\subsection{Logit Regression}

\label{sec:logit}

In some parts of my analysis, I employ logit regression when analyzing a binary outcome--I'll briefly explain that empirical method here. This is not a primary empirical method I employ to answer my question, but instead a model I use to analyze selection into kindergarten, attrition, and robustness checks for binary outcomes such as school completion or school attendance.

The logit model differs from OLS in that it restricts the estimated probability from a function to between 0 and 1, improving our ability to interpret model estimates. It differs from the probit model in that it uses a logistic cumulative distribution function (c.d.f.), rather than a standard normal. Additionally, the logit model allows the use of fixed-effects, which I employ to analyze within-household selection. To be clear, when I present coefficients resulting from a logit model, those are the \textit{average} marginal effects (MEs)--calculated after model estimation.

The logit model fits the following model:

\begin{gather}

\text{Pr}(y\_i = 1 | \mathbf{x}\_i) = \Lambda(\mathbf{x}\_i' \beta)

\end{gather}

where $\Lambda(.)$ is the logistic c.d.f., so that $\Lambda(z) = \frac{e^z}{1+e^z}$.

\subsection{Non-Random Attrition}

\label{sec:attrition}

As I discuss thoroughly in \nameref{sec:data}, my sample suffers from attrition as I merge data from across different waves of the Indonesian Family Life Survey (IFLS). In this section, I examine that attrition and check whether it introduces bias into my estimates of kindergarten's effects.\footnote{ In this section, when I discuss attrition, I am referring to total attrition, i.e., whether an individual was \textit{ever} lost to attrition--whether that be from merging across waves of the IFLS or from missing observations.}

Following the approach of Baulch and Quisumbing (2011), I used the set of covariates employed in the regression analysis as ell as variables collected on the quality of the initial childhood interview in 1997.\footnote{ The two interview quality questions were the interviewer's evaluation of the interviewed child's attention (on a scale of 1-5) and the interviewer's evaluation of the respondent's answers' accuracy. The two are measures are highly correlated (their pair-wise correlation is 0.8941) and both have a significant, positive coefficient when predicting total attrition. I had to exclude some covariates because of a lack of variation in attrition; in particular, I excluded province fixed-effects, the two-parent household dummy, the birth order variable, and general health categorical variable. None of these variables were significant predictors of attrition in the first place. On another note, I employ logit regression rather than probit regression to ensure consistency with my analysis of switching and selection into kindergarten--on the other hand, Baulch and Quisimbing employ the latter.} \citep{Baulch2011} There is variation in the resultant inverse probability weight; the mean weight is 1.07 and the median is 0.993. A basic t-test confirms that the mean weight is not 0, with a 95\% confidence interval of (1.06, 1.08).

I applied the inverse probability weight to both the OLS and IV models, focusing on years of educatino.\footnote{ Since these inverse probability weights are calculated at the individual level, they're vary within household and between siblings, thus disallowing me from using them in the mother fixed-effects model. If attrition is not significant at the individual level, then I would suspect it isn't significant at the household level when switching attrition--which occurs at a significantly higher rate--is accounted for (which I do in \nameref{sec:switching}.} For the OLS model, my results are not different; the inclusion inverse probability weights slightly raises the point estimate of kindergarten's effect from 0.66 to 0.65, and makes it slightly less precise with a standard error of 0.15, rather than the previous standard error of 0.14. The IV estimates, on the other hand, are more different; kindergarten's estimated effect on educational attainment is now 1.64, rather than 1.40--and the standard error is unchanged. Comparing the non-weighted model to the weighted models, there are no significant divergences in the covariates' coefficients. Because of my inability to apply inverse probability weights to the fixed-effects model and the absence of a significant difference for the OLS and IV estimates, I do not include inverse probability weights for the main results presented in this paper.

\subsection{Fixed Effects - Switching Attrition}

\label{sec:switching}

As briefly discussed in \nameref{sec:mom\_fe}, the employment of fixed-effects places additional stringent restrictions on my sample--resulting in attrition of 76.83\% from my full sample to the `switching' sample employed in the fixed-effects estimation.\footnote{ To be transparent about this reduction in sample size, I've ensured that the sample size indicated at the foot of the regression tables with fixed-effects results \textit{only} includes the switching sample.}

I've also conducted basic analysis and diagnosis of the effects of this attrition. To begin, I've constructed a diagnosis scatterplot, Figure ~\ref{fig:scatter\_switching}, to visualize selection into the switching sample. This graph allows us to make several observations. Province fixed-effects--which I employ for my OLS estimates--does not result in any attrition, since each group has a non-zero standard deviation in kindergarten attendance. On the other hand, there are many families, denoted by blue circles, with a standard deviation of zero in kindergarten attendance and that are thus attrited from the sample. Nonetheless, these families still have a great deal of variance in educational attainment.

I've also prepared a table capturing the summary statistics for all variables--the controls as well as outcome variables, since all observations in the table are non-attrited, seen in Table ~\ref{table:switching\_stats}. There appears to be some selection bias into the switching sample--it appears that the switching sample is more educated than the non-switching sample, for example. I follow this with a logit model, employing my full set of covariates and province fixed-effects. The only significant predictors of switching status at a 95\% confidence level are kindergarten attendance (a positive effect), whether a household has two parents (a negative effect), a household's expenditures in 1997 (a negative effect), and a household's expenditures in 2000 (a positive effect).\footnote{ Some province dummies are positive significant, though none are significantly negative.} It makes sense kindergarten has significant positive margin effects; after all, some kindergarten attendance within a household is strictly necessary for switching status. Disregarding the expenditure variables (which appear to counteract one another) and kindergarten attendance, I conduct a joint significance test that the marginal effects for the most important variables for determining educational attainment are all equal to 0--and fail to reject the null hypothesis that none are significant.\footnote{ Specifically, I test the AMEs for household head's years of education, mother's years of education, urban (1997, 2000, and 2007), number of kindergartens per kecamatan (1990 and 2000), and kindergarten spillover from older siblings.} These results tentatively suggest that the attrition--despite its high rate--from the full sample to the switching sample does not bias my fixed-effects results.

On the other hand, the process of selection into kindergarten appears to be biased by this attrition. I run two separate logit regressions, one with only my switching sample and the other only with the non-switching sample. The switching sample logit reveals negative, significant AMEs for the kindergarten spillover variable and the total number of visits to a doctor--an inversion of both my theoretical expectations as well as the empirical findings from both the OLS and IV estimations of selection into kindergarten for the full sample. Additionally, the non-switching sample reflects strong positive effects for both the household head's years of education as well as the mother's--while the switching sample only shows significance for the latter, albeit with three times as large of a standard error. This is emblematic of the core problem with this attrition--even if, as my analysis above suggests, there is not a clear bias in switching attrition, the sheer decrease in sample size results in imprecision and loss of efficiency in my estimates that cloud my results. This, in conjunction with the correlated error term issue I discuss later, motivate my use of Instrumental Variable (IV) estimation.