords: st0335, mdraws, mvnp(), redpace, Halton draws, random effects, simulated multivariate probit

1 Introduction

Dynamic models fit the effect of the past outcome on the current one, for example, the previous years' labor-market position (employed, unemployed) on today's labor-market position. Estimations can be biased when not taking individual-specific effects into account, which is done by including a time-invariant error term (Heckman 1981a). Furthermore, the initial conditions might be correlated with the time-invariant error term and, therefore, endogenous, referred to in the literature as the "initial condition problem" (Heckman 1981b).

Different estimation techniques take care of both aspects. If the outcome variable is binary, nonlinear models such as probit and logit can be applied. For probit models, there are two different commands, `redprob` and `redpace`, both written by Stewart (2006a,b).[1] The first command is calculated using (adaptive) Gauss–Hermite quadratures, and the second one, using maximum simulated likelihood (MSL). A limitation of the `redprob` command is that it needs a balanced panel. While the `redpace` help file states that the command requires a balanced panel, it can be applied for unbalanced panels, although gaps are not allowed.

This article focuses on simulated estimations. The `mvnp() egen` function (Cappellari and Jenkins 2006a) uses the Geweke–Hajivassilou–Keane (GHK) (Keane 1994) simulator to calculate multivariate normal probabilities. For simulation, one needs to take random

1. In the case of a categorical variable, Haan and Uhlendorff (2006) have presented the implementation of a multinomial random-effects logit estimator for Stata. Hole (2007) has written a command for mixed logits by using maximum simulated likelihood.

draws from multivariate normal density to simulate multivariate normal probabilities.[2] One command that derives quasi-random numbers, also referred to as Halton draws, is `mdraws`, written by Cappellari and Jenkins (2003, 2005, 2006b).[3] Halton draws are applied because of their high effectiveness. For example, fewer numbers are needed as compared with pseudorandom numbers. In the estimation, the likelihood is simulated, and the average of these simulations is derived. The principles of MSL are described in Train (2009).

The main advantage of applying simulated multivariate normal probabilities is that the link between the time points in a dynamic model can be directly adjusted to the researcher's purpose by specifying the variance–covariance matrix accordingly. The first illustration presents the basic estimation technique of a dynamic process with an unbalanced panel (including gaps) based on simulated multivariate normal probabilities. The second illustration shows how to specify the variance–covariance matrix to allow for autocorrelated errors. The advantages of this estimation technique, when compared with the `redpace` command, are high accuracy and improved stability. Furthermore, indications that computational time can be saved are found in the model with autocorrelated errors.

The remainder of the article is organized as follows: The first part presents the underlying econometric structure of a simulated multivariate random-effects probit model. The first illustration shows the Stata routine for an unbalanced panel with gaps and an empirical example based on artificial data. Robustness checks of the simulation properties are performed, including a comparison of the estimation results with the `redpace` command. The second illustration extends the model by allowing for autocorrelated error terms. The last part concludes.

2 Multivariate random-effects probit model

The latent variable y^*_{it} is specified by

$$y^*_{it} = \gamma y_{i(t-1)} + x'_{it}\boldsymbol{\beta} + \alpha_i + u_{it}$$

2. Halton sequences can also be generated using Mata (see Drukker and Gates [2006]).
3. A detailed discussion of the underlying commands can be found here.

with $i = 1, \ldots, N$ indicating the individuals and $t = 2, \ldots, T$ indicating the time periods. Explanatory variables are the lagged dependent variable y_{it-1} and the exogenous regressors x'_{it}. α_i is the time-invariant error term, and u_{it} is a time-specific idiosyncratic shock. It is assumed that the time-invariant error term and the explanatory variables are uncorrelated.[4] It is also assumed that the idiosyncratic shock follows a standard normal distribution $u_{it} \sim N(0, 1)$ and that the time-invariant error term is independent and identically distributed $\alpha_i \sim N(0, \sigma^2_\alpha)$. In most panels, the number of observed periods is small; therefore, in these cases, asymptotics are on N alone. The observed binary outcome variable is defined as

$$y_{it} = \begin{cases} 1 & \text{if } y^*_{it} > 0 \\ 0 & \text{else} \end{cases}$$

The composite error term is $\nu_{it} = \alpha_i + u_{it}$.[5] Because the composite error term incorporates the time-invariant error term, the composite error term is correlated over time. Therefore, the composite error term takes the following equicorrelation structure over time,

$$\text{corr}(\nu_{it}, \nu_{is}) = \sigma^2_\alpha$$

with $t, s = 2, \ldots, T$ and $t \neq s$. Because the outcome in the initial period might not be randomly distributed and T is thus endogenous, the proposition of Heckman (1981b) is followed by estimating a static equation for the first period[6]

$$y^*_{i1} = z'_{i1}\pi + \epsilon_i$$

with $i = 1, \ldots, N$. z_{i1} contains the explanatory variables x_{it} and exogenous instruments that have an effect on the outcome in only the initial period. It is assumed that the time-invariant error term of the subsequent periods α_i is correlated with the error term in the initial period in the following manner:

$$\epsilon_i = \theta\alpha_i + u_{i1}$$

For the idiosyncratic shock in the initial period, the following normalization is chosen: $u_{i1} \sim N(0, 1)$. The correlation of the composite error term between the initial period and the subsequent ones is

$$\text{corr}(\epsilon_i, \nu_{it}) = \theta\sigma^2_\alpha$$

4. This assumption might easily be violated. Alternatively, Mundlak–Chamberlain decomposition could be applied by including the time mean of the explanatory variables on the right side of the equation system (see Mundlak [1978] and Chamberlain [1984]).
5. Note that the composite error term is not standard normal distributed; thus coefficients must be adjusted when they are interpreted (see Arulampalam [1999]).
6. Wooldridge (2005) proposed an alternative approach by including the outcome of the initial period as an exogenous regressor in periods $t \geq 2$.

where $t = 2, \ldots, T$. The variance–covariance matrix of order $T \times T$ of the equation system now takes the following form:[7]

$$\mathbf{\Omega} = \begin{pmatrix} \theta^2\sigma_\alpha^2 + 1 & & & & \\ \theta\sigma_\alpha^2 & \sigma_\alpha^2 + 1 & & & \\ \theta\sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\alpha^2 + 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \theta\sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 + 1 \end{pmatrix}$$

Following the approach of Cappellari and Jenkins (2006a), a multivariate probit model is applied. By this means, the likelihood contribution of each individual is[8]

$$\begin{aligned} \Phi_{iT} = (&k_{i1} z'_{i1}\pi, k_{i2} x'_{i2}\beta, \ldots, k_{iT} x'_{iT}\beta, \\ &k_{i1}k_{i2}\Omega_{2,1}, k_{i1}k_{i3}\Omega_{3,1}, \ldots, k_{iT-1}k_{iT}\Omega_{T,T-1}) \end{aligned}$$

Φ_{iT} is the cumulative multivariate normal distribution function of order T. $\Omega_{p,q}$ refers to row p and column q of the variance–covariance matrix $\mathbf{\Omega}$. The model has T levels of explanatory variables and $T(T-1)/2$ covariance parameters. There are T sign variables k_{it}, where

$$k_{it} = \begin{cases} 1 & \text{if } y_{it} = 1 \\ -1 & \text{else} \end{cases}$$

Hence, the log likelihood to be maximized is the sum of the individual log-likelihood contributions

$$\ln L = \sum_{i=1}^{N} \ln\Phi_{iT}(\boldsymbol{\mu}; \mathbf{\Omega})$$

where $\boldsymbol{\mu} = (k_{i1} z'_{i1}\pi, \ldots, k_{iT} x'_{iT}\beta)$ and $\mathbf{\Omega} = (k_{i1}k_{i2}\Omega_{2,1}, \ldots, k_{iT-1}k_{iT}\Omega_{T,T-1})$. For deriving the likelihood, multivariate normal probability functions of order T are required. In Stata, only the bivariate normal distribution function exists. Using Halton draws that are derived by the command `mdraws`, simulated multivariate normal probabilities are generated by the `mvnp() egen` function. The total number of generated Halton draws is R, and with each draw $r \in (1, \ldots, R)$, multivariate normal probabilities are simulated, and the average of these simulations is derived. Hence, the logarithm of the simulated likelihood is

$$\ln SL = \frac{1}{R} \sum_{r=1}^{R} \sum_{i=1}^{N} \ln\Phi_{iT}^{r}(\boldsymbol{\mu}; \mathbf{\Omega})$$

In the first illustration, the implementation of the Stata routine[9] for a simulated multivariate random-effects probit model is presented with the help of simulated data. The advantage of simulated data is that the accuracy of the estimator can be derived. In the second step, a robustness check is applied, and the performance of the estimator

7. Note that the variance–covariance matrix is symmetric, so only the lower triangular part is shown.
8. To simplify notation, the lagged dependent variable is incorporated into x'_{it}.
9. The `mvnp() egen` function and the `mdraws` command must be installed.

is compared with the performance of the `redpace` command because both estimators are based on simulated likelihoods. In the second illustration, the Stata routine for simulated multivariate random-effects probit models with autocorrelated errors is presented.

3 Illustrations

3.1 Illustration 1: Simulated multivariate random-effects probit model for unbalanced panels

An artificial dataset containing 1,000 individuals is created. Each individual is identified by the variable `id`. A panel containing five periods is constructed by expanding the existing dataset by the value of five. The variable `tper` identifies the period for each individual, running from $t = 1, \ldots, 5$. With the command `drawnorm`, the time-invariant error term (`alpha`), explanatory (`x1`, `x2`, `x3`) and instrumental variables (`Instrument`), the idiosyncratic shock (`u_i`), and a variable called `Random` are generated. The time-invariant error term has a normalization of $\sim N(0, 2)$, and all other variables are standard normal distributed, that is, $\sim N(0, 1)$. The variable `Random` is a temporary identifier that helps to construct an unbalanced panel with gaps. With the function `normal()`, the normal distributed numbers are transformed into numbers within the range 0 and 1. For each individual, the values of the variable `Random` and `alpha` for the time points $t = 2, \ldots, 5$ are replaced by the initial value.

```
. set obs 1000
obs was 0, now 1000
. set seed 987654321
. generate id=_n
. expand 5
(4000 observations created)
. by id, sort: generate tper=_n
. matrix m = (0,0,0,0,0,0,0)
. matrix sd = (sqrt(2),1,1,1,1,1,1)
. drawnorm alpha Instrument x1 x2 x3 u_i Random, n(5000) means(m) sds(sd)
> seed(987654321)
. replace Random=normal(Random)
(5000 real changes made)
. sort id tper
. by id: replace alpha=alpha[1]
(4000 real changes made)
. by id: replace Random=Random[1]
(4000 real changes made)
```

The latent variable y^* is constructed in the following manner:

$$y^*_{i1} = 0.7 + 0.35x_1 + 0.66x_2 + 0.25x_3 + 1.5x_{\text{Instrument}} + \theta\alpha_i + u_{i1}$$

Here $x_{\text{Instrument}}$ is an instrumental variable that will affect the outcome of the initial period but not the subsequent ones. For the initial period, it is assumed that θ takes the value 1. For the subsequent periods, $t = 2, \ldots, 5$, the following relationship is defined:

$$y^*_{it} = 0.35 + 0.46y_{t-1} + 0.25x_1 + 0.75x_2 + 0.55x_3 + \alpha_i + u_{it}$$

The observable variable y_{it} becomes 1 if $y^*_{it} > 0$ and 0 otherwise. In addition, the variable `ylag` is generated, which takes the value of the outcome variable of the previous period.

```
. sort id (tper)
. local theta=1
. by id: generate ystar=.35*x1 + .66*x2 + .25*x3 + 1.5*Instrument + .7 +
> `theta´*alpha + u_i if _n==1
(4000 missing values generated)
. by id: generate y=cond(ystar>0,1,0) if _n==1
(4000 missing values generated)
. sort id (tper)
. forvalues i=2/5 {
  2. by id: replace ystar =.25*x1 + .75*x2 + .55*x3 + .46*y[_n-1] + .35  +
> alpha + u_i if _n==`i´
  3. by id: replace y=cond(ystar>0,1,0) if _n==`i´
  4. }
  (output omitted)
. sort id (tper)
. by id: generate ylag=cond(_n>1,y[_n-1],.)
(1000 missing values generated)
```

The fifth time point of those observations is dropped if the value of the variable `Random` exceeds 0.85, and the fourth and the fifth time point are dropped if `Random` is below 0.10. To construct a gap within the time sequence of the observation, the routine drops the third time point when the value of the variable `Random` is between 0.25 and 0.30.[10] The variable `nwave` contains information about the number of waves of which an individual is part and ranges between 3 and 5.

```
. drop if tper==5 & Random>.85
(151 observations deleted)
. drop if tper>=4 & Random<.10
(160 observations deleted)
. drop if tper==3 & Random>.25 & Random<=.30
(56 observations deleted)
. by id (tper), sort: generate nwave=_N
```

In Stata, only the cumulative bivariate normal distribution is implemented. To calculate multivariate normal probabilities, one applies the approach of Cappellari and

10. All thresholds are picked arbitrarily.

Jenkins (2006a). Multivariate normal probabilities are simulated by either Halton quasi-random or pseudorandom sequences using the GHK simulator. For multivariate normal distributions of order T, `mvnp()` returns the joint cumulative distribution of multivariate normal probabilities. In this simulation, 100 Halton quasi-random draws are generated.[11] The number of periods determines the dimensions of integration. Though an unbalanced sample is applied, five dimensions are generated for each observation. A prefix is used to identify the generated random numbers for the simulation of the multivariate normal probabilities later. The random numbers are based on prime numbers in the following manner (for details of the method, see Cappellari and Jenkins [2006a, 2003, 2005, 2006b] and Train [2009]):

```
. matrix p=(2,3,5,7,11)
. global dr = 100
. global T_max=5
. global T_min=3
. mdraws, neq(5) draws($dr) prefix(z) primes(p) burn(15)
Created 100 Halton draws per equation for 5 dimensions. Number of initial
draws dropped per dimension = 15 . Primes used:
   2   3   5   7  11
```

The number of random draws is saved in the macro `($dr)`. Additionally, 2 globals are generated, with `T_max` referring to the maximum number of periods, here `5`, and `T_min` referring to the minimum number of periods, here `3`.[12] Both globals are needed later in the Stata syntax.

The implementation of the Stata code is based on an extension of the approach of Cappellari and Jenkins (2006a).[13]

11. Halton quasi-random numbers instead of pseudorandom numbers are applied because of greater accuracy; see Cappellari and Jenkins (2006a) and Train (2009).
12. The group with several periods of 4 also contains those observations with a gap in their time sequence.
13. A helpful introduction into maximum likelihood estimation can be found in Gould, Pitblado, and Poi (2010).

```
program define mpheckman_d0
        args todo b lnf
        tempname sigma theta
        tempvar beta pi lnsigma lntheta T fi fi6 fi5 fi4 fi3 FF
        mleval `beta´     = `b´, eq(1)
        mleval `pi´       = `b´, eq(2)
        mleval `lnsigma´  = `b´, eq(3)  scalar
        mleval `lntheta´  = `b´, eq(4)  scalar

        scalar `sigma´=(exp(`lnsigma´))^2
        scalar `theta´=exp(`lntheta´)

        qui: {
        by idcode: generate double `T´ = (_n == _N)
        sort idcode (year)
        tempvar k1 zb1
        by idcode: generate double `k1´ = (2*$ML_y1[1]) - 1
        by idcode: generate double `zb1´ = `pi´[1]
        forvalues r = 2/$T_max {
                tempvar k`r´ xb`r´
                by idcode: generate double `k`r´´ = (2*$ML_y1[`r´]) - 1
                by idcode: generate double `xb`r´´ = `beta´[`r´]
        }

        forvalues s=$T_min/$T_max {
                tempname V`s´ C`s´
        }

        mat `V$T_max´=I($T_max)*(`sigma´+1)
        mat `V$T_max´[1,1]=(`theta´^2)*`sigma´+1

        forvalues row=2/$T_max {
                mat `V$T_max´[`row´,1] = (`theta´*`sigma´)
                mat `V$T_max´[1,`row´] = `V$T_max´[`row´,1]
                local s = `row´-1
                forvalues col=2/`s´ {
                        mat `V$T_max´[`row´,`col´] = `sigma´
                        mat `V$T_max´[`col´,`row´] = `V$T_max´[`row´,`col´]
                }
        }

        forvalues r = $T_min/$T_max {
                mat `V`r´´ = `V$T_max´[1..`r´,1..`r´]
                mat `C`r´´ = cholesky(`V`r´´)
        }
        egen double `fi5´ = mvnp(`zb1´ `xb2´ `xb3´ `xb4´ `xb5´) if nwave==5, /*
        */ chol(`C5´) dr($dr) prefix(z) signs(`k1´ `k2´ `k3´ `k4´ `k5´) adoonly
        egen double `fi4´ = mvnp(`zb1´ `xb2´ `xb3´ `xb4´) if nwave==4, /*
        */ chol(`C4´) dr($dr) prefix(z) signs(`k1´ `k2´ `k3´ `k4´) adoonly
        egen double `fi3´ = mvnp(`zb1´ `xb2´ `xb3´) if nwave==3, /*
        */ chol(`C3´) dr($dr) prefix(z) signs(`k1´ `k2´ `k3´)  adoonly

        generate double `fi´=cond(nwave==5,`fi5´,cond(nwave==4,`fi4´,`fi3´))
        generate double `FF´ = cond(!`T´,0,ln(`fi´))
        }
        mlsum `lnf´ = `FF´ if `T´
        if (`todo´==0 | `lnf´>=.) exit

end
```

In the first part of the maximum likelihood model, the utilized parameters are specified, where `beta` refers to the explanatory variables for the time period $t = 2, \ldots, T$, and `pi` refers to the explanatory variables of the initial period. The parameters used for the specification of the variance–covariance matrix, θ and σ_α, are included in logarithmic form. Therefore, the exponential function is applied; hence, the second expression is taken in addition to the square.

Time-specific sign variables $(k_1, \ldots, k_5)$ and explanatory variables $(zb_1, xb_2, \ldots, xb_5)$ are defined afterward. The explanatory variables referring to the initial period are labeled as zb_1 and the subsequent ones as xb_t with $t \in (2, \ldots, 5)$.

Thereafter, the variance–covariance matrix, with respect to the full time periods $\mathbf{V}_5$, is defined.[14] The elements of the main diagonal are $\sigma_\alpha^2 + 1$ for the time period $t \geq 2$; they are $\theta^2\sigma_\alpha^2 + 1$ for the initial period $t = 1$. In the following loop, the remaining elements of the matrix are defined. Those covariances correlated with the initial period are defined as $\theta\sigma_\alpha^2$; the remaining covariances are defined as σ_α^2. In the next loop, those observations with $T < 5$ submatrices from the variance–covariance matrix $\mathbf{V}_5$ are extracted. Then the Cholesky decompositions of the variance–covariance matrix, $\mathbf{C}_3 \ldots \mathbf{C}_5$, are derived by the matrix function `cholesky()`.

The Cholesky decomposed variance–covariance matrices are needed when multivariate normal probabilities are calculated with the `mvnp() egen` function. The Cholesky decomposition of the variance–covariance matrix is $\mathbf{\Omega} = \mathbf{V} \times \mathbf{V}'$. When `mvnp()` is applied, the explanatory variables, the Cholesky decomposition of the variance–covariance matrix (`chol()`), and the sign variables (`signs()`) must be inserted. The link to the generated random numbers is `prefix(z)`, the number of draws used by `dr($dr)`. For lower computational time, the calculation of the probabilities is restricted to those observations with the corresponding number of time periods [constraint `nwave==`T, where $T \in (3, \ldots, 5)$]. Finally, the logarithm of the calculated probabilities is summed up over each observation.

Because of the application of the two macros, `T_max` and `T_min`, the described syntax can easily be extended to datasets with different settings, for example, with longer periods. However, the calculation of the probabilities and the sum over each observation must still be adjusted accordingly. Generalization of this procedure is an open research task. An important advantage of this procedure is that the variance–covariance matrix can be easily adjusted, for example, by taking autocorrelated errors into account (see *Illustration 2*).

14. It is the same procedure as in the `redpace` command.

Estimation time can be reduced by starting the estimation with feasible initial values. In this example, we use the parameters of a probit model that does not account for random effects:

```
. qui: probit y ylag x1 x2 x3 if tper>1
. matrix b0=e(b)
. qui: probit y x1 x2 x3 Instrument if tper==1
. matrix b1=e(b)
. matrix b12 = (-.5,-.5)
. matrix b0 = (b0 , b1 , b12)
```

Note that the instrumental variable for the initial period is `Instrument`, which is not included in the subsequent periods as an explanatory variable. Furthermore, the lagged dependent variable `ylag` is not on the right hand of the initial-period equation system. Consequently, this variable contains missing values for the initial period. Hence, each observation has one variable with missing values. To stop Stata from eliminating these observations with missing values, we insert the option `missing` at the end of the `ml model`.

```
. ml model d0 mpheckman_d0 (y: y = ylag x1 x2 x3) (Init_Period: y = x1 x2 x3 In
> strument) /lnsigma /lntheta, title(Multivariate RE Probit, $dr Halton draws)
> missing
. ml init b0, copy
. ml max
  (output omitted)
Multivariate RE Probit, 100 Halton draws          Number of obs   =       4633
                                                  Wald chi2(4)    =     452.75
Log likelihood = -2078.5332                       Prob > chi2     =     0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
y						
ylag	.4676766	.0820586	5.70	0.000	.3068447	.6285085
x1	.3017219	.0360704	8.36	0.000	.2310251	.3724187
x2	.7494927	.0433351	17.30	0.000	.6645574	.834428
x3	.5721053	.0394884	14.49	0.000	.4947094	.6495012
_cons	.3166202	.0826484	3.83	0.000	.1546324	.4786081
Init_Period						
x1	.3691082	.0719819	5.13	0.000	.2280263	.5101901
x2	.6834951	.0836506	8.17	0.000	.5195428	.8474473
x3	.34158	.0726513	4.70	0.000	.199186	.4839739
Instrument	1.495445	.1384893	10.80	0.000	1.224011	1.766879
_cons	.6932924	.0923201	7.51	0.000	.5123483	.8742365
lnsigma						
_cons	.3614243	.0688798	5.25	0.000	.2264223	.4964262
lntheta						
_cons	-.0729095	.1373428	-0.53	0.596	-.3420965	.1962774

Looking at the output, we can see that the values of the estimated coefficients for the explanatory variables are close to the true values. The null hypothesis—it states the estimated coefficient is equal to the true value—is not rejected in any case. To derive σ_α^2 and θ, we must transform both variables. When we apply the command `_diparm`, the first derivative of the function must be obtained.

```
. _diparm lnsigma, function((exp(@))^2) deriv(2*(exp(@))*(exp(@)))
> label("Sigma^2") prob
     Sigma^2 |   2.060294   .2838253     7.26   0.000      1.57278    2.698922
. _diparm lntheta, function(exp(@)) deriv(exp(@)) label("Theta") prob
       Theta |   .9296849   .1276855     7.28   0.000     .7102797    1.216864
```

As can be seen, the estimated variance of α_i is close to the true value. The results also indicate that α_i is correlated with the initial conditions because $\theta = 0$ is strongly rejected. Hence, the initial conditions cannot be taken as exogenous. Also θ is very close to 1 (the true value), which implies that the impact of α_i is not significantly different from the impact in the subsequent periods $t \geq 2$.

The simulation of multivariate normal probabilities is based on Halton quasi-random numbers. The robustness of the estimation technique is checked by applying different sets of primes to generate Halton draws. Five primes in the range between $2, \ldots, 97$ are picked randomly, following a suggestion of Stewart (2006b) constructing random seeds:[15]

```
. forvalues r=1/10 {
  2. primes 100, clear
  3. set seed 987654321
  4. generate long s=int((runiform()+10-_n)*10000000)
  5. global d=s[`r´]
  6. set seed $d
  7. generate Random_`r´=runiform()
  8. sort Random_`r´
  9. qui: keep if _n<=5
 10. sort prime
 11. matrix p_`r´=(prime[1],prime[2],prime[3],prime[4],prime[5])
 12. }
```

15. The `primes` command (Kolenikov 2005) must be installed.

Altogether, 10 estimations are run; they use a different set of primes to simulate multivariate normal probabilities. The results can be found in table 1.

Table 1. Multivariate random-effects probit model (different sets of primes)

		Multivariate random-effects probit† Halton quasi-random numbers‡				
		(1)	(2)	(3)	(4)	(5)
Dynamic sequence ($t > 1$)						
y_{t-1}	0.460	0.465	0.466	0.470	0.465	0.476
		(0.082)	(0.082)	(0.082)	(0.082)	(0.082)
x_1	0.250	0.302	0.301	0.301	0.302	0.299
		(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
x_2	0.750	0.750	0.749	0.750	0.747	0.747
		(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
x_3	0.550	0.574	0.573	0.572	0.573	0.569
		(0.040)	(0.040)	(0.040)	(0.039)	(0.039)
constant	0.350	0.319	0.318	0.316	0.323	0.309
		(0.083)	(0.083)	(0.083)	(0.083)	(0.082)
Initial period ($t = 1$)						
x_1	0.350	0.369	0.370	0.369	0.377	0.370
		(0.072)	(0.072)	(0.072)	(0.073)	(0.072)
x_2	0.660	0.680	0.686	0.685	0.694	0.686
		(0.083)	(0.084)	(0.084)	(0.085)	(0.084)
x_3	0.250	0.341	0.343	0.341	0.349	0.340
		(0.072)	(0.073)	(0.073)	(0.074)	(0.073)
$x_{\text{Instrument}}$	1.500	1.490	1.497	1.495	1.520	1.497
		(0.137)	(0.138)	(0.140)	(0.142)	(0.141)
constant	0.700	0.691	0.694	0.693	0.707	0.695
		(0.092)	(0.092)	(0.093)	(0.094)	(0.093)
σ_α^2	2.000	2.067	2.069	2.052	2.061	2.021
		(0.284)	(0.286)	(0.283)	(0.284)	(0.277)
θ	1.000	0.924	0.930	0.931	0.951	0.941
		(0.125)	(0.127)	(0.130)	(0.130)	(0.131)
lnSL	–	−2077.769	−2078.334	−2078.834	−2076.100	−2079.220
Observations	–	4,633	4,633	4,633	4,633	4,633
Prime numbers	–	7,13, 29,31,67	2,7, 23,31,83	2,23 41,79,83	17,19, 43,59,79	3,37, 43,47,83

Continued on next page

		Multivariate random-effects probit† Halton quasi-random numbers‡				
		(6)	(7)	(8)	(9)	(10)
Dynamic sequence ($t > 1$)						
y_{t-1}	0.460	0.467	0.469	0.470	0.466	0.462
		(0.082)	(0.082)	(0.082)	(0.082)	(0.082)
x_1	0.250	0.302	0.300	0.300	0.302	0.304
		(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
x_2	0.750	0.751	0.747	0.748	0.750	0.753
		(0.043)	(0.043)	(0.043)	(0.043)	(0.044)
x_3	0.550	0.574	0.570	0.570	0.573	0.575
		(0.040)	(0.039)	(0.039)	(0.040)	(0.040)
constant	0.350	0.319	0.315	0.313	0.319	0.327
		(0.083)	(0.083)	(0.082)	(0.083)	(0.083)
Initial period ($t = 1$)						
x_1	0.350	0.370	0.369	0.372	0.371	0.370
		(0.072)	(0.072)	(0.073)	(0.072)	(0.072)
x_2	0.660	0.683	0.681	0.689	0.685	0.679
		(0.083)	(0.083)	(0.085)	(0.084)	(0.083)
x_3	0.250	0.341	0.341	0.344	0.342	0.341
		(0.072)	(0.072)	(0.073)	(0.073)	(0.072)
$x_{\text{Instrument}}$	1.500	1.494	1.489	1.507	1.500	1.488
		(0.137)	(0.137)	(0.141)	(0.139)	(0.135)
constant	0.700	0.693	0.690	0.699	0.695	0.691
		(0.092)	(0.092)	(0.093)	(0.093)	(0.091)
σ_α^2	2.000	2.072	2.050	2.039	2.064	2.108
		(0.284)	(0.283)	(0.281)	(0.285)	(0.290)
θ	1.000	0.924	0.925	0.946	0.932	0.912
		(0.126)	(0.127)	(0.131)	(0.129)	(0.123)
lnSL	–	−2077.593	−2080.571	−2079.387	−2077.690	−2076.939
Observations	–	4,633	4,633	4,633	4,633	4,633
Prime numbers	–	3,17, 31,59,97	19,23, 53,67,71	2,3, 19,73,79	13,23, 31,43,59	5,59, 61,89,97

† Standard error in brackets. ‡ For each estimation, 100 Halton quasi-random numbers are applied.

The estimation results of the multivariate random-effects probit model are now compared with those of the `redpace` command, which was written by Stewart (2006b) and uses Halton quasi-random numbers and additional pseudorandom numbers. Because the latter command cannot be applied when the time sequence of the observation contains gaps, a dataset identical to the one above is generated but without dropping time periods. Thus the new dataset contains 5,000 observations. The latent variables are constructed analogously to those in the previous section.

Table 2. Comparison of random-effects models[†]

Coefficients	True values	Multivariate random-effects probit Halton quasi-random numbers			Random-effects probit (`redpace`) Pseudorandom numbers[‡]			Halton quasi-random numbers		
Dynamic sequence ($t > 1$)										
y_{t-1}	0.460	0.493	0.491	0.491	0.514	0.498	0.495	0.489	0.492	0.489
		(0.077)	(0.078)	(0.077)	(0.077)	(0.077)	(0.078)	(0.078)	(0.078)	(0.078)
x_1	0.250	0.290	0.290	0.290	0.283	0.288	0.291	0.291	0.290	0.291
		(0.034)	(0.034)	(0.034)	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
x_2	0.750	0.745	0.749	0.750	0.735	0.746	0.749	0.749	0.748	0.750
		(0.040)	(0.041)	(0.041)	(0.040)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
x_3	0.550	0.577	0.580	0.581	0.571	0.580	0.581	0.580	0.580	0.581
		(0.037)	(0.038)	(0.038)	(0.037)	(0.037)	(0.038)	(0.038)	(0.038)	(0.038)
constant	0.350	0.304	0.307	0.306	0.280	0.300	0.294	0.306	0.306	0.308
		(0.079)	(0.079)	(0.079)	(0.077)	(0.079)	(0.079)	(0.079)	(0.079)	(0.079)
Initial period ($t = 1$)										
x_1	0.350	0.365	0.373	0.370	0.358	0.367	0.369	0.367	0.372	0.371
		(0.071)	(0.073)	(0.072)	(0.070)	(0.072)	(0.071)	(0.071)	(0.072)	(0.072)
x_2	0.660	0.687	0.701	0.695	0.674	0.693	0.694	0.691	0.696	0.696
		(0.083)	(0.085)	(0.084)	(0.081)	(0.084)	(0.083)	(0.083)	(0.085)	(0.084)
x_3	0.250	0.335	0.341	0.339	0.327	0.339	0.336	0.337	0.338	0.339
		(0.072)	(0.073)	(0.072)	(0.070)	(0.072)	(0.072)	(0.072)	(0.072)	(0.072)
$x_{\text{Instrument}}$	1.500	1.501	1.537	1.522	1.493	1.533	1.525	1.505	1.522	1.523
		(0.138)	(0.143)	(0.140)	(0.133)	(0.140)	(0.137)	(.137)	(0.141)	(0.140)
constant	0.700	0.691	0.708	0.700	0.676	0.701	0.697	0.691	0.701	0.701
		(0.092)	(0.094)	(0.093)	(0.089)	(0.093)	(0.091)	(0.091)	(0.093)	(0.093)
σ_α^2	2.000	2.004	2.051	2.062	–	–	–	–	–	–
		(0.257)	(0.266)	(0.267)						
λ	0.667	–	–	–	0.658	0.673	0.673	0.675	0.672	0.673
					(0.029)	(0.029)	(0.028)	(0.028)	(0.029)	(0.028)
θ	1.000	0.938	0.963	0.945	0.943	0.954	0.951	0.928	0.950	0.948
		(0.126)	(0.129)	(0.126)	(0.123)	(0.126)	(0.123)	(0.123)	(0.128)	(0.126)
$\ln SL$	–	−2218.83	−2216.17	−2216.08	−2223.77	−2218.42	−2217.28	−2216.08	−2217.14	−2216.02
Observations	–	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
Numbers	–	20	50	100	20	50	100	20	50	100

[†] Standard error in brackets. [‡] Seed is 81234567, the default.

To compare both techniques, we run the estimations on the basis of 20, 50, and 100 draws. In the multivariate random-effects model, Halton quasi-random numbers are applied; in the `redpace` command, additional pseudorandom numbers are used. Estimation results can be found in table 2.

The command `redpace` does not report the coefficient σ_α^2 but the proportion of the time-invariant variance on the composite error term

$$\lambda = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2}$$

As can be seen in table 2, when 100 draws are applied, all estimators derive similar coefficients that are close to the true values. This is independent of the type of random numbers used. Also, in all estimations, none of the null hypotheses—the estimated value of a coefficient is equal to the true value—is rejected. Furthermore, the log likelihoods are close to each other, especially when Halton quasi-random numbers are applied.

In addition, table 2 shows that if the estimation is run with 50 Halton draws, the difference in the log likelihood compared with the estimation with 100 Halton draws is slightly smaller in the multivariate random-effects probit case. The higher accuracy of the multivariate probit estimator becomes apparent when comparing the log likelihood of the estimations run with 20 and 100 Halton draws ($\ln SL = -2218.829$) for the multivariate probit model compared with the estimations based on pseudorandom numbers ($\ln SL = -2223.770$). These results indicate some better estimation fit of the simulated multivariate random-effects probit model compared with the `redpace` command, in which the simulation is based on pseudorandom numbers.

3.2 Illustration 2: Extending to autocorrelated errors

An advantage of applying multivariate probit models is that the link between the time points in the dynamic model can be directly adjusted by specifying the variance–covariance matrix. As an extension to the previous model, it is now assumed that the idiosyncratic shock is autocorrelated so that it follows a first-order autoregressive process:[16]

$$u_{it} = \delta u_{it-1} + \epsilon_{it}$$

16. Another extension might regard a moving-average process or an autocorrelated error process of higher order.

The generalized variance–covariance matrix takes on the following form:

$$\mathbf{\Omega} = \begin{pmatrix} \theta^2\sigma_\alpha^2+1 & & & & & \\ \theta\sigma_\alpha^2+\delta & \sigma_\alpha^2+1 & & & & \\ \theta\sigma_\alpha^2+\delta^2 & \sigma_\alpha^2+\delta & \sigma_\alpha^2+1 & & & \\ \theta\sigma_\alpha^2+\delta^3 & \sigma_\alpha^2+\delta^2 & \sigma_\alpha^2+\delta & \sigma_\alpha^2+1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \theta\sigma_\alpha^2+\delta^{T-1} & \sigma_\alpha^2+\delta^{T-2} & \sigma_\alpha^2+\delta^{T-3} & \sigma_\alpha^2+\delta^{T-4} & \dots & \sigma_\alpha^2+1 \end{pmatrix}$$

The implementation of a multivariate random-effects probit model with autocorrelated errors is illustrated by following the empirical approach of Stewart (2006b). U.S. data from the National Longitudinal Survey of Youth concerning young women are used to investigate the state dependence of union membership.[17] Following Stewart (2006b), those observations from 1978 onward, excluding 1983, are used. Additional variables that identify the time period (`tper`) and the number of time periods an observation is in the sample (`nwave`) are generated. As an excluding restriction, those observations that are sample members for the complete sequence are kept. Also the lagged dependent variable—`Lunion`—is generated.

```
. webuse union
(NLS Women 14-24 in 1968)
. drop if year<78
(10136 observations deleted)
. drop if year==83
(2194 observations deleted)
. bys idcode (year): gen nwave=_N
. bys idcode (year): gen tper=_n
. keep if nwave==6
(9076 observations deleted)
. by idcode (year), sort: gen Lunion = union[_n-1]
(799 missing values generated)
```

The Stata routine of *Illustration 1* must be extended by introducing the parameter ρ, which refers to the autocorrelated error term. This parameter will be integrated into the Stata syntax as the inverse hyperbolic tangent of ρ. The variance–covariance matrix must be adjusted according to the adjusted $\mathbf{\Omega}$. Because a balanced panel is applied, the multivariate normal distribution function of order $T = 6$ must be simulated; hence, the probabilities are calculated only one time.

17. A description of the data and of the variables can be found in Stewart (2006b).

```
cap prog drop mpheckman_d0
program define mpheckman_d0
        args todo b lnf
        tempname sigma theta rho
        tempvar beta pi lnsigma lntheta trho T fi FF V C
        mleval `beta´       = `b´, eq(1)
        mleval `pi´         = `b´, eq(2)
        mleval `lnsigma´    = `b´, eq(3) scalar
        mleval `lntheta´    = `b´, eq(4) scalar
        mleval `trho´       = `b´, eq(5) scalar

        scalar `sigma´=(exp(`lnsigma´))^2
        scalar `theta´=exp(`lntheta´)
        scalar `rho´  =tanh(`trho´)

        qui: {
        sort idcode (year)
        tempvar k1 zb1
        by idcode: generate double `k1´ = (2*$ML_y1[1]) - 1
        by idcode: generate double `zb1´ = `pi´[1]
        forvalues r = 2/6 {
                tempvar k`r´ xb`r´
                by idcode: generate double `k`r´´ = (2*$ML_y1[`r´]) - 1
                by idcode: generate double `xb`r´´ = `beta´[`r´]
        }

        by idcode: generate double `T´ = (_n == _N)

        mat `V´=I(6)*(`sigma´+1)
        mat `V´[1,1]=(`theta´^2)*`sigma´+1

        forvalues row=2/6 {
                mat `V´[`row´,1] = (`theta´*`sigma´ + `rho´^(`row´-1))
                mat `V´[1,`row´] = `V´[`row´,1]
                local r1 = `row´-1
                forvalues col=2/`r1´ {
                        mat `V´[`row´,`col´] = `sigma´ + `rho´^(`row´-`col´)
                        mat `V´[`col´,`row´] = `V´[`row´,`col´]
                }
        }

        mat `C´ = cholesky(`V´)

        egen `fi´ = mvnp(`zb1´ `xb2´ `xb3´ `xb4´ `xb5´ `xb6´), chol(`C´) /*
        */ dr($dr) prefix(z) signs(`k1´ `k2´ `k3´ `k4´ `k5´ `k6´) adoonly

        generate double `FF´ = cond(!`T´,0,ln(`fi´))
        }
        mlsum `lnf´ = `FF´ if `T´
        if (`todo´==0 | `lnf´>=.) exit

end
```

Estimation results are compared with those of the `redpace` command. Stewart (2006b) uses a GHK simulator to fit a random-effects probit model with autocorrelated errors. Hereby, the probability of a sequence is calculated as the product of recursively defined conditional probabilities (Stewart 2006b). In the multivariate probit model, the probabilities are calculated directly, but the multivariate cumulative distributions must be simulated. Note that the `redpace` command does not report σ_α^2 but instead λ, which shows the attribution of the individual specific error-term on the composite error term.

To reduce computational time, we derive initial values by fitting two probit models, the first one referring to the dynamic sequence ($t \geq 2$) and the second one to the initial period. Note that the variable `not_smsa` is integrated as an instrumental variable for the initial period.

```
. qui: probit union Lunion age grade south if tper>1
. matrix b0=e(b)
. qui: probit union age grade south not_smsa if tper==1
. matrix b1=e(b)
. matrix b12 = (-.5,-.5,-.2)
. matrix b0 = (b0 , b1 , b12)
```

For the simulation of the multivariate normal probabilities, Halton pseudorandom numbers must be generated. In the beginning, 20 Halton draws are chosen and successively increased. Then the multivariate random-effects probit model with autocorrelated errors is fit. The option `search(off)` is applied so that observations are not dropped from the estimation sample.

```
. matrix p=(2,3,5,7,11,13)
. mdraws, neq(6) draws(20) prefix(z) primes(p) burn(15)
Created 20 Halton draws per equation for 6 dimensions. Number of initial
draws dropped per dimension = 15 . Primes used:
   2   3   5   7  11  13
. global dr = r(n_draws)
. ml model d0 mpheckman_d0 (`y1´: `y1´ = `a´) (Init_Period: `y1´ = `b´)
> /lnsigma /lntheta /trho, title(Multivariate AR1 Probit, $dr Halton draws)
> missing
. ml init b0, copy
. ml max, search(off)
  (output omitted)
```

The estimation results of the multivariate random-effects probit model with autocorrelated errors can be found in table 3. Estimations are run on the basis of 20, 50, and 100 Halton draws.[18] In the case of the `redpace` command, simulations based on 20, 50, 100, and 500 pseudorandom numbers are applied and reported in the last four columns of table 3. The `redpace` command has substantial problems when Halton draws are applied instead of pseudorandom numbers. Estimations based on 20, 50, or 100 Halton draws did not converge and interrupted after several iterations.

18. The parameter ρ is derived by `_diparm trho, tanh prob`.

Table 3. Random effects with autocorrelated errors[†]

Coefficients	Multivariate random-effects probit			Random-effects probit (`redpace`)			
	Halton quasi-random numbers			Pseudorandom numbers[‡]			
Dynamic sequence ($t > 1$)							
`Lunion`	1.319	1.344	1.325	1.376	1.341	1.297	1.322
	(0.152)	(0.147)	(0.154)	(0.147)	(0.154)	(0.161)	(0.154)
`age`	−0.023	−0.023	−0.023	−0.021	−0.023	−0.024	−0.023
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
`grade`	−0.038	−0.036	−0.037	−0.033	−0.035	−0.037	−0.036
	(0.020)	(0.020)	(0.020)	(0.019)	(0.019)	(0.020)	(0.020)
`south`	−0.382	−0.372	−0.374	−0.341	−0.360	−0.381	−0.370
	(0.099)	(0.097)	(0.099)	(0.092)	(0.097)	(0.101)	(0.099)
`constant`	0.068	0.057	0.076	−0.035	0.044	0.101	0.080
	(0.401)	(0.399)	(0.400)	(0.380)	(0.394)	(0.406)	(0.400)
Initial period ($t = 1$)							
`age`	0.009	0.009	0.010	0.013	0.011	0.011	0.011
	(0.025)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
`grade`	−0.014	−0.014	−0.013	−0.012	−0.014	−0.012	−0.013
	(0.034)	(0.033)	(0.034)	(0.033)	(0.033)	(0.033)	(0.033)
`south`	−0.768	−0.767	−0.760	−0.731	−0.749	−0.760	−0.755
	(0.171)	(0.169)	(0.168)	(0.165)	(0.166)	(0.167)	(0.167)
`not_smsa`	−0.414	−0.415	−0.418	−0.418	−0.417	−0.414	−0.420
	(0.166)	(0.167)	(0.167)	(0.166)	(0.166)	(0.166)	(0.166)
`constant`	−0.843	−0.851	−0.866	−0.968	−0.887	−0.931	−0.891
	(0.864)	(0.857)	(0.853)	(0.845)	(0.846)	(0.849)	(0.848)
σ_α^2	1.100	1.057	1.080	–	–	–	–
	(0.306)	(0.290)	(0.307)				
λ	–	–	–	0.479	0.502	0.526	0.519
				(0.070)	(0.072)	(0.072)	(0.071)
θ	1.265	1.265	1.239	1.296	1.251	1.211	1.227
	(0.219)	(0.217)	(0.216)	(0.224)	(0.222)	(0.213)	(0.214)
ρ	−0.329	−0.347	−0.338	−0.328	−0.329	−0.321	−0.338
	(0.057)	(0.055)	(0.058)	(0.053)	(0.057)	(0.061)	(0.058)
lnSL	−1856.328	−1854.106	−1854.579	−1859.836	−1858.151	−1855.833	−1854.062
Observations	4,794	4,794	4,794	4,794	4,794	4,794	4,794
Numbers	20	50	100	20	50	100	500

[†] Standard error in brackets.
[‡] Seed is 945430778, using starting values from Stewart (2006b) for the estimation with 500 draws.

As can be seen in table 3, the log likelihood of the multivariate random-effects probit model with autocorrelated errors changes only slightly when using 100 Halton quasi-random numbers instead of 50. In addition, table 3 shows that the idiosyncratic shocks are significantly negatively correlated. The findings go along with those of the `redpace` command, especially when 500 pseudorandom numbers are applied. A noticeable change in the log likelihood can be found with the `redpace` command when the numbers of the pseudorandom draws are increased from 20 or 50 to 100. Again, in the simulated multivariate random-effects probit model, accuracy can already be found at a low level of Halton draws, which helps save computational time. Also, when one applies

Halton draws, indications of greater stability can be found in the simulated multivariate random-effects probit model because convergence is reached at a low number of draws; in contrast, the `redpace` command has difficulties in converging.

4 Conclusion

The main feature of this estimation technique is that the time link in the dynamic model can be directly adjusted by specifying the variance–covariance matrix. In the first illustration, the Stata routine of a simulated multivariate random-effects probit model for unbalanced panels with gaps is presented. The robustness check shows that this estimation strategy produces similar results when compared with the existing `redpace` command. In the second illustration, the specification of the variance–covariance matrix is described to allow for autocorrelated errors. In sum, the advantages of the proposed estimator when compared with the `redpace` command are higher accuracy and improved estimation stability. Additionally, regarding the model with autocorrelated errors, computational time can be saved because a lower number of Halton draws is needed for the estimation.

5 Acknowledgments

I thank Editor Joseph Newton and an anonymous referee for helpful comments.

6 References

Arulampalam, W. 1999. A note on estimated coefficients in random effects probit models. *Oxford Bulletin of Economics and Statistics* 61: 597–602.

Cappellari, L., and S. P. Jenkins. 2003. Multivariate probit regression using simulated maximum likelihood. *Stata Journal* 3: 278–294.

———. 2005. Software update: st0045_1: Multivariate probit regression using simulated maximum likelihood. *Stata Journal* 5: 285.

———. 2006a. Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation. *Stata Journal* 6: 156–189.

———. 2006b. Software update: st0045_2: Multivariate probit regression using simulated maximum likelihood. *Stata Journal* 6: 284.

Chamberlain, G. 1984. Panel data. In *Handbook of Econometrics*, ed. Z. Griliches and M. D. Intriligator, vol. 2, 1247–1318. Amsterdam: Elsevier.

Drukker, D. M., and R. Gates. 2006. Generating Halton sequences using Mata. *Stata Journal* 6: 214–228.

Gould, W., J. Pitblado, and B. Poi. 2010. *Maximum Likelihood Estimation with Stata*. 4th ed. College Station, TX: Stata Press.

Haan, P., and A. Uhlendorff. 2006. Estimation of multinomial logit models with unobserved heterogeneity using maximum simulated likelihood. *Stata Journal* 6: 229–245.

Heckman, J. J. 1981a. Heterogeneity and state dependence. In *Studies in Labor Markets*, ed. S. Rosen, 91–140. Chicago: University of Chicago Press.

———. 1981b. The incidental parameters problem and the problem of initial conditions in estimating a discrete time—Discrete data stochastic process. In *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. F. Manski and D. McFadden, 179–195. Cambridge: MIT Press.

Hole, A. R. 2007. Fitting mixed logit models by using maximum simulated likelihood. *Stata Journal* 7: 388–401.

Keane, M. P. 1994. A computationally practical simulation estimator for panel data. *Econometrica* 62: 95–116.

Kolenikov, S. 2005. primes: Stata module to generate prime numbers. Statistical Software Components S456504, Department of Economics, Boston College. http://ideas.repec.org/c/boc/bocode/s456504.html.

Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica* 46: 69–85.

Stewart, M. B. 2006a. Heckman estimator of the random effects dynamic probit model. http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/stewart/stata.

———. 2006b. Maximum simulated likelihood estimation of random-effects dynamic probit models with autocorrelated errors. *Stata Journal* 6: 256–272.

Train, K. E. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge: Cambridge University Press.

Wooldridge, J. M. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20: 39–54.

About the author

Alexander Plum is a research assistant at the Chair of Public Economics at Otto von Guericke University Magdeburg. His main research interests are labor economics and applied econometrics.