Technical Working Paper

Your resource for urban reproductive health

MLE Measurement, Learning & Evaluation Project for the Urban Reproductive Health Initiative

A Guide to Longitudinal Program Impact Evaluation

Gustavo Angeles*

Christopher Cronin

David K. Guilkey

Peter M. Lance*

Brian Sullivan

* All authors are affiliated with the Measurement, Learning & Evaluation Project. In addition, these two authors are affiliated with the MEASURE Evaluation project at Carolina Population Center, University of North Carolina at Chapel Hill.

This guide is a collaborative effort of the Measurement, Learning & Evaluation (MLE) Project of the Urban Reproductive Health Initiative and the MEASURE Evaluation project, both of which are based in the Carolina Population Center at the University of North Carolina at Chapel Hill.

The MLE Project is the evaluation component of the Urban Reproductive Health Initiative, a multi-country program in India, Kenya, Nigeria and Senegal that aims to improve the health of the urban poor. The goal of the project is to promote evidence-based decision making in the design of integrated family planning and reproductive health interventions. MEASURE Evaluation provides technical leadership through collaboration at local, national and global levels to advance the field of global health monitoring and evaluation.

The opinions expressed in this guide are those of the authors and do not necessarily reflect the views of the MLE Project's donor organization, the Bill & Melinda Gates Foundation; or MEASURE Evaluation's donor agency, the U.S. Agency for International Development; or the U.S. government. Additional information about this guide may be obtained from:

Ilene S. Speizer, Technical Deputy Director, Principal Investigator
Measurement, Learning & Evaluation Project
Carolina Population Center, UNC Chapel Hill
CB#8120, University Square
123 West Franklin Street
Chapel Hill, NC 27516-2524
E-mail: speizer@email.unc.edu

Information about the Urban Reproductive Health Initiative and the MLE project may be obtained at:

www.urbanreproductivehealth.org.

Information about the MEASURE Evaluation project may be obtained at:

http://www.cpc.unc.edu

Recommended citation:

Angeles G, Cronin C, Guilkey DK, Lance PM, Sullivan B. *A Guide to Longitudinal Program Impact Evaluation*. Chapel Hill, NC: Measurement, Learning & Evaluation Project; 2014.

Table of Contents

List of Abbreviations

CEA	census enumeration area
CPR	contraceptive prevalence rate
EA	enumeration area
FD	first differences
FE	fixed effects
GLLAMM	Generalized Linear Latent And Mixed Models (software)
GLS	generalized least squares
GTZ	Deutsche Gesellschaft für Internationale Zusammena
IFLS	Indonesian Family Life Survey
MLSFH	Malawi Longitudinal Study of Families and Health
MDICP	Malawi Diffusion and Ideational Change Project
OLS	ordinary least squares
MLE Project	Measurement, Learning & Evaluation Project for the Urban Reproductive Health Initiative
TA	traditional authorities
TMCSM	Traditional Methods of Child Spacing in Malawi survey
Urban RH Initiative	Urban Reproductive Health Initiative

Introduction

The MEASURE Evaluation project and the Measurement, Learning & Evaluation (MLE) Project are both tasked with the evaluation of the impact of programs. In many cases, the data that inform these evaluations are longitudinal in nature. In other words, the evaluations exploit data that involve the repeated measurement of outcomes, behaviors and characteristics of individuals, families, or communities.

Consider, for instance, the MLE Project. It is tasked with evaluating the 2009 reproductive health strategy of the Bill & Melinda Gates Foundation as implemented through the Urban Reproductive Health Initiative (Urban RH Initiative), which seeks to increase contraceptive use in large cities in India, Kenya, Senegal and Nigeria with a particular focus on poor and vulnerable populations. Key evaluation questions for the MLE project include:

- Did the project achieve a significant increase in the contraceptive prevalence rate (CPR) in the selected urban areas?
- Which supply side interventions contributed to increased contraceptive use?
- Which demand and behavior change interventions contributed to increased contraceptive use?
- Which private-sector interventions contributed to increased contraceptive use?

To answer these questions, longitudinal surveys of women comprised of three waves (baseline, mid-term and endline) are in progress with baseline and mid-term data already gathered in all four countries and the endline underway in India. The sample sizes of these waves range from approximately 18,000 in India and Nigeria to between 5,000 and 10,000 in Kenya and Senegal.

The nature of longitudinal data, whereby units of observation are observed more than once over time, opens the door to powerful statistical tools for impact evaluation but also creates some complications about which the analyst needs to be aware. This manual is intended to serve as a practical guide for implementing longitudinal estimation procedures. It thus fills a key technical gap for both the MLE Project and MEASURE Evaluation project, and potentially for other organizations conducting impact evaluations with longitudinal data.

The guide first explains the advantages of gathering longitudinal data and then describes methods of analysis for use with longitudinal data, including assumptions required by each method. We demonstrate that the applications of these methods to longitudinal data can yield the statistically correct estimates of program impact required to answer the key evaluation questions such as those that we have listed for the MLE project. The methods discussion is supplemented with examples from the literature and with empirical examples demonstrated with Stata.

In section I, we present a brief overview of experimental methods and relate the two basic types of experimental design to the two basic types of survey data: cross sectional and longitudinal. In section II, we present longitudinal methods that are appropriate when the program interventions or “treatments” can be assumed to be randomly assigned as defined in that section. Section III presents methods that are appropriate when the assumption of random assignment is not correct. Section IV presents extensions of longitudinal methods to multilevel models and we conclude in section V.

I. Longitudinal Data and Experimental Design

To motivate the use of longitudinal data for program impact analysis, we discuss two common data designs employed in evaluation research. These provide a framework for considering the advantages of longitudinal data. To many, experimental methods represent a "gold standard" in evaluation research, albeit one that can be difficult to achieve in practice (due to other factors including program allocation priorities, ethical considerations, selective participation, and selective attrition). Experimental methods are straightforward: randomly assign individuals into treatment and comparison groups. Randomization insures that the treatment and comparison groups are alike on average in terms of their characteristics. Any differences in outcomes between the two groups after treatment can thus be ascribed to the treatment, since the experience of treatment is the only way that the treatment and comparison groups systematically differ. We refer to data drawn from populations subjected to such a randomization process as experimental data.

Such straightforward inference breaks down in the non-experimental case, whereby individuals are not randomly sorted into the treatment and comparison groups. Rather, they are systematically sorted into one of these two groups. Often, this reflects their individual choices. In other cases, program managers may drive the process as, for instance, they select certain types of villages or neighborhoods when they allocate a program. In either case, the result of such systematic selection into the treatment and comparison groups is that the individuals in the two typically systematically differ in terms of their characteristics. Thus, they no longer differ simply by the experience of exposure to the program: they also do so by other characteristics. This complication makes it impossible to ascribe simple differences in average outcomes between the two groups to the experience of treatment alone.

To begin with, we consider what we mean when we speak of estimating program impact or treatment effects (two terms often used interchangeably in program impact evaluation), and how we gauge whether it is a decent estimate. Our objective is to capture average program impact. Suppose, for instance, that Y_i^1 is the outcome that an individual experiences when he or she participates in a program and that Y_i^0 is the outcome that an individual experiences when he or she does not participate in that program. Consider the impact parameter

$$\alpha = E\left(Y_i^1 - Y_i^0\right)$$

where the expectation $E(\cdot)$ is taken across the population of interest (i.e., those for whom we want to learn something about program impact). This is often referred to as the average treatment effect or expected program impact, and is of central interest in this manual. That said, this is not the only potential impact parameter. Consider, for instance,

$$E\left(Y_i^1 - Y_i^0 | T_i = 1\right)$$

where T_i i is a binary indicator of program impact that equals 1 if an individual participates and 0 otherwise. This program impact parameter is often referred to as the average effect of treatment on the treated. It captures program impact for those who actually participated in the program.

If we could observe both outcomes for the members of a sample of N individuals randomly drawn from (and hence representative of) the population of interest, then we could estimate the average treatment effect in that sample by:

$$\frac{\sum_{i=1}^{N}\left(Y_i^1 - Y_i^0\right)}{N}$$

Unfortunately, we cannot observe <u>both</u> $Y_i^1$and Y_i^0 for any given individual since we cannot observe an individual both when he or she does, and does not, participate in a program.

Another important issue is how we evaluate the success of an estimator in terms of the statistical “appropriateness’’ or “correctness’’ of the estimates it produces. There are two main concepts of “statistical correctness’’:

1. **Unbiasedness:** An estimator is unbiased if its expectation equals the true population value of a parameter. Suppose that we have an estimator of program impact that produces estimates represented by $\hat{\alpha}$. $\hat{\alpha}$ is an unbiased estimate of the average treatment effect if

   $$E(\hat{\alpha}) = \alpha$$

   where α is the true value of the parameter for the population of interest. Suppose, for instance, that we drew many samples of the same size (*N*) and formed an estimate $\hat{\alpha}$ for each sample with this estimator of program impact. If the average of all of those estimates was α then the estimator is unbiased. An intuitive way of thinking about this is that an estimator is unbiased if it is right on average in terms of the population parameter it seeks to estimate. Note that unbiasedness is a property that does not depend on sample size;

2. **Consistency:** An estimator is consistent if the probability distribution for the estimates it might yield becomes concentrated on the true population parameter value as sample size becomes increasingly large. In other words, an estimator is consistent if as the sample size becomes large it becomes increasingly unlikely that the estimate it produces differs from the true population value. Consistency is a property that depends on a large sample size.

These concepts give us two distinct but quite useful ways of thinking about “statistical correctness”.

Having established this background, the first of the designs popular to impact evaluation that we consider is the **post-test only control group design**. This design considers outcomes among those exposed to the treatment and a comparison group not exposed to that treatment who serve as controls. However, outcomes are observed for these two groups only after the treatment has been implemented.

To give a concrete example, suppose that the treatment group hears a family planning message, while some comparison group does not do so. Generally speaking, we can think of some population sorted into treatment and comparison subpopulations. The resulting subpopulations are mutually exclusive (i.e., no individual belongs to both) and collectively exhaustive (i.e., all members of the population under consideration are in either the treatment or comparison subpopulations).

Define the variable Yi as follows:

$$Y_i = \begin{cases} 1, if\ individual\ i\ uses\ contraception \\ 0, otherwise \end{cases}$$

where *i* is some individual from the population under consideration or a sample of treatment and comparison individuals taken from that population. For concreteness, we assume in the discussion below that *Y* indicates contraceptive use. The program or treatment we wish to evaluate is hearing or recall of a family planning message on contraceptive use.

If, following the application of treatment to the treatment subpopulation, we observe *Yi* for random samples of size N_1 (from the treatment subpopulation) and N_2 (from the comparison subpopulation), one estimator that might potentially be used for determining the impact of hearing the family planning message on contraceptive use would be:

$$\hat{\alpha} = \frac{1}{N_1}\sum_{i=1}^{N_1} Y_i - \frac{1}{N_2}\sum_{j=1}^{N_2} Y_j \qquad (1)$$

This estimates program impact by a simple difference of mean outcomes between the treatment and comparison samples after the application of treatment to the former. This is an estimator since it is based on random samples from larger populations.

Let us first consider this estimator when assignment of individuals from the population of interest to the treatment and comparison subpopulations is random (i.e., experimental data). In this case, we would expect that the process of randomly sorting individuals to the two subpopulations would insure that the average characteristics of the individuals in the two subpopulations would be the same on average: all of the personal attributes, environmental circumstances and other characteristics that might influence contraceptive use (i.e., the outcome of interest) would, on average, have essentially the same distribution in the treatment and comparison subpopulations. In that case, the two subpopulations would differ only by the experience of receiving the family planning message (i.e., the treatment) and any post-treatment differences in contraceptive use could be ascribed to that message.

Thus, with experimental data, the post-test only control group design can yield an unbiased estimate of treatment impact simply by comparing the mean outcomes experienced between the treatment and comparison groups. However, this typically is not the case with the non-experimental setting, for which you have non-random sorting into the treatment and comparison subpopulations. When treatment status is non-randomly determined, those who are treated and those who are not typically differ in terms of their average characteristics. This means that any differences in mean outcomes between samples from the two subpopulations might reflect not only the experience of treatment, but differences in their characteristics as well.

The post-test only control group design can also lend itself to regression modeling. Consider the model:

$$Y_i = \delta + \alpha \cdot T_i + \varepsilon_i \qquad (2)$$

where

$$T_i = \begin{cases} 1, if\ individual\ i\ is\ exposed\ to\ the\ treatment \\ 0, otherwise. \end{cases}$$

The error term, $\varepsilon_{i,}$, represents other factors that influence contraceptive use. One can think of the error term in this regression model as containing, along with a completely random element, all of the other factors that shape the outcome Y_i. We call these background characteristics. They might contain the attributes of the individual (such as age or education), their environmental circumstances, or perhaps even exposure to other unrelated programs that might influence contraceptive use.

As we have seen, randomization leads to the same types of individuals (in terms of background characteristics) among the treatment and comparison groups. Put differently, it insures that treatment status T_i is uncorrelated with any background characteristics that might be in ε_i . This implies that:

$$E(T_i\varepsilon_i) = corr(T_i, \varepsilon_i) = 0.$$

As it turns out, this is the key condition for insuring that straightforward regression estimation procedures will yield an unbiased estimate of α. The reason for this is intuitively straightforward. When T_i is uncorrelated with the background characteristics relegated to ε_i, it can serve in the regression process as a control only for exposure to treatment, which is what it should do. It is not forced to serve a double role: to control for exposure to treatment and as a proxy for any background characteristic relegated to $\varepsilon_{i.}$ Regression estimation would presumably be done with samples (for instance, concatenated subsamples from the treatment and comparison subpopulations), yielding an estimate $\hat{\alpha}$ of the program impact parameter α.

One key drawback of the post-test only comparison design with experimental data is that it does not permit one to test whether randomization of the population into the treatment and control groups was indeed successful. This is not a small consideration: in practice, successful randomization can be difficult to achieve since people are quite purposeful and thus can often defy their assignment. It should not be taken for granted.

Before moving on to the case of non-random assignment, we consider the possibility of adding additional controls X_i for background characteristics to the model in the face of experimental data:

$$Y_i = \delta + \alpha \cdot T_i + \beta \cdot X_i + \varepsilon_i \qquad (3)$$

The inclusion of these covariates (age, education and socioeconomic status, for example) reduces the variance represented by the error term and increases the precision of the estimate of program impact. To be sure, the estimate of α generated with experimental data is statistically correct (i.e., unbiased and consistent) with or without the extra controls X_i. However the inclusion of the additional controls should serve to improve the precision of the estimate of α.

When treatment status is not randomly determined (i.e., when we have non-experimental data), individuals are selected for or, in many instances, select themselves (i.e., self-select) into treatment status. This selection process tends to be systematic and guided by the background characteristics of those individuals. As we have seen, the result is that the distribution of these background characteristics differs between the treatment and comparison populations. For instance, the treatment population might be older, less educated, or poorer than the comparison population (or, alternatively, younger, more educated and richer).

This implies that background characteristics are correlated with treatment status. However, these background characteristics are contained in the error term of the basic regression model (model [2]), which implies that:

$$E(T_i \varepsilon_i) = corr(T_i, \varepsilon_i) \neq 0$$

When correlation between the two exists, T_i must serve not only as a control for itself but also as a proxy for background characteristics that influence both treatment status T_i (whether one hears the family planning message) and Y_i (contraception) but are relegated to the unobservable ε_i. Thus, the key condition for unbiased estimation of program impact (essentially, that T_i serves empirically as a control only for the treatment status of individual *i*) no longer holds.

One might hope that adding controls for background characteristics X_i as regressors, as done in model (3), would solve the problem. The idea behind this strategy would be that by moving characteristics out of the error term and adding them as regressors we might again render treatment status uncorrelated with the new error term, thus insuring unbiased estimation of the treatment effect. However, this strategy will only work if we can add as a regressor every characteristic correlated with both treatment status and the outcome of interest. Unfortunately, it seems unlikely that we will be able to do so in practice. Our main limitation is that we likely do not have measures for all characteristics influencing treatment status and the outcome. Most likely, several of those characteristics will remain unobserved by the evaluator. Adding regressors might, or might not, reduce bias to the estimates of program impact, but it will not eliminate it.

We now turn to the other commonly employed data design for impact evaluation studies, the ***pre- and post-test control group design***. This design expands on the post-test only comparison design by the inclusion of samples from the treatment and comparison populations observed before treatment occurs, as well as after it has happened.

We begin by considering the implications of this design for the experimental data case. First, we introduce a subscript to key variables that indexes time. For instance, re-define Y_i as Y_{ti} where the *t* subscript captures time *t*, which takes on two values: *t*=1, which corresponds to the pre-test period (before treatment has occurred) and *t*=2 (after it has occurred). With this extension, we can define Y_{ti} and X_{ti} as the outcome and characteristics, respectively, of individual *i* at time *t*. Suppose as well that representative samples of individuals for whom T_i, Y_{ti} and X_{ti} are observed at times *t*=1 and t=2 are selected from the treatment and comparison populations.

Under this data design, program impact can be measured via the difference-in-differences estimator. To begin with, define

$$T\hat{D} = \frac{1}{N_1}\sum_{i=1}^{N_1} Y_{2i} - \frac{1}{N_1}\sum_{i=1}^{N_1} Y_{1i} \tag{4}$$

TD measures the change in average contraceptive use for the treatment group between the pre- and post-treatment observation intervals. Further, define

$$C\hat{D} = \frac{1}{N_2}\sum_{j=1}^{N_2} Y_{2j} - \frac{1}{N_2}\sum_{j=1}^{N_2} Y_{1j} \tag{5}$$

This measures the change in average contraceptive use for the control group between the pre- and post-treatment observational intervals. Differences in differences estimates program impact as follows:

$$\hat{\alpha} = \widehat{TD} - \widehat{CD} \tag{6}$$

The difference-in-differences estimator hence estimates program impact essentially as the difference in time trends in average outcomes (contraceptive prevalence, in this discussion) between the treatment and control populations (as reflected by representative estimates from each).

There is a regression analogue to the program impact estimator in equation (6):

$$Y_{ti} = \delta + T_{ti}\alpha + \mu_i + \varepsilon_{ti} \tag{7}$$

where μ_i represents unmeasured fixed characteristics of individuals and ε_{it} represents unmeasured time-varying characteristics and *t* denotes pre (*t*=1) and post (*t*=2) time periods. In this regression framework, for time period 1, T_{1i}=0 for all individuals and for time period 2, T_{2i}=1 for an individual in the treatment group and T_{2i}=0 for individuals in the control group. This specification essentially assumes that there is no trend to *Y* common to the treatment and non-treatment populations (a trivial extension can capture this).

Now suppose that even though individuals are randomly assigned to the treatment and control groups, a higher proportion of individuals motivated to limit family size end up in the treatment group (i.e., an individual's level of motivation is correlated with receiving the treatment). If motivation is an unobserved fixed characteristic (i.e., it is part of μ_i), then:

$$Y_{2i} - Y_{1i} = (T_{2i} - T_{1i})\alpha + \varepsilon_{2i} - \varepsilon_{1i}$$

or

$$Y_{2i} - Y_{1i} = T_{2i}\alpha + \varepsilon_{2i} - \varepsilon_{1i}$$

since T_1=0 for all individuals prior to treatment.

This differenced equation makes clear that the source of possible bias (sometimes referred to assignment bias) is eliminated and ordinary least squares can again be used to measure program impact. This estimate is, once again, identical to the "differences-in-differences" estimate presented above.

As with the post-test only experimental design, one can add covariates to the regression model:

$$Y_{it} = \delta + T_{it}\alpha + X_{it}\beta + \mu_i + \varepsilon_{it}$$

Just like the model without covariates, we can examine change in the outcome to eliminate as a potential source of bias:

$$Y_{i2} - Y_{i1} = (T_{i2} - T_{i1})\alpha + (X_{i2} - X_{i1})\beta + \varepsilon_{i2} - \varepsilon_{i1}$$

As with the post-test only design, the inclusion of the *X* variables enhances precision to the estimate of the treatment effect but does not affect bias because of random assignment in the experimental design and the additional control for assignment bias that is introduced by differencing. Note that *X*'s that do not vary through time (variables such as race, for example) would be dropped from the model.

The above discussion makes clear that the pre-test/post-test design represents a significant improvement over the post-test only design in terms of reducing possible sources of bias. One can think of the post-test only design as using cross sectional data since we only observed the treatment and control groups a single time. The pre-test/post-test design is longitudinal since we observe each individual twice (pre and post treatment) and then take advantage of these multiple measures to develop a better estimate of program impact. In the sections below, we will show how longitudinal data in a non-experimental setting still maintains many of the bias reducing properties that we see above. In fact, this is the reason that the well-known Princeton University economist Angus Deaton made the following statement (Deaton, 1997):

> *When our data contain repeated observations on each individual, the resulting panel data open up a number of possibilities that are not available in the single cross section. In particular, the opportunity to compare the same individual under different circumstances permits the possibility of using that individual as his or her own control, so that we can come closer to the ideal experimental situation.*

II. Longitudinal Methods for Models with Exogenous Explanatory Variables

As we have seen in the previous section, with experimental data the treatment and control groups are randomly assigned. As a result, both observed and unobserved characteristics of the individuals in the sample are uncorrelated with whether or not they actually receive the treatment. In non-experimental settings, this is almost never the case. Expanding on the example introduced in the previous section, suppose that the treatment is whether or not an individual heard a family planning message on the radio. This is, for instance, most likely correlated with the individual's observed socioeconomic status and whether or not he or she owns a radio. In this case, the inclusion of control variables is essential if one wants to obtain an unbiased estimate of the treatment effect. The treatment may also be correlated with unobserved characteristics of the individuals. For example, respondents with a high level of motivation to limit family size may be more likely to recall a family planning message than those who are less motivated.

In this section, we restrict our attention to the case in which the treatment is only correlated with observed variables that are measured in the survey data set and not with unobserved variables. This same assumption is made for all the observed variables as well and so all explanatory variables are exogenous. We will explore the case in which both the treatment and the control variables are correlated with unobservables in section III.

Section II is divided into methods that are appropriate for continuous, binary and categorical outcome variables.

A. *Continuous Dependent Variable*

The basic form for the statistical model is essentially the same as the one presented above for the pre-test/post-test experimental design with control variables except that we do not include a time specific dummy variable, since its inclusion would not alter any of the estimation methods but it would add to the notational complexity:

$$Y_{ti} = X_{ti}\beta + T_{ti}\alpha + Z_i\delta + \mu_i + \varepsilon_{ti}. \qquad (8)$$

where

- Y_{ti} is a continuous outcome for individual i at time t;
- X_{ti} are time-varying explanatory variables (asset ownership and possibly education, for example);
- T_{ti} is the time-varying program participation variable or treatment variable;
- Z_i are time-invariant regressors (race and sex, for example);
- μ_i is a time-invariant source of unobserved heterogeneity; and
- ε_{ti} represents time-varying unobserved heterogeneity.

Finally, i indexes the N individuals in the sample (i.e., the sample contains i=1,…,N individuals) and there are t=1,2,…,M_i observations for the i^{th} individual.

While this model has a similarity to the basic regression model presented during the discussion of the pre-test/post-test comparison design in the preceding section, there are important differences. For present purposes, one important consideration is the longitudinal nature of the data. In the earlier discussion we focused on the case of two observations (one "pre" and one "post") for each individual. Thus we had a circumstance of a two-period balanced panel: every individual was observed twice. The current framework allows far greater flexibility in that it affords the possibility of multiple time periods per individual (beyond two) and that different individuals might have different numbers of observations M_i (i.e., an unbalanced panel). Another important distinction is that we now introduce more explicitly complex control variables. In particular we now make a clear distinction between permanent or fixed and time-varying variables.

Just as in the pre-test/post-test model, the longitudinal model allows for two types of unobservables; μ_i captures unobserved individual level characteristics that influence the outcome and that do not vary through time. In the context of our example, it might be an individual's level of motivation to limit family size. The term ε_{ti} captures time-varying factors which could include for example, contacts with individuals in the community that might have influence on the respondent's decision about the outcome.

The key statistical assumptions are as follows:

1. $E(X_{ti}\mu_i) = E(T_{ti}\mu_i) = E(Z_i\mu_i) = 0$ (there is no correlation between observed variables and the time invariant error);
2. $E(X_{ti}\varepsilon_{ti}) = E(T_{ti}\varepsilon_{ti}) = E(Z_i\varepsilon_{ti}) = 0$ (there is no correlation between observed variables and the time-varying error);
3. $E(\mu_i) = E(\varepsilon_{ti}) = E(\varepsilon_{ti}\mu_i) = 0$ (the error terms have mean zero and are not correlated with each other);
4. $Var(\mu_i) = \sigma_\mu^2$ for all individuals, and $Var(\varepsilon_{ti}) = \sigma_\varepsilon^2$ for all observations.

The fact that the unobserved μ_i is constant across all M_i observations for individual i means that even after controlling for correlation due to observed characteristics, the set of observations on individual i are correlated with each other. As long as assumptions 1 and 2 hold, this dependence means that standard multivariate regression methods will yield correct point estimates of all the estimated coefficients including the coefficient that captures treatment impact (α), but estimates of their standard errors will be biased downwards and the level of significance of the results will thus be overstated. That said, there are methods that can improve the efficiency of the estimates by explicitly accounting for the dependence of the error terms for each individual.

Given the assumptions, the following estimation methods will yield statistically correct estimates of the treatment effect and all methods can be implemented in Stata data analysis software:

1. **Ordinary Least Squares (OLS):** OLS is an unbiased and consistent estimator. The OLS standard errors are incorrect due to the correlation of observations for the same individual, but correct standard errors can easily be obtained (for instance, via the "cluster" option in Stata).
2. **Feasible generalized least squares (GLS):** The feasible GLS or random effects estimator ("xtreg" in Stata with the "re" option) is consistent and is more asymptotically efficient than OLS since it explicitly takes the dependence of observations from the same individual into account. It is also asymptotically equivalent to the maximum likelihood estimator if one makes the additional assumption that both components of the error term are normally distributed.
3. **Maximum Likelihood Estimator:** The maximum likelihood estimator must make a specific distributional assumption about the error terms. The most common assumption is that both error components (time-varying and time-invariant errors) are normally distributed and this is the assumption used by Stata ("xtreg" in Stata with the "mle" option). If the distributional assumption is correct, maximum likelihood is the asymptotically efficient estimator for all of the model's parameters and is thus the optimal estimator. An advantage of maximum likelihood is that it provides standard errors for σ_μ^2 and σ_ε^2. Thus, all the

information that is needed for a simple and direct test of the null hypothesis that $\sigma_{\mu}^{2} = 0$ is available from the standard Stata output from the “xtreg” command with “mle” option. This test is of obvious interest since OLS standard errors would be correct if the null hypothesis is true.

Stata Examples

Appendix A provides details on the two longitudinal data sets from Indonesia and Malawi that are used throughout this guide to provide empirical examples. Both data sets have four waves of data. In the family planning literature, it is difficult to find continuous dependent variables and so we use ideal family size as a less than perfect example of a continuous variable. For both data sets, we use relatively simple models with a few basic socioeconomic variables, community variables and policy related variables and we apply each of the estimation methods discussed above. We present the Stata command and the associated output along with a brief discussion of the results.

In example 1, the policy variable for Malawi, `fpmess`, is whether or not women heard a family planning message from an agent, clinic, or the radio. This variable has a negative effect on ideal family size. The other individual control variables behave as expected: women with higher levels of education desire fewer children. Employment increases the desired number of children, as do Catholic, Protestant and traditional religious affiliations. Whether or not the household is Muslim has little effect in this model.

Example 1: Ordinary Least Squares without Corrected Standard Errors

Malawi

```
. regress ideal age $education $work $religion $fptype

      Source |       SS       df       MS              Number of obs =    4238
-------------+------------------------------           F(  9,  4228) =   41.30
       Model |  2097.60002     9  233.066669           Prob > F      =  0.0000
    Residual |  23857.3967  4228  5.64271445           R-squared     =  0.0808
-------------+------------------------------           Adj R-squared =  0.0789
       Total |  25954.9967  4237  6.12579577           Root MSE      =  2.3754

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0559602   .0032689    17.12   0.000     .0495514     .062369
        edu1 |  -.1638498   .0884946    -1.85   0.064    -.3373458    .0096461
        edu2 |  -.6939161   .1668292    -4.16   0.000    -1.020989   -.3668432
    employed |   .3297308   .0929446     3.55   0.000     .1475105     .511951
    catholic |   .3421634   .1249518     2.74   0.006     .0971922    .5871346
  protestant |   .2852989   .1099207     2.60   0.009     .0697965    .5008013
      muslim |  -.0502296   .1263129    -0.40   0.691    -.2978693    .1974101
 traditional |   .3910266   .1331555     2.94   0.003     .1299718    .6520813
      fpmess |  -.2495788   .0857384    -2.91   0.004    -.4176712   -.0814865
       _cons |   2.382247   .1981519    12.02   0.000     1.993765    2.770728
------------------------------------------------------------------------------
```

The two policy variables that are included in example 2 are the number of posyandus (a type of health facility) and whether or not a midwife is present in the community (these are represented, respectively, by the variables `num_pos_0` and `midwife_0` in the output). We see that both variables have the intended negative effect on ideal family size. Furthermore, the individual level control variables typically behave as one would expect: women with higher levels of education and who can write desire fewer children while older women desire more. Husband's education has little effect in this model as does whether or not the household is Muslim. We do get the rather surprising result that urban residence has a positive effect but this is after controlling for other community variables such as proximity to the capital and availability of public transport. Note that some women responded "up to God" to the ideal family size question and these women were dropped in these results and the results that follow.

Example 2: Ordinary Least Squares without Corrected Standard Errors

Indonesia

```
. reg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0

      Source |       SS       df       MS              Number of obs =   19389
-------------+------------------------------           F( 23, 19365) =   59.37
       Model |  2888.80985    23  125.600428           Prob > F      =  0.0000
    Residual |  40969.6996 19365  2.11565709           R-squared     =  0.0659
-------------+------------------------------           Adj R-squared =  0.0648
       Total |  43858.5095 19388  2.26214718           Root MSE      =  1.4545

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |  -.0055111   .0023317    -2.36   0.018    -.0100815   -.0009407
   midwife_0 |  -.1970529    .023036    -8.55   0.000    -.2422054   -.1519005
        edu1 |   .0760984   .0465203     1.64   0.102    -.0150854    .1672821
        edu2 |  -.2197894   .0624338    -3.52   0.000     -.342165   -.0974138
        edu3 |  -.3869874   .0878118    -4.41   0.000     -.559106   -.2148688
       hedu1 |   .0487344   .0543071     0.90   0.370    -.0577123    .1551811
       hedu2 |   .0537594   .0670156     0.80   0.422    -.0775971    .1851158
       hedu3 |   .1223556   .0875321     1.40   0.162    -.0492148     .293926
         age |   .0295686   .0012787    23.12   0.000     .0270622     .032075
    work_any |  -.0232795   .0424905    -0.55   0.584    -.1065645    .0600054
   work_most |   -.077604   .0425403    -1.82   0.068    -.1609866    .0057787
      smoker |  -.1203682    .074411    -1.62   0.106    -.2662202    .0254837
    goodhlth |  -.0975583   .0331038    -2.95   0.003    -.1624446    -.032672
      muslim |   .0060752   .0337928     0.18   0.857    -.0601616     .072312
    electric |  -.2809094   .0347688    -8.08   0.000    -.3490593   -.2127595
        read |  -.0516365   .0692704    -0.75   0.456    -.1874125    .0841395
       write |  -.2062215   .0643881    -3.20   0.001    -.3324277   -.0800153
       urban |   .0973887   .0268457     3.63   0.000     .0447688    .1500086
    pubtrans |   .1252388   .0249363     5.02   0.000     .0763615    .1741161
      caphrs |   .0184083   .0038876     4.74   0.000     .0107883    .0260282
     fhead50 |   -.031062   .0305249    -1.02   0.309    -.0908935    .0287695
     fhead75 |  -.1302276   .0330637    -3.94   0.000    -.1950353   -.0654198
    fhead100 |  -.1309725   .0426041    -3.07   0.002    -.2144803   -.0474647
       _cons |   2.567558   .0909989    28.22   0.000     2.389192    2.745923
------------------------------------------------------------------------------
```

Example 3: Ordinary Least Squares with Corrected Standard Errors

Malawi

```
. regress ideal age $education $work $religion $fptype, cluster(respondentid)

Linear regression                                      Number of obs =    4238
                                                       F(  9,  2298) =   30.85
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0808
                                                       Root MSE      =  2.3754

                        (Std. Err. adjusted for 2299 clusters in respondentid)
------------------------------------------------------------------------------
             |               Robust
   ideal_num |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0559602   .0039803    14.06   0.000     .0481549    .0637655
        edu1 |  -.1638498   .0982007    -1.67   0.095    -.3564212    .0287215
        edu2 |  -.6939161   .1228516    -5.65   0.000    -.9348277   -.4530045
    employed |   .3297308   .0990388     3.33   0.001     .1355159    .5239456
    catholic |   .3421634   .1233548     2.77   0.006      .100265    .5840617
  protestant |   .2852989   .1037026     2.75   0.006     .0819384    .4886594
      muslim |  -.0502296   .1317247    -0.38   0.703    -.3085413    .2080821
 traditional |   .3910266   .1241582     3.15   0.002     .1475528    .6345004
      fpmess |  -.2495788   .0881682    -2.83   0.005    -.4224764   -.0766813
       _cons |   2.382247   .2045721    11.65   0.000     1.981081    2.783412
------------------------------------------------------------------------------
```

The only difference in the Stata command from simple OLS is the addition of the cluster option (`, cluster(respondentid)`) with the individual identification variable `respondentid` that corrects the standard errors for the fact that we have up to four observations on each individual. Note that the point estimates of the coefficients are the same as OLS but the standard errors are typically larger. This is the expected result. First, the point estimates remain the same since the cluster option is a post-estimation modification (meaning that it influences standard error estimation, but only after point estimation is complete). The standard error estimates generally grow larger because the correlation of errors at the individual level left the misimpression in the first regression (which ignored possible correlation of errors) of more independent variation in Y between observations than was actually the case.

Example 4: Ordinary Least Squares with Corrected Standard Errors

Indonesia

```
. reg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, cluster(ind_id)

Linear regression                                      Number of obs =   19389
                                                       F( 23,  9176) =   32.60
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0659
                                                       Root MSE      =  1.4545

                              (Std. Err. adjusted for 9177 clusters in ind_id)
------------------------------------------------------------------------------
             |               Robust
       ideal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |  -.0055111   .0027054    -2.04   0.042    -.0108143   -.0002079
   midwife_0 |  -.1970529   .0249921    -7.88   0.000    -.2460429   -.1480629
        edu1 |   .0760984   .0697681     1.09   0.275    -.0606627    .2128595
        edu2 |  -.2197894   .0806431    -2.73   0.006    -.3778677   -.0617111
        edu3 |  -.3869874   .0933107    -4.15   0.000     -.569897   -.2040777
       hedu1 |   .0487344   .0844446     0.58   0.564    -.1167958    .2142647
       hedu2 |   .0537594   .0922057     0.58   0.560    -.1269842     .234503
       hedu3 |   .1223556    .101718     1.20   0.229    -.0770343    .3217455
         age |   .0295686   .0017081    17.31   0.000     .0262203    .0329168
    work_any |  -.0232795    .042198    -0.55   0.581     -.105997    .0594379
   work_most |   -.077604   .0422594    -1.84   0.066    -.1604418    .0052338
      smoker |  -.1203682   .0938639    -1.28   0.200    -.3043623    .0636259
     goodhlth |  -.0975583   .0371644    -2.63   0.009    -.1704088   -.0247078
      muslim |   .0060752   .0469293     0.13   0.897    -.0859167    .0980671
    electric |  -.2809094    .046467    -6.05   0.000    -.3719951   -.1898238
        read |  -.0516365   .0852799    -0.61   0.545    -.2188041    .1155311
       write |  -.2062215   .0763106    -2.70   0.007    -.3558072   -.0566358
       urban |   .0973887   .0360714     2.70   0.007     .0266808    .1680966
    pubtrans |   .1252388   .0288692     4.34   0.000     .0686487    .1818289
      caphrs |   .0184083   .0055172     3.34   0.001     .0075933    .0292233
     fhead50 |   -.031062   .0398814    -0.78   0.436    -.1092385    .0471145
     fhead75 |  -.1302276   .0421483    -3.09   0.002    -.2128477   -.0476075
    fhead100 |  -.1309725   .0542255    -2.42   0.016    -.2372666   -.0246784
       _cons |   2.567558    .126679    20.27   0.000     2.319239    2.815877
------------------------------------------------------------------------------
```

The only difference in the Stata command from simple OLS is the addition of the cluster option (`, cluster(ind_id)`) with the individual identification variable (`ind_id`) that corrects the standard errors for the fact that we have up to four observations on each individual. Note that the point estimates of the coefficients are the same as OLS but the standard errors are typically larger. Once again, this is the expected result.

Example 5: Generalized Least Squares

Malawi

```
. xtset respondentid year
       panel variable:  respondentid (unbalanced)
        time variable:  year, 1998 to 2006, but with gaps
                delta:  1 unit

. xtreg ideal age $education $work $religion $fptype, re

Random-effects GLS regression                   Number of obs      =      4238
Group variable: respondentid                    Number of groups   =      2299

R-sq:  within  = 0.0339                         Obs per group: min =         1
       between = 0.1110                                        avg =       1.8
       overall = 0.0808                                        max =         3

                                                Wald chi2(9)       =    362.99
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0559889   .0033012    16.96   0.000     .0495187     .062459
        edu1 |  -.1561887   .0898491    -1.74   0.082    -.3322897    .0199123
        edu2 |  -.6829545   .1690722    -4.04   0.000     -1.01433   -.3515791
    employed |   .3221119   .0927377     3.47   0.001     .1403494    .5038744
    catholic |   .3431308   .1259217     2.72   0.006     .0963288    .5899328
  protestant |   .2951811    .109523     2.70   0.007     .0805199    .5098423
      muslim |  -.0319364    .127479    -0.25   0.802    -.2817907    .2179179
 traditional |   .3960529   .1321707     3.00   0.003     .1370031    .6551026
      fpmess |  -.2599821   .0852319    -3.05   0.002    -.4270336   -.0929306
       _cons |   2.379147   .1993811    11.93   0.000     1.988368    2.769927
-------------+----------------------------------------------------------------
     sigma_u |  .50455906
     sigma_e |  2.2562679
         rho |  .04762664   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

The GLS estimator is asymptotically efficient relative to OLS and we see that, in this model, the family planning message variable remains significant. The standard errors for this model are close to the standard errors of the OLS model. In addition to the estimated coefficients, the bottom of the output displays the estimated standard deviations of the two error components

$$(\sigma_\mu \text{ and } \sigma_\varepsilon)$$

along with the fraction of the total error variance due to the fixed component of the error (rho).

We see that in this case

$$\sigma_\mu^2$$

accounts for a smaller portion of the variance than

$$\sigma_\varepsilon^2 .$$

Although the output does not present standard errors for these two estimated parameters, the MLE estimator presented below does so that significance tests can be performed easily.

Example 6: Generalized Least Squares

Indonesia

```
. xtset ind_id year
       panel variable:  ind_id (unbalanced)
        time variable:  year, 1 to 4, but with gaps
                delta:  1 unit

. xtreg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, re

Random-effects GLS regression                   Number of obs      =     19389
Group variable: ind_id                          Number of groups   =      9177

R-sq:  within  = 0.0036                         Obs per group: min =         1
       between = 0.0983                                        avg =       2.1
       overall = 0.0638                                        max =         4

                                                Wald chi2(23)      =    894.87
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |  -.0030393   .0024176    -1.26   0.209    -.0077778    .0016991
   midwife_0 |  -.2061709   .0202944   -10.16   0.000    -.2459471   -.1663947
        edu1 |   .0640566   .0436742     1.47   0.142    -.0215431    .1496564
        edu2 |  -.2498483   .0593777    -4.21   0.000    -.3662264   -.1334701
        edu3 |  -.3532148   .0837347    -4.22   0.000    -.5173319   -.1890977
       hedu1 |  -.0546946   .0511274    -1.07   0.285    -.1549026    .0455133
       hedu2 |  -.0453687   .0616465    -0.74   0.462    -.1661936    .0754562
       hedu3 |  -.0084377   .0793606    -0.11   0.915    -.1639817    .1471063
         age |   .0257827   .0013142    19.62   0.000     .0232069    .0283586
    work_any |   .0571344   .0354044     1.61   0.107     -.012257    .1265257
   work_most |  -.0940141   .0354472    -2.65   0.008    -.1634894   -.0245388
      smoker |  -.1040005   .0725307    -1.43   0.152     -.246158    .0381571
    goodhlth |   -.015107   .0278207    -0.54   0.587    -.0696346    .0394205
      muslim |  -.0648428   .0413517    -1.57   0.117    -.1458907    .0162051
    electric |  -.2809922   .0313469    -8.96   0.000     -.342431   -.2195534
        read |   -.033661   .0579943    -0.58   0.562    -.1473278    .0800058
       write |  -.1300231   .0529805    -2.45   0.014    -.2338629   -.0261833
       urban |   .0644273   .0304641     2.11   0.034     .0047188    .1241358
    pubtrans |   .0650889   .0223274     2.92   0.004     .0213281    .1088497
      caphrs |   .0133892   .0041273     3.24   0.001     .0052998    .0214786
     fhead50 |   -.052519   .0302413    -1.74   0.082    -.1117908    .0067528
     fhead75 |  -.1154232   .0344954    -3.35   0.001    -.1830329   -.0478135
    fhead100 |  -.1338502   .0432379    -3.10   0.002    -.2185948   -.0491055
       _cons |   2.718067   .0907418    29.95   0.000     2.540216    2.895917
-------------+----------------------------------------------------------------
     sigma_u |  1.0057005
     sigma_e |  .97075394
         rho |   .5176759   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

The effect of midwives is stronger while most of the control variables have coefficients and standard errors that are close to the OLS with corrected standard error results.

Example 7: Maximum Likelihood

Malawi

```
. xtreg ideal age $education $work $religion $fptype, mle


Random-effects ML regression                    Number of obs      =      4238
Group variable: respondentid                    Number of groups   =      2299

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       1.8
                                                               max =         3

                                                LR chi2(9)         =    331.74
Log likelihood  =   -9650.05                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0559808   .0033595    16.66   0.000     .0493964    .0625652
        edu1 |  -.1444928   .0923053    -1.57   0.117    -.3254078    .0364221
        edu2 |  -.6641885   .1733724    -3.83   0.000    -1.003992   -.3243848
    employed |   .3092389   .0923408     3.35   0.001     .1282542    .4902236
    catholic |   .3439155   .1277197     2.69   0.007     .0935894    .5942415
  protestant |   .3117987   .1088624     2.86   0.004     .0984324     .525165
      muslim |  -.0028391   .1298911    -0.02   0.983     -.257421    .2517428
 traditional |   .4035109   .1304711     3.09   0.002     .1477922    .6592296
      fpmess |   -.279926   .0845275    -3.31   0.001    -.4455969   -.1142551
       _cons |   2.379004     .20166    11.80   0.000     1.983758     2.77425
-------------+----------------------------------------------------------------
    /sigma_u |    .871373   .0657721                      .7515443    1.010308
    /sigma_e |   2.205339   .0323947                      2.142752    2.269754
         rho |   .1350376   .0197593                      .1000571    .1775993
------------------------------------------------------------------------------
Likelihood-ratio test of sigma_u=0: chibar2(01)=   50.10 Prob>=chibar2 = 0.000
```

We see little difference in the estimated coefficients and standard errors for GLS and MLE in this case, perhaps indicating that the assumption that the error terms are normally distributed, as is the case for MLE, is reasonable for these data. The advantage we have with the MLE estimator is that we get standard errors for the error components and a test of the null hypothesis that

$$\sigma_{\mu} = 0 .$$

The result of the test provides strong evidence of importance of unobserved time invariant variables for this model.

Example 8: Maximum Likelihood Indonesia

```
. xtreg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, mle

Random-effects ML regression                    Number of obs      =     19389
Group variable: ind_id                          Number of groups   =      9177

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       2.1
                                                               max =         4

                                                LR chi2(23)        =    818.42
Log likelihood  = -32347.994                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |  -.0028755   .0024272    -1.18   0.236    -.0076326    .0018817
   midwife_0 |  -.2058716   .0201376   -10.22   0.000    -.2453406   -.1664026
        edu1 |   .0625607   .0434589     1.44   0.150    -.0226173    .1477386
        edu2 |  -.2524512   .0592593    -4.26   0.000    -.3685974   -.1363051
        edu3 |  -.3509845   .0836618    -4.20   0.000    -.5149587   -.1870103
       hedu1 |  -.0603185   .0508507    -1.19   0.236    -.1599839     .039347
       hedu2 |  -.0521487   .0612583    -0.85   0.395    -.1722127    .0679153
       hedu3 |  -.0174879    .078834    -0.22   0.824    -.1719997    .1370239
         age |   .0253115   .0013253    19.10   0.000      .022714     .027909
    work_any |   .0622693   .0350504     1.78   0.076    -.0064283    .1309669
   work_most |  -.0955572   .0350676    -2.72   0.006    -.1642885    -.026826
      smoker |  -.1041001   .0724169    -1.44   0.151    -.2460346    .0378344
    goodhlth |  -.0103777   .0275482    -0.38   0.706    -.0643712    .0436159
      muslim |  -.0694443   .0422091    -1.65   0.100    -.1521726     .013284
    electric |  -.2788177   .0311277    -8.96   0.000    -.3398269   -.2178086
        read |  -.0332986   .0573227    -0.58   0.561    -.1456491    .0790518
       write |  -.1243028   .0523342    -2.38   0.018    -.2268759   -.0217297
       urban |   .0620617   .0308448     2.01   0.044     .0016069    .1225164
    pubtrans |   .0612281   .0221803     2.76   0.006     .0177555    .1047007
      caphrs |    .013002   .0041536     3.13   0.002      .004861    .0211429
     fhead50 |  -.0540001   .0302326    -1.79   0.074     -.113255    .0052548
     fhead75 |  -.1146147   .0346289    -3.31   0.001    -.1824862   -.0467433
    fhead100 |  -.1318138    .043356    -3.04   0.002      -.21679   -.0468375
       _cons |   2.734052   .0910106    30.04   0.000     2.555674    2.912429
-------------+----------------------------------------------------------------
    /sigma_u |   1.066418   .0114837                      1.044146    1.089165
    /sigma_e |   .9638885   .0065525                      .9511309    .9768171
         rho |    .550371   .0069366                      .5367487    .5639343
------------------------------------------------------------------------------
Likelihood-ratio test of sigma_u=0: chibar2(01)= 4833.04 Prob>=chibar2 = 0.000
```

We again see little difference in the estimated coefficients and standard errors for GLS and MLE in this case, perhaps indicating that the assumption that the error terms are normally distributed, as is the case for MLE, is not harmful for this data in the sense of representing a serious misspecification error. Once again the test of the null hypothesis that

$$\sigma_{\mu} = 0$$

provides strong evidence of importance of unobserved time invariant variables for this model.

B. *Binary and Categorical Dependent Variables*

We now turn to the case of a binary dependent variable. In terms of the error distribution, we consider two scenarios: 1) both error components follow a normal distribution, which leads to probit-type models; and 2) the time-varying error follows a negative extreme value distribution while the time-invariant error is normally distributed, leading to logit-type models. We also consider extensions that do not make a specific distributional assumption about

$$\mu_i$$

but instead use a Heckman and Singer (1984) style discrete factor approach.

Probit or Logit

The basic form for the statistical model is a slight variation on equation (8):

$$Y_{ti}^* = X_{ti}\beta + T_{ti}\alpha + Z_i\delta + \mu_i + \varepsilon_{ti}. \qquad (9)$$

where Y_{ti}^* is the latent variable underlying the observed variable Y_{ti} which is equal to 1 when the latent variable is positive and 0 otherwise. We further assume that assumptions 1-4 above hold.

Given these assumptions, the following estimation methods, all of which can be implemented in Stata, will yield statistically unbiased estimates of the treatment effect:

1. **Probit or logit:** Simple probit or logit is a consistent estimator – probit assumes normality for the error terms while logit assumes that the error term is the difference of two negative extreme value distributed random variables. The standard errors for both methods are incorrect due to the correlation in observations for the same individual, but correct standard errors can easily be obtained (cluster option in Stata).
2. **Random Effects Probit or Logit**: These are the maximum likelihood estimators. They are consistent and asymptotically efficient under the model assumptions. We provide some details on the estimation methods since they will provide a better understanding of the options for this estimator and establish how this estimator is related to the semi-parametric estimator.

The maximum likelihood estimator maximizes the joint probability of the observed random sample. To build the likelihood function, note that under the normality assumption for probit:

$$P(Y_{ti} = 1 | \mu_i) = \Phi(X_{ti}\beta + T_{ti}\alpha + Z_i\delta + \mu_i). \quad (10)$$

where Φ is the standard normal cumulative distribution function. The corresponding expression for logit is:

$$P(Y_{ti} = 1 | \mu_i) = \frac{e^{X_{ti}\beta + T_{ti}\alpha + Z_i\delta + \mu_i}}{1 + e^{X_{ti}\beta + T_{ti}\alpha + Z_i\delta + \mu_i}}. \qquad (11)$$

This implies that the joint probability of the observed set of Y’s for individual i conditional on μ_i is:

$$A_i(\mu_i) = \prod_{t=1}^{M_i} P(Y_{ti} = 1 | \mu_i)^{Y_{ti}} (1 - P(Y_{ti} = 1 | \mu_i))^{1-Y_{ti}}. \quad (12)$$

The unconditional joint probability is obtained by integration:

$$A_i = \int_{-\infty}^{\infty} A_i(\mu_i) d\mu_i. \qquad (13)$$

This integral is typically approximated:

$$A_i \approx \sum_{k=1}^{K} P(\mu_k) A_i(\mu_k). \qquad (14)$$

where K is the number of points of support used in the approximation. Typically more accuracy is obtained with larger K. Given (14), the likelihood function is:

$$L = \prod_{i=1}^{N} A_i. \qquad (15)$$

The likelihood function is maximized with respect to the regression coefficients in (9) which are identified up to an unknown positive scale factor along with one additional parameter that is related to the variances of both error components:

$$\rho = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\varepsilon^2}. \qquad (16)$$

where ρ is the correlation between the error terms for two observations for the same individual. Note that, since the dependent variable is binary,

$$\sigma_\mu^2 \text{ and } \sigma_\varepsilon^2$$

are not separately identified and Stata imposes the normalization that

$$\sigma_\varepsilon^2 = 1.$$

Thus, as is true for any probit or logit model, coefficient estimates are identified to scale and we interpret the sign and significance of the coefficient which are not affect by scale.

Of course, the key part of the estimation is the evaluation of (14). If one assumes normally distributed error terms, then Hermite integration can be used with known weights

(the $P(\mu_k)$)

and points

(the μ_k).

If the true probabilities are small (equations [10] and [11]) or if there are a large number of observations per individual then K may have to be larger than the Stata default of 12.

An alternative to the normality assumption is to use a discrete factor approximation, which does not make any specific distributional assumption about the

$$\mu_i\text{'s}$$

but instead attempts to estimate the optimal mass points points and weights (for a given number of mass points K) for a discrete approximation to the true distribution of the

$$\mu_i\text{'s}.$$

This method is semi-parametric since specific distributional assumptions are still made about the other error component (leading either to a probit or logit based regression model). The discrete factor model was introduced by Heckman and Singer (1984) and has been evaluated with Monte Carlo experiments by Mroz (1999) and Guilkey and Lance (ND). The Monte Carlo work shows that the discrete factor method works reasonably well when the true distribution of the error term is normal but is far superior to methods that assume normality when the true error term's distribution is not normal – even for fairly mild departures from normality.

Mroz suggests increasing K (the number of points of support) until the log likelihood function increases by less than the number of additional parameters. Note that if there is a constant in the model, one of the points of support is set to zero since the constant term and all the points of support are not separately identified. In addition, the restriction that the weights sum to one is imposed. In these examples, we restrict the number of points of support to three. However, in actual practice, one should follow Mroz's suggestion for determining the number of points to use.

In the Stata examples below, the dependent variable is a zero/one indicator for use of contraception.

Stata Examples

Probit and Logit without Corrected Standard Errors

Example 9: Indonesia

Probit

```
. probit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community

Iteration 0:   log likelihood = -13551.564
Iteration 1:   log likelihood = -13207.034
Iteration 2:   log likelihood = -13206.714
Iteration 3:   log likelihood = -13206.714

Probit regression                                 Number of obs   =      20000
                                                  LR chi2(23)     =     689.70
                                                  Prob > chi2     =     0.0000
Log likelihood = -13206.714                       Pseudo R2       =     0.0254

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0109369   .0020704     5.28   0.000      .006879    .0149949
   midwife_0 |  -.0538529   .0199294    -2.70   0.007    -.0929138    -.014792
        edu1 |   .1712404   .0399162     4.29   0.000     .0930062    .2494747
        edu2 |  -.0158299   .0539504    -0.29   0.769    -.1215706    .0899109
        edu3 |  -.1050954   .0761855    -1.38   0.168    -.2544163    .0442254
       hedu1 |   .1708949   .0462358     3.70   0.000     .0802745    .2615153
       hedu2 |   .2684045   .0576111     4.66   0.000     .1554888    .3813202
       hedu3 |   .3578141   .0758883     4.72   0.000     .2090757    .5065526
         age |  -.0163741   .0011002   -14.88   0.000    -.0185305   -.0142178
    work_any |   .1594098   .0372256     4.28   0.000     .0864489    .2323707
   work_most |  -.0640241   .0372647    -1.72   0.086    -.1370616    .0090134
      smoker |  -.2856446   .0651997    -4.38   0.000    -.4134336   -.1578555
    goodhlth |   .1953823   .0284801     6.86   0.000     .1395624    .2512022
      muslim |  -.0234799   .0293651    -0.80   0.424    -.0810344    .0340747
    electric |  -.0188835   .0299369    -0.63   0.528    -.0775588    .0397918
        read |   .0839892   .0598674     1.40   0.161    -.0333487    .2013271
       write |   .0147675   .0557239     0.27   0.791    -.0944493    .1239844
       urban |   .0798585   .0232891     3.43   0.001     .0342128    .1255042
    pubtrans |  -.0560268   .0215556    -2.60   0.009    -.0982749   -.0137786
      caphrs |  -.0078833   .0033393    -2.36   0.018    -.0144283   -.0013384
     fhead50 |  -.0787679   .0262193    -3.00   0.003    -.1301567   -.0273791
     fhead75 |   .0300348   .0286076     1.05   0.294     -.026035    .0861047
    fhead100 |  -.0893216   .0369476    -2.42   0.016    -.1617375   -.0169056
       _cons |   .2344628    .077722     3.02   0.003     .0821304    .3867952
------------------------------------------------------------------------------
```

In probit, as with logit, we can interpret the sign and significance of the results. We see in this case that the number of posyandus is a positive and significant predictor of current contraceptive use while the number of midwives has a significant, negative effect which may be considered perverse. This could be due to program targeting – placing midwives in high need areas. Many of the control variables are significant and of the theoretically correct or anticipated sign. The interpretation of the results for both probit and logit estimators can be aided by the calculation of marginal effects through the use of the `margins` post estimation command in Stata. The types of analysis that can be done using the margins command are quite extensive and will not be covered here.

Example 10: Logit

```
. logit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community
```

```
Logistic regression                               Number of obs   =      20000
                                                  LR chi2(23)     =     693.30
                                                  Prob > chi2     =     0.0000
Log likelihood = -13204.916                       Pseudo R2       =     0.0256
```

cont_use	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
num_pos_0	.0178849	.0033975	5.26	0.000	.011226	.0245439
midwife_0	-.0871613	.0322669	-2.70	0.007	-.1504032	-.0239194
edu1	.2732954	.0639525	4.27	0.000	.1479508	.39864
edu2	-.0300828	.0869595	-0.35	0.729	-.2005203	.1403547
edu3	-.1739462	.1233162	-1.41	0.158	-.4156415	.067749
hedu1	.2711507	.074288	3.65	0.000	.1255489	.4167524
hedu2	.4294358	.0929697	4.62	0.000	.2472186	.6116531
hedu3	.5764629	.1230008	4.69	0.000	.3353857	.81754
age	-.0268319	.0017924	-14.97	0.000	-.0303449	-.023319
work_any	.2550064	.0603414	4.23	0.000	.1367394	.3732734
work_most	-.0996337	.0604092	-1.65	0.099	-.2180336	.0187662
smoker	-.455311	.104563	-4.35	0.000	-.6602507	-.2503713
goodhlth	.3135413	.0457627	6.85	0.000	.2238481	.4032345
muslim	-.0380366	.04756	-0.80	0.424	-.1312525	.0551794
electric	-.0298566	.0483442	-0.62	0.537	-.1246095	.0648963
read	.1326402	.0961124	1.38	0.168	-.0557367	.321017
write	.0256461	.089539	0.29	0.775	-.1498471	.2011394
urban	.1290811	.0376749	3.43	0.001	.0552396	.2029226
pubtrans	-.0907196	.034928	-2.60	0.009	-.1591772	-.022262
caphrs	-.0128743	.0053777	-2.39	0.017	-.0234144	-.0023342
fhead50	-.1277503	.0423011	-3.02	0.003	-.210659	-.0448416
fhead75	.0478089	.0463394	1.03	0.302	-.0430146	.1386325
fhead100	-.1462684	.0598709	-2.44	0.015	-.2636133	-.0289236
_cons	.4012711	.1256459	3.19	0.001	.1550096	.6475325

Point estimates from probit and logit models are scaled differently and so we cannot directly compare the point estimates of coefficients between the two types of regression models. However, the "z" statistic is a ratio of the estimated coefficient and its estimated standard error. Both are scaled by the same unkown constant and so the scaling divides out of the ratio. If you compare "z" statistics, we see that we get very similar results for logit and probit for virtually every variable.

Probit and Logit with Corrected Standard Errors

We see the expected result that the coefficient estimates are identical to probit with uncorrected standard errors. The "z" statistics are typically smaller.

Example 11: Indonesia

Probit

```
. probit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community, cluster(ind_id)

Probit regression                                 Number of obs   =      20000
                                                  Wald chi2(23)   =     518.68
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -13206.714                 Pseudo R2       =     0.0254

                              (Std. Err. adjusted for 9351 clusters in ind_id)
------------------------------------------------------------------------------
             |               Robust
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0109369    .002306     4.74   0.000     .0064173    .0154566
   midwife_0 |  -.0538529   .0212225    -2.54   0.011    -.0954483   -.0122575
        edu1 |   .1712404   .0454157     3.77   0.000     .0822272    .2602536
        edu2 |  -.0158299   .0599914    -0.26   0.792     -.133411    .1017512
        edu3 |  -.1050954   .0849752    -1.24   0.216    -.2716438     .061453
       hedu1 |   .1708949    .051919     3.29   0.001     .0691354    .2726543
       hedu2 |   .2684045   .0632872     4.24   0.000     .1443639     .392445
       hedu3 |   .3578141   .0822166     4.35   0.000     .1966725    .5189557
         age |  -.0163741   .0012732   -12.86   0.000    -.0188695   -.0138787
    work_any |   .1594098   .0370064     4.31   0.000     .0868785    .2319411
   work_most |  -.0640241   .0374718    -1.71   0.088    -.1374675    .0094193
      smoker |  -.2856446   .0717635    -3.98   0.000    -.4262985   -.1449906
     goodhlth |   .1953823   .0286817     6.81   0.000     .1391672    .2515974
      muslim |  -.0234799   .0355719    -0.66   0.509    -.0931995    .0462397
    electric |  -.0188835   .0323575    -0.58   0.559    -.0823031    .0445361
        read |   .0839892   .0627234     1.34   0.181    -.0389464    .2069249
       write |   .0147675   .0564454     0.26   0.794    -.0958634    .1253985
       urban |   .0798585   .0273926     2.92   0.004     .0261701    .1335469
    pubtrans |  -.0560268    .022953    -2.44   0.015    -.1010137   -.0110398
      caphrs |  -.0078833   .0036744    -2.15   0.032     -.015085   -.0006816
     fhead50 |  -.0787679   .0299259    -2.63   0.008    -.1374216   -.0201141
     fhead75 |   .0300348   .0330576     0.91   0.364    -.0347568    .0948265
    fhead100 |  -.0893216   .0414157    -2.16   0.031    -.1704948   -.0081483
       _cons |   .2344628   .0889082     2.64   0.008     .0602059    .4087197
------------------------------------------------------------------------------
```

Example 12: Logit

```
. logit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community, cluster(ind_id)


Logistic regression                               Number of obs   =      20000
                                                  Wald chi2(23)   =     510.52
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -13204.916                 Pseudo R2       =     0.0256

                              (Std. Err. adjusted for 9351 clusters in ind_id)
------------------------------------------------------------------------------
             |               Robust
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0178849   .0037906     4.72   0.000     .0104556    .0253143
   midwife_0 |  -.0871613   .0343626    -2.54   0.011    -.1545108   -.0198118
        edu1 |   .2732954   .0725995     3.76   0.000     .1310031    .4155878
        edu2 |  -.0300828   .0965212    -0.31   0.755    -.2192608    .1590953
        edu3 |  -.1739462   .1376078    -1.26   0.206    -.4436525      .09576
       hedu1 |   .2711507   .0832271     3.26   0.001     .1080286    .4342727
       hedu2 |   .4294358   .1018927     4.21   0.000     .2297297     .629142
       hedu3 |   .5764629   .1331344     4.33   0.000     .3155243    .8374014
         age |  -.0268319   .0020804   -12.90   0.000    -.0309094   -.0227544
    work_any |   .2550064   .0598092     4.26   0.000     .1377826    .3722303
   work_most |  -.0996337   .0605915    -1.64   0.100    -.2183908    .0191234
      smoker |   -.455311   .1146241    -3.97   0.000      -.67997    -.230652
    goodhlth |   .3135413   .0460637     6.81   0.000     .2232581    .4038245
      muslim |  -.0380366   .0577319    -0.66   0.510    -.1511891     .075116
    electric |  -.0298566   .0522581    -0.57   0.568    -.1322805    .0725673
        read |   .1326402   .1004165     1.32   0.187    -.0641725    .3294528
       write |   .0256461   .0904471     0.28   0.777    -.1516269    .2029192
       urban |   .1290811    .044341     2.91   0.004     .0421743    .2159879
    pubtrans |  -.0907196   .0372271    -2.44   0.015    -.1636833   -.0177559
      caphrs |  -.0128743   .0059114    -2.18   0.029    -.0244605   -.0012881
     fhead50 |  -.1277503   .0483122    -2.64   0.008    -.2224404   -.0330601
     fhead75 |   .0478089   .0535777     0.89   0.372    -.0572015    .1528194
    fhead100 |  -.1462684   .0670922    -2.18   0.029    -.2777667   -.0147702
       _cons |   .4012711   .1438964     2.79   0.005     .1192392    .6833029
------------------------------------------------------------------------------
```

We get the expected result that the coefficient estimates are identical to logit with uncorrected standard errors but the "z" statistics are typically smaller.

Random Effects Probit and Logit with Normal Error Assumption

In the two sets of results that follow, the following Stata command was used to define the individual and time identifying variables:

```
xtset ind_id year
```

"`xtset`" is a required pre-requisite command for estimation of the Stata "xt" family of commands. Essentially, "`xtset`" is where one formally informs Stata of the key parameters for "`xt`" commands, for example the individual identifier variable or the time variable. For both sets of random effects results we used the default number of quadrature points (12).

Example 13: Indonesia

Probit

```
. xtprobit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community, re

Random-effects probit regression                Number of obs      =     20000
Group variable: ind_id                          Number of groups   =      9351

Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       2.1
                                                               max =         4

                                                Wald chi2(23)      =    481.59
Log likelihood  = -12586.652                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0115763   .0030638     3.78   0.000     .0055714    .0175811
   midwife_0 |  -.0705085   .0271081    -2.60   0.009    -.1236394   -.0173776
        edu1 |   .2512627   .0568511     4.42   0.000     .1398367    .3626887
        edu2 |  -.0194227   .0769048    -0.25   0.801    -.1701534    .1313079
        edu3 |  -.1299522    .107899    -1.20   0.228    -.3414304     .081526
       hedu1 |   .1606739   .0663152     2.42   0.015     .0306985    .2906494
       hedu2 |   .2925679   .0811508     3.61   0.000     .1335154    .4516205
       hedu3 |   .4106312   .1052508     3.90   0.000     .2043433     .616919
         age |  -.0219594   .0016474   -13.33   0.000    -.0251882   -.0187306
    work_any |   .1943832    .048518     4.01   0.000     .0992897    .2894767
   work_most |  -.0728665   .0485648    -1.50   0.134    -.1680518    .0223189
      smoker |  -.2735474    .094066    -2.91   0.004    -.4579133   -.0891815
    goodhlth |   .2336692   .0374362     6.24   0.000     .1602957    .3070427
      muslim |  -.0008929   .0477509    -0.02   0.985    -.0944829    .0926971
    electric |  -.0622471   .0415269    -1.50   0.134    -.1436384    .0191442
        read |   .1202474   .0794428     1.51   0.130    -.0354576    .2759525
       write |   .0000375   .0729884     0.00   1.000    -.1430171    .1430922
       urban |   .1287606   .0364688     3.53   0.000     .0572831     .200238
    pubtrans |  -.0754734   .0295184    -2.56   0.011    -.1333283   -.0176184
      caphrs |  -.0113017   .0050453    -2.24   0.025    -.0211903   -.0014132
     fhead50 |  -.0679126    .038314    -1.77   0.076    -.1430067    .0071815
     fhead75 |   .0606825   .0428779     1.42   0.157    -.0233567    .1447217
    fhead100 |  -.0611929   .0542834    -1.13   0.260    -.1675864    .0452007
       _cons |   .3336153   .1139158     2.93   0.003     .1103445    .5568861
-------------+----------------------------------------------------------------
    /lnsig2u |  -.0954713   .0535037                     -.2003367    .0093941
-------------+----------------------------------------------------------------
     sigma_u |   .9533858   .0255049                      .9046851    1.004708
         rho |   .4761503   .0133455                      .4500827    .5023485
------------------------------------------------------------------------------
Likelihood-ratio test of rho=0: chibar2(01) =  1240.12 Prob >= chibar2 = 0.000
```

Because of scale differences, we cannot even compare the point estimates of the coefficients of the random effects probit to the point estimates of the coefficients for simple probit. However, we can still compare the "z" statistics. We see that they are roughly comparable to the "z" statistics for probit with corrected standard errors. If the assumption of normality is correct for both error components, we know that random effects probit is asymptotically efficient. However, we see little gain in this particular application even though a test of the null hypothesis that

$$\rho = 0$$

is strongly rejected.

Example 14: Logit

```
. xtlogit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community, re


Random-effects logistic regression               Number of obs      =     20000
Group variable: ind_id                           Number of groups   =      9351

Random effects u_i ~ Gaussian                    Obs per group: min =         1
                                                                avg =       2.1
                                                                max =         4

                                                 Wald chi2(23)      =    473.65
Log likelihood  =  -12583.57                     Prob > chi2        =    0.0000

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |    .019734   .0052248     3.78   0.000     .0094936    .0299744
   midwife_0 |  -.1202664   .0460992    -2.61   0.009    -.2106192   -.0299136
        edu1 |   .4251938   .0962956     4.42   0.000      .236458    .6139296
        edu2 |   -.035901   .1304789    -0.28   0.783    -.2916349    .2198329
        edu3 |  -.2257908   .1835479    -1.23   0.219    -.5855379    .1339564
       hedu1 |   .2705757   .1125454     2.40   0.016     .0499907    .4911608
       hedu2 |   .4939675   .1379043     3.58   0.000       .22368    .7642549
       hedu3 |   .6974616   .1792575     3.89   0.000     .3461233      1.0488
         age |  -.0376674   .0028162   -13.38   0.000    -.0431871   -.0321477
    work_any |   .3284708   .0825756     3.98   0.000     .1666257    .4903159
   work_most |  -.1230638   .0826313    -1.49   0.136    -.2850181    .0388906
      smoker |  -.4641005   .1592931    -2.91   0.004    -.7763092   -.1518918
    goodhlth |   .3943995   .0636068     6.20   0.000     .2697324    .5190665
      muslim |  -.0014422    .081195    -0.02   0.986    -.1605814    .1576971
    electric |   -.106157   .0706988    -1.50   0.133    -.2447241      .03241
        read |   .2052651   .1344833     1.53   0.127    -.0583173    .4688475
       write |  -.0006182   .1235841    -0.01   0.996    -.2428386    .2416022
       urban |   .2176307   .0620372     3.51   0.000       .09604    .3392214
    pubtrans |  -.1293711   .0502274    -2.58   0.010     -.227815   -.0309272
      caphrs |  -.0193324   .0085684    -2.26   0.024    -.0361262   -.0025386
     fhead50 |  -.1156362   .0651028    -1.78   0.076    -.2432354     .011963
     fhead75 |   .1034831   .0729442     1.42   0.156    -.0394848    .2464511
    fhead100 |  -.1041968   .0923139    -1.13   0.259    -.2851288    .0767352
       _cons |   .5900322   .1937374     3.05   0.002     .2103139    .9697505
-------------+----------------------------------------------------------------
    /lnsig2u |   .9570903    .055905                      .8475186    1.066662
-------------+----------------------------------------------------------------
     sigma_u |   1.613725   .0451076                      1.527694    1.704601
         rho |   .4418254    .013787                      .4150011    .4689935
------------------------------------------------------------------------------
Likelihood-ratio test of rho=0: chibar2(01) =  1242.69 Prob >= chibar2 = 0.000
```

We see very minor differences between this set of results and the results for simple logit with corrected standard errors. We also see little substantive difference in the results for logit and probit.

Random Effects Probit and Logit with the Stata Implementation of the Discrete Factor Approximation Approach

Once again, the point estimates of the coefficients cannot be compared due to differences in scaling. In the case of the Stata version of the discrete factor model,

Example 15: Indonesia

Probit

```
. gllamm cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 /*
> */ $individual $community, i(ind_id) family(binom) link(probit) nip(3) ip(f) trace
dot

number of level 1 units = 20000
number of level 2 units = 9351

Condition Number = 2664.3203

gllamm model

log likelihood = -12584.159
------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0121095   .0030955     3.91   0.000     .0060423    .0181766
   midwife_0 |  -.0691841   .0271633    -2.55   0.011    -.1224232   -.0159449
        edu1 |    .256545   .0580227     4.42   0.000     .1428226    .3702674
        edu2 |  -.0102305   .0778819    -0.13   0.895    -.1628762    .1424152
        edu3 |   -.129959   .1092196    -1.19   0.234    -.3440255    .0841074
       hedu1 |   .1512052   .0682464     2.22   0.027     .0174447    .2849656
       hedu2 |   .2785999   .0827962     3.36   0.001     .1163222    .4408775
       hedu3 |   .4041642   .1066379     3.79   0.000     .1951577    .6131707
         age |  -.0218221   .0016723   -13.05   0.000    -.0250998   -.0185444
    work_any |   .1930435   .0486714     3.97   0.000     .0976493    .2884376
   work_most |  -.0718741   .0487405    -1.47   0.140    -.1674036    .0236555
      smoker |   -.275149   .0946751    -2.91   0.004    -.4607088   -.0895893
     goodhlth |   .2338546   .0377361     6.20   0.000     .1598932    .3078159
      muslim |   .0018042   .0479845     0.04   0.970    -.0922438    .0958522
    electric |  -.0632245   .0416404    -1.52   0.129    -.1448383    .0183893
        read |     .12036   .0795332     1.51   0.130    -.0355222    .2762422
       write |  -.0031475   .0731878    -0.04   0.966    -.1465931     .140298
       urban |   .1289513   .0363347     3.55   0.000     .0577366     .200166
    pubtrans |  -.0717043   .0295539    -2.43   0.015    -.1296288   -.0137798
      caphrs |   -.010509   .0050382    -2.09   0.037    -.0203837   -.0006343
     fhead50 |  -.0712806   .0384083    -1.86   0.063    -.1465594    .0039982
     fhead75 |   .0547927   .0430998     1.27   0.204    -.0296813    .1392667
    fhead100 |  -.0708628   .0548103    -1.29   0.196    -.1782891    .0365634
       _cons |   .4000858   .1596458     2.51   0.012     .0871857    .7129858
------------------------------------------------------------------------------
Probabilities and locations of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    loc1: -1.2135, 2.2586, .18991
  var(1): 1.1028223
    prob: 0.3143, 0.1215, 0.5642
------------------------------------------------------------------------------
```

the normalization imposed is that the expectation of the mass point distribution is zero. The best way to compare the effects of variables from the discrete factor results to random effects probit and logit is by using simulation methods. We do see that in terms of the "z" statistics that the discrete factor results and the random effects probit and logit results are similar. This may be an indication that the underlying assumption of normality is not unreasonable in this application in the sense of representing a serious misspecification.

As with all probit/logit comparisons, scale is different and so we compare in terms of sign and

Example 16: Logit

```
. gllamm cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 /*
> */ $individual $community, i(ind_id) family(binom) link(logit) nip(3) ip(f) trace
dot

number of level 1 units = 20000
number of level 2 units = 9351

Condition Number = 1380.9801

gllamm model

log likelihood = -12582.852

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0205155   .0053198     3.86   0.000     .0100888    .0309421
   midwife_0 |  -.1191791   .0461745    -2.58   0.010    -.2096794   -.0286787
        edu1 |   .4327188   .0974286     4.44   0.000     .2417623    .6236753
        edu2 |  -.0247845   .1313739    -0.19   0.850    -.2822726    .2327036
        edu3 |  -.2146267    .184338    -1.16   0.244    -.5759225    .1466691
       hedu1 |   .2730537   .1134983     2.41   0.016     .0506011    .4955062
       hedu2 |    .490124   .1388749     3.53   0.000     .2179341    .7623139
       hedu3 |   .7003171   .1802491     3.89   0.000     .3470354    1.053599
         age |  -.0375087   .0028557   -13.13   0.000    -.0431059   -.0319116
    work_any |   .3263321   .0828994     3.94   0.000     .1638523    .4888119
   work_most |  -.1169797   .0830437    -1.41   0.159    -.2797424    .0457829
      smoker |  -.4684548   .1598765    -2.93   0.003    -.7818071   -.1551026
    goodhlth |   .3957469    .063534     6.23   0.000     .2712225    .5202713
      muslim |   .0005657    .081547     0.01   0.994    -.1592634    .1603948
    electric |  -.1086617   .0708982    -1.53   0.125    -.2476197    .0302962
        read |    .205385   .1344244     1.53   0.127     -.058082    .4688521
       write |  -.0047697   .1237713    -0.04   0.969    -.2473569    .2378176
       urban |   .2159044   .0618427     3.49   0.000     .0946949    .3371138
    pubtrans |  -.1236216   .0503124    -2.46   0.014    -.2222321    -.025011
      caphrs |  -.0189458     .00861    -2.20   0.028    -.0358211   -.0020706
     fhead50 |  -.1190166   .0652343    -1.82   0.068    -.2468734    .0088402
     fhead75 |   .0983816   .0735143     1.34   0.181    -.0457037    .2424669
    fhead100 |  -.1162186   .0935374    -1.24   0.214    -.2995485    .0671113
       _cons |   .5881314    .204967     2.87   0.004     .1864036    .9898593
------------------------------------------------------------------------------
Probabilities and locations of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    loc1: -2.0083, 2.4055, .17703
  var(1): 2.3626656
    prob: 0.2891, 0.2041, 0.5068
------------------------------------------------------------------------------
```

significance and we see that the substantive results for the two estimation methods are quite consistent. However, a better comparison between the results could be done using the `margins` post estimation command.

Multinomial Logit

Extensions of binary dependent variable models to cases in which the dependent variable is categorical are straightforward. However, the logit-based extension is much more computationally manageable than probit-based alternatives, which require simulation methods when the number of categories for the dependent variable is more than three and these simulation methods can be very computationally intensive. Therefore, in this section, we only consider the logit-based extension.

Consider a dependent variable, Y_{ti}, that takes on G possible values (g=1,2,…,G). If it is assumed that the time-varying error term follows the negative extreme value distribution, then:

$$P(Y_{ti} = 1|\mu_i) = \frac{1}{1 + \sum_{h=2}^{G} e^{X_{ti}\beta_h + T_{ti}\alpha_h + Z_i\delta_h + \rho_h\mu_i}} \tag{17}$$

and for g=2,3,…,G:

$$P(Y_{ti} = g|\mu_i) = \frac{e^{X_{ti}\beta_g + T_{ti}\alpha_g + Z_i\delta_g + \rho_g\mu_i}}{1 + \sum_{h=2}^{G} e^{X_{ti}\beta_h + T_{ti}\alpha_h + Z_i\delta_h + \rho_h\mu_i}} \tag{18}$$

Given (17) and (18), it is straightforward to write the model in the following familiar format for g=2,3,…,G:

$$\ln\left[\frac{P(Y_{ti} = g \mid \mu_i)}{P(Y_{ti} = 1 \mid \mu_i)}\right] = e^{X_{ti}\beta_g + T_{ti}\alpha_g + Z_i\delta_g + \rho_g\mu_i}. \tag{19}$$

Given the assumptions for the model, the following estimators give statistically correct (i.e., consistent) parameter estimates, and hence behavioral parameters of interest, such as the marginal effect of treatment on the probability of being in each outcome category:

1. **Multinomial Logit:** Simple multinomial logit is a consistent estimator. he standard errors are incorrect due to the correlation in observations for the same individual, but correct standard errors can easily be obtained (cluster option in Stata).
2. **Parametric and Semi-Parametric Random Effects Multinomial Logit**: These are the maximum likelihood estimators. They are consistent and asymptotically efficient under the model assumptions.

It is straightforward to use (17) and (18) to write multinomial logit extensions of equations (12) though (15) and so that a random effects maximum likelihood estimator can be used either assuming normality for the time-invariant error or using the discrete factor model for a semi-parametric estimator.

Stata Examples

Example 17: Indonesia

Multinomial Logit without Corrected Standard Error

```
. mlogit new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual
$community, base(3)

Multinomial logistic regression                   Number of obs   =      19989
                                                  LR chi2(38)     =     841.25
                                                  Prob > chi2     =     0.0000
Log likelihood =  -15085.67                       Pseudo R2       =     0.0271

------------------------------------------------------------------------------
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |   .0173876   .0034314     5.07   0.000     .0106622    .0241131
   midwife_0 |  -.0534229   .0353398    -1.51   0.131    -.1226877    .0158418
  del_post_0 |  -.0810706   .0370866    -2.19   0.029     -.153759   -.0083822
  med_post_0 |  -.0148349   .0495479    -0.30   0.765     -.111947    .0822773
         age |  -.0284746   .0017951   -15.86   0.000     -.031993   -.0249562
    work_any |    .248673   .0607439     4.09   0.000     .1296171    .3677289
   work_most |  -.1006145   .0608558    -1.65   0.098    -.2198897    .0186607
      smoker |   -.490984   .1062579    -4.62   0.000    -.6992456   -.2827225
    goodhlth |   .3234357   .0463135     6.98   0.000      .232663    .4142084
      muslim |   .0357844   .0480173     0.75   0.456    -.0583279    .1298967
    electric |  -.0173764   .0482078    -0.36   0.719    -.1118619    .0771091
        read |   .2798236   .0928182     3.01   0.003     .0979034    .4617439
       write |   .0374617   .0894746     0.42   0.675    -.1379053    .2128287
       urban |   .0935631   .0376155     2.49   0.013      .019838    .1672882
    pubtrans |  -.0986045   .0350799    -2.81   0.005    -.1673598   -.0298491
      caphrs |   -.014734   .0054474    -2.70   0.007    -.0254108   -.0040572
     fhead50 |  -.1314345   .0425862    -3.09   0.002    -.2149018   -.0479671
     fhead75 |   .0447068   .0465214     0.96   0.337    -.0464734     .135887
    fhead100 |   -.175815   .0604816    -2.91   0.004    -.2943568   -.0572732
       _cons |   .7233915   .1130516     6.40   0.000     .5018143    .9449686
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |   .0157647   .0088031     1.79   0.073     -.001489    .0330185
   midwife_0 |  -.0748726   .1134741    -0.66   0.509    -.2972778    .1475326
  del_post_0 |  -.2286236   .1371954    -1.67   0.096    -.4975216    .0402744
  med_post_0 |   .2279939   .1487726     1.53   0.125     -.063595    .5195827
         age |   .0177922   .0057564     3.09   0.002     .0065099    .0290745
    work_any |   .3448375   .1887716     1.83   0.068     -.025148     .714823
   work_most |  -.1592836   .1869351    -0.85   0.394    -.5256696    .2071025
      smoker |  -.1814589   .3038884    -0.60   0.550    -.7770692    .4141513
    goodhlth |   .0452051   .1408074     0.32   0.748    -.2307723    .3211825
      muslim |  -.7289283   .1209649    -6.03   0.000    -.9660151   -.4918416
    electric |   -.056446   .1861924    -0.30   0.762    -.4213763    .3084843
        read |   .3222572   .3416637     0.94   0.346    -.3473913    .9919057
       write |   .2974415   .3256701     0.91   0.361    -.3408602    .9357432
       urban |   .5271656   .1273432     4.14   0.000     .2775776    .7767536
    pubtrans |   .1914594   .1336173     1.43   0.152    -.0704256    .4533445
      caphrs |   .0290562   .0152733     1.90   0.057    -.0008789    .0589913
     fhead50 |  -.0575051    .156341    -0.37   0.713    -.3639278    .2489175
     fhead75 |  -.0448876   .1658717    -0.27   0.787      -.36999    .2802149
    fhead100 |   .3061924   .1859207     1.65   0.100    -.0582055    .6705903
       _cons |  -4.169341   .3802088   -10.97   0.000    -4.914537   -3.424146
-------------+----------------------------------------------------------------
No_Method    |  (base outcome)--------------------------------------------
------------------------------------------------------------------------------
```

Multinomial logit with a large number of categories for the dependent variable generates a large number of results. For this example, we limited the categories to three to limit the output: the dependent variable is contraceptive use, and the possible choices the individual can make are no contraceptive method (the "baseline category"), modern contraceptive use and traditional contraceptive use.

Example 18: Multinomial Logit with Corrected Standard Error

```
. mlogit new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual $community,
cluster(ind_id) base(3)

Multinomial logistic regression                   Number of obs   =      19989
                                                  Wald chi2(38)   =     689.72
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood =  -15085.67                 Pseudo R2       =     0.0271

                              (Std. Err. adjusted for 9350 clusters in ind_id)
------------------------------------------------------------------------------
             |               Robust
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |   .0173876   .0038716     4.49   0.000     .0097994    .0249759
   midwife_0 |  -.0534229   .0375556    -1.42   0.155    -.1270306    .0201848
  del_post_0 |  -.0810706   .0393947    -2.06   0.040    -.1582828   -.0038583
  med_post_0 |  -.0148349    .050507    -0.29   0.769    -.1138267     .084157
         age |  -.0284746   .0021005   -13.56   0.000    -.0325914   -.0243578
    work_any |    .248673   .0604646     4.11   0.000     .1301646    .3671815
   work_most |  -.1006145     .06131    -1.64   0.101     -.22078     .0195509
      smoker |   -.490984   .1172498    -4.19   0.000    -.7207894   -.2611786
    goodhlth |   .3234357   .0467966     6.91   0.000      .231716    .4151554
      muslim |   .0357844   .0595799     0.60   0.548     -.08099     .1525588
    electric |  -.0173764   .0522569    -0.33   0.739     -.119798    .0850452
        read |   .2798236   .0968637     2.89   0.004     .0899743     .469673
       write |   .0374617   .0906952     0.41   0.680    -.1402976     .215221
       urban |   .0935631   .0448799     2.08   0.037     .0056001     .181526
    pubtrans |  -.0986045   .0375293    -2.63   0.009    -.1721606   -.0250483
      caphrs |   -.014734   .0060881    -2.42   0.016    -.0266664   -.0028016
     fhead50 |  -.1314345   .0489298    -2.69   0.007    -.2273351   -.0355339
     fhead75 |   .0447068   .0541813     0.83   0.409    -.0614865    .1509002
    fhead100 |   -.175815   .0686465    -2.56   0.010    -.3103597   -.0412703
       _cons |   .7233915   .1286599     5.62   0.000     .4712226    .9755603
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |   .0157647   .0085134     1.85   0.064    -.0009213    .0324508
   midwife_0 |  -.0748726    .115961    -0.65   0.518     -.302152    .1524068
  del_post_0 |  -.2286236   .1392587    -1.64   0.101    -.5015656    .0443184
  med_post_0 |   .2279939   .1527515     1.49   0.136    -.0713937    .5273814
         age |   .0177922   .0056121     3.17   0.002     .0067927    .0287917
    work_any |   .3448375   .1922801     1.79   0.073    -.0320246    .7216996
   work_most |  -.1592836   .1932703    -0.82   0.410    -.5380864    .2195192
      smoker |  -.1814589   .3237857    -0.56   0.575    -.8160671    .4531493
    goodhlth |   .0452051   .1401215     0.32   0.747     -.229428    .3198381
      muslim |  -.7289283   .1323059    -5.51   0.000    -.9882432   -.4696135
    electric |   -.056446   .1889804    -0.30   0.765    -.4268407    .3139488
        read |   .3222572   .3267316     0.99   0.324    -.3181249    .9626392
       write |   .2974415   .3086179     0.96   0.335    -.3074385    .9023216
       urban |   .5271656   .1434954     3.67   0.000     .2459199    .8084113
    pubtrans |   .1914594   .1393387     1.37   0.169    -.0816394    .4645583
      caphrs |   .0290562   .0179441     1.62   0.105    -.0061136     .064226
     fhead50 |  -.0575051   .1654955    -0.35   0.728    -.3818704    .2668601
     fhead75 |  -.0448876    .174712    -0.26   0.797    -.3873167    .2975416
    fhead100 |   .3061924   .1937001     1.58   0.114    -.0734528    .6858376
       _cons |  -4.169341   .3899908   -10.69   0.000    -4.933709   -3.404973
-------------+----------------------------------------------------------------
No Method    |  (base outcome)
```

The multinomial logit results with corrected standard errors have the same point estimates of the regression parameters but typically larger standard errors and smaller "z" statistics. The reason is much the same as in earlier applications of the corrected standard errors: the correct standard errors involve recognizing that there is less information present in the sample than would be understood to be the case under more naïve estimates of the standard errors which assume complete independence of observations.

Example 20: Random Effects Multinomial Logit with Normal Error Assumption

```
gllamm new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual $community if
method!=95, i(ind_id) family(binomial) link(mlogit) b(3) nip(20) ip(g) trace dot

gllamm model

log likelihood = -14454.757

------------------------------------------------------------------------------
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |   .0190517   .0052616     3.62   0.000     .0087391    .0293642
   midwife_0 |  -.0669888   .0501144    -1.34   0.181    -.1652112    .0312336
  del_post_0 |   -.145471    .053698    -2.71   0.007    -.2507171   -.0402248
  med_post_0 |   .0089145   .0681895     0.13   0.896    -.1247345    .1425635
         age |   -.039486    .002814   -14.03   0.000    -.0450014   -.0339705
    work_any |   .3188672   .0830006     3.84   0.000      .156189    .4815455
   work_most |    -.12106   .0831287    -1.46   0.145    -.2839893    .0418693
      smoker |  -.5011374   .1611052    -3.11   0.002    -.8168977    -.185377
    goodhlth |   .4050086   .0641778     6.31   0.000     .2792224    .5307948
      muslim |   .0871847   .0816352     1.07   0.286    -.0728173    .2471866
    electric |  -.0952896   .0706664    -1.35   0.178    -.2337932    .0432141
        read |   .3788079   .1306218     2.90   0.004     .1227938    .6348219
       write |   .0183154   .1236428     0.15   0.882    -.2240201     .260651
       urban |    .162416   .0618416     2.63   0.009     .0412087    .2836233
    pubtrans |  -.1380943   .0504852    -2.74   0.006    -.2370435   -.0391452
      caphrs |  -.0220487   .0086462    -2.55   0.011    -.0389949   -.0051025
     fhead50 |  -.1188477   .0655249    -1.81   0.070    -.2472741    .0095788
     fhead75 |   .1060594   .0733177     1.45   0.148    -.0376407    .2497594
    fhead100 |  -.1327621    .093046    -1.43   0.154     -.315129    .0496047
       _cons |   1.008322   .1729042     5.83   0.000     .6694356    1.347208
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |    .016971   .0097543     1.74   0.082    -.0021472    .0360891
   midwife_0 |  -.0907898   .1193974    -0.76   0.447    -.3248045    .1432248
  del_post_0 |  -.3027018   .1429596    -2.12   0.034    -.5828975   -.0225061
  med_post_0 |   .2534894   .1565259     1.62   0.105    -.0532958    .5602745
         age |   .0091857    .006314     1.45   0.146    -.0031894    .0215608
    work_any |   .4131892   .1974391     2.09   0.036     .0262157    .8001628
   work_most |   -.195297   .1956532    -1.00   0.318    -.5787702    .1881763
      smoker |  -.1608926   .3269778    -0.49   0.623    -.8017574    .4799722
    goodhlth |   .1209426    .147982     0.82   0.414    -.1690967    .4109819
      muslim |  -.6871284   .1383365    -4.97   0.000     -.958263   -.4159938
    electric |  -.1449156   .1938751    -0.75   0.455    -.5249037    .2350726
        read |    .423401   .3553102     1.19   0.233    -.2729942    1.119796
       write |    .270902   .3383261     0.80   0.423    -.3922049    .9340089
       urban |    .594295    .136957     4.34   0.000     .3258641    .8627259
    pubtrans |   .1608431   .1386913     1.16   0.246    -.1109869    .4326731
      caphrs |   .0231539   .0166496     1.39   0.164    -.0094787    .0557865
     fhead50 |  -.0490571   .1645788    -0.30   0.766    -.3716256    .2735115
     fhead75 |   .0044929   .1758678     0.03   0.980    -.3402017    .3491874
    fhead100 |   .3388345    .199772     1.70   0.090    -.0527115    .7303805
       _cons |  -3.929608    .402972    -9.75   0.000    -4.719419   -3.139798
------------------------------------------------------------------------------
Variances and covariances of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    var(1): 2.6513581 (.14795414)
------------------------------------------------------------------------------
```

There are several important points to glean from these results. First, we used 20 interpolation points rather than the default value of 12 so that we could get more accurate results. The Generalized Linear Latent And Mixed Models (GLLAMM) estimator imposes the restriction that $\rho_2 = \rho_3$, which means the effect of the time-invariant individual-level unobservables is the same for all comparisons. This restriction is not necessary – varying effects are identified in this model just as the effects of observable regressors are allowed to have different effects across comparisons. Nevertheless, we do see that by comparison that the estimation results are quite similar to the results for multinomial logit with corrected standard errors in terms of sign and significance.

Example 21: Random Effects Multinomial Logit with the Discrete Factor Approximation

```
. gllamm new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual $community if method!=95,
i(ind_id) family(binomial) link(mlogit) b(3) nip(3) ip(f) trace dot

gllamm model
log likelihood = -14453.85
------------------------------------------------------------------------------
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |    .019989   .0053161     3.76   0.000     .0095697    .0304084
   midwife_0 |  -.0611109   .0499966    -1.22   0.222    -.1591024    .0368806
  del_post_0 |   -.149761   .0535767    -2.80   0.005    -.2547694   -.0447527
  med_post_0 |   .0053797   .0679972     0.08   0.937    -.1278924    .1386518
         age |  -.0395301   .0028464   -13.89   0.000    -.0451089   -.0339513
    work_any |   .3176317   .0832865     3.81   0.000     .1543933    .4808702
   work_most |  -.1156429   .0834422    -1.39   0.166    -.2791867    .0479008
      smoker |  -.5088546   .1572613    -3.24   0.001    -.8170811    -.200628
    goodhlth |   .4059272   .0639884     6.34   0.000     .2805121    .5313422
      muslim |   .0888798   .0815994     1.09   0.276    -.0710521    .2488116
    electric |  -.1008509   .0707992    -1.42   0.154    -.2396147     .037913
        read |   .3677506   .1304627     2.82   0.005     .1120483    .6234528
       write |   .0127391   .1237447     0.10   0.918     -.229796    .2552743
       urban |   .1636063   .0614539     2.66   0.008     .0431588    .2840537
    pubtrans |  -.1310189     .05046    -2.60   0.009    -.2299187   -.0321192
      caphrs |  -.0215443   .0084652    -2.55   0.011    -.0381357   -.0049528
     fhead50 |  -.1235393   .0652168    -1.89   0.058    -.2513617    .0042832
     fhead75 |   .0968273   .0732249     1.32   0.186    -.0466908    .2403454
    fhead100 |  -.1496543   .0934108    -1.60   0.109    -.3327361    .0334275
       _cons |   1.678753   15.99491     0.10   0.916     -29.6707     33.0282
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |   .0179595   .0097879     1.83   0.067    -.0012246    .0371435
   midwife_0 |  -.0843515   .1193608    -0.71   0.480    -.3182944    .1495914
  del_post_0 |  -.3073886    .142927    -2.15   0.032    -.5875204   -.0272569
  med_post_0 |   .2496661   .1564945     1.60   0.111    -.0570576    .5563898
         age |   .0091067   .0063283     1.44   0.150    -.0032966    .0215101
    work_any |   .4115542   .1975848     2.08   0.037     .0242951    .7988134
   work_most |  -.1891342   .1958105    -0.97   0.334    -.5729157    .1946473
      smoker |  -.1707497    .325839    -0.52   0.600    -.8093825     .467883
    goodhlth |   .1215186   .1479266     0.82   0.411    -.1684121    .4114493
      muslim |  -.6852552   .1384349    -4.95   0.000    -.9565826   -.4139277
    electric |  -.1503879    .193861    -0.78   0.438    -.5303484    .2295726
        read |   .4130398   .3551663     1.16   0.245    -.2830735    1.109153
       write |   .2651068   .3382971     0.78   0.433    -.3979434     .928157
       urban |   .5965677   .1367702     4.36   0.000     .3285031    .8646323
    pubtrans |   .1673123   .1386651     1.21   0.228    -.1044663    .4390909
      caphrs |   .0236671    .016549     1.43   0.153    -.0087685    .0561026
     fhead50 |   -.053016   .1644108    -0.32   0.747    -.3752552    .2692231
     fhead75 |  -.0046758   .1758791    -0.03   0.979    -.3493924    .3400409
    fhead100 |    .320879   .1999711     1.60   0.109    -.0710571    .7128151
       _cons |  -3.258997   15.99971    -0.20   0.839    -34.61785    28.09986
------------------------------------------------------------------------------
Probabilities and locations of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    loc1: -2.4885, 8.6742, -.15526
  var(1): 10.156409
    prob: 0.3382, 0.107, 0.5548
```

The GLLAMM version of the discrete factor model also imposes the restriction that the variances are the same across comparisons (which is again an unnecessary restriction). We see very comparable results for the two versions of the random effects multinomial logit which probably indicates that for this particular example, the assumption of normality is not a serious misspecification error.

III. Longitudinal Methods for Models When the Treatment is Endogenous

In this section, we relax the assumption that the treatment is exogenous. Typically, the assumption that the treatment variable is endogenous is more realistic in a non-experimental setting. If the treatment is whether or not a respondent heard a family planning message and the outcome is the respondent's ideal family size, one could make the argument that simple methods that do not account for the endogeneity of hearing such a message could be biased and the direction of bias could be in either direction. For instance, an individual who is receptive to limiting family size may be more likely to remember hearing a message that encourages smaller family sizes – this would tend to bias the treatment effect upwards. On the other hand, if the government targeted the messages to needy areas where large families are the norm, the treatment effect could be biased downwards since women who are less likely to heed the message would be more likely to have heard a message. Similarly, if the treatment is access to facilities that provide family planning, one could get biased results for the effect of the treatment on contraceptive use if, for example, the program is targeted to high need areas where the women would be less likely to take advantage of access to these types of facilities.

Allowing the treatment effect to be endogenous requires that we modify the statistical assumptions:

1. $E(X_{ti}\mu_i) \neq E(T_{ti}\mu_i) \neq E(Z_i\mu_i) \neq 0$ (there is correlation between observed variables and the time-invariant error);
2. $E(X_{ti}\varepsilon_{ti}) = E(T_{ti}\varepsilon_{ti}) = E(Z_i\varepsilon_{ti}) = 0$ (there is no correlation between observed variables and the time-varying error);
3. $E(\mu_i) = E(\varepsilon_{ti}) = E(\varepsilon_{ti}\mu_i) = 0$ (the error terms have mean zero and are not correlated with each other);
4. $Var(\mu_i) = \sigma_\mu^2$ for all individuals; and $Var(\varepsilon_{ti}) = \sigma_\varepsilon^2$ for all observations.

The only change from section II is for assumption 1 (nonetheless, we provide the entire list of assumptions for clarity). The key change that we make in assumption 1 is to allow for correlation between the treatment variable and the time invariant error term. However, the estimation methods that we discuss below can achieve unbiased results even if some or all of the other variables in the model are also correlated with the time-invariant error.

A. Continuous Dependent Variable

The estimation methods are:

1. **Fixed Effects (FE):** FE is an unbiased estimator. To define the estimator, first average equation (8) over individuals to obtain the "between" regression equation:

$$\overline{Y}_i = \overline{X}_i\beta + \overline{T}_i\alpha + Z_i\delta + \mu_i + \overline{\varepsilon}_i \tag{20}$$

Subtract equation (20) from equation (8):

$$Y_{ti} - \overline{Y}_i = (X_{ti} - \overline{X}_i)\beta + (T_{ti} - \overline{T}_i)\alpha + \varepsilon_{ti} - \overline{\varepsilon}_i \tag{21}$$

This is commonly referred to as the "within regression equation". In the within regression, the source of correlation in the error terms (the μ's) is removed. OLS can thus be applied to the within regression equation. Note that the cost is very high in terms of lost degrees of freedom: if there are 1,000 individuals, then 1,000 degrees of freedom are lost in the averaging. The reason is that one observation per individual is effectively tied down by the need to insure that the sum of the deviations from the mean over time for the outcome and regressors for each individual equals zero.

2. **First Differences (FD):** This estimator solves the endogeneity problem in a manner similar to fixed effects by differencing equation (8) as follows:

$$Y_{ti} - Y_{t-1,i} =$$

$$(X_{ti} - X_{t-1,i})\beta + (T_{ti} - T_{t-1,i})\alpha + \varepsilon_{ti} - \varepsilon_{t-1,i} \quad (22)$$

Again, the source of the correlation is removed and OLS provides a consistent estimator. This consistency is achieved at a high cost in terms of lost degrees of freedom since one time period must be dropped in order to do the differencing. However, this cost is no greater than with the FE estimator.

Both FE and FD estimators have the advantage of obtaining a statistically correct estimate of the treatment effect without resorting to instrumental variables methods (discussed below) that may rely on questionable exclusion restrictions. Thus, both are unbiased estimators of program impact under the same circumstances and the choice between them usually comes down to the comparative precision of their estimates. The FD estimator is generally preferred on this basis in instances where the ε's are correlated over time for each individual. However, before resorting to these methods, the researcher should test to see whether or not there is empirical support for the assumption that the treatment effect is correlated with the time invariant error. The Hausman-Taylor test (Hausman & Taylor, 1981; Hausman, 1978) provides a test for exogeneity by testing whether or not the estimates of β change substantially between the fixed effects and the maximum likelihood estimator that assumes an exogenous treatment effect.

The null hypothesis of the Hausman-Taylor test is that the treatment effect is exogenous. If the null is true, the FE and maximum likelihood estimators should give similar results. However, if the null is false, the results will be different and only FE will yield a consistent estimator. Thus, if the results change substantially, this is evidence that endogeneity is a problem and that first differences or fixed effects should be used.

Stata Examples

We see that, in general, the statistical significance of the coefficient estimates is reduced in the fixed effects estimations relative to the random effects results. This is the expected result due to the large loss in degrees of freedom. However, we see that exposure to a family planning message still has a significant and negative effect on ideal family size – the desired result. Education, employment and the Catholic and traditional religion variables are all insignificant in this model.

Example 22: Fixed Effects

Malawi

```
. xtreg ideal age $education $work $religion $fptype, fe

Fixed-effects (within) regression               Number of obs      =      4238
Group variable: respondentid                    Number of groups   =      2299

R-sq:  within  = 0.0413                         Obs per group: min =         1
       between = 0.0753                                        avg =       1.8
       overall = 0.0622                                        max =         3

                                                F(9,1930)          =      9.24
corr(u_i, Xb)  = -0.0074                        Prob > F           =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0482504   .0074202     6.50   0.000      .033698    .0628028
        edu1 |  -.1467206    .229127    -0.64   0.522    -.5960831    .3026419
        edu2 |  -.0141241   .5389645    -0.03   0.979    -1.071138     1.04289
    employed |   .1874763   .1343902     1.40   0.163     -.076089    .4510417
    catholic |   .1635772   .2754351     0.59   0.553    -.3766044    .7037589
  protestant |   .4224988   .1669302     2.53   0.011     .0951164    .7498813
      muslim |   .4385648   .4831697     0.91   0.364    -.5090248    1.386154
 traditional |   .3376532   .1846207     1.83   0.068    -.0244238    .6997303
      fpmess |  -.6487476   .1261282    -5.14   0.000    -.8961096   -.4013857
       _cons |   2.911137   .3971029     7.33   0.000     2.132342    3.689933
-------------+----------------------------------------------------------------
     sigma_u |  1.9601946
     sigma_e |  2.2562679
         rho |   .4301261   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(2298, 1930) =     1.20          Prob > F = 0.0000
------------------------------------------------------------------------------
```

In general, the statistical significance of the coefficient estimates is reduced in the fixed effects estimations relative to the random effects results. Once again, this is an expected result due to the large loss in degrees of freedom. However, we see that the presence of a midwife still has a significant and negative effect on ideal family size – the desired result. The number of posyandus has a positive effect in the fixed effects results with a p-value of 0.09. However, the point estimate of the coefficient is quite small.

Example 23: Indonesia

```
. xtreg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, fe

Fixed-effects (within) regression               Number of obs      =     19389
Group variable: ind_id                          Number of groups   =      9177

R-sq:  within  = 0.0107                         Obs per group: min =         1
       between = 0.0072                                        avg =       2.1
       overall = 0.0044                                        max =         4

                                                F(23,10189)        =      4.81
corr(u_i, Xb)  = -0.0823                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0067896   .0040511     1.68   0.094    -.0011513    .0147306
   midwife_0 |   -.150851   .0254811    -5.92   0.000    -.2007989   -.1009031
        edu1 |    .105857   .0573394     1.85   0.065    -.0065395    .2182536
        edu2 |  -.1076671   .1026222    -1.05   0.294    -.3088269    .0934927
        edu3 |  -.0109061   .1490359    -0.07   0.942    -.3030458    .2812336
       hedu1 |  -.0719232   .0636732    -1.13   0.259    -.1967353    .0528889
       hedu2 |  -.1149624   .0760482    -1.51   0.131    -.2640317     .034107
       hedu3 |  -.1318419   .0980391    -1.34   0.179    -.3240178     .060334
         age |   .0024369   .0023766     1.03   0.305    -.0022216    .0070954
    work_any |   .1789616   .0416408     4.30   0.000     .0973374    .2605857
   work_most |  -.1496873   .0413381    -3.62   0.000    -.2307182   -.0686565
      smoker |  -.1423897   .0984873    -1.45   0.148    -.3354443    .0506649
    goodhlth |   .0499408   .0324426     1.54   0.124     -.013653    .1135346
      muslim |  -.3799097   .1759995    -2.16   0.031    -.7249033   -.0349161
    electric |  -.1476358   .0391097    -3.77   0.000    -.2242985   -.0709732
        read |   .0216131   .0668014     0.32   0.746    -.1093308     .152557
       write |  -.0002823   .0596983    -0.00   0.996    -.1173028    .1167382
       urban |   .0592473   .0725455     0.82   0.414    -.0829562    .2014509
    pubtrans |   -.011627   .0273987    -0.42   0.671    -.0653339    .0420798
      caphrs |  -.0030822   .0070283    -0.44   0.661     -.016859    .0106946
     fhead50 |  -.0301402   .0431383    -0.70   0.485    -.1146997    .0544194
     fhead75 |   .0343107   .0556093     0.62   0.537    -.0746945    .1433158
    fhead100 |   .1525193     .07152     2.13   0.033     .0123261    .2927126
       _cons |   3.325014   .1929856    17.23   0.000     2.946724    3.703304
-------------+----------------------------------------------------------------
     sigma_u |  1.3334556
     sigma_e |  .97075394
         rho |   .6536024   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(9176, 10189) =     3.63         Prob > F = 0.0000
```

Example 24: First Differences

Malawi

```
. xtivreg ideal age $education $work $religion (fpmess = contraception_meth), fd reg

First-differenced IV regression
Group variable:    respondentid                    Number of obs      =      1812
Time variable:     year2                           Number of groups   =      1276

R-sq:  within  = 0.0097                            Obs per group: min =         1
       between = 0.1860                                           avg =       1.5
       overall = 0.0085                                           max =         2

                                                   Wald chi2(9)       =     15.34
corr(u_i, Xb)  = -0.1206                           Prob > chi2        =    0.0821

------------------------------------------------------------------------------
 D.ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      fpmess |
         D1. |  -.1099379   .1789231    -0.61   0.539    -.4606207     .240745
             |
         age |
         D1. |    .018861   .0090677     2.08   0.038     .0010886    .0366335
             |
        edu1 |
         D1. |  -.0126092   .2385562    -0.05   0.958    -.4801707    .4549522
             |
        edu2 |
         D1. |   .2500164   .5362902     0.47   0.641     -.801093    1.301126
             |
    employed |
         D1. |   -.078331    .140411    -0.56   0.577    -.3535315    .1968695
             |
    catholic |
         D1. |   .1018224   .2913085     0.35   0.727    -.4691318    .6727766
             |
  protestant |
         D1. |   .5244873   .1718842     3.05   0.002     .1876005    .8613741
             |
      muslim |
         D1. |   1.220235   .5691658     2.14   0.032     .1046901    2.335779
             |
 traditional |
         D1. |   .1692409   .1897316     0.89   0.372    -.2026263    .5411081
             |
       _cons |   .4020244   .1066796     3.77   0.000     .1929362    .6111126
-------------+----------------------------------------------------------------
     sigma_u |  2.1256451
     sigma_e |  3.1450473
         rho |  .31356449   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:   fpmess
Instruments:    age edu1 edu2 employed catholic protestant muslim traditional fpmess
------------------------------------------------------------------------------
```

In order to estimate a model in first differences in Stata, one uses the xtivreg with the "reg' or "regress" option. This option overrides the use of instrumental variables. However, instruments still must be specified even though they are not used as instruments. The first difference results are substantially different from the fixed effects results, as the family planning message and education variables are no longer significant. The time period loss created by the first difference estimator cuts the number of observations from the fixed effects by more than half.

Example 25: Indonesia

```
. xtivreg ideal midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community (num_pos_0=cont_meth) if god==0, fd reg

First-differenced IV regression
Group variable:    ind_id                      Number of obs      =      9805
Time variable:     year                        Number of groups   =      5545
R-sq:  within  = 0.0187                        Obs per group: min =         1
       between = 0.1664                                       avg =       1.8
       overall = 0.0016                                       max =         3
                                               Wald chi2(23)      =    145.51
corr(u_i, Xb)  = -0.1424                       Prob > chi2        =    0.0000
------------------------------------------------------------------------------
     D.ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |
         D1. |   .0028664   .0040737     0.70   0.482    -.0051179    .0108506
   midwife_0 |
         D1. |  -.1753167   .0244562    -7.17   0.000      -.22325   -.1273833
        edu1 |
         D1. |   .1948845   .0548281     3.55   0.000     .0874234    .3023456
        edu2 |
         D1. |   .1345734   .0997964     1.35   0.178    -.0610239    .3301707
        edu3 |
         D1. |   .2567408   .1490602     1.72   0.085    -.0354118    .5488933
       hedu1 |
         D1. |  -.0058228   .0622117    -0.09   0.925    -.1277556    .1161099
       hedu2 |
         D1. |  -.1434721   .0763147    -1.88   0.060    -.2930463     .006102
       hedu3 |
         D1. |  -.2058475   .0995346    -2.07   0.039    -.4009317   -.0107632
         age |
         D1. |   .0307326   .0067859     4.53   0.000     .0174325    .0440328
    work_any |
         D1. |   .1648989   .0402612     4.10   0.000     .0859883    .2438095
   work_most |
         D1. |  -.1366391   .0400864    -3.41   0.001     -.215207   -.0580711
       smoker |
         D1. |   -.105327   .0950191    -1.11   0.268    -.2915611     .080907
    goodhlth |
         D1. |   .0357445    .030415     1.18   0.240    -.0238679    .0953568
      muslim |
         D1. |  -.2744924   .1784668    -1.54   0.124    -.6242809     .075296
    electric |
         D1. |  -.1085738   .0400558    -2.71   0.007    -.1870817   -.0300659
        read |
         D1. |  -.0132184   .0640084    -0.21   0.836    -.1386726    .1122357
       write |
         D1. |   .0248241   .0564856     0.44   0.660    -.0858857    .1355339
       urban |
         D1. |  -.0209795   .0784894    -0.27   0.789    -.1748159    .1328569
    pubtrans |
         D1. |   .0101175   .0256848     0.39   0.694    -.0402238    .0604588
      caphrs |
         D1. |   .0036469   .0065796     0.55   0.579    -.0092488    .0165427
     fhead50 |
         D1. |  -.0041504   .0441422    -0.09   0.925    -.0906676    .0823668
     fhead75 |
         D1. |   .0255665   .0568658     0.45   0.653    -.0858884    .1370214
    fhead100 |
         D1. |    .165295   .0719034     2.30   0.022     .0243669    .3062231
       _cons |    -.17695   .0335118    -5.28   0.000     -.242632    -.111268
-------------+----------------------------------------------------------------
     sigma_u |  1.4944408
     sigma_e |  1.2481342
         rho |  .58908995   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:   num_pos_0
Instruments:    midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 age work_any work_most smoker goodhlth muslim electric
read write urban pubtrans caphrs fhead50 fhead75
                fhead100 num_pos_0
------------------------------------------------------------------------------
```

In this instance, the first difference and fixed effects results are similar.

Example 26: Hausman-Taylor Test

Malawi

```
. hausman consistent efficient, equations(1:1)

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |   consistent   efficient      Difference          S.E.
-------------+----------------------------------------------------------------
         age |    .0482504     .0559808       -.0077304        .0066161
        edu1 |   -.1467206    -.1444928       -.0022278        .2097115
        edu2 |   -.0141241    -.6641885        .6500644        .5103183
    employed |    .1874763     .3092389       -.1217626        .0976417
    catholic |    .1635772     .3439155       -.1803382        .2440331
  protestant |    .4224988     .3117987        .1107002        .1265491
      muslim |    .4385648    -.0028391        .4414038         .465383
 traditional |    .3376532     .4035109       -.0658576        .1306219
      fpmess |   -.6487476     -.279926       -.3688216        .0936132
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       22.58
                Prob>chi2 =      0.0072
--------------------------------------------------------------------------
```

The fixed effects results were stored in "consistent" while the maximum likelihood results were stored in "efficient"using the `estimates store` command. The results of the test strongly indicate that they are statistically significant differences in the two sets of results. This means one or more of the right-hand-side variables is correlated with the time invariant error terms. This does not necessarily mean that the treatment variables are the source of correlation. Fixed effects is a "brute force" method to correct for endogeneity bias, by which we mean that it addresses any endogeneity bias and not necessarily endogeneity bias associated with any particular variable (such as treatment status). There are more targeted methods such as instrumental variables that may have a higher level of efficiency but these methods are beyond the scope of this paper. The results are suggestive of endogeneity.

Example 27: Indonesia

```
. hausman consistent efficient, equations(1:1)

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |   consistent   efficient      Difference          S.E.
-------------+----------------------------------------------------------------
   num_pos_0 |     .0067896    -.0028755        .0096651        .0032435
   midwife_0 |     -.150851    -.2058716        .0550206        .0156129
        edu1 |      .105857     .0625607        .0432963        .0374049
        edu2 |    -.1076671    -.2524512        .1447841        .0837834
        edu3 |    -.0109061    -.3509845        .3400784        .1233385
       hedu1 |    -.0719232    -.0603185       -.0116047        .0383209
       hedu2 |    -.1149624    -.0521487       -.0628137        .0450638
       hedu3 |    -.1318419    -.0174879        -.114354        .0582826
         age |     .0024369     .0253115       -.0228746        .0019727
    work_any |     .1789616     .0622693        .1166922        .0224816
   work_most |    -.1496873    -.0955572       -.0541301        .0218884
      smoker |    -.1423897    -.1041001       -.0382896        .0667499
    goodhlth |     .0499408    -.0103777        .0603184        .0171352
      muslim |    -.3799097    -.0694443       -.3104654        .1708631
    electric |    -.1476358    -.2788177        .1311819        .0236777
        read |     .0216131    -.0332986        .0549117        .0343007
       write |    -.0002823    -.1243028        .1240205        .0287233
       urban |     .0592473     .0620617       -.0028143        .0656616
    pubtrans |     -.011627     .0612281       -.0728552        .0160849
      caphrs |    -.0030822      .013002       -.0160842        .0056696
     fhead50 |    -.0301402    -.0540001          .02386        .0307718
     fhead75 |     .0343107    -.1146147        .1489254        .0435113
    fhead100 |     .1525193    -.1318138        .2843331        .0568802
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                 chi2(23) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      234.77
                Prob>chi2 =      0.0000
```

B. Binary Dependent Variable

There is no statistically correct fixed effects probit (Heckman & Macurdy, 1980, present a fixed effects Tobit estimator that is also inconsistent). There is a consistent fixed effects estimator for logit. It is based on the fact that

$$e^a e^b = e^{a+b}$$

and involves conditional maximum likelihood. Suppose that we have two observations for individual *i*. Then $Y_{1i} + Y_{2i}$ can only take on one of three value, depending on the contraceptive use choice of individual *i* across the two time periods: 0, 1, or 2. Furthermore, by definition

$$P(Y_{1i} = 0 \; and \; Y_{2i} = 0 \mid Y_{1i} + Y_{2i} = 0) = 1$$

and

$$P(Y_{1i} = 1 \; and \; Y_{2i} = 1 \mid Y_{1i} + Y_{2i} = 2) = 1$$

These two equations establish that individuals that do not change their contraceptive choice over time provide no information (conditional on knowing the sum of their binary choices over time regarding contraception). For individuals that do change, we have:

$$P(Y_{1i} = 0 \; and \; Y_{2i} = 1 \mid Y_{1i} + Y_{2i} = 1) = \frac{P(Y_{1i} = 0 \; and \; Y_{2i} = 1)}{P(Y_{1i} = 0 \; and \; Y_{2i} = 1) + P(Y_{1i} = 1 \; and \; Y_{2i} = 0)}$$

Given this, it is easy to show that:

$$P(Y_{1i} = 0 \; and \; Y_{2i} = 1 | Y_{1i} + Y_{2i} = 1) = \frac{e^{(X_{2i} - X_{1i})\beta + (T_{2i} - T_{1i})\alpha}}{1 + e^{(X_{2i} - X_{1i})\beta + (T_{2i} - T_{1i})\alpha}} \qquad (23)$$

and

$$P(Y_{1i} = 1 \; and \; Y_{2i} = 0 | Y_{1i} + Y_{2i} = 1) = 1 - P(Y_{1i} = 0 \; and \; Y_{2i} = 1 | Y_{1i} + Y_{2i} = 1) \qquad (24)$$

Using (23) and (24), we can specify the likelihood function but the estimation method is to simply estimate a binary logit model for the two outcomes conditional on individuals who change states as a function of the time-varying independent variables in differences.

Extensions to data with more than two time periods are straightforward and presented in the Stata manual.

Stata Examples

Example 28: Logit Fixed Effects

Indonesia

```
. xtlogit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community, fe

Conditional fixed-effects logistic regression   Number of obs      =      7770
Group variable: ind_id                          Number of groups   =      2550

                                                Obs per group: min =         2
                                                               avg =       3.0
                                                               max =         4

                                                LR chi2(23)        =    138.40
Log likelihood  = -2779.8826                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0021844   .0099269     0.22   0.826     -.017272    .0216408
   midwife_0 |  -.0848474   .0657373    -1.29   0.197    -.2136901    .0439953
        edu1 |   .0369176   .1433714     0.26   0.797    -.2440853    .3179205
        edu2 |  -.5218017   .2740664    -1.90   0.057    -1.058962    .0153586
        edu3 |   .1106149   .3839324     0.29   0.773    -.6418787    .8631085
       hedu1 |   -.260134   .1672891    -1.55   0.120    -.5880145    .0677465
       hedu2 |  -.1632933   .2003679    -0.81   0.415    -.5560072    .2294205
       hedu3 |   .0327835   .2559052     0.13   0.898    -.4687815    .5343485
         age |   -.039268   .0059091    -6.65   0.000    -.0508496   -.0276864
    work_any |   .1811883   .1056377     1.72   0.086    -.0258577    .3882344
   work_most |   .0116205   .1045868     0.11   0.912     -.193366    .2166069
      smoker |    .254633   .2582713     0.99   0.324    -.2515693    .7608354
    goodhlth |   .2816237   .0808208     3.48   0.000     .1232179    .4400296
      muslim |   .8214671   .4859159     1.69   0.091    -.1309105    1.773845
    electric |  -.2748318   .1004257    -2.74   0.006    -.4716626   -.0780011
        read |   -.008263   .1764357    -0.05   0.963    -.3540707    .3375446
       write |  -.1896957   .1565802    -1.21   0.226    -.4965874    .1171959
       urban |   .2664627   .1738726     1.53   0.125    -.0743214    .6072468
    pubtrans |  -.0948852   .0687808    -1.38   0.168    -.2296932    .0399227
      caphrs |   .0068509   .0183796     0.37   0.709    -.0291723    .0428742
     fhead50 |   .0457876   .1095415     0.42   0.676    -.1689098     .260485
     fhead75 |   .1717072    .137577     1.25   0.212    -.0979386    .4413531
    fhead100 |   .1844447   .1775273     1.04   0.299    -.1635024    .5323917
------------------------------------------------------------------------------
```

As was true for the continuous dependent variable case, the large loss in degrees of freedom again causes a large decrease in the precision of the parameter estimates. We also see that neither policy variable is significant in the fixed effects results while both were significant in the random effects results although the midwife effect was counter-intuitive.

The Hausman-Taylor test strongly rejects the null hypothesis of no correlation between the regressors and the time invariant error term. It is also clear that the two policy variables changed a great deal between the fixed and random effects results and likely contributed to this rejection. However, more sophisticated methods would have to be used to pinpoint the source or sources of the endogeneity bias.

Example 29: Hausman-Taylor Test

```
. hausman consistent efficient

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |   consistent   efficient      Difference          S.E.
-------------+----------------------------------------------------------------
   num_pos_0 |    .0021844      .019734       -.0175496        .0084407
   midwife_0 |   -.0848474    -.1202664         .035419        .0468642
        edu1 |    .0369176     .4251938       -.3882762        .1062193
        edu2 |   -.5218017     -.035901       -.4859007        .2410138
        edu3 |    .1106149    -.2257908        .3364056        .3372154
       hedu1 |    -.260134     .2705757       -.5307097        .1237706
       hedu2 |   -.1632933     .4939675       -.6572608        .1453606
       hedu3 |    .0327835     .6974616       -.6646781        .1826314
         age |    -.039268    -.0376674       -.0016006        .0051948
    work_any |    .1811883     .3284708       -.1472825        .0658832
   work_most |    .0116205    -.1230638        .1346843         .064113
      smoker |     .254633    -.4641005        .7187336        .2032972
     goodhlth |    .2816237     .3943995       -.1127757        .0498615
      muslim |    .8214671    -.0014422        .8229093        .4790841
    electric |   -.2748318     -.106157       -.1686748        .0713232
        read |    -.008263     .2052651       -.2135282        .1142095
       write |   -.1896957    -.0006182       -.1890776        .0961475
       urban |    .2664627     .2176307         .048832        .1624287
    pubtrans |   -.0948852    -.1293711        .0344859        .0469895
      caphrs |    .0068509    -.0193324        .0261834        .0162601
     fhead50 |    .0457876    -.1156362        .1614237        .0880963
     fhead75 |    .1717072     .1034831        .0682241        .1166472
    fhead100 |    .1844447    -.1041968        .2886415         .151638
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtlogit
            B = inconsistent under Ha, efficient under Ho; obtained from xtlogit

    Test:  Ho:  difference in coefficients not systematic

                 chi2(23) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      113.14
                Prob>chi2 =      0.0000
```

IV. Extensions to Multilevel Models

To this point, the estimation methods that have been discussed assume that we select a simple random sample of individuals completely randomly (i.e., individuals are randomly and directly selected, as opposed to their selection being the last step in some multi-stage selection process) from some population (a country, state, or city, for example) and then interview these individuals at multiple points in time. This may be a reasonable approximation to the truth in many cases. However, in many other cases some type of cluster or enumeration area based sampling method is used, making it necessary to take survey design into account so that correct statistical inferences can be made. The basic idea is that classical standard errors based on standard default formulas presume simple independent random and direct (i.e., without any intervening multistage selection process) sampling of individuals, in which case the number of individuals in the sample is indicative of the true number of observations worth of information contained in the sample. Any design that deviates from that simple framework has implications for the true degree of information in the sample (given the actual number of observations) and standard errors must be computed by alternative means (i.e., modified formulas) to account for this.

Suppose, for instance, that we selected a sample of 1,000 individuals from some population of interest. Simple random sampling would imply randomly selecting the 1,000 individuals from a complete list (i.e., a sampling frame) of all individuals in that population. In this case the 1,000 individuals were completely independently selected and hence contribute 1,000 observations worth of true variation/information.

However, such an exhaustive list of all of the individuals in a population is rarely available in practice. More typically, a multi-stage process is pursued. For instance, the first stage might involve random selection of 10 communities. Once these communities are selected, all individuals in them could be enumerated, and then 100 individuals could be selected within each selected community from the list of individuals in that community. In this multi-stage process, the selection of individuals is not completely independent across the population. Instead, the selected individuals are "clustered" in communities. This implies that their behaviors, characteristics and outcomes might be correlated to the extent that there are community-level norms or circumstances that exert a common influence on those variables. However, this implies that the 1,000 individuals selected in this manner do not really contribute 1,000 truly independent observations worth of information/variation since the information/variation is correlated among individuals in each community.

The default standard errors produced by statistical packages essentially assume that the individuals were directly and independently selected (i.e., simple independent random sampling). Modified standard errors need to be explicitly indicated to account for factors like the loss of information associated with multi-stage cluster sampling.

In our exposition below, we will use "community" to refer to the cluster or enumeration area. The form for the statistical model is modified as follows:

$$Y_{tij} = X_{tij}\beta + T_{tij}\alpha + Z_{ij}\delta + C_{tj}\gamma + \lambda_j + \mu_{ij} + \varepsilon_{tij} \qquad (25)$$

where:

Y_{tij} : continuous outcome for individual *i* from community *j* at time *t*;

X_{tij} : time-varying, individual level explanatory variables (asset ownership and possibly education, for example);

T_{tij}: time-varying program variable or treatment variable. The treatment variable could be at the community level rather than at the individual level – the presence of a community organizer in a community, for example;

Z_{ij}: time-invariant, individual-level regressors (race and sex, for example);

C_{tj}: time-varying, community-level regressors (presence of a health facility or pharmacy in the community, for example);

λ_j: time-invariant community level unobserved heterogeneity;

μ_{ij}: time-invariant, individual level unobserved heterogeneity;

ε_{tij}: time-varying, individual level unobserved heterogeneity.

The communities, individuals and time periods are indexed, respectively, by:

j=1,2,…J (communities);
i=1,2,…,N_j (persons in community j);
t=1,2,…,M_{ij} (observations per person).

This is thus a framework that allows for a tremendous amount of rich variation, but also suggested some degree of concentration of information at the community and individual levels.

A. *Methods for Models for When the Treatment is Exogenous*

The key statistical assumptions are as follows:

1. $E(X_{tij}\lambda_j) = E(T_{tij}\lambda_j) = E(Z_{ij}\lambda_j) = E(C_{tj}\lambda_j) = 0$ (there is no correlation between observed variables and the community level error);
2. $E(X_{tij}\mu_{ij}) = E(T_{tij}\mu_{ij}) = E(Z_{ij}\mu_{ij}) = E(C_{tj}\mu_{ij}) = 0$ (there is no correlation between observed variables and the time invariant, individual level error);
3. $E(X_{tij}\varepsilon_{tij}) = E(T_{tij}\varepsilon_{tij}) = E(Z_{ij}\varepsilon_{tij}) = E(C_{tj}\varepsilon_{tij}) = 0$ (there is no correlation between observed variables and the time-varying, individual level error);
4. $E(\lambda_j) = E(\mu_{ij}) = E(\varepsilon_{tij}) = E(\lambda_j\mu_{ij}) = E(\lambda_j\varepsilon_{tij}) = E(\varepsilon_{tij}\mu_{ij}) = 0$ (the error terms have mean zero and are not correlated with each other);
5. $Var(\lambda_j) = \sigma_\lambda^2$, $Var(\mu_i) = \sigma_\mu^2$, and $Var(\varepsilon_{ti}) = \sigma_\varepsilon^2$

Using the assumptions, the error term correlation presented in equation (25) is now generalized to:

$$\rho_1 = \frac{\sigma_\lambda^2}{\sigma_\lambda^2 + \sigma_\mu^2 + \sigma_\varepsilon^2} \tag{26}$$

and

$$\rho_2 = \frac{\sigma_\lambda^2 + \sigma_\mu^2}{\sigma_\lambda^2 + \sigma_\mu^2 + \sigma_\varepsilon^2} \tag{27}$$

where ρ_1 and ρ_2 are the error term correlations for observations for individuals in the same community and same individual, respectively. Just as in the basic longitudinal data model, these correlations mean that standard multivariate regression methods will yield correct point estimates of all the estimated coefficients including the coefficient $\hat{\alpha}$ that captures treatment impact (α) but the standard errors of the estimated coefficients will be biased downwards (meaning that the level of significance of the results will be overstated). However, there are methods that explicitly model the dependencies in all the observations to obtain improved estimates (smaller variance) of the model's parameters.

Given the assumptions, the following estimation methods will yield statistically correct (i.e., unbiased and consistent) point estimates of the treatment effect and all methods can be implemented in Stata:

1. **Ordinary Least Squares (OLS)**: Given the assumptions, OLS is an unbiased and consistent estimator. The OLS standard errors are incorrect due to the correlation in observations for the same individual, but correct standard errors can easily be obtained by using the cluster option with regress in Stata. It is important to cluster at the highest level of aggregation – the community in this case. By doing so, correlation at the individual level is also controlled (see Angeles, Guilkey & Mroz, 2005);

2. **Maximum Likelihood Estimator**: The maximum likelihood estimator must make a specific distributional assumption about the error terms. The most common assumption is that both error components (time-varying and time-invariant errors) are normally distributed and this is the assumption used by Stata (`xtmixed` in Stata with the `mle` option or GLLAMM in Stata with the normal distribution specified). `xtmixed` and GLLAMM use different optimization methods with `xtmixed` using the iterative EM algorithm and GLLAMM using Newton-Raphson. The EM algorithm is typically faster than Newton-Raphson and is even more so in this case since the GLLAMM implementation uses numerical rather than analytic derivatives. Newton-Raphson is typically more accurate especially in the region of the likelihood function that is close to the maximum. If the distributional assumption is correct, maximum likelihood is the asymptotically efficient estimator for all of the model's parameters and is thus the optimal estimator. An advantage of maximum likelihood is it provides standard errors for

$$\sigma_\lambda^2,\ \sigma_\mu^2 \text{ and } \sigma_\varepsilon^2$$

and so all the information that is needed for a simple and direct test of the null hypothesis that

$$\sigma_\lambda^2 = \sigma_\mu^2 = 0$$

is available.

Thus, while there are numerous options for obtaining unbiased and consistent point estimates, not all will yield correct standard error estimates.

Stata Examples

Example 30: Ordinary Least Squares without Corrected Standard Errors
Malawi

```
. regress ideal age $education $work $religion $fptype

      Source |       SS       df       MS              Number of obs =    4238
-------------+------------------------------           F(  9,  4228) =   41.30
       Model |  2097.60002     9  233.066669           Prob > F      =  0.0000
    Residual |  23857.3967  4228  5.64271445           R-squared     =  0.0808
-------------+------------------------------           Adj R-squared =  0.0789
       Total |  25954.9967  4237  6.12579577           Root MSE      =  2.3754

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0559602   .0032689    17.12   0.000     .0495514     .062369
        edu1 |  -.1638498   .0884946    -1.85   0.064    -.3373458    .0096461
        edu2 |  -.6939161   .1668292    -4.16   0.000    -1.020989   -.3668432
    employed |   .3297308   .0929446     3.55   0.000     .1475105     .511951
    catholic |   .3421634   .1249518     2.74   0.006     .0971922    .5871346
  protestant |   .2852989   .1099207     2.60   0.009     .0697965    .5008013
      muslim |  -.0502296   .1263129    -0.40   0.691    -.2978693    .1974101
 traditional |   .3910266   .1331555     2.94   0.003     .1299718    .6520813
      fpmess |  -.2495788   .0857384    -2.91   0.004    -.4176712   -.0814865
       _cons |   2.382247   .1981519    12.02   0.000     1.993765    2.770728
------------------------------------------------------------------------------
```

These results are the same as the OLS results with uncorrected standard errors presented in section II and are only duplicated here to make comparisons to the estimators that follow easier.

Example 31: Indonesia

```
. reg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0

      Source |       SS       df       MS              Number of obs =   19389
-------------+------------------------------           F( 23, 19365) =   59.37
       Model |  2888.80985    23  125.600428           Prob > F      =  0.0000
    Residual |  40969.6996 19365  2.11565709           R-squared     =  0.0659
-------------+------------------------------           Adj R-squared =  0.0648
       Total |  43858.5095 19388  2.26214718           Root MSE      =  1.4545

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |  -.0055111   .0023317    -2.36   0.018    -.0100815   -.0009407
   midwife_0 |  -.1970529    .023036    -8.55   0.000    -.2422054   -.1519005
        edu1 |   .0760984   .0465203     1.64   0.102    -.0150854    .1672821
        edu2 |  -.2197894   .0624338    -3.52   0.000     -.342165   -.0974138
        edu3 |  -.3869874   .0878118    -4.41   0.000     -.559106   -.2148688
       hedu1 |   .0487344   .0543071     0.90   0.370    -.0577123    .1551811
       hedu2 |   .0537594   .0670156     0.80   0.422    -.0775971    .1851158
       hedu3 |   .1223556   .0875321     1.40   0.162    -.0492148     .293926
         age |   .0295686   .0012787    23.12   0.000     .0270622     .032075
    work_any |  -.0232795   .0424905    -0.55   0.584    -.1065645    .0600054
   work_most |   -.077604   .0425403    -1.82   0.068    -.1609866    .0057787
      smoker |  -.1203682    .074411    -1.62   0.106    -.2662202    .0254837
    goodhlth |  -.0975583   .0331038    -2.95   0.003    -.1624446    -.032672
      muslim |   .0060752   .0337928     0.18   0.857    -.0601616     .072312
    electric |  -.2809094   .0347688    -8.08   0.000    -.3490593   -.2127595
        read |  -.0516365   .0692704    -0.75   0.456    -.1874125    .0841395
       write |  -.2062215   .0643881    -3.20   0.001    -.3324277   -.0800153
       urban |   .0973887   .0268457     3.63   0.000     .0447688    .1500086
    pubtrans |   .1252388   .0249363     5.02   0.000     .0763615    .1741161
      caphrs |   .0184083   .0038876     4.74   0.000     .0107883    .0260282
     fhead50 |   -.031062   .0305249    -1.02   0.309    -.0908935    .0287695
     fhead75 |  -.1302276   .0330637    -3.94   0.000    -.1950353   -.0654198
    fhead100 |  -.1309725   .0426041    -3.07   0.002    -.2144803   -.0474647
       _cons |   2.567558   .0909989    28.22   0.000     2.389192    2.745923
------------------------------------------------------------------------------
```

Again, these results are the same as the OLS results with uncorrected standard errors presented in section II and are only duplicated here to make comparisons to the estimators that follow easier.

Example 32: Ordinary Least Squares with Corrected Standard Errors

Malawi

```
. regress ideal age $education $work $religion $fptype, cluster(tribenum)

Linear regression                                      Number of obs =    4042
                                                       F(  7,     8) =       .
                                                       Prob > F      =       .
                                                       R-squared     =  0.0812
                                                       Root MSE      =  2.3856

                               (Std. Err. adjusted for 9 clusters in tribenum)
------------------------------------------------------------------------------
             |               Robust
   ideal_num |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0566304   .0064673     8.76   0.000     .0417168     .071544
        edu1 |  -.1666549   .1626844    -1.02   0.336    -.5418058     .208496
        edu2 |  -.7137374   .2065618    -3.46   0.009     -1.19007   -.2374051
    employed |   .3099572   .0593094     5.23   0.001     .1731894     .446725
    catholic |    .369504   .1991063     1.86   0.101     -.089636    .8286441
  protestant |   .3198403   .1269629     2.52   0.036     .0270633    .6126174
      muslim |  -.0264247   .0819372    -0.32   0.755    -.2153723    .1625229
 traditional |   .4136603   .1357412     3.05   0.016     .1006405    .7266801
      fpmess |  -.2865623   .0530556    -5.40   0.001    -.4089088   -.1642158
       _cons |   2.378067   .3846794     6.18   0.000     1.490995    3.265139
------------------------------------------------------------------------------
```

In this model, the standard error correction is at the outermost level, which is community (`tribenum`). If you compare these results to the results presented in section II where we only corrected at the individual level, you see that these standard errors are typically larger and the t statistics smaller – in some cases quite a bit smaller. This clearly indicates the importance of obtaining statistically correct standard errors so that valid inferences can be made. This (cluster correction at the outermost level of likely correlation) is the standard correct practice for models such as the present circumstance, where correlation of observations can operate at several different levels. The community level is, in the present circumstance, the outermost level of potential correlation because all individual time observations are within the respective communities from which individuals are observed/sampled.

Example 33: Indonesia

```
. reg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, cluster(com_id)

Linear regression                                      Number of obs =   19389
                                                       F( 23,    312) =   24.25
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0659
                                                       Root MSE      =  1.4545

                               (Std. Err. adjusted for 313 clusters in com_id)
------------------------------------------------------------------------------
             |               Robust
       ideal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |  -.0055111   .0053549    -1.03   0.304    -.0160473    .0050251
   midwife_0 |  -.1970529   .0464788    -4.24   0.000    -.2885044   -.1056014
        edu1 |   .0760984   .0809352     0.94   0.348    -.0831495    .2353462
        edu2 |  -.2197894   .0890406    -2.47   0.014    -.3949853   -.0445935
        edu3 |  -.3869874   .0997917    -3.88   0.000    -.5833373   -.1906375
       hedu1 |   .0487344   .0956399     0.51   0.611    -.1394464    .2369152
       hedu2 |   .0537594   .1005272     0.53   0.593    -.1440375    .2515563
       hedu3 |   .1223556   .1079102     1.13   0.258    -.0899682    .3346794
         age |   .0295686   .0018526    15.96   0.000     .0259234    .0332137
    work_any |  -.0232795   .0448386    -0.52   0.604    -.1115037    .0649446
   work_most |   -.077604   .0455946    -1.70   0.090    -.1673157    .0121077
      smoker |  -.1203682   .1033152    -1.17   0.245    -.3236508    .0829144
    goodhlth |  -.0975583   .0404172    -2.41   0.016    -.1770831   -.0180335
      muslim |   .0060752    .088226     0.07   0.945     -.167518    .1796683
    electric |  -.2809094   .0913273    -3.08   0.002    -.4606046   -.1012142
        read |  -.0516365   .0840896    -0.61   0.540     -.217091     .113818
       write |  -.2062215   .0767997    -2.69   0.008    -.3573323   -.0551107
       urban |   .0973887   .0663756     1.47   0.143    -.0332118    .2279891
    pubtrans |   .1252388   .0631616     1.98   0.048     .0009622    .2495154
      caphrs |   .0184083   .0102285     1.80   0.073    -.0017173    .0385338
     fhead50 |   -.031062   .0923331    -0.34   0.737    -.2127363    .1506123
     fhead75 |  -.1302276   .0967767    -1.35   0.179    -.3206451      .06019
    fhead100 |  -.1309725    .110666    -1.18   0.238    -.3487185    .0867735
       _cons |   2.567558   .2300797    11.16   0.000     2.114854    3.020262
------------------------------------------------------------------------------
```

The coefficient estimates are unchanged from the standard OLS results but the standard errors are typically larger and the t statistics smaller. As we have explained, this is because the `cluster()` option is a post-estimation correction to the standard error estimates and hence will not affect the parameter estimate values themselves. The standard error correction is at the outermost level which is community (`com_id`).

Example 34: Maximum Likelihood using xtmixed

Malawi

```
. xtmixed ideal age $education $work $religion $fptype || tribenum: || respondentid:

Mixed-effects REML regression                   Number of obs      =      4042

-----------------------------------------------------------
                |   No. of       Observations per Group
 Group Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
       tribenum |        9         16      449.1       1305
   respondentid |     2414          1        1.7          3
-----------------------------------------------------------

                                                Wald chi2(9)       =    341.33
Log restricted-likelihood = -9246.4523          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0576779   .0034896    16.53   0.000     .0508385    .0645173
        edu1 |  -.1244445   .0980701    -1.27   0.204    -.3166582    .0677693
        edu2 |  -.6059863   .1862655    -3.25   0.001    -.9710599   -.2409126
    employed |   .2983846   .0945989     3.15   0.002     .1129741    .4837951
    catholic |   .3617523   .1398686     2.59   0.010     .0876148    .6358899
  protestant |   .3454279   .1211591     2.85   0.004     .1079604    .5828954
      muslim |   .2438693   .1951087     1.25   0.211    -.1385368    .6262754
 traditional |   .4321477   .1437094     3.01   0.003     .1504824    .7138131
      fpmess |  -.2920399   .0901779    -3.24   0.001    -.4687853   -.1152944
       _cons |   2.285109   .2484676     9.20   0.000     1.798121    2.772096
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
tribenum: Identity           |
                   sd(_cons) |   .2673575   .1419285      .0944553    .7567604
-----------------------------+------------------------------------------------
respondentid: Identity       |
                   sd(_cons) |   .7734143   .0801427      .6312611    .9475789
-----------------------------+------------------------------------------------
                sd(Residual) |    2.25124   .0350209      2.183636    2.320937
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(2) =    32.12   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

The maximum likelihood estimator is more efficient than OLS with corrected standard errors and we see that the presence of a midwife has a stronger effect in the setting of maximum likelihood estimation. As is always the case with maximum likelihood, this method is only asymptotically efficient when the distributional assumptions are correct. It is less robust than OLS. The advantage of maximum likelihood is that we obtain estimates of the error standard deviations and we see that both are strongly significant with the standard deviation of the individual random effect approximately twice the size of the standard deviation of the community random effect.

Example 35: Indonesia

```
. xtmixed ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0 || com_id: || ind_id:

Mixed-effects ML regression                     Number of obs      =     19389
-----------------------------------------------------------
                |   No. of       Observations per Group
 Group Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
         com_id |      313          6       61.9        135
         ind_id |     9216          1        2.1          4
-----------------------------------------------------------
                                                Wald chi2(23)      =    914.48
Log likelihood = -31979.226                     Prob > chi2        =    0.0000
------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0018884   .0031395     0.60   0.548    -.0042648    .0080416
   midwife_0 |  -.2318394   .0219262   -10.57   0.000    -.2748141   -.1888648
        edu1 |   .0690483   .0428206     1.61   0.107    -.0148785    .1529752
        edu2 |  -.2671518   .0584393    -4.57   0.000    -.3816908   -.1526129
        edu3 |  -.3345629   .0821795    -4.07   0.000    -.4956319    -.173494
       hedu1 |  -.0522089   .0502289    -1.04   0.299    -.1506558    .0462379
       hedu2 |  -.0456639   .0606641    -0.75   0.452    -.1645634    .0732356
       hedu3 |  -.0123991   .0780003    -0.16   0.874    -.1652768    .1404786
         age |   .0265972   .0013036    20.40   0.000     .0240422    .0291523
    work_any |   .0667326   .0346843     1.92   0.054    -.0012473    .1347125
   work_most |  -.0854164   .0346993    -2.46   0.014    -.1534258   -.0174071
      smoker |  -.1237776   .0711992    -1.74   0.082    -.2633255    .0157703
    goodhlth |   .0280553   .0272719     1.03   0.304    -.0253965    .0815072
      muslim |  -.0344901   .0576769    -0.60   0.550    -.1475347    .0785545
    electric |  -.2610365   .0321138    -8.13   0.000    -.3239784   -.1980946
        read |  -.0471661    .056725    -0.83   0.406    -.1583451    .0640129
       write |  -.1101406    .051786    -2.13   0.033    -.2116394   -.0086418
       urban |  -.0162026   .0468487    -0.35   0.729    -.1080245    .0756192
    pubtrans |   .0165636   .0244526     0.68   0.498    -.0313627    .0644899
      caphrs |   .0100769    .005538     1.82   0.069    -.0007773    .0209311
     fhead50 |  -.0981817   .0362879    -2.71   0.007    -.1693047   -.0270588
     fhead75 |  -.1368424   .0442086    -3.10   0.002    -.2234895   -.0501952
    fhead100 |   -.149778   .0546152    -2.74   0.006     -.256822   -.0427341
       _cons |   2.694524   .1060392    25.41   0.000     2.486691    2.902357
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
com_id: Identity             |
                   sd(_cons) |   .4576729    .023304      .4142031    .5057047
-----------------------------+------------------------------------------------
ind_id: Identity             |
                   sd(_cons) |   .9618635   .0112975      .9399738     .984263
-----------------------------+------------------------------------------------
                sd(Residual) |   .9637987   .0065419      .9510616    .9767063
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(2) =  5570.58   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

Example 36: Maximum Likelihood using GLLAMM

Malawi

```
. gllamm ideal age $education $work $religion $fptype, i(respondentid tribenum)
trace
gllamm model

log likelihood = -9229.6658

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0580788   .0034714    16.73   0.000      .051275    .0648825
        edu1 |  -.0981826   .0955619    -1.03   0.304    -.2854806    .0891153
        edu2 |  -.5574689   .1831681    -3.04   0.002    -.9164717    -.198466
    employed |   .3016004    .094436     3.19   0.001     .1165094    .4866915
    catholic |   .3488177   .1393605     2.50   0.012     .0756762    .6219592
  protestant |   .3447714   .1210923     2.85   0.004     .1074349     .582108
      muslim |   .1623425   .1491066     1.09   0.276     -.129901     .454586
 traditional |   .4301557   .1434832     3.00   0.003     .1489339    .7113776
      fpmess |   -.289314    .090088    -3.21   0.001    -.4658832   -.1127447
       _cons |   2.261899   .2226111    10.16   0.000     1.825589    2.698208
------------------------------------------------------------------------------

Variance at level 1
------------------------------------------------------------------------------

  5.0578804 (.15720003)

Variances and covariances of random effects
------------------------------------------------------------------------------


***level 2 (respondentid)

    var(1): .59358035 (.12327622)

***level 3 (tribenum)

    var(1): .06758045 (.04238911)
------------------------------------------------------------------------------
```

GLLAMM reports variances for the error components rather than standard deviations of the error components as is done for `xtmixed`. In this case, we get the reassuring result that GLLAMM and xtmixed give similar results. If there were differences, one would typically expect GLLAMM to be more accurate, given its estimation procedure. However, one can always check to see which method yields the largest value for the likelihood function and in this case GLLAMM achieves a smaller negative value for the log of the likelihood function. On this basis, one may have slightly more confidence in the results for GLLAMM.

Example 37: Indonesia

```
. gllamm ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, i(ind_id com_id) trace dot

gllamm model

log likelihood = -31985.329

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |    .001867   .0030293     0.62   0.538    -.0040703    .0078044
   midwife_0 |  -.2335017   .0217259   -10.75   0.000    -.2760837   -.1909196
        edu1 |   .0738165   .0426193     1.73   0.083    -.0097158    .1573488
        edu2 |  -.2662508   .0583332    -4.56   0.000    -.3805818   -.1519197
        edu3 |  -.3288973   .0822209    -4.00   0.000    -.4900473   -.1677473
       hedu1 |  -.0524967   .0502602    -1.04   0.296    -.1510049    .0460115
       hedu2 |  -.0443849   .0606938    -0.73   0.465    -.1633426    .0745728
       hedu3 |  -.0143394   .0781628    -0.18   0.854    -.1675356    .1388568
         age |   .0266747   .0013043    20.45   0.000     .0241183    .0292312
    work_any |   .0651418   .0347274     1.88   0.061    -.0029225    .1332062
   work_most |  -.0843534   .0347229    -2.43   0.015     -.152409   -.0162977
      smoker |  -.1231328   .0712433    -1.73   0.084    -.2627671    .0165015
    goodhlth |   .0301819   .0272812     1.11   0.269    -.0232882    .0836521
      muslim |  -.0378647    .055832    -0.68   0.498    -.1472935     .071564
    electric |  -.2627684   .0320659    -8.19   0.000    -.3256165   -.1999204
        read |  -.0511021   .0566529    -0.90   0.367    -.1621396    .0599355
       write |  -.1089663   .0517495    -2.11   0.035    -.2103934   -.0075392
       urban |  -.0176774    .043115    -0.41   0.682    -.1021812    .0668263
    pubtrans |   .0107687   .0240028     0.45   0.654     -.036276    .0578134
      caphrs |   .0083568    .005119     1.63   0.103    -.0016763    .0183899
     fhead50 |  -.0930674   .0351675    -2.65   0.008    -.1619944   -.0241403
     fhead75 |   -.135226   .0425813    -3.18   0.001    -.2186839   -.0517682
    fhead100 |  -.1492223   .0536186    -2.78   0.005    -.2543129   -.0441316
       _cons |    2.69711   .1014748    26.58   0.000     2.498223    2.895997
------------------------------------------------------------------------------
Variance at level 1
------------------------------------------------------------------------------
  .93286448 (.01260304)
Variances and covariances of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    var(1): .92697946 (.02160505)
***level 3 (com_id)
    var(1): .20553819 (.01656457)
------------------------------------------------------------------------------
```

In this case `xtmixed` achieves a smaller negative value for the log of the likelihood function. On this basis, one may have slightly more confidence in the results for `xtmixed`.

Probit or Logit

The basic form for the statistical model is a slight variation on equation (25):

$$Y^*_{tij} = X_{tij}\beta + T_{tij}\alpha + Z_{ij}\delta + C_{tj}\gamma + \lambda_j + \mu_{ij} + \varepsilon_{tij} \qquad (28)$$

where Y^*_{tij} is a latent variable. The corresponding observed variable is Y_{tij} equals to 1 when the latent variable is positive and 0 otherwise. We further assume that assumptions 1-5 for this section hold.

Given the assumptions, the following estimation methods will yield statistically correct estimates of the treatment effect and all methods can be implemented in Stata:

1. **Probit or logit:** Simple probit or logit is a consistent estimator – probit assumes normality for the error term while logit assumes the difference of two negative extreme value random variables. The standard errors for both methods are incorrect due to the correlation in observations for the same individual, but correct standard errors can easily be obtained by using the cluster option in Stata with clustering at the community level.
2. **Random Effects Probit or Logit**: These are the maximum likelihood estimators. They are consistent and asymptotically efficient under the model assumptions. We provide some details on the estimation methods since they will provide a better understanding of the options for this estimator and establish how this estimator is related to the semi-parametric estimator.

The maximum likelihood estimator maximizes the joint probability of the observed random sample. To build the likelihood function, note that under the normality assumption for probit:

$$P\left(Y_{tij} = 1 \middle| \lambda_j, \mu_{ij}\right) =$$

$$\Phi X_{tij}\beta + T_{tij}\alpha + Z_{ij}\delta + C_{tj}\gamma + \lambda_j + \mu_{ij} \qquad (29)$$

where Φ is the standard normal cumulative distribution function. The corresponding expression for logit is:

$$P\left(Y_{tij} = 1 \middle| \lambda_j, \mu_{ij}\right) = \frac{e^{X_{tij}\beta + T_{tij}\alpha + Z_{ij}\delta + C_{tj}\gamma + \lambda_j + \mu_{ij}}}{1 + e^{X_{tij}\beta + T_{tij}\alpha + Z_{ij}\delta + C_{tj}\gamma + \lambda_j + \mu_{ij}}} \qquad (30)$$

This implies that the joint probability of the observed set of *Y*'s for individual *i* from community *j* conditional on λ_j and μ_{ij} is:

$$A_{ij}(\lambda_j, \mu_{ij}) = \prod_{t=1}^{M_{ij}} P(Y_{tij} = 1 \mid \lambda_j, \mu_{ij})^{Y_{tij}} (1 - P(Y_{tij} = 1 \mid \lambda_j, \mu_{ij}))^{1-Y_{tij}} \qquad (31)$$

The joint probability conditional on λ_j is obtained by integration:

$$A_j(\lambda_j) = \int_{-\infty}^{\infty} A_{ij}(\lambda_j, \mu_{ij}) d\mu_{ij} \qquad (32)$$

The unconditional joint probability for all individuals in community j is:

$$A_j = \int_{-\infty}^{\infty} A_j(\lambda_j) d\lambda_j \qquad (33)$$

Which leads to the following likelihood function:

$$L = \prod_{j=1}^{J} A_j \qquad (34)$$

The likelihood function is maximized with respect to the regression coefficients in (28), σ^2_λ

and

$$\sigma_{\mu}^{2}$$

which are identified up to an unknown positive scale factor. As in the basic longitudinal model for random effects logit and probit, if we assume normality for the error terms, Hermite integration can be used to evaluate the integrals in (32) and (33). However, the semi-parametric discrete factor model can also be used.

Stata Examples

As we have seen in previous sections, the result for probit and logit are very similar except for a scale factor. Therefore, we only present logit results below.

The results with uncorrected standard errors are identical to the results reported in section II. They are reproduced here to make comparisons

Example 38: Logit without Corrected Standard Errors

Indonesia

```
. logit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community


Logistic regression                               Number of obs   =      20000
                                                  LR chi2(23)     =     693.30
                                                  Prob > chi2     =     0.0000
Log likelihood = -13204.916                       Pseudo R2       =     0.0256

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0178849   .0033975     5.26   0.000      .011226    .0245439
   midwife_0 |  -.0871613   .0322669    -2.70   0.007    -.1504032   -.0239194
        edu1 |   .2732954   .0639525     4.27   0.000     .1479508      .39864
        edu2 |  -.0300828   .0869595    -0.35   0.729    -.2005203    .1403547
        edu3 |  -.1739462   .1233162    -1.41   0.158    -.4156415     .067749
       hedu1 |   .2711507    .074288     3.65   0.000     .1255489    .4167524
       hedu2 |   .4294358   .0929697     4.62   0.000     .2472186    .6116531
       hedu3 |   .5764629   .1230008     4.69   0.000     .3353857      .81754
         age |  -.0268319   .0017924   -14.97   0.000    -.0303449    -.023319
    work_any |   .2550064   .0603414     4.23   0.000     .1367394    .3732734
   work_most |  -.0996337   .0604092    -1.65   0.099    -.2180336    .0187662
      smoker |   -.455311    .104563    -4.35   0.000    -.6602507   -.2503713
    goodhlth |   .3135413   .0457627     6.85   0.000     .2238481    .4032345
      muslim |  -.0380366     .04756    -0.80   0.424    -.1312525    .0551794
    electric |  -.0298566   .0483442    -0.62   0.537    -.1246095    .0648963
        read |   .1326402   .0961124     1.38   0.168    -.0557367     .321017
       write |   .0256461    .089539     0.29   0.775    -.1498471    .2011394
       urban |   .1290811   .0376749     3.43   0.001     .0552396    .2029226
    pubtrans |  -.0907196    .034928    -2.60   0.009    -.1591772    -.022262
      caphrs |  -.0128743   .0053777    -2.39   0.017    -.0234144   -.0023342
     fhead50 |  -.1277503   .0423011    -3.02   0.003     -.210659   -.0448416
     fhead75 |   .0478089   .0463394     1.03   0.302    -.0430146    .1386325
    fhead100 |  -.1462684   .0598709    -2.44   0.015    -.2636133   -.0289236
       _cons |   .4012711   .1256459     3.19   0.001     .1550096    .6475325
------------------------------------------------------------------------------
```

to the results that follow easier.

Example 39: Logit with Corrected Standard Errors

```
. logit cont_use num_pos_0 midwife_0 edu1 edu2 edu3 /*
> */ hedu1 hedu2 hedu3 $individual $community, cluster(com_id)


Logistic regression                               Number of obs   =      20000
                                                  Wald chi2(23)   =     454.58
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -13204.916                 Pseudo R2       =     0.0256

                               (Std. Err. adjusted for 313 clusters in com_id)
------------------------------------------------------------------------------
             |               Robust
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0178849   .0046025     3.89   0.000     .0088642    .0269056
   midwife_0 |  -.0871613   .0545063    -1.60   0.110    -.1939916     .019669
        edu1 |   .2732954    .077306     3.54   0.000     .1217785    .4248123
        edu2 |  -.0300828   .1057886    -0.28   0.776    -.2374247    .1772592
        edu3 |  -.1739462   .1435462    -1.21   0.226    -.4552915    .1073991
       hedu1 |   .2711507   .1066646     2.54   0.011     .0620918    .4802095
       hedu2 |   .4294358   .1204749     3.56   0.000     .1933094    .6655622
       hedu3 |   .5764629   .1571482     3.67   0.000     .2684581    .8844677
         age |  -.0268319   .0024237   -11.07   0.000    -.0315823   -.0220816
    work_any |   .2550064   .0633424     4.03   0.000     .1308576    .3791553
   work_most |  -.0996337   .0662439    -1.50   0.133    -.2294693    .0302019
      smoker |   -.455311   .1182048    -3.85   0.000    -.6869883   -.2236338
    goodhlth |   .3135413   .0458539     6.84   0.000     .2236693    .4034133
      muslim |  -.0380366   .1237048    -0.31   0.758    -.2804936    .2044204
    electric |  -.0298566   .0735657    -0.41   0.685    -.1740428    .1143296
        read |   .1326402   .0979395     1.35   0.176    -.0593178    .3245981
       write |   .0256461   .0829582     0.31   0.757     -.136949    .1882413
       urban |   .1290811   .0678632     1.90   0.057    -.0039284    .2620905
    pubtrans |  -.0907196   .0566701    -1.60   0.109     -.201791    .0203518
      caphrs |  -.0128743   .0102608    -1.25   0.210    -.0329851    .0072365
     fhead50 |  -.1277503   .0809985    -1.58   0.115    -.2865045    .0310039
     fhead75 |   .0478089   .0858153     0.56   0.577    -.1203859    .2160038
    fhead100 |  -.1462684   .0896719    -1.63   0.103    -.3220221    .0294852
       _cons |   .4012711   .2504652     1.60   0.109    -.0896316    .8921737
------------------------------------------------------------------------------
```

The results with standard errors corrected at the community level can be compared to the results directly above and the results in section II where the correction was done at the individual level. We get the expected result that the standard errors are typically larger and the z-statistics typically smaller than with either of the previous results.

Example 40: Random Effects Logit with Normal Error Distribution Assumption

```
. gllamm cont_use num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community, i(ind_id com_id) family(binomial) link(logit) trace dot
log likelihood = -12484.249

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0111639   .0064408     1.73   0.083    -.0014597    .0237876
   midwife_0 |  -.1592398   .0509604    -3.12   0.002    -.2591204   -.0593592
        edu1 |   .4454926    .095869     4.65   0.000     .2575929    .6333924
        edu2 |   .0026777   .1298396     0.02   0.984    -.2518033    .2571587
        edu3 |  -.1877655   .1820352    -1.03   0.302    -.5445479    .1690169
       hedu1 |   .2472319   .1128426     2.19   0.028     .0260645    .4683992
       hedu2 |   .4583718   .1380799     3.32   0.001     .1877402    .7290034
       hedu3 |   .6401823   .1788969     3.58   0.000     .2895509    .9908138
         age |  -.0369026   .0028108   -13.13   0.000    -.0424117   -.0313936
    work_any |   .3251286   .0822671     3.95   0.000     .1638881    .4863691
   work_most |   -.134646   .0822925    -1.64   0.102    -.2959363    .0266444
      smoker |  -.4898388   .1577137    -3.11   0.002    -.7989519   -.1807257
    goodhlth |   .3653778   .0633404     5.77   0.000      .241233    .4895227
      muslim |   .0478536   .1043098     0.46   0.646    -.1565898    .2522971
    electric |  -.1489546   .0744712    -2.00   0.045    -.2949154   -.0029937
        read |   .1939256   .1337497     1.45   0.147     -.068219    .4560703
       write |  -.0007308   .1228252    -0.01   0.995    -.2414638    .2400021
       urban |   .2082559   .0854518     2.44   0.015     .0407735    .3757383
    pubtrans |  -.1118989   .0563445    -1.99   0.047     -.222332   -.0014657
      caphrs |  -.0169883   .0113566    -1.50   0.135    -.0392468    .0052703
     fhead50 |  -.0764323   .0794214    -0.96   0.336    -.2320953    .0792307
     fhead75 |   .1110832   .0928649     1.20   0.232    -.0709287    .2930951
    fhead100 |   .0311331   .1145303     0.27   0.786    -.1933421    .2556083
       _cons |   .6214719    .216684     2.87   0.004     .1967791    1.046165
------------------------------------------------------------------------------


Variances and covariances of random effects
------------------------------------------------------------------------------


***level 2 (ind_id)

    var(1): 2.2516484 (.13341685)

***level 3 (com_id)

    var(1): .30919036 (.04130403)
------------------------------------------------------------------------------
```

The random effects logit estimator is asymptotically efficient if the distributional assumptions are correct. This estimator produces estimators of the components of the error variance,

$$\sigma_{\lambda}^{2} \text{ and } \sigma_{\mu}^{2},$$

and we see that both are significantly different from zero at all standard levels of significance.

Example 41: Random effects Logit with the Discrete Factor Approximation

```
. gllamm cont_use num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community, i(ind_id com_id) family(binomial) link(logit) nip(3)
ip(f) trace dot

gllamm model

log likelihood = -12480.709

------------------------------------------------------------------------------
    cont_use |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0100653   .0063997     1.57   0.116    -.0024778    .0226084
   midwife_0 |  -.1530332   .0503824    -3.04   0.002    -.2517808   -.0542856
        edu1 |   .4190875   .0966563     4.34   0.000     .2296445    .6085304
        edu2 |  -.0258532    .130366    -0.20   0.843    -.2813658    .2296595
        edu3 |  -.2160236   .1822944    -1.19   0.236     -.573314    .1412667
       hedu1 |   .2446239   .1130467     2.16   0.030     .0230564    .4661914
       hedu2 |   .4579841   .1381257     3.32   0.001     .1872628    .7287055
       hedu3 |    .636729   .1788354     3.56   0.000     .2862181      .98724
         age |  -.0372525   .0028367   -13.13   0.000    -.0428122   -.0316927
    work_any |    .331776   .0823565     4.03   0.000     .1703604    .4931917
   work_most |  -.1360706   .0824103    -1.65   0.099    -.2975919    .0254507
      smoker |  -.4927903    .157679    -3.13   0.002    -.8018354   -.1837451
    goodhlth |   .3676312   .0634175     5.80   0.000     .2433352    .4919271
      muslim |  -.1012803   .0968394    -1.05   0.296    -.2910821    .0885214
    electric |  -.1527156   .0741504    -2.06   0.039    -.2980478   -.0073834
        read |     .18642   .1334903     1.40   0.163    -.0752162    .4480562
       write |   .0069645   .1229326     0.06   0.955    -.2339791     .247908
       urban |   .1911718    .083146     2.30   0.021     .0282087    .3541349
    pubtrans |  -.1046454   .0548362    -1.91   0.056    -.2121224    .0028315
      caphrs |  -.0264635    .010092    -2.62   0.009    -.0462435   -.0066835
     fhead50 |   -.097332   .0744662    -1.31   0.191     -.243283    .0486191
     fhead75 |   .1123744    .088559     1.27   0.204     -.061198    .2859468
    fhead100 |   .0231176   .1155935     0.20   0.841    -.2034414    .2496766
       _cons |   .8485291    .224656     3.78   0.000     .4082114    1.288847
------------------------------------------------------------------------------
Probabilities and locations of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)

    loc1: -1.7747, 2.6609, .45272
  var(1): 2.0741904
    prob: 0.3319, 0.1297, 0.5384

***level 3 (com_id)

    loc1: -1.8543, .6082, -.28707
  var(1): .31062982
    prob: 0.0352, 0.3824, 0.5824
------------------------------------------------------------------------------
```

The results for random effects logit assuming normality and the discrete factor estimator are quite similar in this case. This would indicate that the normality assumption is not inappropriate for this particular problem.

Multinomial Logit

The multinomial logit model can also be extended to a multilevel setting. Again, consider a dependent variable, Y_{tij}, that takes on G possible values (g=1,2,…,G). Then:

$$P\left(Y_{tij} = 1 \middle| \lambda_j, \mu_{ij}\right) = \frac{1}{1 + \sum_{h=2}^{G} e^{X_{tij}\beta_h + T_{tij}\alpha_h + Z_{ij}\delta_h + C_{tj}\gamma_h + \phi_h\lambda_j + \rho_h\mu_{ij}}} \quad (35)$$

and for g=2,3,…,G:

$$P\left(Y_{tij} = g \middle| \lambda_j, \mu_{ij}\right) = \frac{e^{X_{tij}\beta_g + T_{tij}\alpha_g + Z_{ij}\delta_g + C_{tj}\gamma_g + \phi_g\lambda_j + \rho_g\mu_{ij}}}{1 + \sum_{h=2}^{G} e^{X_{tij}\beta_h + T_{tij}\alpha_h + Z_{ij}\delta_h + C_{tj}\gamma_h + \phi_h\lambda_j + \rho_h\mu_{ij}}} \quad (36)$$

Given the assumptions for the model, the following estimators give statistically correct parameter estimates:

1. **Multinomial Logit:** Simple multinomial logit is a consistent estimator. The standard errors are incorrect due to the correlation in observations for the same individual, but correct standard errors can easily be obtained using the cluster option in Stata with clustering at the communit level.
2. **Parametric and Semi-Parametric Random Effects Multinomial Logit**: These are the maximum likelihood estimators. They are consistent and asymptotically efficient under the model assumptions.

It is straightforward to use (35) and (36) to write multinomial logit extensions of equations (31) through (34). A random effects maximum likelihood estimator can be used either assuming normality for the time-invariant error or using the discrete factor model for a semi-parametric estimator.

Stata Examples

Example 42: Multinomial Logit without Corrected Standard Errors

Indonesia

```
. mlogit new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual
$community, base(3)
Multinomial logistic regression                   Number of obs   =      19989
                                                  LR chi2(38)     =     841.25
                                                  Prob > chi2     =     0.0000
Log likelihood =  -15085.67                       Pseudo R2       =     0.0271
------------------------------------------------------------------------------
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |   .0173876   .0034314     5.07   0.000     .0106622    .0241131
   midwife_0 |  -.0534229   .0353398    -1.51   0.131    -.1226877    .0158418
  del_post_0 |  -.0810706   .0370866    -2.19   0.029     -.153759   -.0083822
  med_post_0 |  -.0148349   .0495479    -0.30   0.765     -.111947    .0822773
         age |  -.0284746   .0017951   -15.86   0.000     -.031993   -.0249562
    work_any |    .248673   .0607439     4.09   0.000     .1296171    .3677289
   work_most |  -.1006145   .0608558    -1.65   0.098    -.2198897    .0186607
      smoker |   -.490984   .1062579    -4.62   0.000    -.6992456   -.2827225
    goodhlth |   .3234357   .0463135     6.98   0.000      .232663    .4142084
      muslim |   .0357844   .0480173     0.75   0.456    -.0583279    .1298967
    electric |  -.0173764   .0482078    -0.36   0.719    -.1118619    .0771091
        read |   .2798236   .0928182     3.01   0.003     .0979034    .4617439
       write |   .0374617   .0894746     0.42   0.675    -.1379053    .2128287
       urban |   .0935631   .0376155     2.49   0.013      .019838    .1672882
    pubtrans |  -.0986045   .0350799    -2.81   0.005    -.1673598   -.0298491
      caphrs |   -.014734   .0054474    -2.70   0.007    -.0254108   -.0040572
     fhead50 |  -.1314345   .0425862    -3.09   0.002    -.2149018   -.0479671
     fhead75 |   .0447068   .0465214     0.96   0.337    -.0464734     .135887
    fhead100 |   -.175815   .0604816    -2.91   0.004    -.2943568   -.0572732
       _cons |   .7233915   .1130516     6.40   0.000     .5018143    .9449686
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |   .0157647   .0088031     1.79   0.073     -.001489    .0330185
   midwife_0 |  -.0748726   .1134741    -0.66   0.509    -.2972778    .1475326
  del_post_0 |  -.2286236   .1371954    -1.67   0.096    -.4975216    .0402744
  med_post_0 |   .2279939   .1487726     1.53   0.125     -.063595    .5195827
         age |   .0177922   .0057564     3.09   0.002     .0065099    .0290745
    work_any |   .3448375   .1887716     1.83   0.068     -.025148     .714823
   work_most |  -.1592836   .1869351    -0.85   0.394    -.5256696    .2071025
      smoker |  -.1814589   .3038884    -0.60   0.550    -.7770692    .4141513
    goodhlth |   .0452051   .1408074     0.32   0.748    -.2307723    .3211825
      muslim |  -.7289283   .1209649    -6.03   0.000    -.9660151   -.4918416
    electric |   -.056446   .1861924    -0.30   0.762    -.4213763    .3084843
        read |   .3222572   .3416637     0.94   0.346    -.3473913    .9919057
       write |   .2974415   .3256701     0.91   0.361    -.3408602    .9357432
       urban |   .5271656   .1273432     4.14   0.000     .2775776    .7767536
    pubtrans |   .1914594   .1336173     1.43   0.152    -.0704256    .4533445
      caphrs |   .0290562   .0152733     1.90   0.057    -.0008789    .0589913
     fhead50 |  -.0575051    .156341    -0.37   0.713    -.3639278    .2489175
     fhead75 |  -.0448876   .1658717    -0.27   0.787      -.36999    .2802149
    fhead100 |   .3061924   .1859207     1.65   0.100    -.0582055    .6705903
       _cons |  -4.169341   .3802088   -10.97   0.000    -4.914537   -3.424146
-------------+----------------------------------------------------------------
No_Method    |  (base outcome)
------------------------------------------------------------------------------
```

These results are the same as reported in section II. They are simply reproduced here for ease of comparison with the results that follow.

Example 43: Multinomial Logit with Corrected Standard Errors

Indonesia

```
. mlogit new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual $community,
cluster(com_id) base(3)

Multinomial logistic regression                   Number of obs   =      19989
                                                  Wald chi2(38)   =     562.94
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood =  -15085.67                 Pseudo R2       =     0.0271

                               (Std. Err. adjusted for 313 clusters in com_id)
------------------------------------------------------------------------------
             |               Robust
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |   .0173876   .0046113     3.77   0.000     .0083497    .0264256
   midwife_0 |  -.0534229   .0592882    -0.90   0.368    -.1696257    .0627798
  del_post_0 |  -.0810706   .0608804    -1.33   0.183     -.200394    .0382529
  med_post_0 |  -.0148349   .0795554    -0.19   0.852    -.1707605    .1410908
         age |  -.0284746   .0024843   -11.46   0.000    -.0333438   -.0236054
    work_any |    .248673   .0646154     3.85   0.000     .1220292    .3753168
   work_most |  -.1006145    .067699    -1.49   0.137    -.2333021    .0320731
      smoker |   -.490984   .1228072    -4.00   0.000    -.7316818   -.2502863
    goodhlth |   .3234357   .0454732     7.11   0.000     .2343098    .4125616
      muslim |   .0357844   .1291705     0.28   0.782     -.217385    .2889538
    electric |  -.0173764   .0763732    -0.23   0.820    -.1670652    .1323124
        read |   .2798236   .0986263     2.84   0.005     .0865196    .4731277
       write |   .0374617   .0833939     0.45   0.653    -.1259872    .2009107
       urban |   .0935631   .0700761     1.34   0.182    -.0437837    .2309098
    pubtrans |  -.0986045   .0581232    -1.70   0.090    -.2125238    .0153149
      caphrs |   -.014734    .011577    -1.27   0.203    -.0374246    .0079566
     fhead50 |  -.1314345   .0841499    -1.56   0.118    -.2963652    .0334962
     fhead75 |   .0447068   .0864077     0.52   0.605    -.1246492    .2140628
    fhead100 |   -.175815    .091898    -1.91   0.056    -.3559319    .0043019
       _cons |   .7233915   .2177933     3.32   0.001     .2965245    1.150258
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |   .0157647    .008358     1.89   0.059    -.0006167    .0321462
   midwife_0 |  -.0748726   .1484342    -0.50   0.614    -.3657982     .216053
  del_post_0 |  -.2286236   .1639779    -1.39   0.163    -.5500144    .0927672
  med_post_0 |   .2279939   .2043256     1.12   0.264     -.172477    .6284647
         age |   .0177922   .0055436     3.21   0.001     .0069269    .0286575
    work_any |   .3448375   .1842017     1.87   0.061    -.0161912    .7058663
   work_most |  -.1592836   .1817256    -0.88   0.381    -.5154592    .1968921
      smoker |  -.1814589   .3249111    -0.56   0.577     -.818273    .4553552
    goodhlth |   .0452051   .1528803     0.30   0.767    -.2544348     .344845
      muslim |  -.7289283   .1495037    -4.88   0.000     -1.02195   -.4359065
    electric |   -.056446   .2006096    -0.28   0.778    -.4496335    .3367416
        read |   .3222572   .3071438     1.05   0.294    -.2797335    .9242479
       write |   .2974415   .2712772     1.10   0.273     -.234252     .829135
       urban |   .5271656    .187616     2.81   0.005      .159445    .8948861
    pubtrans |   .1914594   .1630145     1.17   0.240     -.128043    .5109619
      caphrs |   .0290562   .0209581     1.39   0.166     -.012021    .0701334
     fhead50 |  -.0575051   .1939925    -0.30   0.767    -.4377235    .3227132
     fhead75 |  -.0448876    .197795    -0.23   0.820    -.4325586    .3427835
    fhead100 |   .3061924   .2228782     1.37   0.170    -.1306409    .7430256
       _cons |  -4.169341   .4144337   -10.06   0.000    -4.981616   -3.357066
-------------+----------------------------------------------------------------
No_Method    |  (base outcome)
------------------------------------------------------------------------------
```

These results with the correction at community level provide the most conservative estimates of the estimated standard errors. They are typically larger than the uncorrected results. They are also larger than the results in section II where the correction was done at the individual level.

Example 44: Random Effects Multinomial Logit with Normal Error Distribution Assumption

Indonesia

```
. gllamm new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual $community if method!=95,
i(ind_id com_id) family(binomial) link(mlogit) b(3) nip(20) ip(g) trace dot

gllamm model

log likelihood = -14352.692

------------------------------------------------------------------------------
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Modern       |
   num_pos_0 |   .0109118   .0065233     1.67   0.094    -.0018737    .0236972
   midwife_0 |  -.0968136   .0550463    -1.76   0.079    -.2047023    .0110751
  del_post_0 |  -.1887622   .0611133    -3.09   0.002     -.308542   -.0689824
  med_post_0 |    .057981   .0718603     0.81   0.420    -.0828627    .1988247
         age |  -.0385936    .002797   -13.80   0.000    -.0440756   -.0331117
    work_any |   .3165967   .0826726     3.83   0.000     .1545615     .478632
   work_most |  -.1365583   .0827699    -1.65   0.099    -.2987843    .0256678
      smoker |  -.5266573   .1592775    -3.31   0.001    -.8388354   -.2144792
    goodhlth |   .3771808   .0638705     5.91   0.000     .2519968    .5023647
      muslim |   .1296622   .1050825     1.23   0.217    -.0762957    .3356201
    electric |  -.1381372   .0744398    -1.86   0.063    -.2840365    .0077621
        read |   .3716783   .1300876     2.86   0.004     .1167112    .6266454
       write |   .0180669   .1229545     0.15   0.883    -.2229195    .2590533
       urban |   .1518452     .08541     1.78   0.075    -.0155553    .3192457
    pubtrans |  -.1178523   .0566115    -2.08   0.037    -.2288088   -.0068959
      caphrs |  -.0202995   .0113263    -1.79   0.073    -.0424986    .0018996
     fhead50 |  -.0730519   .0795946    -0.92   0.359    -.2290546    .0829507
     fhead75 |   .1208059    .092875     1.30   0.193    -.0612257    .3028376
    fhead100 |   .0066835   .1149009     0.06   0.954    -.2185182    .2318852
       _cons |    1.02463     .19629     5.22   0.000     .6399087    1.409352
-------------+----------------------------------------------------------------
Traditional  |
   num_pos_0 |   .0087279   .0105089     0.83   0.406    -.0118693     .029325
   midwife_0 |  -.1208983   .1216322    -0.99   0.320     -.359293    .1174965
  del_post_0 |  -.3442737   .1459699    -2.36   0.018    -.6303696   -.0581779
  med_post_0 |   .2997737   .1582257     1.89   0.058    -.0103431    .6098904
         age |   .0099862   .0063023     1.58   0.113    -.0023661    .0223386
    work_any |   .4090092   .1973093     2.07   0.038     .0222901    .7957284
   work_most |  -.2097923   .1954893    -1.07   0.283    -.5929442    .1733597
      smoker |  -.1795677   .3260321    -0.55   0.582    -.8185789    .4594435
    goodhlth |   .0926379   .1478445     0.63   0.531    -.1971321    .3824078
      muslim |  -.6526979   .1524549    -4.28   0.000    -.9515039   -.3538918
    electric |  -.1877292   .1953502    -0.96   0.337    -.5706086    .1951502
        read |   .4176465   .3547309     1.18   0.239    -.2776133    1.112906
       write |   .2726309   .3376742     0.81   0.419    -.3891983    .9344601
       urban |   .5835688   .1491585     3.91   0.000     .2912234    .8759141
    pubtrans |   .1833154   .1411046     1.30   0.194    -.0932446    .4598754
      caphrs |   .0253203   .0180311     1.40   0.160    -.0100199    .0606605
     fhead50 |  -.0088956   .1706305    -0.05   0.958    -.3433252    .3255341
     fhead75 |   .0092387    .184764     0.05   0.960     -.352892    .3713695
    fhead100 |   .4705968   .2108105     2.23   0.026     .0574158    .8837778
       _cons |  -3.901104   .4129904    -9.45   0.000     -4.71055   -3.091658
------------------------------------------------------------------------------

Variances and covariances of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    var(1): 2.2918416 (.13661713)
***level 3 (com_id)
    var(1): .3246459 (.0439855)
------------------------------------------------------------------------------
```

The GLLAMM version of the random effect multinomial logit model imposes the restriction that $\phi_2 = \phi_3$ and $\rho_2 = \rho_3$. These restrictions are theoretically not necessary since all four of the parameters are separately identified. We note that there are some differences in the estimated effects of the policy variables between multinomial logit with corrected standard errors and random effect multinomial logit assuming normality. In this case, one would probably have more confidence in multinomial logit with corrected standard errors since it does not impose two restrictions that may or may not be valid. The only way to test for their validity would be to estimate an unrestricted version of the model but this estimator is not available in Stata.

Example 45: Random Effects Multinomial Logit with the Discrete Factor Approximation

Indonesia

```
. gllamm new_method num_pos_0 midwife_0 del_post_0 med_post_0 $individual $community if method!=95, i(ind_id
com_id) family(binomial) link(mlogit) b(3) nip(3) ip(f) trace dot

gllamm model
log likelihood = -14348.878
------------------------------------------------------------------------------
  new_method |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
c1           |
   num_pos_0 |   .0097989   .0066013     1.48   0.138    -.0031395    .0227373
   midwife_0 |   -.087596   .0542142    -1.62   0.106    -.1938538    .0186619
  del_post_0 |  -.1741277   .0599081    -2.91   0.004    -.2915454    -.05671
  med_post_0 |   .0572019   .0702988     0.81   0.416    -.0805813    .1949851
         age |  -.0388158   .0028121   -13.80   0.000    -.0443273   -.0333043
    work_any |   .3195184   .0828348     3.86   0.000     .1571651    .4818717
   work_most |  -.1345533    .082861    -1.62   0.104    -.2969579    .0278512
      smoker |  -.5180349   .1563459    -3.31   0.001    -.8244671   -.2116026
    goodhlth |   .3766591   .0636693     5.92   0.000     .2518695    .5014487
      muslim |  -.0048056   .0966515    -0.05   0.960    -.1942392    .1846279
    electric |  -.1423141   .0740982    -1.92   0.055    -.2875438    .0029157
        read |   .3429242   .1295737     2.65   0.008     .0889645     .596884
       write |   .0281743   .1226218     0.23   0.818    -.2121599    .2685085
       urban |   .1361302   .0826929     1.65   0.100    -.0259449    .2982054
    pubtrans |   -.110912   .0545349    -2.03   0.042    -.2177985   -.0040256
      caphrs |  -.0311154   .0100428    -3.10   0.002    -.0507988   -.0114319
     fhead50 |  -.1023736   .0739946    -1.38   0.167    -.2474003    .0426531
     fhead75 |   .1157823   .0879069     1.32   0.188    -.0565119    .2880766
    fhead100 |  -.0083937   .1151798    -0.07   0.942     -.234142    .2173546
       _cons |   1.262629   .2072949     6.09   0.000     .8563379    1.668919
-------------+----------------------------------------------------------------
c2           |
   num_pos_0 |    .007671   .0105641     0.73   0.468    -.0130342    .0283762
   midwife_0 |  -.1112011   .1212333    -0.92   0.359     -.348814    .1264118
  del_post_0 |  -.3301175   .1455185    -2.27   0.023    -.6153285   -.0449065
  med_post_0 |   .2979693   .1576263     1.89   0.059    -.0109726    .6069112
         age |   .0098373   .0063131     1.56   0.119    -.0025362    .0222107
    work_any |   .4134911   .1973254     2.10   0.036     .0267404    .8002417
   work_most |  -.2086882   .1954918    -1.07   0.286     -.591845    .1744686
      smoker |  -.1720206    .324994    -0.53   0.597    -.8089971    .4649559
    goodhlth |   .0931015   .1477788     0.63   0.529    -.1965397    .3827427
      muslim |  -.7820475   .1474926    -5.30   0.000    -1.071128   -.4929672
    electric |  -.1934405   .1953123    -0.99   0.322    -.5762455    .1893645
        read |   .3912092   .3546196     1.10   0.270    -.3038326    1.086251
       write |   .2822392   .3376206     0.84   0.403     -.379485    .9439634
       urban |   .5659104   .1478284     3.83   0.000     .2761721    .8556487
    pubtrans |   .1896825   .1402808     1.35   0.176    -.0852627    .4646278
      caphrs |   .0148664   .0172979     0.86   0.390    -.0190369    .0487697
     fhead50 |  -.0398071   .1683454    -0.24   0.813     -.369758    .2901437
     fhead75 |   .0056661   .1825662     0.03   0.975    -.3521571    .3634894
    fhead100 |   .4564185   .2112735     2.16   0.031       .04233     .870507
       _cons |  -3.671866   .4186656    -8.77   0.000    -4.492436   -2.851297
------------------------------------------------------------------------------
Probabilities and locations of random effects
------------------------------------------------------------------------------
***level 2 (ind_id)
    loc1: -1.7236, 3.6596, .6133
  var(1): 2.3207744
    prob: 0.3629, 0.0771, 0.56
***level 3 (com_id)
    loc1: -1.8874, .64094, -.27983
  var(1): .31709849
    prob: 0.0339, 0.3631, 0.603
```

In terms of the z-statistics for the two policy variables, the discrete factor results are similar to the random effects results with normality. While the discrete factor model is less restrictive than random effects with normality, it is imposes restrictions on the variances that may or may not be valid. Therefore, multinomial logit with corrected standard errors is probably still the best method to use, given the limitations of the software.

B. *Methods for Models for When the Treatment is Correlated with the Community Level Error*

Programs or treatments are typically not randomly assigned to individuals (as they would under an experimental design). In fact, in many cases, the treatment may be targeted in a way that could introduce bias if not properly controlled. For example, a program may be targeted to high need areas – adding family planning clinics to areas with high fertility. If this is not taken into account, one would expect program impact to be biased downwards. On the other hand, if programs are introduced into areas where the government thinks individuals are more receptive to program activities, the bias may go in the opposite direction.

The key statistical assumptions are modified as follows to allow for program targeting:

1. $E(X_{tij}\lambda_j) = E(Z_{ij}\lambda_j) = E(C_{tj}\lambda_j) = 0$ and $E(T_{tij}\lambda_j) \neq 0.$
   (there is no correlation between observed individual level and community level variables and the community level error but we allow correlation between the treatment variable and the community level error);
2. $E(X_{tij}\mu_{ij}) = E(T_{tij}\mu_{ij}) = E(Z_{ij}\mu_{ij}) = E(C_{tj}\mu_{ij}) = 0$

   (there is no correlation between observed variables and the time-invariant, individual level error);
3. $E(X_{tij}e_{tij}) = E(T_{tij}e_{tij}) = E(Z_{ij}e_{tij}) = E(C_{tij}e_{tij}) = 0$
   (there is no correlation between observed variables and the time-varying, individual level error);
4. $E(\lambda_j) = E(\mu_{ij}) = E(\varepsilon_{tij}) = E(\lambda_j\mu_{ij}) = E(\lambda_j\varepsilon_{tij}) = E(\varepsilon_{tij}\mu_{ij}) = 0$
   (the error terms have mean zero and are not correlated with each other);
5. $Var(\lambda_j) = \sigma_\lambda^2$,
   $Var(\mu_i) = \sigma_\mu^2$,
   and,
   $Var(\varepsilon_{ti}) = \sigma_\varepsilon^2$.

The most straightforward procedure to correct for program targeting is to use community fixed effects. This approach can be implemented by simply adding community dummies to (8), (9), (25) or (28) depending on the form of the dependent variable (continuous, binary or categorical) and dropping the λ_j from the model. This procedure amounts to using a non-parametric specification for the community level heterogeneity. With the community level heterogeneity explicitly controlled for, all of the methods laid out in section II are now directly applicable. Note that this method would also control for correlation between the community level unobservable and any observed right-hand-side variable. It is thus quite flexible.

A more advanced method would try to explicitly model the program targeting. For example, if the program involved the strategic placement of family planning clinics, one would need to include an equation that explicitly modeled the placement of these facilities and this equation would then be jointly estimated with the equation of primary interest. This is the approach that was taken by Angeles, Guilkey and Mroz (1998) who compared it to the fixed effects (community dummy) approach and found that the two approaches yielded similar results. As expected, since it is less costly in terms of degrees of freedom, the joint estimation strategy yielded parameter estimates with smaller standard errors. However, this becomes less of an issue as sample sizes increase.

Since adding community dummies sacrifices a considerable number of degrees of freedom, it is important to test explicitly whether or not they are necessary – i.e. whether or not the treatment variable is actually correlated with the community level unobservable
(λ_j).
A Hausman test involves estimating the random effect specification of the model and the fixed effects specification. Under the null hypothesis of no correlation between the treatment effect and community level unobservables, these two sets of estimates both consistently estimate the models parameters and the null is rejected if there is significant differences between them.

Stata Examples

We make the procedure that is used explicit in the following Stata examples.

Example 46: Continuous Dependent Variable Example

Malawi

```
. xi: xtreg ideal age $education $work $religion $fptype i.tribenum, re
i.tribenum        _Itribenum_1-9      (naturally coded; _Itribenum_1 omitted)

Random-effects GLS regression                   Number of obs      =      4042
Group variable: respondentid                    Number of groups   =      2200

R-sq:  within  = 0.0371                         Obs per group: min =         1
       between = 0.1191                                        avg =       1.8
       overall = 0.0872                                        max =         3

                                                Wald chi2(17)      =    375.84
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0570896    .003451    16.54   0.000     .0503258    .0638534
        edu1 |  -.1472511    .097365    -1.51   0.130    -.3380831    .0435808
        edu2 |  -.6255278    .184578    -3.39   0.001    -.9872941   -.2637615
    employed |   .2998759   .0947612     3.16   0.002     .1141473    .4856046
    catholic |   .3520524   .1390493     2.53   0.011     .0795207    .6245841
  protestant |    .329537   .1214708     2.71   0.007     .0914586    .5676154
      muslim |   .3408631   .2114927     1.61   0.107    -.0736551    .7553812
 traditional |   .4280003    .144841     2.95   0.003      .144117    .7118835
      fpmess |  -.2851824   .0904986    -3.15   0.002    -.4625563   -.1078085
_Itribenum_2 |   .4853205   .1943701     2.50   0.013      .104362    .8662789
_Itribenum_3 |   .1917301    .243634     0.79   0.431    -.2857838     .669244
_Itribenum_4 |   .2847516   .2044198     1.39   0.164    -.1159038     .685407
_Itribenum_5 |   .3571505   .2462037     1.45   0.147    -.1253998    .8397009
_Itribenum_6 |   1.185205   .6218108     1.91   0.057    -.0335214    2.403932
_Itribenum_7 |   1.033567    .638523     1.62   0.106    -.2179153    2.285049
_Itribenum_8 |   .9993179   .3045411     3.28   0.001     .4024282    1.596207
_Itribenum_9 |  -.2378167   .2928369    -0.81   0.417    -.8117665    .3361331
       _cons |   1.969858   .2742675     7.18   0.000     1.432304    2.507413
-------------+----------------------------------------------------------------
     sigma_u |  .49209551
     sigma_e |  2.2735377
         rho |  .04475182   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

(Continues next page)

(Continues from previous page)

```
. estimates store consistent

. xtmixed ideal age $education $work $religion $fptype || tribenum: || respondentid:
Mixed-effects REML regression                   Number of obs      =      4042

-----------------------------------------------------------
                |   No. of       Observations per Group
 Group Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
       tribenum |        9         16      449.1       1305
   respondentid |     2414          1        1.7          3
-----------------------------------------------------------

                                                Wald chi2(9)       =    341.33
Log restricted-likelihood = -9246.4523          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0576779   .0034896    16.53   0.000     .0508385    .0645173
        edu1 |  -.1244445   .0980701    -1.27   0.204    -.3166582    .0677693
        edu2 |  -.6059863   .1862655    -3.25   0.001    -.9710599   -.2409126
    employed |   .2983846   .0945989     3.15   0.002     .1129741    .4837951
    catholic |   .3617523   .1398686     2.59   0.010     .0876148    .6358899
  protestant |   .3454279   .1211591     2.85   0.004     .1079604    .5828954
      muslim |   .2438693   .1951087     1.25   0.211    -.1385368    .6262754
 traditional |   .4321477   .1437094     3.01   0.003     .1504824    .7138131
      fpmess |  -.2920399   .0901779    -3.24   0.001    -.4687853   -.1152944
       _cons |   2.285109   .2484676     9.20   0.000     1.798121    2.772096
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
tribenum: Identity           |
                   sd(_cons) |   .2673575   .1419285      .0944553    .7567604
-----------------------------+------------------------------------------------
respondentid: Identity       |
                   sd(_cons) |   .7734143   .0801427      .6312611    .9475789
-----------------------------+------------------------------------------------
                sd(Residual) |    2.25124   .0350209      2.183636    2.320937
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(2) =    32.12   Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

(Continues next page)

(Continues from previous page)

```
. estimates store efficient

. hausman consistent efficient, equations(1:1)

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |   consistent   efficient      Difference          S.E.
-------------+----------------------------------------------------------------
         age |     .0570896     .0576779       -.0005883               .
        edu1 |    -.1472511    -.1244445       -.0228067               .
        edu2 |    -.6255278    -.6059863       -.0195416               .
    employed |     .2998759     .2983846        .0014913        .0055438
    catholic |     .3520524     .3617523       -.0096999               .
  protestant |      .329537     .3454279       -.0158909        .0086963
      muslim |     .3408631     .2438693        .0969938        .0816196
 traditional |     .4280003     .4321477       -.0041475          .01807
     fpmess |    -.2851824    -.2920399        .0068575        .0076114
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
          B = inconsistent under Ha, efficient under Ho; obtained from xtmixed

    Test:  Ho:  difference in coefficients not systematic

                  chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =    -3.24    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test
```

The first set of results includes community dummies using the "xi" command. We do not present all of the dummies to save space. These results are consistent whether or not there is correlation between the explanatory variables and the community-level error term. The second set of results is efficient if the null hypothesis of no correlation between the explanatory variables and the community-level error term is true. The last set of results is for the Hausman test that compares the common parameter estimates for the consistent and efficient estimators. In this case, the overall Hausman test itself fails in the sense that the test statistic fails to meet the asymptotic assumptions of the test, as indicated by the error. This is indicated by a χ^2 (`chi2`) statistic with a negative value (`chi2=-3.24`), which is outside of the domain of the χ^2 distribution.

A disadvantage of the Stata form of the Hausman test is it does not allow one to test subsets of coefficients. The `suest` procedure in Stata can be used to form a joint covariance matrix for parameter estimates from two sets of estimation procedures and, once one obtains this covariance matrix, tests across subsets of parameters can be performed. Unfortunately, all the information needed for the `suest` procedure to calculate the joint covariance matrix is not available for all estimators in Stata, including many of the ones used here.

Example 47: Indonesia

```
. xi: xtreg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community i.com_id if god==0, re
i.com_id          _Icom_id_1201-7316  (naturally coded; _Icom_id_1201 omitted)

Random-effects GLS regression                   Number of obs      =     19389
Group variable: ind_id                          Number of groups   =      9177

R-sq:  within  = 0.0048                         Obs per group: min =         1
       between = 0.2348                                        avg =       2.1
       overall = 0.1900                                        max =         4

                                                Wald chi2(335)     =   2793.78
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0076671   .0037621     2.04   0.042     .0002934    .0150408
   midwife_0 |  -.2423821   .0231855   -10.45   0.000    -.2878249   -.1969394
        edu1 |   .0738649   .0430932     1.71   0.087    -.0105963    .1583261
        edu2 |  -.2693007   .0587843    -4.58   0.000    -.3845157   -.1540856
        edu3 |  -.3262645   .0826179    -3.95   0.000    -.4881927   -.1643364
       hedu1 |  -.0415077   .0506429    -0.82   0.412    -.1407659    .0577505
       hedu2 |  -.0325009   .0612638    -0.53   0.596    -.1525757     .087574
       hedu3 |  -.0013419   .0788013    -0.02   0.986    -.1557897    .1531059
         age |   .0270549   .0013072    20.70   0.000     .0244929     .029617
    work_any |     .06281   .0351315     1.79   0.074    -.0060465    .1316666
   work_most |  -.0790619   .0351257    -2.25   0.024    -.1479069   -.0102168
      smoker |  -.1304029   .0715639    -1.82   0.068    -.2706655    .0098598
    goodhlth |   .0367173    .027582     1.33   0.183    -.0173425    .0907771
      muslim |   .0135949   .0660936     0.21   0.837    -.1159462     .143136
    electric |  -.2508337   .0328363    -7.64   0.000    -.3151917   -.1864758
        read |  -.0496951   .0574094    -0.87   0.387    -.1622155    .0628252
       write |  -.1105581   .0524539    -2.11   0.035    -.2133659   -.0077502
       urban |  -.0735001   .0638054    -1.15   0.249    -.1985564    .0515562
    pubtrans |   .0061531   .0255866     0.24   0.810    -.0439957     .056302
      caphrs |   .0027204   .0065324     0.42   0.677    -.0100829    .0155238
     fhead50 |  -.1143434   .0393807    -2.90   0.004    -.1915281   -.0371587
     fhead75 |   -.132316   .0494383    -2.68   0.007    -.2292133   -.0354187
    fhead100 |  -.1392386   .0625732    -2.23   0.026    -.2618798   -.0165974
_Icom_id_1202 |   2.470994   .6879829     3.59   0.000     1.122573    3.819416
_Icom_id_1203 |   .7592294    .384202     1.98   0.048     .0062073    1.512252
.
.
.
_Icom_id_7315 |  -.8097318   .4101102    -1.97   0.048    -1.613533   -.0059306
_Icom_id_7316 |  -.7554136   .4072923    -1.85   0.064    -1.553692    .0428646
       _cons |   3.476628   .3480407     9.99   0.000     2.794481    4.158775
-------------+----------------------------------------------------------------
     sigma_u |  .90119612
     sigma_e |  .97144476
         rho |  .46253957   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

(Continues next page)

(Continues from previous page)

```
. estimates store consistent
.
. * OLS w/ community and individual random effects
. xtmixed ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0 || com_id: || ind_id:

Mixed-effects ML regression                     Number of obs      =     19389

-----------------------------------------------------------
                |   No. of       Observations per Group
 Group Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
         com_id |      313          6       61.9        135
         ind_id |     9216          1        2.1          4
-----------------------------------------------------------
                                                Wald chi2(23)      =    914.48
Log likelihood = -31979.226                     Prob > chi2        =    0.0000
------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0018884   .0031395     0.60   0.548    -.0042648    .0080416
   midwife_0 |  -.2318394   .0219262   -10.57   0.000    -.2748141   -.1888648
        edu1 |   .0690483   .0428206     1.61   0.107    -.0148785    .1529752
        edu2 |  -.2671518   .0584393    -4.57   0.000    -.3816908   -.1526129
        edu3 |  -.3345629   .0821795    -4.07   0.000    -.4956319    -.173494
       hedu1 |  -.0522089   .0502289    -1.04   0.299    -.1506558    .0462379
       hedu2 |  -.0456639   .0606641    -0.75   0.452    -.1645634    .0732356
       hedu3 |  -.0123991   .0780003    -0.16   0.874    -.1652768    .1404786
         age |   .0265972   .0013036    20.40   0.000     .0240422    .0291523
    work_any |   .0667326   .0346843     1.92   0.054    -.0012473    .1347125
   work_most |  -.0854164   .0346993    -2.46   0.014    -.1534258   -.0174071
      smoker |  -.1237776   .0711992    -1.74   0.082    -.2633255    .0157703
    goodhlth |   .0280553   .0272719     1.03   0.304    -.0253965    .0815072
      muslim |  -.0344901   .0576769    -0.60   0.550    -.1475347    .0785545
    electric |  -.2610365   .0321138    -8.13   0.000    -.3239784   -.1980946
        read |  -.0471661    .056725    -0.83   0.406    -.1583451    .0640129
       write |  -.1101406    .051786    -2.13   0.033    -.2116394   -.0086418
       urban |  -.0162026   .0468487    -0.35   0.729    -.1080245    .0756192
    pubtrans |   .0165636   .0244526     0.68   0.498    -.0313627    .0644899
      caphrs |   .0100769    .005538     1.82   0.069    -.0007773    .0209311
     fhead50 |  -.0981817   .0362879    -2.71   0.007    -.1693047   -.0270588
     fhead75 |  -.1368424   .0442086    -3.10   0.002    -.2234895   -.0501952
    fhead100 |   -.149778   .0546152    -2.74   0.006     -.256822   -.0427341
       _cons |   2.694524   .1060392    25.41   0.000     2.486691    2.902357
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
com_id: Identity             |
                   sd(_cons) |   .4576729    .023304      .4142031    .5057047
-----------------------------+------------------------------------------------
ind_id: Identity             |
                   sd(_cons) |   .9618635   .0112975      .9399738     .984263
-----------------------------+------------------------------------------------
                sd(Residual) |   .9637987   .0065419      .9510616    .9767063
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(2) =  5570.58   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

(Continues next page)

(Continues from previous page)

```
. estimate store efficient


. hausman consistent efficient, equations(1:1)

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |   consistent   efficient      Difference          S.E.
-------------+----------------------------------------------------------------
   num_pos_0 |    .0076671     .0018884         .0057787         .0020731
   midwife_0 |   -.2423821    -.2318394        -.0105427          .007537
        edu1 |    .0738649     .0690483         .0048166         .0048397
        edu2 |   -.2693007    -.2671518        -.0021488         .0063589
        edu3 |   -.3262645    -.3345629         .0082984         .0084996
       hedu1 |   -.0415077    -.0522089         .0107012         .0064619
       hedu2 |   -.0325009    -.0456639          .013163         .0085508
       hedu3 |   -.0013419    -.0123991         .0110572         .0112075
         age |    .0270549     .0265972         .0004577         .0000968
    work_any |      .06281     .0667326        -.0039226         .0055882
   work_most |   -.0790619    -.0854164         .0063546         .0054563
      smoker |   -.1304029    -.1237776        -.0066253         .0072152
    goodhlth |    .0367173     .0280553          .008662         .0041249
      muslim |    .0135949    -.0344901          .048085          .032276
    electric |   -.2508337    -.2610365         .0102028         .0068502
        read |   -.0496951    -.0471661         -.002529         .0088378
       write |   -.1105581    -.1101406        -.0004174          .008344
       urban |   -.0735001    -.0162026        -.0572975         .0433165
    pubtrans |    .0061531     .0165636        -.0104105         .0075328
      caphrs |    .0027204     .0100769        -.0073565         .0034646
     fhead50 |   -.1143434    -.0981817        -.0161617         .0152979
     fhead75 |    -.132316    -.1368424         .0045264         .0221303
    fhead100 |   -.1392386     -.149778         .0105395         .0305382
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
          B = inconsistent under Ha, efficient under Ho; obtained from xtmixed

    Test:  Ho:  difference in coefficients not systematic

                 chi2(23) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       23.18
                Prob>chi2 =      0.4500
                (V_b-V_B is not positive definite)
```

In this case, the null hypothesis is not rejected at any reasonable level of significance. However, the results are suspect since the asymptotic covariance matrix for the differences in coefficients is not positive definite. This cannot happen asymptotically (i.e., as sample size gets very large), but can happen in any finite sample size.

C. *Methods for Models for When the Treatment is Correlated with the Community Level Error and the Time-Invariant Individual Level Error*

The key statistical assumptions are modified as follows to allow for program targeting:

1. $E(X_{tij}\lambda_j) = E(Z_{ij}\lambda_j) = E(C_{tj}\lambda_j) = 0$ and $E(T_{tij}\lambda_j) \neq 0$.
   (there is no correlation between observed individual level and community level variables and the community level error but we allow correlation between the treatment variable and the community level error);
2. $E(X_{tij}\mu_{ij}) = E(Z_{ij}\mu_{ij}) = E(C_{tj}\mu_{ij}) = 0$ and $E(T_{tij}\mu_{ij}) \neq 0$
   (there is no correlation between observed individual level and community level variables and the time-invariant individual level error but we allow correlation between the treatment variable and the time-invariant level error);
3. $E(X_{tij}\varepsilon_{tij}) = E(T_{tij}\varepsilon_{tij}) = E(Z_{ij}\varepsilon_{tij}) = E(C_{tj}\varepsilon_{tij}) = 0$
   (there is no correlation between observed variables and the time-varying, individual level error);
4. $E(\lambda_j) = E(\mu_{ij}) = E(\varepsilon_{tij}) = E(\lambda_j\mu_{ij}) = E(\lambda_j\varepsilon_{tij}) = E(\varepsilon_{tij}\mu_{ij}) = 0$
   (the error terms have mean zero and are not correlated with each other);
5. $Var(\lambda_j) = \sigma_\lambda^2$
   $Var(\mu_i) = \sigma_\mu^2$,
   and
   $Var(\varepsilon_{ti}) = \sigma_\varepsilon^2$.

We have already discussed circumstances where one would expect the treatment variables to be correlated with the time-invariant individual level error in section III. The solution is also the same. Using the continuous dependent variable case as an example, we can define the "between" equation for equation (25):

$$\bar{Y}_{ij} = \bar{X}_{ij}\beta + \bar{T}_{ij}\alpha + Z_{ij}\delta + \bar{C}_j\gamma + \lambda_j + \mu_{ij} + \bar{\varepsilon}_{ij} \quad (37)$$

And subtract equation (37) from equation (25):

$$Y_{tij} - \bar{Y}_{ij} = (X_{tij} - \bar{X}_{ij})\beta + (T_{tij} - \bar{T}_{ij})\alpha + (C_{tj} - \bar{C}_j)\gamma + (\varepsilon_{tij} - \bar{\varepsilon}_{ij}) \quad (38)$$

Equation (38) can now be estimated by OLS since both the time-invariant community and individual-level error terms subtract out (this, of course, assumes that individuals do not change communities).

The fixed effects logit procedure can be extended to this case in an exactly an analogous manner.

Stata Examples

Example 48: Malawi

```
. xtreg ideal age $education $work $religion $fptype, fe

Fixed-effects (within) regression               Number of obs      =      4238
Group variable: respondentid                    Number of groups   =      2299

R-sq:  within  = 0.0413                         Obs per group: min =         1
       between = 0.0753                                        avg =       1.8
       overall = 0.0622                                        max =         3

                                                F(9,1930)          =      9.24
corr(u_i, Xb)  = -0.0074                        Prob > F           =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0482504   .0074202     6.50   0.000      .033698    .0628028
        edu1 |  -.1467206    .229127    -0.64   0.522    -.5960831    .3026419
        edu2 |  -.0141241   .5389645    -0.03   0.979    -1.071138     1.04289
    employed |   .1874763   .1343902     1.40   0.163     -.076089    .4510417
    catholic |   .1635772   .2754351     0.59   0.553    -.3766044    .7037589
  protestant |   .4224988   .1669302     2.53   0.011     .0951164    .7498813
      muslim |   .4385648   .4831697     0.91   0.364    -.5090248    1.386154
 traditional |   .3376532   .1846207     1.83   0.068    -.0244238    .6997303
      fpmess |  -.6487476   .1261282    -5.14   0.000    -.8961096   -.4013857
       _cons |   2.911137   .3971029     7.33   0.000     2.132342    3.689933
-------------+----------------------------------------------------------------
     sigma_u |  1.9601946
     sigma_e |  2.2562679
         rho |   .4301261   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(2298, 1930) =     1.20          Prob > F = 0.0000
```

(Continues next page)

(Continues from previous page)

```
. estimates store consistent

. xtmixed ideal age $education $work $religion $fptype || tribenum: || respondentid:

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0:   log restricted-likelihood =  -9255.742
Iteration 1:   log restricted-likelihood = -9246.4529
Iteration 2:   log restricted-likelihood = -9246.4523

Computing standard errors:

Mixed-effects REML regression                   Number of obs      =      4042

-----------------------------------------------------------
                |   No. of       Observations per Group
 Group Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
       tribenum |        9         16      449.1       1305
   respondentid |     2414          1        1.7          3
-----------------------------------------------------------

                                                Wald chi2(9)       =    341.33
Log restricted-likelihood = -9246.4523          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   ideal_num |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0576779   .0034896    16.53   0.000     .0508385    .0645173
        edu1 |  -.1244445   .0980701    -1.27   0.204    -.3166582    .0677693
        edu2 |  -.6059863   .1862655    -3.25   0.001    -.9710599   -.2409126
    employed |   .2983846   .0945989     3.15   0.002     .1129741    .4837951
    catholic |   .3617523   .1398686     2.59   0.010     .0876148    .6358899
  protestant |   .3454279   .1211591     2.85   0.004     .1079604    .5828954
      muslim |   .2438693   .1951087     1.25   0.211    -.1385368    .6262754
 traditional |   .4321477   .1437094     3.01   0.003     .1504824    .7138131
      fpmess |  -.2920399   .0901779    -3.24   0.001    -.4687853   -.1152944
       _cons |   2.285109   .2484676     9.20   0.000     1.798121    2.772096
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
tribenum: Identity           |
                   sd(_cons) |   .2673575   .1419285      .0944553    .7567604
-----------------------------+------------------------------------------------
respondentid: Identity       |
                   sd(_cons) |   .7734143   .0801427      .6312611    .9475789
-----------------------------+------------------------------------------------
                sd(Residual) |    2.25124   .0350209      2.183636    2.320937
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(2) =    32.12   Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

The comparison of the consistent fixed effects estimator and the asymptotically efficient random effects estimator under the null hypothesis strongly rejects the null hypothesis of no correlation between the explanatory variables and the time invariant individual and community error components.

Example 49: Indonesia

```
. xtreg ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0, fe

Fixed-effects (within) regression               Number of obs      =     19389
Group variable: ind_id                          Number of groups   =      9177

R-sq:  within  = 0.0107                         Obs per group: min =         1
       between = 0.0072                                        avg =       2.1
       overall = 0.0044                                        max =         4

                                                F(23,10189)        =      4.81
corr(u_i, Xb)  = -0.0823                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0067896   .0040511     1.68   0.094    -.0011513    .0147306
   midwife_0 |   -.150851   .0254811    -5.92   0.000    -.2007989   -.1009031
        edu1 |    .105857   .0573394     1.85   0.065    -.0065395    .2182536
        edu2 |  -.1076671   .1026222    -1.05   0.294    -.3088269    .0934927
        edu3 |  -.0109061   .1490359    -0.07   0.942    -.3030458    .2812336
       hedu1 |  -.0719232   .0636732    -1.13   0.259    -.1967353    .0528889
       hedu2 |  -.1149624   .0760482    -1.51   0.131    -.2640317     .034107
       hedu3 |  -.1318419   .0980391    -1.34   0.179    -.3240178     .060334
         age |   .0024369   .0023766     1.03   0.305    -.0022216    .0070954
    work_any |   .1789616   .0416408     4.30   0.000     .0973374    .2605857
   work_most |  -.1496873   .0413381    -3.62   0.000    -.2307182   -.0686565
      smoker |  -.1423897   .0984873    -1.45   0.148    -.3354443    .0506649
    goodhlth |   .0499408   .0324426     1.54   0.124     -.013653    .1135346
      muslim |  -.3799097   .1759995    -2.16   0.031    -.7249033   -.0349161
    electric |  -.1476358   .0391097    -3.77   0.000    -.2242985   -.0709732
        read |   .0216131   .0668014     0.32   0.746    -.1093308     .152557
       write |  -.0002823   .0596983    -0.00   0.996    -.1173028    .1167382
       urban |   .0592473   .0725455     0.82   0.414    -.0829562    .2014509
    pubtrans |   -.011627   .0273987    -0.42   0.671    -.0653339    .0420798
      caphrs |  -.0030822   .0070283    -0.44   0.661     -.016859    .0106946
     fhead50 |  -.0301402   .0431383    -0.70   0.485    -.1146997    .0544194
     fhead75 |   .0343107   .0556093     0.62   0.537    -.0746945    .1433158
    fhead100 |   .1525193     .07152     2.13   0.033     .0123261    .2927126
       _cons |   3.325014   .1929856    17.23   0.000     2.946724    3.703304
-------------+----------------------------------------------------------------
     sigma_u |  1.3334556
     sigma_e |  .97075394
         rho |   .6536024   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(9176, 10189) =     3.63         Prob > F = 0.0000
```

(Continues next page)

(Continues from previous page)

```
. estimates store consistent

. * Hausman-Taylor Test
. xtmixed ideal num_pos_0 midwife_0 edu1 edu2 edu3 hedu1 hedu2 hedu3 /*
> */ $individual $community if god==0 || com_id: || ind_id:

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0:   log likelihood =  -31979.23
Iteration 1:   log likelihood = -31979.226
Iteration 2:   log likelihood = -31979.226

Computing standard errors:

Mixed-effects ML regression                     Number of obs      =     19389
-----------------------------------------------------------
                |   No. of       Observations per Group
 Group Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
         com_id |      313          6       61.9        135
         ind_id |     9216          1        2.1          4
-----------------------------------------------------------
                                                Wald chi2(23)      =    914.48
Log likelihood = -31979.226                     Prob > chi2        =    0.0000
------------------------------------------------------------------------------
       ideal |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   num_pos_0 |   .0018884   .0031395     0.60   0.548    -.0042648    .0080416
   midwife_0 |  -.2318394   .0219262   -10.57   0.000    -.2748141   -.1888648
        edu1 |   .0690483   .0428206     1.61   0.107    -.0148785    .1529752
        edu2 |  -.2671518   .0584393    -4.57   0.000    -.3816908   -.1526129
        edu3 |  -.3345629   .0821795    -4.07   0.000    -.4956319    -.173494
       hedu1 |  -.0522089   .0502289    -1.04   0.299    -.1506558    .0462379
       hedu2 |  -.0456639   .0606641    -0.75   0.452    -.1645634    .0732356
       hedu3 |  -.0123991   .0780003    -0.16   0.874    -.1652768    .1404786
         age |   .0265972   .0013036    20.40   0.000     .0240422    .0291523
    work_any |   .0667326   .0346843     1.92   0.054    -.0012473    .1347125
   work_most |  -.0854164   .0346993    -2.46   0.014    -.1534258   -.0174071
      smoker |  -.1237776   .0711992    -1.74   0.082    -.2633255    .0157703
    goodhlth |   .0280553   .0272719     1.03   0.304    -.0253965    .0815072
      muslim |  -.0344901   .0576769    -0.60   0.550    -.1475347    .0785545
    electric |  -.2610365   .0321138    -8.13   0.000    -.3239784   -.1980946
        read |  -.0471661    .056725    -0.83   0.406    -.1583451    .0640129
       write |  -.1101406    .051786    -2.13   0.033    -.2116394   -.0086418
       urban |  -.0162026   .0468487    -0.35   0.729    -.1080245    .0756192
    pubtrans |   .0165636   .0244526     0.68   0.498    -.0313627    .0644899
      caphrs |   .0100769    .005538     1.82   0.069    -.0007773    .0209311
     fhead50 |  -.0981817   .0362879    -2.71   0.007    -.1693047   -.0270588
     fhead75 |  -.1368424   .0442086    -3.10   0.002    -.2234895   -.0501952
    fhead100 |   -.149778   .0546152    -2.74   0.006     -.256822   -.0427341
       _cons |   2.694524   .1060392    25.41   0.000     2.486691    2.902357
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
com_id: Identity             |
                   sd(_cons) |   .4576729    .023304      .4142031    .5057047
-----------------------------+------------------------------------------------
ind_id: Identity             |
                   sd(_cons) |   .9618635   .0112975      .9399738     .984263
-----------------------------+------------------------------------------------
                sd(Residual) |   .9637987   .0065419      .9510616    .9767063
----------------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(2) =  5570.58   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

(Continues next page)

(Continues from previous page)

```
. estimate store efficient

.
. hausman consistent efficient, equations(1:1)

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |   consistent   efficient      Difference          S.E.
-------------+----------------------------------------------------------------
   num_pos_0 |     .0067896     .0018884        .0049012        .0025603
   midwife_0 |     -.150851    -.2318394        .0809884        .0129817
        edu1 |      .105857     .0690483        .0368087         .038134
        edu2 |    -.1076671    -.2671518        .1594847        .0843574
        edu3 |    -.0109061    -.3345629        .3236568        .1243311
       hedu1 |    -.0719232    -.0522089       -.0197142        .0391323
       hedu2 |    -.1149624    -.0456639       -.0692985        .0458605
       hedu3 |    -.1318419    -.0123991       -.1194428        .0593938
         age |     .0024369     .0265972       -.0241603        .0019871
    work_any |     .1789616     .0667326         .112229        .0230425
   work_most |    -.1496873    -.0854164       -.0642709        .0224677
      smoker |    -.1423897    -.1237776       -.0186121        .0680472
    goodhlth |     .0499408     .0280553        .0218854        .0175717
      muslim |    -.3799097    -.0344901       -.3454196        .1662804
    electric |    -.1476358    -.2610365        .1134006         .022322
        read |     .0216131    -.0471661        .0687792        .0352803
       write |    -.0002823    -.1101406        .1098583        .0297002
       urban |     .0592473    -.0162026          .07545          .05539
    pubtrans |     -.011627     .0165636       -.0281907        .0123595
      caphrs |    -.0030822     .0100769       -.0131591        .0043276
     fhead50 |    -.0301402    -.0981817        .0680416         .023326
     fhead75 |     .0343107    -.1368424         .171153        .0337342
    fhead100 |     .1525193     -.149778        .3022974        .0461766
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
          B = inconsistent under Ha, efficient under Ho; obtained from xtmixed

    Test:  Ho:  difference in coefficients not systematic

                 chi2(23) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      203.26
                Prob>chi2 =      0.0000
```

Once again, the null hypothesis of no correlation between the explanatory variables and the time invariant individual and community error components is rejected. In this instance, as per the p-value it is strongly rejected.

V. Conclusion

Longitudinal data can be a powerful tool for program evaluation. The purpose of this guide is to motivate the collection of longitudinal data by showing its similarities to well-accepted pre-test/post test experimental design and then provide a large set of examples. The large number of examples is necessary so that we can cover the wide range of types of outcome variables that are collected in evaluation research ranging from continuous variables to unordered categorical variables. In this guide, we restricted the set of estimation procedures considered to those that could easily be implemented in Stata statistical software that is widely used and widely available. However, this restriction does impose some limits on the flexibility of the estimators that we can consider. We will address these limitations in future research.

References

Angeles G, Guilkey DK, Mroz T. Purposive program placement and the estimation of program effects: the impact of family planning programs in Tanzania. *J American Stat Assoc*. 1998;93(443):884-889.

Angeles G, D. Guilkey DK,. Mroz T. The effects of education and family planning programs on fertility in Indonesia. *Econ Devel Cultural Change.* 2005;54(1):165-201.

Angeles G, D. Guilkey DK,. Mroz T. The impact of community-level variables on individual-level outcomes: theoretical results and applications. *Sociological Methods & Research.* 2005;34(1):76-121.

Deaton A. *The Analysis of Household Surveys: a Microeconometric Approach to Development Policy*. Baltimore, MD: The World Bank, Johns Hopkins University Press; 1997:105-106.

Guilkey DK, Lance PM. Program impact estimation with binary outcome variables: Monte Carlo results for alternative estimators and empirical examples. In Sickles R, Horrace W (eds.) Festschrift in honor of Peter Schmidt. ND; unpublished draft.

Hausman J, Taylor W. Panel data and unobservable individual effects. *Econometrica* 1981;49(6):1377-1398.

Hausman J. Specification tests in econometrics. *Econometrica.* 1978;46:1251-1271.

Heckman JJ,. Macurdy TE. A life cycle model of female labour supply. *Rev Econ Studies.* 1980;XLVII:47-74.

Heckman JJ, Singer B. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*. 1984;52(2): 271-320.

Mroz TA. Discrete factor approximations in simultaneous equation models: estimating the impact of a dummy endogenous variable on a continuous outcome. *J Econometrics.* 1999;92(2):233-274..

Appendix A: Overview of Data Sets Used in Stata Examples

This appendix contains a brief overview of the two longitudinal data sets used in the Stata examples.

A. *Indonesia*

Longitudinal data from the Indonesian Family Life Survey (IFLS) contain four waves (1993, 1997, 2000 and 2007) of individual and family-level information on fertility, health, education, migration and employment. Community and facility (health care providers and schools) level data supplement the household data.

The sample was drawn from 13 of Indonesia's then 27 provinces (there are now 33). These 13 provinces (North, West and South Sumatra; West, Central and East Java; Lampung; DKI Jakarta; DI Yogyakarta; Bali; West Nusa Tenggara; South Kalimantan; and South Sulawesi) contained roughly 83 percent of the nation's population. These provinces were selected in order to maximize representation of the population, capture the cultural and socioeconomic diversity of Indonesia, and be cost effective given the size and terrain of the country. Enumeration areas (EAs) were randomly selected from the 13 provinces using the 1993 SUSENAS sampling frame, with over sampling of urban EAs and EAs in smaller provinces (321 EAs were selected in total). Households were randomly selected by field teams within each EA.

While many household members were interviewed as part of the IFLS, only married women between age 15 and 49 were questioned about contraceptive use and ideal family size. Because marriage is required for inclusion in this sub-sample, divorce is responsible for some sample attrition and new marriages lead some women to enter the sub-sample in later waves. New households were also added in each year to combat traditional survey attrition (due to death and migration outside an EA of study). A summary of usable woman-year observations can be seen in table A1.

Households are linked to EA-level data in each round. EA variables used in this guide were gathered during a group interview with the village leader, one or two members of his staff and one or two members of the village elders advisory board. Questions pertaining to the availability of health facilities were presented to this group and the head of the village women's group. In this guide, discrepancies between the two respondents are settled by simple averages.

Table A2 summarizes individual and community level information for the 20,000 woman-year observations used in this guide. The final sample omits all observations with missing variable values that could not be reasonably inferred from other entries in the panel. Only 362 observations are lost due to missing information.

Table A1: Indonesian Family Life Survey Longitudinal Sample Size

Initial Participation Cohort	Survey Year				
	1993	1997	2000	2007	Total
Wave 1 Cohort	3520	2873	2684	1498	10575
Wave 2 Cohort		2207	1742	1152	5101
Wave 3 Cohort			1466	933	2399
Wave 4 Cohort				2287	2287
Total observations					20362

Table A2: IFLS Summary Statistics

	mean	s.d.
Dependent Variables		
ideal number of kids	3.009	1.504
contraceptive use	0.588	0.492
Individual-level Independent Variables		
highest education grade school	0.669	0.470
highest education high school	0.169	0.375
highest education college	0.049	0.217
husband highest education grade school	0.697	0.460
husband highest education high school	0.192	0.394
husband highest education college	0.061	0.239
age	34.099	8.842
work >1 hour	0.519	0.500
full time employment	0.449	0.497
smoker	0.020	0.140
good health	0.887	0.316
Muslim	0.891	0.312
electricity in home	0.877	0.329
can read	0.820	0.384
can write	0.801	0.400
Community-level Independent Variables		
number of posyandus	7.507	6.251
has village midwife	0.588	0.492
has village delivery post	0.275	0.447
urban	0.462	0.499
public transportation	0.738	0.440
hours travel time to capital	3.246	2.915
population in 25-50 percent	0.249	0.433
population in 50-75 percent	0.251	0.434
population in 75-100 percent	0.249	0.433
Observations	20,000	

The paper examines three dependent variables in estimation: ideal number of children, contraceptive use and method of contraception. A women's ideal number of children is determined by summing her current number of children and her response to the question "how many more children do you wish to have?" Note that this sample mean is calculated from 19,389 responses, instead of 20,000, because 611 responded that they would "leave it up to God." These individuals are coded as missing in analysis. Contraceptive method is a categorical variable and is summarized in table A3. Individual level independent variables are fairly straight-forward. Age has been corrected for standard misreporting, common in longitudinal studies. Working greater than one hour is a dummy variable created if the woman reported working (for income) more than one hour in the prior week. Full time employment is a dummy variable indicating those who claim that working was their primary activity in the prior week. Good health is a dummy variable set equal to one if the woman reported her health as being "good" or "very good."

Table A3: Contraceptive Method

	observations	percent
condom	168	0.84
pill	2983	14.91
injection	4784	23.92
diaphragm/IUD/Norplant	2429	12.15
sterilization	919	4.59
traditional method	466	2.33
missing	11	0.06
no use	8240	41.20
total	20000	100.00

Community (EA) level variables are also used throughout the study. A Posyandu is a government run facility that specializes in family planning programs, contraceptive distribution and child health. Public transportation is a dummy variable indicating villages with public 3-4 wheeled or motor boat transportation. Travel time to the province capital does vary over time for each EA as roads and waterways changed over the 11 years studied. Population is approximated using number of family heads in each village and is described by a set of dummy variables in estimation.

B. Malawi

The Malawi Longitudinal Study of Families and Health (MLSFH), formerly known as the Malawi Diffusion and Ideational Change Project (MDICP), is led by a group of researchers in the Population Studies Center at the University of Pennsylvania. The Republic of Malawi, situated in southeast Africa, is over 45,000 square miles, with a population of almost 14 million. Malawi is known to have low life expectancy, high rates of infant mortality and a strong prevalence of HIV/AIDS.

The longitudinal study was originally formed to examine the role of social interactions in changing the attitudes and behaviors of the population in Malawi. It was formed in parallel to a similar study for Kenya Diffusion and Ideational Change Project. There are four waves of data that comprise the MLSFH, however, only the first two waves (1998 and 2001) were publicly available for use in this project. The focus on these two waves was on understanding the relationship between social interactions and both the diffusion of knowledge of AIDS and the acceptance or rejection of modern contraceptive methods.

The MDICP was conducted in three distinctive districts of Malawi, one in each of the three regions of the country: Rumphi District, located in the Northern region; Mchinji District, located in the Central region; and Balaka District, located in the Southern region. The survey contains individual-level information on income, education, family planning practices and demography for women of sampled households and their husbands. A substantial portion of the survey is devoted to questions regarding the participants peers, or people that the participant knows in their village.

The sampling strategy adopted for the three regions differ from one another. In Mchinji and Rumphi districts the sample was designed to cover census enumeration areas (CEAs) included in the 1988 Traditional Methods of Child Spacing in Malawi (TMCSM) survey. Some of the CEAs had fewer ever married women than desired, in which case three neighboring CEAs were added to the sample. In each district a cluster sampling strategy was used with a total of 145 villages randomly selected. Before the fieldwork began, a list of households containing residents in each village was compiled. A sample of eligible women was then randomly selected from the household list. Since villages varied in size, sampling fractions were used that were inversely proportional to village populations, such that a higher proportion of eligible women in the smaller villages was sampled.

The sampling strategy was different for the Balaka district in the southern region of Malawi (Chiradzulu District). The sampling in this region was designed following a baseline survey conducted by the German aid agency Deutsche Gesellschaft für Internationale Zusammena (GTZ), with 1098 women and men in 1993. The target sample for Balaka was 500 women and this was met by implementing a random one in four sampling procedure of women in the region. The sample was expected to yield about 90 women who were also interviewed by GTZ and about 75 men, however, 260 women and 125 men were over-sampled.

Ultimately, the sampling strategy was not designed to be representative of the national population in Malawi, but the sample characteristics end up closely matching those of the Malawi Demographic and Health Survey, a representative survey administered by the National Statistical Office of Malawi. The first survey wave, carried out in 1998, interviewed 1541 ever-married women of childbearing age and 1198. In the summer of 2001, the second round of the survey was administered to the same sample and new spouses of old respondents with a change in marital status were also interviewed. The second wave saw 1571 women and 1097 men interviewed, including 186 new wives and 28 new husbands.

For the purposes of this guide, the sample contains married women interviewed during either of the survey waves. Observations with variables that had missing values were dropped from the sample. Table A2 shows summary statistics for the main variables of interest. There were no women who said that their ideal number of children was "up to God" and, with the exception of a handful of husbands, no individuals in the sample had any university level education. As a result, these variables are excluded from our analysis. Village level data were not publicly available from the Web site and, as a result, the only location indicators are those for traditional authorities (TA). The four TAs present in the data are Kalembo, Mkanda, Mwahenga and Mwankhun.

The contraceptive method variable asks a woman if she currently uses a particular form of contraception. The variable contains six different forms of contraception: 1, condom; 2, pill; 3, injection; 4, IUD; 5, sterilization; and 6, traditional methods. Women were asked if they had heard a family planning message from the radio, a hospital/clinic, or from a visit from an agent. There are indicators for these message variables, as well as a constructed `fpmess` variable that indicates whether or not any of the three messages were received by the woman. Table A4 shows the frequency by which different methods of contraception were used.

Table A4: Frequency of Contraceptive Methods

	Frequency	Percent
no contraception	1,425	65.88
condom	23	1.06
pill	127	5.87
injection	302	13.96
IUD	2	0.09
sterilization	54	2.50
traditional methods	230	1.63