Methodological Challenges in Studying the Impact of Domestic Violence on Children's Human Capital: An Application to Colombia

Ragui Assaad, Greta Friedemann-Sánchez, and Deborah Levison*
Humphrey School of Public Affairs
University of Minnesota

August 2013

Working Paper No. 2013-08
https://doi.org/10.18128/MPC2013-08

*Authors' names listed alphabetically

ABSTRACT

This paper examines the effects of intimate-partner violence (IPV) against the mother on the educational outcomes of her children ages 6-14. We explore the potential non-random selection of children into situations where they are exposed to IPV using non-parametric matching methods and parametric instrumental variables methods. The analyses of Colombia's 2005 DHS (N= 21,827) indicate that mother's exposure to IPV reduces children's school attendance by 1.2 to 2.7 percentage points, depending on methodology, substantial when compared to the 6.7 percent average non-attendance rate. It reduces unconditional grade advancement by 2.1 to 2.8 percentage points, which should be compared to an average non-advancement rate of 8 percent. It reduces grade advancement conditional on staying in school by 1.5 to 1.8 percentage points, relative to an average non-advancement rate of 4.4 percent. The effect of mother's IPV on the probability of drop-out in the past year is not statistically significant, but it lowers grade attainment conditional on current attendance by 0.06 to 0.12 years and unconditional years of education completed by about 0.10 years.

Kew Words: Domestic violence, intimate partner violence, education, children, outcomes, Colombia.

INTRODUCTION

Due to its endemic nature, intimate partner violence (IPV) – violence perpetrated by a male against his female partner – is increasingly being recognized as a human development problem worldwide (Kishor and Johnson 2004). IPV is also known as domestic violence, spousal abuse, and wife battery.[1] IPV is a multidimensional phenomenon that includes physical, emotional, and sexual violence, as well as stalking. The prevalence of physical IPV in different parts of the world has been estimated to range between 13 and 61% (Garcia-Moreno 2006). Our interest is in Colombia, where, as we discuss below, the level of physical intimate partner violence is among the highest in the world.

IPV has cascading negative effects on the economic wellbeing (Renzetti 2009), physical health (Matthew et al. 1996) and mental health (DeJonghe et al. 2008) of individual victims, as well as on the incidence of unintended pregnancy (Pallitto and O'Campo 2004). Furthermore, IPV has negative consequences not only for the woman subjected to violence but also for the human development of her

[1] The term "domestic violence" typically includes violence between household members. It can be female to-male violence, childhood maltreatment, or between siblings. IPV is restricted to male-to-female violence.

children (Evans et al. 2008). For instance, IPV is highly predictive of poor child nutrition (Heaton and Forste 2008) and poor cognitive, emotional and behavioral outcomes (Kirtzmann et al. 2003). The objective of this paper is to explore the effects of women being subjected to IPV on multiple measures of their children's educational outcomes in Colombia, including current school attendance, grade advancement in the last year (both unconditional and conditional on staying in school), drop-out in the past year, current grade conditional on still being in school, completed years of education, and grades completed per year of exposure to school.

Why might the violence that children witness affect their educational outcomes? Remarkably little has been written on this, and – as we conclude after reviewing related literature while discussing our results – the causal pathways are not well documented. Still, common sense suggests that children who witness violence against their mothers will be distracted in school, less able to focus on their school work and learn to their potential. They may be tired from being kept awake by violence or worry about it. They may themselves be beaten, and they may be kept home from school while bruised. While we cannot model violence against children directly – since children are often beaten for poor performance in school, possible reverse causality creates a serious impediment to estimation – it is well known that men who beat their wives or partners are also more likely to beat her/their children (Patel 2011). In this paper we assume that causal pathways exist, but we do not attempt to identify them.

Estimating the effect of IPV on children's outcomes poses methodological challenges because of potential selection and endogeneity problems. There are a large number of possible confounding factors that could affect both IPV status and child outcomes, and there is also the potential for reverse causality if some women experience IPV as a result of their children's performance in school. We therefore argue that overlooking the potential endogeneity and selection issues associated with IPV – as is standard in this literature – is problematic. If the types of confounding or reverse causality described above exist but are

ignored in an analysis, the regression coefficients estimates would likely be biased upward. We use both parametric instrument variable (IV) methods as well as non-parametric matching methods to address the potential endogeneity and selection problems associated with our "treatment" -- intimate partner violence. We test for the endogeneity of the treatment to the child's educational outcomes. Failing to reject exogeneity, we argue that results from non-IV parametric and non-parametric estimation methods that correct for selection on a large number of observables provide consistent estimates of the effects of IPV.

LITERATURE ON THE EFFECTS OF INTIMATE PARTNER VIOLENCE ON CHILDREN

There is a large literature on risk factors for IPV, but less has been written on effects of intimate partner violence on other outcomes, including those for children living in households with intimate partner violence. Although the study of IPV is more extensive for developed nations and specifically for the United States, an area that is understudied in both developed and developing nations is the effect that IPV has on children's educational outcomes; the evidence about this is extremely limited. Among the few studies conducted in poor countries is one from Sri Lanka, which found that children who were directly (watching, hearing, intervening) or indirectly (observing maternal injuries, depression) exposed to IPV at home had poor school attendance and lower academic achievement on average as measured by exam scores (Jayasinghe et al.2009). A study conducted in Brazil found that children 5 to 12 years old who lived with mothers exposed to psychological, physical and sexual IPV were more likely to be among those dropping out of school or failing a school year (Durand et al. 2011). Studies conducted in the United States have found lower reading levels among adolescents who have been exposed to IPV (Thompson and Whimper 2010), lower academic achievement in math and reading for children in elementary and middle school (Kiesel et al. 2011), lower scores on standardized tests for children ages 6 to 17 – especially for girls and children younger than 12 years old (Peek-Asa et al. 2007) – and more grade repetition and truancy among children 6 to 15 years old (Emery, 2011).

This paper addresses two gaps in the literature. First, it provides evidence about the effects of IPV on child education. Second, we take seriously the issues of non-random selection and endogeneity, which are overlooked in much of the literature on consequences of IPV. Emery (2011) is the only study of child educational outcomes we have found that addresses the potential endogeneity of IPV. Emery uses Chicago panel data and estimates fixed-effect models to separate the effects of IPV from effects of child abuse and selection bias. His approach is not available to us, as we do not have panel data. In our study, endogeneity and selection bias may arise due to at least two types of circumstances. The incidence of IPV could be related to confounding variables that affect both IPV and other household outcomes. For instance, an alcoholic father may not only beat his wife but also make the home environment difficult for study. It may also arise because of reverse causality: The outcome variable may provide the rationale for violence, as when the poor performance of children in school leads to a man abusing his partner. Addressing endogeneity and selection bias is challenging in practice. In this paper we employ instrumental variable and matching methods, both of which rely on identifying assumptions.

IPV AND EDUCATION IN COLOMBIA

Colombia is a multiethnic democratic republic with a population of about 42 million individuals as of 2005. Roughly 70% of the population lives in urban areas. Armed conflict between the government, paramilitary groups and guerrilla groups has been going on for 40 years. The use of violence in multiple social contexts is so wide-reaching that some scholars argue it is contagious (Sánchez 2007). Recent evidence shows that 44.3% of women who are conflict-displaced have been physically abused by their intimate partners (Sanchez Lara et al. 2008).

Colombia has one of the highest physical IPV prevalence rates in the world and, after Peru, the highest in Latin America (Kishor and Johnson, 2004). In a study using Colombia's 2005 Demographic Health Survey, 40% of women reported having ever experienced any type of physical violence, whereas

22% reported it for the last 12 months (Friedemann-Sánchez and Lovatón 2012). The prevalence of a woman having *ever* experienced severe forms of physical violence (threatened or attacked with a knife or a fire arm, strangled or burned, raped) is 16.6%. Sexual assault (11.7%) constitutes the most common severe form of violence. Being pushed or shaken is the most frequently reported among the less severe forms of violence (34%). The life-time and past-year rates of emotional abuse are even higher than the rates of physical violence at 66.4% and 52.3% respectively. This same study reveals that among the women who experienced IPV, 13.05% had bones broken, 23.7% reported having suicidal thoughts, and over one-third reported loss of productivity at work or in their studies.

What factors predict IPV in Colombia? Living in an urban environment, cohabitating with a partner, being younger, and having a larger number of children are all predictors of an increased probability of experiencing IPV (ibid). However, the highest probability of experiencing IPV is associated with the maltreatment of the woman's partner when he was a child (ibid). There is robust evidence for developed countries (Whitfield, Anda, Dube, & Felitti, 2003), also reported for a few developing regions like India (Martinet al. 2002), that childhood exposure to violence between parents is a risk factor for becoming a victim and/or perpetrator of violence later in life. The intergenerational perpetuation of intimate partner violence, along with its consequences related to multidimensional deprivation, suggests that domestic violence can contribute to intergenerational poverty traps (Evans et al. 2008).

According to Colombia's 2005 population census, the literacy rate among individuals 15 years old and older was 91.6% (91.3% for men and 91.8% for women). A comparison of literacy rates since 1964 (when 75% of men and 71.1% of women reported they were literate) shows that they have been steadily improving for both men and women. In 2005, literacy rates were slightly higher in Colombia than in those of other Andean countries like Peru (88.6% in 2005) and Bolivia (87.2% in 2003) but were comparable to the rates of Venezuela (93.4% in 2003) and Ecuador (92.5% in 2003). According to the 2005 census, 78%

of 5-6 year olds, 92% of 7-11 year olds, and 77% of 12-17 year-old children were enrolled in school; individuals between 15 and 24 years old had on average 9 years of formal schooling (DANE, 2005). A 2010 national study reports that of all children in elementary and secondary school, 77.6% were enrolled in public schools and 5.5% received tuition subsidies; 75.6% of urban children were enrolled (DANE, 2011). There is almost equal distribution by gender of current enrollment rates in primary (51% boys, 48% girls) and secondary (49% boys, 51% girls) education (ibid).

Our results using the 2005 Demographic and Health Survey (DHS) for Colombia show similar levels of educational outcomes to those estimated by DANE (See Tables 1 and 3). We find that 93% of the 6-14 year-olds in our sample were attending school at the time of the survey. Over 91% of those who were in school the previous year advanced to the next grade; among those in school in both the previous year and the survey year, over 95% advanced. Only 2% dropped out in the year prior to the survey.

ESTIMATION STRATEGY

Our estimation strategy consists of first estimating parametric regression models that examine the effect of IPV, controlling for a large number of individual, household-level and community-level covariates, on the assumption that IPV is exogenous to child outcomes. We then estimate non-parametric matching models that also assume the conditional exogeneity of treatment, i.e., that selection into IPV status is based solely on observable characteristics. Non-parametric matching models have two distinct advantages over regression-based models: they do not assume any *a priori* functional form for the relationship between IPV and the child's educational outcome, and they rely on matching the treatment observations with a closely matched set of control observations rather than using all the observations in the sample in the estimation, some of which are simply not comparable to those experiencing IPV.

Given that both these methods assume exogeneity of treatment, we also estimate both linear and non-linear parametric instrumental variable (IV) models to test for the exogeneity of IPV. Such an IV

strategy crucially depends on the validity of the instruments selected.[2] Our instrument of choice relates to the mother's partner's experience of violence when he was a child (that is, whether or not he was regularly beaten as a child). Previous studies (Friedemann-Sánchez and Lovatón 2012) have shown that the partner's childhood experience of violence is a powerful predictor of IPV. We argue that, additionally, this variable satisfies all the necessary conditions for a valid instrument. We posit that with the inclusion of appropriate controls for household socioeconomic status and social context, the mother's partner childhood experience of violence is excludable from the child's educational outcome equation.[3] We also argue that because this instrument is determined at a much earlier time, it is independent of both the mother's IPV status and the child's educational outcomes. Finally, we claim that this instrument satisfies the monotonicity assumption, in the sense that a man's experience of violence as a child is likely to either not affect his chances of perpetuating violence himself or to increase it, but not to decrease it.

We considered using another instrument, namely the mother's childhood experience of witnessing IPV among her own parents. While this instrument probably satisfies the exclusion restriction and the

[2] For readers new to this methodology: An instrumental variable is a proxy for the endogenous treatment of interest (in this case IPV). A valid instrument is a variable that is correlated with the treatment (IPV) but uncorrelated with any other determinants of the (child's human capital) outcome. That is, it only affects the outcome through its effect on the treatment. This condition is referred to as an exclusion restriction. When causal effects of the treatment are heterogeneous, i.e., when they differ across individuals, two additional assumptions are necessary: (i) that the instrument is exogenous, that is independent of the treatment and the outcome (the exogeneity assumption), and (ii) that its effect on the treatment is monotonic, that is, while some individuals' treatment status may not be affected by the instrument, all those that are affected are affected in the same way (the monotonicity assumption). Subject to these assumptions, the IV method yields the effect of the treatment for those individuals whose treatment status changes when the instrument changes value (the compliers), what is known as the local average treatment effect (LATE). This could be different from the average treatment effect (ATE), which would also include the effect for individuals whose treatment status is not affected by the instrument (the always treated and the never treated). See Chapter 4 of Angrist and Pischke (2009) for a more detailed discussion of IV estimation.

[3] It is possible that the mother's partner childhood experience with violence will more likely to subject children to violence and through that affect their educational performance. Since we cannot differentiate in this paper between violence directed to children or to an intimate partner, we treat both as part and parcel of IPV.

exogeneity assumption, it is less likely to satisfy the monotonicity assumption. A woman who witnesses IPV among her parents may either be more accepting or more resistant to being a victim of IPV as an adult, depending on circumstances. Nonetheless we carry out as set of alternative estimates with this second instrument included as a check on the robustness of our results. Including this additional instrument also allows us to undertake over-identification tests to test the soundness of our exclusion restriction.

Dependent variables. We use seven measures of human capital, all related to child schooling outcomes, as dependent variables. No single variable can perfectly capture human capital accumulation during childhood. In this analysis, we have no information on learning via training outside of formal education, yet this is an important path to the accumulation of skills (Bourdillon et al 2010). Because we expect to learn somewhat different things (discussed below) from alternate measures, our analysis is repeated for each of seven dependent variables that we are able to calculate. These include whether the child is currently attending school or not, has dropped out in the past year or not, has advanced a grade since the previous year or not – both conditional on not dropping out (i.e., repeating the grade only) and unconditionally (either repeating or dropping out) – as well as the child's current grade in school, total number of years of education successfully completed to date, and grades attained per year of exposure to schooling. Because the "years of education" variable is conditional on school entry, we limit the sample in that part of the analysis to children 10-14 year olds instead of 6-14 year olds to avoid problems related to delayed entry. The "grades per year of exposure to school" variable is calculated by dividing grades attained by age minus age of school entry. [4] Since age of school entry is not known, we assume the age of entry to be six. For this dependent variable we use a sample of 10-14 year olds, leaving out the younger children because the effect of measuring age of school entry with error is exaggerated at younger ages

[4] Thanks to David Lam for suggesting this measure.

(since so few grades have been completed). Our sample mean of 0.83 indicates that on average, Colombian children ages 10-14 successfully completed eight-tenths of a year for each year of potential exposure to schooling.

Explanatory Variables. Our main explanatory variable, which we refer to as the treatment, is the potentially endogenous regressor, indicates whether or not the child's mother has experienced physical intimate partner violence in the past 12 months. Women were asked about the following experiences: (i) being pushed or shaken, (ii) hit with a hand, (iii) hit with an object, (iv) bitten, (v) kicked or dragged, (vi) threatened with a knife, (vii) attacked with a knife or firearm, (viii) being subject to an attempt at strangulation or burning, and (ix) being raped. The occurrence of any of these at least once in the past 12 months constitutes IPV by our definition. For the purposes of this paper we did not consider the experience of emotional violence, such as controlling behaviors or threats, to be instances of IPV. Note that while a child's mother's partner may be her husband, he may or may not be the child's biological father. We use the term "partner" to avoid confusion.

Additional individual and household-level controls used in both the child outcome equation and as explanatory variables in the first stage IPV equation include child sex, age, age-squared, whether the child is the son or daughter of the household head, the mother's age and age squared when the child was six, the mother's and her partner's years of schooling, marital and cohabitation status of the mother, the household's migrant status, its wealth quintile, and variables indicating the composition of the household in terms of numbers of female and male children and adults of various ages and sexes and the presence of relatives on either the mother's side or that of her partner. Community-level controls include regional dummy variables and municipality averages for a wealth index, years of education of men and women, the child-woman ratio as a proxy for fertility, the percentage of female-headed households, the percentage of the population living abroad, the percentage of women and men in formal employment, the percentage

of households with piped water and sewage disposal, and the percentage of households cooking with firewood. All these municipality-specific variables were calculated by averaging over the DHS sample in each municipality. Because the partner's information could be missing, we also include two dummy variables indicating whether the partner's education is missing and whether the information on the partner's childhood experience with violence is missing. Although this refers to the instrumental variable, we include this missing indicator in both the first and second stage equations just in case a woman's inability to report such information relates to either the presence or absence of a partner, how well she knows her partner, or other non-excludable aspects of the context

Estimation methods. In the parametric estimation, we select functional forms that are appropriate to each of the outcome variables. For the four binary outcome variables – in school or not, dropped out or not, and advanced a grade or not, both conditional and unconditional on staying in school -- the non-IV parametric model we use is probit and the parametric IV models are IV-probit and IV-regress. For the two count data variables -- current grade and completed years of education – the parametric non-IV model is Poisson and the parametric IV models are IV-Poisson and IV-regress. For the continuous outcome -- grades per year of exposure to school-- the non-IV method is OLS and the IV method is IV-regress.

The nonparametric method we use for all outcome variables is propensity score matching with kernel matching. Propensity score matching methods can theoretically correct for selection into treatment if selection is mainly based on observable characteristics. This is achieved by predicting the probability of selection into treatment, the propensity score, as a function of observables and matching treatment and control observations on the propensity score. However, the propensity score is usually estimated using a Probit or Logit equation with some degree of arbitrariness as to what covariates to include in the model and what functional form to adopt. Boosted regression is an alternative method for selecting the propensity score equation that can significantly improve predictive accuracy. It is a multivariate non-

parametric regression technique that uses an automated, data-adaptive algorithm that can estimate the non-linear relationship between a variable of interest and a large number of covariates (McCaffrey et al. 2004). Boosting produces well-calibrated probability estimates by adding together many simple functions estimated on partitions of the data to obtain a smooth function of a large number of covariates. Boosted models are typically fit iteratively on a portion of the data called the "training data" and then their goodness of fit is tested on the reaming part of the data, referred to as "test data." We present results using a conventional probit approach to estimating the propensity score as well as ones that rely on boosted regression.

We implement boosted regression using the Stata plugin 'boost' (Schonlau 2005). Because such a highly flexible technique runs the risk of over-fitting (that is estimating a model that fits the training data well but that does not generalize to the rest of the data in the sample), there are a number of tuning parameters that must be carefully chosen. The first parameter is the proportion of the data set aside for the training data versus the test data. We use the default in Schonlau's program, which is 80 percent of the sample allocated to the training data. The second parameter is the number of interactions (number of splits in the tree). One split corresponds to a main effect model, two splits to a model with main effects and two-way interactions, etc. Hastie et al. (2001) suggest that two-way interactions are generally not sufficient, but any number of excess of four does not significantly improve the fit of the model. Accordingly, we use three-way interactions as our base estimate and present sensitivity analyses with two and four-way interactions. The third tuning parameter is the shrinkage parameter. Shrinkage means reducing the impact of each additional tree to avoid over-fitting. The smaller the shrinkage parameter is, the less the risk of over-fitting, but the larger the number of iterations must be. We follow the advice of McCaffrey et al. (2004) and use a relatively small shrinkage parameter of 0.0005 to ensure a smooth fit. The fourth parameter is the bagging parameter, which is the fraction of randomly selected observations used for fitting the regression tree at each iteration. We use the program's default value of 0.5. The last

parameter to select is the maximum number of iterations. Schonlau (2005) recommends that the product of the maximum number of iterations and the shrinkage parameter be in the range of 10 and 100. We set the maximum number of iterations at the lower end of this range at 20,000 iterations. Since our treatment variable is binary, we select a logistic distribution. Finally, the covariates we use in the model are the same covariates we used in the parametric model and the conventional matching model.

Once the IPV propensity score is estimated for each child using either probit or boosted regression, different matching methods can be used to match treatment and control observations. For our base estimates, we use kernel matching with the standard Epanechnikov kernel function, but undertake sensitivity analysis of our results using other matching methods, such as uniform and normal kernel and five nearest neighbors.[5]

Because the propensity scores are estimates, analytical standard errors are understated. We therefore report bootstrapped standard errors for all of the matching results. We also report bootstrapped standard errors for the IV-Poisson model since that model is estimated using a two-stage technique rather than by means of a full-information maximum likelihood method. Finally, to account for the fact that children in the same households probably share the same mother, all standard errors reported throughout the paper are based on the assumption that observations are clustered at the household level.

DATA AND SAMPLE

The sample includes all children ages 6-14 living in households where the mother is present, is under the age of 50, and has responded to the domestic violence module in the 2005 Demographic and Health

[5] Once the propensity score has been predicted using the "boost" command or a probit model, we use the user-written STATA ado file PSMATCH2 v4.04 to undertake the matching estimation. See Leuven and Sianesi (2003). Sensitivity analysis for different matching methods is only presented for the boosted regression model. Sensitivity analyses for the conventional matching approach using a probit first stage is available from the authors upon request.

Survey for Colombia. The DHS 2005 sample includes 31,140 children between the ages of 6 and 14. Of those, 23,253 lived with a mother who was between the ages of 15 and 49 and who was selected for interview for the domestic violence module. The final sample includes 21,827 children because of losses due to mothers who could not be safely interviewed in private or who had never been married or in a de facto union. As shown in Table 1, there is a noticeable and highly significant bivariate negative association between the presence of IPV in a household and children's educational outcomes. For example, children in households with IPV have a 1.8 percentage point lower probability of attending school, a 1.9 percentage point lower probability of advancing from one grade to the next and are behind by more than one-third of a year, on average, in terms of current grade attainment. It remains to be seen whether these differences remain after controlling for observables and correcting for selection. In addition, our chosen instrument, the maltreatment of the male partner when he was a child is strongly associated in a bivariate sense with the incidence of IPV in the household (Table 2). The descriptive statistics for our dependent and explanatory variables are shown in Table 3.

RESULTS

As laid out above, we present results on the effect of being exposed to IPV on seven different child education outcomes using both non-parametric methods (boosted matching) and IV and non-IV parametric methods. These results are summarized in Table 4, which shows the marginal effect of IPV exposure on the seven different outcomes for the various methods we consider.[6] What we find, overall, is quite robust: the magnitudes of the effects of IPV are similar across methods, although our instrumental variable results are statistically insignificant. While we feel quite confident about our instrument of choice, the IV methodology is less efficient and is thus less able to capture relatively small effects. Our

[6] Table A1 shows the first stage regressions for the IV models and the propensity score equation for the conventional matching model. Appendix B Tables show the full regressions for the parametric models for each of the outcome variables. Appendix B tables are available from the authors [if not online].

endogeneity tests reported in Table 5 generally indicate that exogeneity cannot be rejected, with the possible exception of completed years of education and grades per year of exposure to schooling. This suggests that both the parametric non-IV and the non-parametric (matching) results are consistent. We favor the results from the non-parametric boosted matching methodology over those from parametric methods because of the much weaker assumptions they make about functional form. In general, our results suggest that IPV adversely affects child schooling. However, it does not appear to do so through increasing drop-out– which, in a high enrollment context like Colombia, might be socially unacceptable – but rather because children exposed to IPV experience delays in their education, reducing their grade attainment.

Estimation Details and Results for School Attendance

In what follows, we begin by discussing the results for the first dependent variable – child's school attendance – from the various estimation methods used and then move to a comparison of the methods and results for the other dependent variables in the next section.

We start with the boosted matching results, which show that the average treatment effect on the treated (ATT) is a reduction in the probability of currently attending school of 1.2 percentage points, which is just shy of being significant at the 5 percent level using the bootstrapped standard errors (See Table 5). The conventional matching result on school attendance is very similar and is significant at the 5 percent level. From this we can conclude that IPV reduces the probability of school attendance for the children whose mothers are subjected to violence by 1.2 percentage points, which means that IPV increases the probability of non-attendance from 6.9 percent for a matched control group to 8.2 percent for the treatment group, a relative increase of nearly 19 percent.

The Probit results produce larger estimates showing a reduction of 2.7 percentage points for a reference child and are significant at the 5 percent level.[7] Given the possible endogeneity of IPV, we estimate an instrumental variable Probit model (IV-Probit) and a linear IV regression (IV-regress). We use the mother's partner childhood experience of violence as an instrument in both models. Both models yield smaller negative and insignificant coefficients for the effect of IPV on school attendance. We use the IV-regress model to conduct a number of tests on the validity of the instruments. The first is whether the instrument has sufficient explanatory power in the first stage equation. As shown in Table 5, the F-statistic for the instruments in the first stage is 190, which is well above the threshold level of 30 that is usually necessary in two-stage least squares models. With one instrument it is not possible to conduct an overidentification test of the exclusion restriction. However, when we include the mother's witnessing of violence as a child as an additional instrument, both the Sargan and Bassman overidentification tests showed that the instruments are excludable from the second stage equation. Finally, we conducted a test of endogeneity, which indicated that exogeneity could not be rejected, lending credence to the non-IV results.

Comparison of Results for All Dependent Variables

The effect of IPV on three other limited dependent variables – grade advancement conditional on staying in school (vs. repeating), recent drop-out, and unconditional grade advancement (vs. repeating or dropping-out) was estimated using the same methods as reported for school attendance. Results are again reported in Table 4. For each of these three dependent variables the IV-Probit and IV-Regress models produced statistically insignificant estimates of the effect of mother's IPV on the child's educational outcome. However the results are similar in magnitude for the most part to those of the probit and

[7] The reference child has all his continuous variables set at the mean and all his dummy variables set at zero.

matching methods. In each case, the test of endogeneity reveals that exogeneity cannot be rejected (see Table 5), so it is appropriate to focus on the results of the Probit and matching models.

Comparing the boosted matching, conventional matching and Probit results, we note that they are generally of the same order of magnitude. The two matching methods produce very similar results and the results from the probit estimation are typically larger probably because matching does a better job in selecting the control group. According to these models, mother's IPV has no discernible effect on the likelihood that children dropped out of school in the year prior to the survey. The two grade advancement measures do, however, seem to be affected. IPV significantly reduces grade advancement conditional on continued attendance at school by 1.5 and 1.8 percentage points, respectively, according to the both matching estimators and the Probit estimator. Overall, in our sample, only four percent of children do not advance from one grade to the next (conditional on staying in school) (See Table 3). IPV therefore increases non-advancement by a relative rate of 34 to 41 percent.

Based on the same estimators, unconditional grade advancement – a combination of the drop-out and unconditional advancement variables – is reduced by 2.1 to 2.8 percentage points. Again, these estimates are statistically significant at higher than the 1 percent level. In this sample, eight percent of children did not advance, instead dropping out or repeating the grade. IPV thus increases non-advancement by a relative rate of 25 to 33 percent. The fact that results differ for conditional and unconditional advancement may indicate that there is, in fact, increased drop-out due to IPV, but that it is imprecisely estimated.

The next two dependent variables, current grade – conditional of still being in school – and completed years of education, are appropriately treated as count data and therefore require different estimation

methods. Results for current grade using a Poisson model[8] and boosted and conventional matching models indicate are negative and statistically significant of IPV. In this case, the parametric model gives the lowest estimate (a reduction of 0.06 grades) and the boosted matching model the highest estimate (a reduction of 0.12 grades) (see Table 4). Alternatively, we may say that 6 to 12 percent of children exposed to IPV are likely to be delayed by one grade. It would be better to look at grade when controlling for age; this is what the grades per year measure, below, does.

Current grade attainment was also estimated using IV-Poisson and IV-Regress. For grade attainment, the IV results were qualitatively different, being positive rather than negative, but all were statistically insignificant.[9] Tests of the endogeneity of IPV with regards to current grade shown in Table 5 rejected exogeneity but only at a p-value of 0.097. This suggest that the matching and Poisson model estimates are still marginally acceptable.

As mentioned above, the analysis of completed years of education used a sample of children ages 10-14 who ever attended school. This sample should capture late-starters who, if they were left out of the analysis, would bias the IPV effect downwards. It also includes drop-outs. Again the boosted matching, conventional matching and Poisson results are statistically significant at at least the 5 percent level (Table 4). The effects of -0.09 to -0.12 indicate that IPV results in the loss of about one-tenth of a year of education. In other words, if a 10-to-14-year-old is exposed to intimate partner violence, there is a 10

[8] We also estimated a negative binomial model. The negative binomial differs from the Poisson model by a parameter alpha, which measures data dispersion. When alpha equals zero, the negative binomial reduces to a Poisson. In our case the estimated alpha was small enough to be effectively zero, so we did do not present results from this model.

[9] Standard errors for the IV-Poisson needed to be bootstrapped because, in the second stage, the analytical standard errors were not corrected for the inclusion of a predicted regressor. Bootstrapped SEs are shown in Table 5.

percent probability that she or he will have one fewer completed years of education as a result of this exposure.

Similar to the case of current grade, the IV-Probit and IV-regress results are qualitatively different and statistically insignificant. The test of endogeneity rejects exogeneity at a p-value of 0.058 (see Table 5) casting some doubt on the matching and Poisson results.

Finally, the grades per year variable provides yet another way of understanding children's schooling attainment by measuring the average number of grades completed per year of exposure to school. As discussed above, we use a sample of children ages 10 to 14. The appropriate regression methodology in this case is Ordinary Least Squares (OLS). Only the OLS results and the conventional boosting are statistically significant at the 5 percent level; they suggest a small decrease of 0.017-0.02 grades per year due to IPV. In contrast, the IV-regress results indicate a positive effect that is statistically significant at the 5 percent level. A test of endogeneity also suggests that exogeneity is rejected at a p-value of 0.008, casting some doubt on the non-IV results.

A Note on Matching Model Methodology

In all but one of the results for IPV discussed above, matching model estimates are statistically significant at the percent level, the exception being recent dropout. We conducted a sensitivity analysis on the choice of matching methodology used to assess the robustness of our results (See Table 6). All the methods test the mean difference in the outcome of treated observations and matched control observations. We discussed above the difference between the conventional way of estimating the propensity score, which uses a Probit regression, and the boosted regression model, which makes use of a non-parametric data-adaptive algorithm. Once the propensity score is estimated, matching is carried out in the same way in both methods.

If kernel matching methods are used, the control observations that are matched to each treated observation are weighted using a kernel function. The kernel function places higher weight on untreated observations that are closer to the treated observation and lower weight on more distant ones (Heckman et al. 1998). The kernel function we use for our base estimate is Epanechnikov (the default). We also conduct sensitivity analyses using Normal and Uniform kernel functions.[10] Five nearest neighbors matching uses the average outcome for the five untreated observations closest to each treated observation, giving each of them equal weight.

We present in Table 6, the results from the conventional matching model using an Epanechnikov kernel function as well as ones from boosted matching using Epanechnikov, normal and uniform kernel functions, as well as five-nearest neighbors matching.[11] Since we also needed to make a decision on the level of interactions to include in the boosted regression model, we also conduct sensitivity analysis on this parameter. Our base model uses a three-level interaction model, but we present in Table 7, results for two and four-level interactions for comparison.

We can see from Tables 6 and 7 that the results for "attending school", "grade advancement or not" and "advance or drop out-repeat" are highly robust to the choice of matching method. The results for "recent drop out", which were statistically insignificant using the Epanechnikov kernel method in the boosted matching model, become larger and statistically significant at the 10 percent level when using conventional matching or a normal kernel function. The results for grade attainment are somewhat less

[10] The kernel function K(u) is a function of the distance measure $u = (p_i - p_j / h)$, where p_i is the propensity score of the treated observation and p_j is the propensity score of the control observation and h is a pre-specified bandwidth (set to a default of 0.06 in PSMatch2). For the Epanechnikov kernel $K(u) \propto (1-u^2)$ if $|u \leq 1|$, $K(u) = 0$, otherwise. The Normal Kernel is based on $K(u) \propto \exp(-u^2/2)$ and uses all untreated observations. The uniform kernel uses equal weights for all observations falling within the bandwidth h. See Sianesi (2001).

[11] Similar sensitivity results on the conventional matching model are available from the authors upon request.

robust, ranging from -0.09 to -0.19 and their statistical significance similarly ranging from 5 percent to under 0.1 percent. Results for years of education are also less robust, ranging from -0.05 and statistically insignificant (five-nearest neighbors and model with four interactions) to -0.1595 and significant at the 0.1 percent level using the normal kernel function. Similarly results for grade per year vary from -0.0075 and statistically insignificant (5-nearest neighbors) to -0.0184 and significant at the 1 percent level (normal kernel).

After matching, balancing tests for all covariates were conducted. Ideally, means of covariates should be the same for the treated and matched untreated observations. In the conventional matching model, all the covariates are balanced for all outcomes, meaning that difference in means tests between the means of all the covariates in the treated and matched control samples were insignificant. The boosted matching model did not perform as well in balancing the covariates, with the number of unbalanced covariates specified in Table 6 for each method and outcome variable. The fact that the boosted model does not match the covariates as well at the mean value does not mean that it necessarily performs worse than the conventional model. A highly non-linear model might do a poorer job at the mean, but a better job matching treated and matched controls at the other points of the distribution.

Other Results

With respect to estimated effects of other variables on the educational outcomes, the results were remarkably consistent across the different parametric estimation methods in terms of signs and levels of significance (See Tables B1 to B7). A few variables stand-out. The effect of the child being female is positive across all dependent variables except drop-out, which is insignificant. As is typical in such analyses, mother's education always has positive effects on educational outcomes, as does partner's education for attendance and grade attainment. Municipality-level averages of female education had the expected effects in most cases (attendance, drop-out and unconditional attainment). Oddly enough,

higher municipality-level education averages for adult males increased drop-out. If the household had migrated, children's educational outcomes were negatively affected. Wealth consistently improved outcomes except in the case of conditional grade advancement, where the effect is insignificant. The household's numbers of 0-5 year-olds had negative effects on all educational outcomes for children 6-14. Even the municipality level child-woman ratio had negative effects on grade advancement (conditional and unconditional) and grade attainment. Clearly, caring for young children is among the responsibilities of 6-14 year-olds in Colombia – except, perhaps, when the household includes multiple women ages 18-64, whose presence has positive effects on children's school attendance and grade attainment.

CONCLUSIONS

Very few studies have considered the effects of family violence on children's human capital formation. The studies that do so almost never take account of the possibility of endogeneity and sample selection bias, as we do using cross-sectional data and as Emery (2011) has done in a rare study using panel data. Although we find that endogeneity is not an issue, for the most part, in this particular instance, it may be a problem in other contexts, seriously biasing results. Even if endogeneity is not an issue, selection can still be a major challenge because of the highly intertwined factors that affect family outcomes. We address this challenge by adopting the most flexible non-parametric methods available to match affected children with comparable controls.

To tackle endogeneity, it is necessary to have a convincing instrument. In addition to much evidence in the literature on the relevance of our instrument of choice – the childhood experience of violence of the mother's partner – it passed all the necessary tests for the validity of an instrument and had a very strong first stage. A series of careful tests do not, however, reject exogeneity (with the possible exception of years of education and grades per year), so we report results from non-parametric boosted and conventional matching models and non-IV parametric models (Probit, Poisson or OLS, depending on the

outcome variable). Although we report results from both the non-parametric matching methods as well as parametric methods, we place more credence on the non-parametric results because they rely upon fewer functional form assumptions. We also present estimates from parametric IV models; the IV and non-IV estimates of the effect of a mother's experience of intimate partner violence with respect to all the binary outcome variables are qualitatively similar, but with varying degrees of precision. They differ qualitatively for the continuous and count data outcomes (current grade, completed years of education and grades per year).

A woman's experience of intimate partner violence has a negative effect on her children's school attendance, decreasing it by 1.2 to 2.7 percentage points, according to the matching and Probit results, respectively. This effect sounds small but is in fact substantial if compared to an average nonattendance rate of 6.9 percent: thus IPV increases non-attendance for the average child by 17 to 39 percent. To our knowledge, the only other study of IPV conducted in a developing country using school attendance as an outcome also found that IPV had a negative effect on education outcomes (Jayasinghe et al. 2009). School attendance has not been used as an education outcome in studies with populations in the United States, probably because attending school is compulsory and compliance is extremely high.

While mother's IPV does not seem to affect the probability that a child dropped out of school in the previous academic year, the two grade advancement measures are negatively affected by IPV. It reduces grade advancement conditional on staying in school by about 1.5 to 1.8 percentage points and 2.1 to 2.8 percentage points, unconditional on staying in school. Again, these effects are substantial as they constitute from 34 to 41 percent of the average conditional non-advancement rate and from 25 to 33 percent of the average unconditional rate, respectively. Studies conducted in Brazil (Durand et al., 2011) and the United States (Emery, 2011) show similar findings.

Effects on current grade (conditional on continued school attendance) might be viewed as showing a cumulative effect of mother's experience of intimate partner violence on their children's educations for children who remain in school. By restricting attention to children in school, however, this measure takes into account the cumulative effect of IPV on grade repetition but not its effect on dropout. Mother's IPV is estimated to reduce current grade by 0.06 to 0.12 grades, or alternatively we may say that 6 to 12 percent of children living with such family violence are likely to be delayed by one grade. When we do not condition on staying in school, years of education completed are reduced by a similar amount. When we again consider grade attainment but controlling for years of exposure to school, the grades per year measure picks up a small negative effect of IPV but the result is not robust to the method used and there appears to be some evidence of endogeneity of IPV to that particular outcome.

Overall, our results imply that intimate partner violence does not have a discernible effect on children's drop-out in any particular school year in the high-enrollment context of Colombia. Instead, it suggests that students miss school sporadically (reducing attendance) and, via undefined mechanisms related to anxiety, exhaustion, pain, and/or a reduced ability to study at home, perform less well in school, leading to a greater likelihood of repeating grades and a lower level of cumulative human capital.

REFERENCES

Angrist, J. and J-S Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Bourdillon, M., Levison, D., Myers, W., & White, B. (2010). *Rights and Wrongs of Children's Work*, New Brunswick NJ: Rutgers University Press.

DeJonghe, E. S., Bogat, G. A., Levendosky, A. A., & von Eye, A. (2008). "Women Survivors of Intimate Partner Violence and Post-Traumatic Stress Disorder: Prediction and Prevention." *Journal of Postgraduate Medicine, 54*:294 - 300.Durand, J. G., Schraiber, L. B., Franca, I., & Barros, C. (2011). "Impact of exposure to intimate partner violence on children's behavior." *Revista De Saude Publica, 45*:355 - 364.

Emery, C. R. (2011). "Controlling for Selection Effects in the Relationship Between Child Behavior Problems and Exposure to Intimate Partner Violence." *Journal of Interpersonal Violence, 26*:1541 - 1558.

---------- (2009). “Stay for the Children? Husband Violence, Marital Stability, and Children’s Behavior Problems.” *Journal of Marriage and Family, 71*:905-916,

Evans, S. E., Davies, C., & DiLillo, D. (2008). "Exposure to Domestic Violence: A Meta-Analysis of Child and Adolescent Outcomes." *Aggression and Violent Behavior, 13*:131 - 140.

Friedemann-Sánchez, Greta, Rodrigo Lovatón Dávila. (2012). “Intimate Partner Violence in Colombia: Who is at Risk?” *Social Forces, 91*: 663-688.

Garcia-Moreno, C., Jansen, H., Ellsberg, M., Heise, L., & Watts, C. H. (2006). "Prevalence of Intimate Partner Violence: Findings from WHO multi-country Study on Women's Health and Domestic Violence." *The Lancet, 368*:1260 - 1269.

Hastie, T., R. Tibshirani and J. Friedman. 2001. *The Elements of Statistical Learning.* New York: Springer-Verlag.

Heaton, T. B. & Forste, R. (2008). "Domestic Violence, Couple Interaction and Children's Health in Latin America." *Journal of Family Violence, 23*: 183 - 193.

Heckman, J. J., Ichimura, H. & Todd, P. E. (1998), “Matching as an Econometric Evaluation Estimator,” *Review of Economic Studies, 65*, 261 - 294.

Huth-Bocks, A. C., Levendosky, A. A., & Semel, M. A. (2001). "The Direct and Indirect Effects of Domestic Violence on Young Children's Intellectual Functioning." *Journal of Family Violence, 16*:269 - 290.

Jayasinghe, S., Jayawardena, P., & Perera, H. (2009). "Influence of Intimate Partner Violence on Behavior, Psychological Status and School Performance in Sri Lanka." *Journal of Family Studies, 15*:274 - 283.

Kiesel, L., Piescher, K., Edleson, J. (2011). Direct and Indirect Child Exposure to Violence: Effects on Academic Performance and Disability. (Mimeo)

Kitzmannn, K. M., Gaylord, N. K., Holt, A. R., & Kenny, E. D. (2003). " Child witnesses to domestic violence: A meta-analytic review." *Journal of consulting and Clinical Psychology,* 71:339 - 352.

Kishor, S. & Johnson, K. (2004). *Profiling Domestic Violence,* Calverton: ORC Macro.

Leuven, E. & Sianesi, B. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html. This version 4.0.4, 10 November 2010.

Martin, S. L., Moracco, K. E., Garro, J., Tsui, A. O., Kupper, L. L., Chase, J. L., & Campbell, J. C. (2002). "Domestic Violence Across Generations: Findings From Northern India." *International Journal of Epidemiology, 31*:560 - 572.

Matthew, P. K., Kapua, F., Soaki, P. J., & Watters, D. A. K. (1996). "Trauma Admissions in the Southern Highlands of Papua New Guinea." *Australian and New Zealand Journal of Surgery, 66*:659 - 663.

McCaffrey, D. F., G. Ridgeway, and A.R. Morral. 2004. “Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies.” Psychological Methods 9(4): 403-425.

Pallitto, C. C. & O'Campo, P. (2004). "The Relationship Between Intimate Partner Violence and Unintended Pregnancy: Analysis of a National Sample from Colombia." *International Family Planning Perspectives, 30*: 165 - 173.

Patel, D., Rapporteur. (2011). "Preventing Violence Against Women: Workshop Summary." Washington DC: Institute of Medicine.

Peek-Asa, C., Maxwell, L., Stromquist, A., Whitten, P., Limbos, M. A., & Merchant, J. (2007). "Does Parental Physical Violence Reduce Children's Standardized Test Score Performance?" *Annals of Epidemiology*, *17*:847 - 853.

Pulerwitz, J. & Barker, G. (2008). "Measuring Attitudes toward Gender Norms among Young Men in Brazil." *Men & Masculinities 10*:322 - 38.

Renzetti, C. M. (2009). "Economic Stress and Domestic Violence." *National Online Resource Center on Violence Against Women*. Available online at http://www.vawnet.org/summary.php?doc_id=2187&find_type=web_desc_AR

Sánchez, F. (2007). *Las Cuentas de la Violencia*, Bogota: Norma.

Sánchez Lara, M., Gomez Lopez, C., Pardo Lugo, J. C., Arenas Garcia, M. C., Da Silva, M., Siado, P., & Concha Jurado, O. (2008). "Promocion y Monitoreo de los Derechos Sexuales y Reproductivos de Mujeres Víctimas de Desplazamiento Forzado con Enfasis en Violencias Intrafamiliar y Sexual." Bogotá: Defensoría del Pueblo.

Schonlau, M. 2005. “Boosted Regression (Boosting): An introductory tutorial and a Stata plugin.” The Stata Journal 5(3): 330-354.

Sianesi, B. (2001). "Implementing Propensity Score Matching Estimation with STATA." Paper presented at the UK STATA users group VII meeting, London. http://fmwww.bc.edu/RePEc/usug2001/psmatch.pdf] accessed 3/22/2012.

Thompson, R. & Whimper, L. A. (2010). "Exposure to Family Violence and Reading Level of Early Adolescents." *Journal of Aggression, Maltreatment and Trauma*, *19*:721 - 733.

Whitfield, C. L., Anda, R. F., Dube, S. R., & Felitti, V. J. (2003). "Violent Childhood Experiences and the Risk of Intimate Partner Violence in Adults." *Journal of Interpersonal Violence*, *18*:166 – 185.

Table 1. Various School Outcomes by Presence of Intimate Partner Violence in Household (weighted)

	Intimate Partner Violence in HH			Sample size*	p-value of t-test of difference in means
	No IPV	IPV	Total		
Attending school or not	93.7	91.9	93.3	21,827	0.0000
Grade advancement or not†	96.0	94.1	95.6	19,683	0.0000
Recent drop out or not	2.18	2.58	2.27	20,433	0.0061
Advance or drop out / repeat	92.3	89.1	91.6	20,429	0.0000
Current grade if still in school	4.31	3.91	4.22	20,370	0.0000
Years of education if ever attended (ages 10-14)	5.03	4.72	4.96	11,689	0.0000
Grades per year of exposure to school (ages 10-14)	0.836	0.789	0.826	11,890	0.0000

* Unless otherwise noted, refers to children 6-14 living with mothers 15-49 who responded to the domestic violence module of the Colombia 2005 DHS.

† Conditional on staying in school

Source: Authors' calculations from Colombia 2005 DHS.

Table 2. Presence of Intimate Partner Violence in Household (%) by Partner Maltreatment as a Child (weighted)*

	Partner Maltreated as a Child		
Intimate Partner Violence in HH	No	Yes	Total
No IPV	81.3	67.2	76.9
IPV	18.7	32.8	23.1
Total	100.0	100.0	100.0
% children whose mothers' partner was maltreated as a child			31.0
Sample size	15,135	6,692	21,827

* Same notes and source as Table 1

Table 3. Descriptive Statistics, Children Ages 6-14, Colombia, DHS 2005 (Weighted)

Variable	Mean	Standard Deviation
Dependent variables:		
school attendance (or not)	0.933	0.250
grade advancement (vs. repeat, conditional on staying in school)	0.956	0.205
recent dropout (or not)	0.023	0.149
grade advancement (vs. drop-out / repeat)	0.916	0.278
current grade if still in school	4.222	2.477
years of education if ever attended (ages 10-14)	4.962	1.854
grades per year of exposure to school (ages 10-14)	0.826	0.267
Instrumental variables:		
partner beaten when a child	0.310	0.463
missing: partner beaten as child	0.122	0.327
mother's parents' violence	0.335	0.472
missing: mother's parents' violence	0.039	0.193
Endogenous regressor:		
mother's intimate partner physical violence	0.310	0.463
Child variables:		
child is female	0.495	0.500
child's age	9.923	2.545
child's age-squared/100	1.050	0.511
child is son/daughter of HH head	0.845	0.362
Mother variables:		
mother's age when child was age 6	31.779	5.806
mother's age squared/100 (when child was 6)	10.436	3.828
mother's years of education completed	7.036	4.147
mother cohabitating (reference category)	0.432	0.495
mother is married	0.364	0.481
mother is widow	0.030	0.171
mother is divorced etc.	0.174	0.379
Partner-of-mother variables:		
partner's years of education completed	6.732	4.458
missing: partner's years of education	0.028	0.165
Household variables:		
HH has migrated	0.148	0.355
wealth quintile 1 (reference category)	0.227	0.419
wealth quintile 2	0.222	0.416
wealth quintile 3	0.206	0.405
wealth quintile 4	0.187	0.390
wealth quintile 5 (richest)	0.158	0.364

Table 3. Descriptive Statistics, Children Ages 6-14, Colombia, DHS 2005 (Weighted) (Cont'd)

Variable	Mean	Standard Deviation
Household composition variables:[§]		
mother has relatives in HH	0.164	0.370
partner has relatives in HH	0.051	0.221
# children ages 0-5	0.642	0.856
# girls ages 6-11	0.725	0.766
# boys ages 6-11	0.726	0.756
# girls ages 12-14	0.335	0.545
# boys ages 12-14	0.365	0.562
# girls ages 15-17	0.161	0.404
# boys ages 15-17	0.179	0.432
# women ages 18-64	1.360	0.693
# men ages 18-64	1.149	0.762
# women ages 65+	0.086	0.292
# men ages 65+	0.069	0.256
Geographic variables:		
rural (vs. urban)	0.309	0.462
Central (reference category)	0.164	0.370
Atlantic region	0.227	0.419
Oriental region	0.194	0.396
Pacific region	0.169	0.374
Bogota region	0.134	0.341
Territories region	0.112	0.315
Municipality-level variables:		
average wealth factor score	-0.021	0.061
average years of education, women 25-64	7.139	1.524
average years of education, men 25-64	7.123	1.726
child-woman ratio (0-4 / f 15-49)	0.386	0.114
% of HHs female-headed	29.687	5.844
% of population living abroad[§§]	1.186	1.361
% of employed women in formal work	23.784	10.333
% of employed men in formal work	20.063	11.411
% HHs with access to piped water	85.092	17.048
% HHs with access to sewer	68.508	26.378
% HHs cooking with firewood etc.	18.836	20.960
Sample size		21,827

Notes: [§]includes the index child

[§§]individuals abroad / individuals present in municipality

Table 4. Marginal Effects of Intimate Partner Violence on Children's Educational Outcomes, Colombia, 2005

Dependent Variable	Boosted Matching		Conventional Matching		Probit		IV Probit	IV Regress	N
Attending school or not	-0.0121		-0.0121		-0.0271		-0.0008	-0.0074	21,827
(SE)	(0.0046)	**	(0.0046)	**	(0.0112)	*	(0.0859)	(0.0351)	
(BSSE)	*(0.0062)*	+	*(0.0054)*	*					
Grade advancement or not†	-0.0154		-0.0152		-0.0184		-0.0269	-0.0124	19,683
(SE)	(0.0044)	***	(0.0043)	***	(0.0056)	***	(0.0396)	(0.0278)	
(BSSE)	*(0.0049)*	**	*(0.0053)*	**					
Recent drop out or not	0.0041		0.0052		0.0099		-0.0020	0.0045	20,433
(SE)	(0.0028)		(0.0028)	+	(0.0067)		(0.0435)	(0.02)	
(BSSE)	*(0.0033)*		*(0.0032)*						
Advance or drop out / repeat	-0.0223		-0.0206		-0.0284		-0.0338	-0.0227	20,429
(SE)	(0.0054)	***	(0.0053)	***	(0.0078)	***	(0.0547)	(0.0352)	
(BSSE)	*(0.0058)*	***	*(0.0058)*	***					
	Boosted Matching		**Conventional Matching**		**Poisson**		**IV Poisson**	**IV Regress**	**N**
Current Grade	-0.1218		-0.0937		-0.0590		0.1525	0.1794	20,370
(SE)	(0.0452)	**	(0.0437)	*	(0.0176)	***	(0.1162)	(0.1523)	
(BSSE)	*(0.0415)*	**	*(0.0454)*	*			*(0.1131)*		
Completed Years of Education ††	-0.104		-0.1216		-0.0918		0.3296	0.3489	11,689
(SE)	(0.0465)	*	(0.0448)	**	(0.0287)	**	(0.2029)	(0.2478)	
(BSSE)	*(0.0489)*	*	*(0.0458)*	**			*(0.2025)*		
	Boosted Matching		**Conventional Matching**		**OLS**			**IV Regress**	**N**
Grades per year §§	-0.0086		-0.0214		-0.0166			0.1015	11,890
(SE)	(0.0069)		(0.0084)	*	(0.0059)	**		(0.0458)	*
(BSSE)	*(0.0081)*		*(0.0100)*	*					

Note: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + p<=0.1

Analytical standard errors (SE) are in parentheses and Bootstrapped standard errors (BSSE) are in parentheses and italics.

SEs are corrected for clustering of children in the same household. Bootstrapped SEs are based on 100 replications.

† Conditional on staying in school.

§ Regressions limited to children 6 to 14 who are currently enrolled in school

†† Regressions limited to children ages 10 to 14 who ever attended school

§§ Regressions limited to children ages 8 to 14.

Table 5. Results of Tests for Goodness of Fit and Endogeneity for IV Methods

	Test Statistic	Attending school or not	Grade advance-ment or not†	Recent drop out or not	Advance or drop out / repeat	Current Grade	Completed Years of Education	Grades per year
First Stage Goodness of Fit	F(1, C-1)†	190.32	187.45	197.35	197.429	196.1	118.82	117.77
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Adj. R-Sq.	0.0678	0.0663	0.0696	0.0673	0.0678	0.0602	0.0607
Test of Endogeneity	F (1, C-1)	0.0230	0.0132	0.0002	0.0032	2.7548	3.60656	7.0508
(H0: treatment is exogenous)	p-value	0.8795	0.9087	0.9896	0.9549	0.0970	0.0576	0.0079
Number of Observations	N	21,827	19,683	20,433	20,429	20,370	11,689	11,890
Number of Clusters	C	13,182	12,365	12,713	12,710	12,620	8,705	8,799

Notes:

All tests are conducted using the IV regress two-stage least squares model.

All tests are adjusted for clustering at the household level.

† C= number of clusters (households)

Table 6. Analysis of the Sensitivity of Results to Using Conventional vs. Boosted Matching and to Different Matching Methodologies in Boosted Matching. Marginal Effects of Intimate Partner Violence on Children's Educational Outcomes

Dependent Variable	**Conventional Matching**			**Boosted Matching**											
	Kernel Epanechnikov			**Kernel** Epanechnikov			**Kernel** Normal			**Kernel** Uniform			**Five Nearest Neighbors**		
Attending school or not	-0.0121	**	§(0)	-0.0121	**	§(9)	-0.0139	**	§(14)	-0.0123	**	§(7)	-0.0139	**	§(13)
(SE)	(0.0046)			(0.0046)			(0.0045)			(0.0046)			(0.0051)		
Grade advancement or not	-0.0152	***	§(0)	-0.0154	***	§(6)	-0.0160	***	§(8)	-0.0155	***	§(2)	-0.0142	**	§(10)
(SE)	(0.0043)			(0.0044)			(0.0043)			(0.0043)			(0.0048)		
Recent drop out or not	0.0052	+	§(0)	0.0041		§(12)	0.0046	+	§(12)	0.0043		§(12)	0.0024		§(17)
(SE)	(0.0028)			(0.0028)			(0.0027)			(0.0028)			(0.0031)		
Advance or drop out / repeat	-0.0208	***	§(0)	-0.0223	***	§(8)	-0.0234	***	§(8)	-0.0224	***	§(6)	-0.0253	***	§(13)
(SE)	(0.0053)			(0.0054)			(0.0053)			(0.0054)			(0.0059)		
Grade Attainment	-0.0937	*	§(0)	-0.1218	**	§(5)	-0.1743	***	§(9)	-0.1324	**	§(3)	-0.1002	*	§(9)
(SE)	(0.0437)			(0.0452)			(0.0433)			(0.0447)			(0.0497)		
Years of Education	-0.1216	**	§(0)	-0.104	*	§(12)	-0.1595	***	§(2)	-0.1141	*	§(9)	-0.0530		§(19)
(SE)	(0.0448)			(0.0465)			(0.0447)			(0.0461)			(0.0530)		
Grades per year	-0.0214	*	§(0)	-0.0086		§(5)	-0.0184	**	§(2)	-0.0105		§(2)	-0.0075		§(12)
(SE)	(0.0084)			(0.0069)			(0.0067)			(0.0069)			(0.0077)		

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

Standard errors (SE) are in parentheses.

All treated and untreated observations are on the common support.

§ Number of covariates that are not balanced are in parentheses.

Table 7. Analysis of the Sensitivity of Boosted Matching Results to Using Different Levels of Interaction in Boosted Regression.

Marginal Effects of Intimate Partner Violence on Children's Educational Outcomes

Dependent Variable	**Three Interactions**			**Two Interactions**			**Four Interactions**		
Attending school or not	-0.0121	**	§(9)	-0.0122	**	§(8)	-0.0120	*	§(8)
(SE)	(0.0046)			(0.0046)			(0.0048)		
Grade advancement or not	-0.0154	***	§(6)	-0.0141	**	§(10)	-0.0160	***	§(12)
(SE)	(0.0044)			(0.0045)			(0.0045)		
Recent drop out or not	0.0041		§(12)	0.0041		§(8)	0.0038		§(13)
(SE)	(0.0028)			(0.0028)			(0.0029)		
Advance or drop out / repeat	-0.0223	***	§(8)	-0.0225	***	§(9)	-0.0210	***	§(10)
(SE)	(0.0054)			(0.0053)			(0.0055)		
Grade Attainment	-0.1218	**	§(5)	-0.1923	***	§(7)	-0.1507	**	§(6)
(SE)	(0.0452)			(0.044)			(0.0463)		
Years of Education	-0.104	*	§(12)	-0.1356	**	§(2)	-0.0523		§(6)
(SE)	(0.0465)			(0.0455)			(0.0485)		
Grades per year	-0.0086		§(5)	-0.0132	+	§(5)	-0.0126	+	§(8)
(SE)	(0.0069)			(0.0068)			(0.0071)		

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

Standard errors (SE) are in parentheses.

All treated and untreated observations are on the common support.

§ Number of covariates that are not balanced are in parentheses.

Appendix Table A1. Coefficient Estimates from First Stage Regressions or IV-Regress and Conventional Matching Models†

Dependent Variable: Whether or not mother experienced physical intimate partner violence in past 12 months

	Linear Probability Model		Conventioal Matching (Probit)	
Instruments				
mother's partner experienced violence as child	0.1282	***	0.4502	***
	(0.0093)		(0.0220)	
Child variables:				
child is female	0.0030		0.0103	
	(0.0050)		(0.0263)	
child's age	-0.0151	+	-0.0551	
	(0.0083)		(0.0345)	
child's age-squared	0.0468		0.1703	
	(0.0413)		(0.1748)	
child is son/daughter of HH head	-0.0309	*	-0.1099	**
	(0.0149)		(0.0370)	
Mother variables:				
mother's age when child was age 6	-0.0049		-0.0034	
	(0.0057)		(0.0172)	
mother's age squared/100 (when child was 6)	-0.0024		-0.0316	
	(0.0085)		(0.0263)	
mother's years of education completed	-0.0032	*	-0.0125	***
	(0.0013)		(0.0034)	
mother is married	-0.0469	***	-0.1983	***
	(0.0090)		(0.0250)	
mother is widow	-0.0030		-0.0124	
	(0.0227)		(0.0617)	
mother is divorced etc.	0.1489	***	0.4725	***
	(0.0140)		(0.0312)	
Partner-of-mother variables:				
partner's years of education completed	-0.0014		-0.0062	+
	(0.0012)		(0.0032)	
missing: partner's years of education	-0.0623	*	-0.2072	***
	(0.0273)		(0.0638)	
missing data on mother's partner childhood experience	0.0872	***	0.3179	***
of violence	(0.0132)		(0.0297)	

Appendix Table A1. Coefficient Estimates from First Stage Regressions or IV-Regress and Conventional Matching Models† (Cont'd)

	Linear Probability Model		Conventioal Matching (Probit)	
Household variables:				
HH has migrated	0.0166		0.0575	*
	(0.0110)		(0.0265)	
wealth quintile 2	0.0101		0.0285	
	(0.0137)		(0.0331)	
wealth quintile 3	0.0107		0.0300	
	(0.0161)		(0.0398)	
wealth quintile 4	-0.0035		-0.0243	
	(0.0179)		(0.0459)	
wealth quintile 5 (richest)	-0.0151		-0.0851	
	(0.0202)		(0.0542)	
Household composition§ variables:				
mother has relatives in HH	-0.0372	*	-0.1291	***
	(0.0154)		(0.0389)	
partner has relatives in HH	-0.0024		0.0006	
	(0.0190)		(0.0499)	
# children ages 0-5	0.0016		0.0032	
	(0.0052)		(0.0124)	
# girls ages 6-11	0.0098		0.0338	*
	(0.0064)		(0.1538)	
# boys ages 6-11	0.0062		0.0202	*
	(0.0064)		(0.0154)	
# girls ages 12-14	0.0035		0.0160	
	(0.0087)		(0.0208)	
# boys ages 12-14	0.0113		0.0427	*
	(0.0083)		(0.0203)	
# girls ages 15-17	0.0132		0.0476	*
	(0.0099)		(0.0240)	
# boys ages 15-17	0.0088		0.0317	
	(0.0096)		(0.0229)	
# women ages 18-64	-0.0115	+	-0.0409	*
	(0.0066)		(0.0184)	
# men ages 18-64	0.0066		0.0227	
	(0.0060)		(0.0153)	
# women ages 65+	-0.0082		-0.0411	
	(0.0154)		(0.0409)	
# men ages 65+	-0.0227		-0.0854	+
	(0.0158)		(0.0446)	

Appendix Table A1. Coefficient Estimates from First Stage Regressions or IV-Regress and Conventional Matching Models† (Cont'd)

	Linear Probability Model		Conventioal Matching (Probit)	
Geographic variables:				
rural (vs. urban)	-0.0177		-0.0674	*
	(0.0127)		(0.0313)	
Atlantic region	-0.0085		-0.0394	
	(0.0126)		(0.0319)	
Oriental region	0.0182		0.0669	+
	(0.0138)		(0.0346)	
Pacific region	0.0159		0.0533	
	(0.0139)		(0.0346)	
Bogota region	0.0408	*	0.1373	**
	(0.0196)		(0.0490)	
Territories region	-0.0730	***	-0.2654	***
	(0.0147)		(0.0398)	
Municipality-level variables:				
average wealth factor score	-0.0629	+	-0.1959	*
	(0.0355)		(0.0884)	
average years of education, women 25-64	0.0073		0.0194	
	(0.0103)		(0.0241)	
average years of education, men 25-64	-0.0024		-0.0007	
	(0.0091)		(0.0212)	
child-woman ratio (0-4 / f 15-49)	0.0375		0.1588	
	(0.0627)		(0.1522)	
% of HHs female-headed	0.0032	***	0.0108	***
	(0.0010)		(0.0022)	
% of population living abroad[§§]	-0.0029		-0.0083	
	(0.0038)		(0.0087)	
% of employed women in formal work	0.0008		0.0026	
	(0.0009)		(0.0022)	
% of employed men in formal work	0.0010		0.0031	
	(0.0009)		(0.0023)	
% HHs with access to piped water	0.0007	+	0.0027	**
	(0.0004)		(0.0010)	
% HHs with access to sewer	-0.0003		-0.0010	
	(0.0004)		(0.0009)	
% HHs cooking with firewood etc.	-0.0005		-0.0016	
	(0.0005)		(0.0013)	
constant	0.2759	*	-0.7672	*
	(0.1321)		(0.3742)	
number of observations	21,827		21,827	

Notes:

† This is the first stage for the IV-regress model for the "in-school regression". The first stage for the IV-Probit and IV Poisson and for other outcome variables is very similar.

§§individuals abroad / individuals present in municipality

*** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

Appendix B

Tables of Full Regressions for Various Educational Outcomes

Table B1. Attending School or Not, Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	Probit	IV Probit	IV Regress
Base Probability for reference person[§]	0.836	0.832	0.933
Endogenous regressor:			
mother's intimate partner physical violence	-0.0271*	-0.0008	-0.0074
	(0.0112)	(0.0859)	(0.0351)
Child variables:			
child is female	0.0397***	0.0403***	0.0208***
	(0.0082)	(0.0089)	(0.0040)
child's age	0.1653***	0.1680***	0.0840***
	(0.0117)	(0.0230)	(0.0060)
child's age-squared/100	-0.8738***	-0.8871***	-0.4510***
	(0.0587)	(0.1191)	(0.0308)
child is son/daughter of HH head	0.0351*	0.0362*	0.0193*
	(0.0160)	(0.0170)	(0.0079)
Mother variables:			
mother's age when child was age 6	0.0072	0.0075	0.0018
	(0.0065)	(0.0067)	(0.0034)
mother's age squared/100 (when child was 6)	-0.0116	-0.0118	-0.0036
	(0.0099)	(0.0101)	(0.0051)
mother's years of education completed	0.0134***	0.0137***	0.0052***
	(0.0017)	(0.0023)	(0.0006)
mother is married	0.0055	0.0066	0.0003
	(0.0106)	(0.0117)	(0.0050)
mother is widow	-0.0079	-0.0081	-0.0020
	(0.0248)	(0.0252)	(0.0126)
mother is divorced etc.	-0.0117	-0.0161	-0.0028
	(0.0148)	(0.0208)	(0.0084)
Partner-of-mother variables:			
partner's years of education completed	0.0043**	0.0044**	0.0013*
	(0.0015)	(0.0016)	(0.0006)
missing: partner's years of education	0.0178	0.0193	0.0032
	(0.0230)	(0.0241)	(0.0151)
missing: partner's childhood violence	-0.0099	-0.0108	-0.0078
	(0.0122)	(0.0131)	(0.0069)
Household variables:			
HH has migrated	-0.0387**	-0.0396**	-0.0144*

	(0.0129)	(0.0136)	(0.0059)
wealth quintile 2	0.0512***	0.0517***	0.0343***
	(0.0130)	(0.0137)	(0.0073)
wealth quintile 3	0.0809***	0.0821***	0.0467***
	(0.0160)	(0.0176)	(0.0078)
wealth quintile 4	0.0754***	0.0767***	0.0424***
	(0.0178)	(0.0193)	(0.0086)
wealth quintile 5 (richest)	0.0573**	0.0586**	0.0311**
	(0.0201)	(0.0214)	(0.0095)
Household composition variables:[§§]			
mother has relatives in HH	0.0001	0.0009	-0.0016
	(0.0175)	(0.0179)	(0.0076)
partner has relatives in HH	0.0065	0.0070	0.0029
	(0.0225)	(0.0228)	(0.0093)
# children ages 0-5	-0.0209***	-0.0213***	-0.0129***
	(0.0051)	(0.0057)	(0.0030)
# girls ages 6-11	-0.0104	-0.0109	-0.0074*
	(0.0062)	(0.0067)	(0.0034)
# boys ages 6-11	0.0033	0.0031	0.0020
	(0.0063)	(0.0065)	(0.0033)
# girls ages 12-14	0.0025	0.0024	0.0062
	(0.0090)	(0.0091)	(0.0042)
# boys ages 12-14	-0.0080	-0.0084	-0.0033
	(0.0089)	(0.0091)	(0.0045)
# girls ages 15-17	0.0069	0.0066	0.0056
	(0.0108)	(0.0110)	(0.0050)
# boys ages 15-17	-0.0034	-0.0036	0.0003
	(0.0096)	(0.0098)	(0.0050)
# women ages 18-64	0.0232**	0.0239*	0.0111***
	(0.0084)	(0.0093)	(0.0033)
# men ages 18-64	0.0022	0.0019	0.0030
	(0.0065)	(0.0066)	(0.0032)
# women ages 65+	0.0274	0.0279	0.0106
	(0.0183)	(0.0191)	(0.0069)
# men ages 65+	0.0101	0.0106	0.0089
	(0.0200)	(0.0206)	(0.0083)
Geographic variables:			
rural (vs. urban)	-0.0139	-0.0137	-0.0162*
	(0.0128)	(0.0130)	(0.0064)
Atlantic region	0.0884***	0.0898***	0.0488***
	(0.0141)	(0.0171)	(0.0070)
Oriental region	0.0987***	0.1004***	0.0666***
	(0.0144)	(0.0170)	(0.0073)

Pacific region	0.0881***	0.0894***	0.0606***
	(0.0137)	(0.0160)	(0.0075)
Bogota region	0.0535**	0.0536**	0.0352***
	(0.0188)	(0.0192)	(0.0091)
Territories region	0.1104***	0.1133***	0.0636***
	(0.0172)	(0.0224)	(0.0077)
Municipality-level variables:			
average wealth factor score	-0.1409***	-0.1419***	-0.0664**
	(0.0368)	(0.0411)	(0.0202)
average years of education, women 25-64	0.0482***	0.0485***	0.0246***
	(0.0093)	(0.0110)	(0.0054)
average years of education, men 25-64	-0.0145	-0.0145	-0.0025
	(0.0082)	(0.0086)	(0.0045)
child-woman ratio (0-4 / f 15-49)	-0.1151	-0.1183	-0.0998**
	(0.0595)	(0.0630)	(0.0326)
% of HHs female-headed	-0.0036***	-0.0037***	-0.0026***
	(0.0009)	(0.0011)	(0.0006)
% of population living abroad§§§	0.0074*	0.0075*	0.0033*
	(0.0032)	(0.0034)	(0.0015)
% of employed women in formal work	0.0012	0.0012	0.0006
	(0.0009)	(0.0009)	(0.0005)
% of employed men in formal work	-0.0012	-0.0012	-0.0010*
	(0.0008)	(0.0008)	(0.0005)
% HHs with access to piped water	0.0004	0.0004	0.0001
	(0.0003)	(0.0004)	(0.0002)
% HHs with access to sewer	-0.0006	-0.0006	-0.0002
	(0.0004)	(0.0004)	(0.0002)
% HHs cooking with firewood etc.	-0.0021***	-0.0021***	-0.0011***
	(0.0005)	(0.0006)	(0.0003)
constant			0.3995
			(0.0752)
number of observations	21,827	21,827	21,827

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

§ For Probit and IV Probit reference person has all his/her continuous characteristics set at their mean and their dummy variables set at zero. For IV regress, all characteristics are set at their mean.

§§ includes the index child

§§§ individuals abroad / individuals present in municipality

Table B2. Grade Advancement (vs. Repeat Grade), Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	Probit	IV Probit	IV Regress
Base Probability for reference person[§]	0.9490	0.9508	0.9559
Endogenous regressor:			
mother's intimate partner physical violence	-0.0184***	-0.0269	-0.0124
	(0.0056)	(0.0396)	(0.0278)
Child variables:			
child is female	0.0189***	0.0184***	0.0198***
	(0.0047)	(0.0050)	(0.0042)
child's age	-0.0028	-0.0028	-0.0035
	(0.0057)	(0.0056)	(0.0050)
child's age-squared/100	0.0200	0.0196	0.0234
	(0.0291)	(0.0285)	(0.0260)
child is son/daughter of HH head	-0.0073	-0.0073	-0.0067
	(0.0068)	(0.0066)	(0.0060)
Mother variables:			
mother's age when child was age 6	0.0035	0.0034	0.0029
	(0.0029)	(0.0029)	(0.0030)
mother's age squared/100 (when child was 6)	-0.0064	-0.0062	-0.0057
	(0.0045)	(0.0044)	(0.0046)
mother's years of education completed	0.0033***	0.0032***	0.0028***
	(0.0007)	(0.0008)	(0.0005)
mother is married	0.0112**	0.0107*	0.0097*
	(0.0043)	(0.0047)	(0.0040)
mother is widow	-0.0058	-0.0057	-0.0053
	(0.0106)	(0.0103)	(0.0109)
mother is divorced etc.	-0.0013	-0.0003	-0.0018
	(0.0057)	(0.0072)	(0.0071)
Partner-of-mother variables:			
partner's years of education completed	0.0010	0.0010	0.0008
	(0.0006)	(0.0006)	(0.0005)
missing: partner's years of education	-0.0103	-0.0108	-0.0132
	(0.0116)	(0.0116)	(0.0135)
missing: partner's childhood violence	-0.0058	-0.0053	-0.0070
	(0.0056)	(0.0057)	(0.0061)
Household variables:			
HH has migrated	-0.0065	-0.0062	-0.0062
	(0.0051)	(0.0051)	(0.0048)
wealth quintile 2	0.0016	0.0016	0.0054
	(0.0054)	(0.0052)	(0.0061)

wealth quintile 3	0.0078	0.0077	0.0110
	(0.0066)	(0.0064)	(0.0069)
wealth quintile 4	0.0089	0.0087	0.0103
	(0.0074)	(0.0073)	(0.0072)
wealth quintile 5 (richest)	0.0158	0.0153	0.0099
	(0.0084)	(0.0085)	(0.0078)
Household composition variables:[§§]			
mother has relatives in HH	-0.0051	-0.0053	-0.0050
	(0.0071)	(0.0070)	(0.0061)
partner has relatives in HH	-0.0149	-0.0146	-0.0122
	(0.0104)	(0.0103)	(0.0083)
# children ages 0-5	-0.0067**	-0.0065**	-0.0085***
	(0.0022)	(0.0023)	(0.0025)
# girls ages 6-11	-0.0038	-0.0037	-0.0041
	(0.0029)	(0.0029)	(0.0030)
# boys ages 6-11	-0.0018	-0.0017	-0.0027
	(0.0027)	(0.0026)	(0.0028)
# girls ages 12-14	-0.0051	-0.0049	-0.0053
	(0.0037)	(0.0037)	(0.0038)
# boys ages 12-14	-0.0121**	-0.0117**	-0.0127**
	(0.0038)	(0.0039)	(0.0040)
# girls ages 15-17	0.0019	0.0020	0.0026
	(0.0043)	(0.0043)	(0.0044)
# boys ages 15-17	-0.0022	-0.0021	-0.0021
	(0.0038)	(0.0038)	(0.0041)
# women ages 18-64	0.0009	0.0008	0.0015
	(0.0033)	(0.0033)	(0.0031)
# men ages 18-64	0.0051	0.0050	0.0047
	(0.0030)	(0.0029)	(0.0028)
# women ages 65+	0.0011	0.0009	0.0014
	(0.0069)	(0.0068)	(0.0058)
# men ages 65+	0.0129	0.0125	0.0122*
	(0.0080)	(0.0081)	(0.0062)
Geographic variables:			
rural (vs. urban)	-0.0075	-0.0075	-0.0079
	(0.0056)	(0.0055)	(0.0055)
Atlantic region	0.0103	0.0100	0.0104
	(0.0057)	(0.0057)	(0.0055)
Oriental region	0.0079	0.0078	0.0084
	(0.0059)	(0.0058)	(0.0056)
Pacific region	-0.0007	-0.0006	-0.0025
	(0.0062)	(0.0061)	(0.0061)

Bogota region	0.0118	0.0117	0.0123
	(0.0085)	(0.0083)	(0.0069)
Territories region	0.0003	-0.0001	0.0014
	(0.0068)	(0.0069)	(0.0065)
Municipality-level variables:			
average wealth factor score	0.0017	0.0016	0.0072
	(0.0154)	(0.0150)	(0.0161)
average years of education, women 25-64	-0.0015	-0.0015	-0.0001
	(0.0042)	(0.0042)	(0.0047)
average years of education, men 25-64	-0.0058	-0.0056	-0.0058
	(0.0038)	(0.0037)	(0.0039)
child-woman ratio (0-4 / f 15-49)	-0.0869**	-0.0843**	-0.0920**
	(0.0293)	(0.0306)	(0.0286)
% of HHs female-headed	0.0002	0.0002	-0.0001
	(0.0004)	(0.0004)	(0.0005)
% of population living abroad§§§	0.0002	0.0002	0.0000
	(0.0014)	(0.0013)	(0.0013)
% of employed women in formal work	0.0003	0.0003	0.0002
	(0.0004)	(0.0004)	(0.0004)
% of employed men in formal work	0.0003	0.0003	0.0001
	(0.0004)	(0.0004)	(0.0004)
% HHs with access to piped water	0.0003	0.0003	0.0004
	(0.0002)	(0.0002)	(0.0002)
% HHs with access to sewer	-0.0001	-0.0001	-0.0002
	(0.0002)	(0.0001)	(0.0002)
% HHs cooking with firewood etc.	0.0002	0.0002	0.0002
	(0.0002)	(0.0002)	(0.0002)
constant			0.9532***
			(0.0626)
number of observations	19,683	19,683	19,683

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + p<=0.1

§ For Probit and IV-Probit reference person has all his/her continuous characteristics set at their mean and their dummy variables set at zero. For IV regress, all characteristics are set at their mean.

§§ includes the index child

§§§ individuals abroad / individuals present in municipality

Table B3. Drop Out or Not, Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	Probit	IV Probit	IV Regress
Base Probability for reference person[§]	0.0455	0.0449	0.0227
Endogenous regressor:			
mother's intimate partner physical violence	0.0099	-0.0020	0.0045
	(0.0067)	(0.0435)	(0.0200)
Child variables:			
child is female	-0.0091	-0.0096	-0.0060*
	(0.0049)	(0.0054)	(0.0024)
child's age	-0.0268**	-0.0283**	-0.0177***
	(0.0091)	(0.0109)	(0.0035)
child's age-squared/100	0.1688***	0.1777**	0.1086***
	(0.0500)	(0.0613)	(0.0187)
child is son/daughter of HH head	-0.0017	-0.0021	-0.0025
	(0.0085)	(0.0091)	(0.0045)
Mother variables:			
mother's age when child was age 6	-0.0017	-0.0018	-0.0005
	(0.0039)	(0.0041)	(0.0019)
mother's age squared/100 (when child was 6)	0.0015	0.0015	0.0006
	(0.0059)	(0.0062)	(0.0028)
mother's years of education completed	-0.0042***	-0.0044**	-0.0016***
	(0.0012)	(0.0015)	(0.0004)
mother is married	-0.0024	-0.0031	-0.0001
	(0.0056)	(0.0063)	(0.0027)
mother is widow	0.0103	0.0107	0.0049
	(0.0145)	(0.0152)	(0.0076)
mother is divorced etc.	0.0149	0.0179	0.0048
	(0.0100)	(0.0155)	(0.0050)
Partner-of-mother variables:			
partner's years of education completed	-0.0010	-0.0011	-0.0002
	(0.0009)	(0.0009)	(0.0003)
missing: partner's years of education	-0.0046	-0.0058	0.0018
	(0.0120)	(0.0129)	(0.0089)
missing: partner's childhood violence	0.0056	0.0063	0.0031
	(0.0075)	(0.0085)	(0.0043)
Household variables:			
HH has migrated	0.0362**	0.0381**	0.0144***
	(0.0114)	(0.0144)	(0.0038)
wealth quintile 2	-0.0173*	-0.0181*	-0.0144***
	(0.0074)	(0.0083)	(0.0042)
wealth quintile 3	-0.0234**	-0.0246*	-0.0167***

	(0.0088)	(0.0103)	(0.0046)
wealth quintile 4	-0.0261**	-0.0276*	-0.0182***
	(0.0098)	(0.0117)	(0.0047)
wealth quintile 5 (richest)	-0.0225*	-0.0239*	-0.0161**
	(0.0102)	(0.0119)	(0.0051)
Household composition variables:[§§]			
mother has relatives in HH	0.0037	0.0035	0.0014
	(0.0104)	(0.0109)	(0.0045)
partner has relatives in HH	-0.0056	-0.0059	-0.0022
	(0.0117)	(0.0124)	(0.0047)
# children ages 0-5	0.0098**	0.0103*	0.0059**
	(0.0034)	(0.0041)	(0.0019)
# girls ages 6-11	-0.0003	-0.0002	0.0009
	(0.0034)	(0.0036)	(0.0019)
# boys ages 6-11	-0.0076	-0.0079	-0.0037*
	(0.0039)	(0.0042)	(0.0018)
# girls ages 12-14	-0.0036	-0.0038	-0.0027
	(0.0051)	(0.0053)	(0.0024)
# boys ages 12-14	0.0018	0.0020	0.0009
	(0.0055)	(0.0058)	(0.0028)
# girls ages 15-17	-0.0087	-0.0089	-0.0037
	(0.0061)	(0.0064)	(0.0028)
# boys ages 15-17	0.0034	0.0037	0.0013
	(0.0052)	(0.0056)	(0.0030)
# women ages 18-64	-0.0048	-0.0052	-0.0024
	(0.0047)	(0.0052)	(0.0020)
# men ages 18-64	0.0002	0.0003	-0.0009
	(0.0034)	(0.0036)	(0.0017)
# women ages 65+	-0.0254*	-0.0267	-0.0087*
	(0.0124)	(0.0144)	(0.0038)
# men ages 65+	0.0060	0.0061	0.0029
	(0.0103)	(0.0108)	(0.0049)
Geographic variables:			
rural (vs. urban)	0.0039	0.0038	0.0041
	(0.0065)	(0.0068)	(0.0033)
Atlantic region	-0.0343***	-0.0363**	-0.0278***
	(0.0097)	(0.0132)	(0.0046)
Oriental region	-0.0298***	-0.0314**	-0.0257***
	(0.0088)	(0.0113)	(0.0050)
Pacific region	-0.0296***	-0.0312**	-0.0243***
	(0.0087)	(0.0112)	*(0.0047)*
Bogota region	-0.0276**	-0.0288*	-0.0226***
	(0.0106)	(0.0121)	(0.0058)

Territories region	-0.0342**	-0.0364**	-0.0282***
	(0.0105)	(0.0141)	(0.0054)
Municipality-level variables:			
average wealth factor score	0.0598*	0.0626*	0.0354**
	(0.0250)	*(0.0287)*	(0.0118)
average years of education, women 25-64	-0.0120*	-0.0125	-0.0058
	(0.0060)	(0.0067)	(0.0031)
average years of education, men 25-64	0.0131*	0.0137*	0.0056
	(0.0060)	(0.0067)	(0.0030)
child-woman ratio (0-4 / f 15-49)	0.0134	0.0149	0.0157
	(0.0351)	(0.0378)	(0.0200)
% of HHs female-headed	-0.0001	-0.0001	-0.0001
	(0.0005)	(0.0006)	(0.0003)
% of population living abroad§§§	-0.0038*	-0.0040	-0.0025***
	(0.0019)	(0.0021)	(0.0007)
% of employed women in formal work	-0.0008	-0.0009	-0.0006*
	(0.0004)	(0.0005)	(0.0002)
% of employed men in formal work	0.0003	0.0004	0.0003
	(0.0005)	(0.0005)	(0.0003)
% HHs with access to piped water	-0.0002	-0.0002	-0.0001
	(0.0002)	(0.0002)	(0.0001)
% HHs with access to sewer	-0.0002	-0.0002	-0.0001
	(0.0002)	(0.0002)	(0.0001)
% HHs cooking with firewood etc.	0.0013**	0.0013**	0.0007***
	(0.0004)	(0.0005)	(0.0002)
constant			0.1557***
			(0.0428)
number of observations	20,433	20,433	20,433

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + p<=0.1

§ For Probit and IV Probit reference person has all his/her continuous characteristics set at their mean and their dummy variables set at zero. For IV regress, all characteristics are set at their mean.

§§ includes the index child

§§§ individuals abroad / individuals present in municipality

Table B4. Grade Advancement vs. Drop Out or Repeat, Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	Probit	IV Probit (2nd stage)	IV Regress
Base Probability	0.8868	0.8879	0.915
Endogenous regressor:			
mother's intimate partner physical violence	-0.0284***	-0.0338	-0.0227
	(0.0078)	(0.0547)	(0.0352)
Child variables:			
child is female	0.0347***	0.0345***	0.0280***
	(0.0067)	(0.0071)	(0.0050)
child's age	0.1119***	0.1111***	0.0873***
	(0.0136)	(0.0157)	(0.0073)
child's age-squared/100	-0.5250***	-0.5212***	-0.4095***
	(0.0659)	(0.0751)	(0.0364)
child is son/daughter of HH head	-0.0072	-0.0073	-0.0052
	(0.0103)	(0.0103)	(0.0078)
Mother variables:			
mother's age when child was age 6	0.0077	0.0076	0.0058
	(0.0045)	(0.0045)	(0.0038)
mother's age squared/100 (when child was 6)	-0.0126	-0.0125	-0.0099
	(0.0068)	(0.0069)	(0.0058)
mother's years of education completed	0.0074***	0.0073***	0.0051***
	(0.0011)	(0.0013)	(0.0007)
mother is married	0.0188**	0.0185**	0.0127*
	(0.0066)	(0.0072)	(0.0050)
mother is widow	-0.0090	-0.0089	-0.0073
	(0.0156)	(0.0155)	(0.0127)
mother is divorced etc.	-0.0094	-0.0085	-0.0059
	(0.0094)	(0.0125)	(0.0090)
Partner-of-mother variables:			
partner's years of education completed	0.0016	0.0016	0.0009
	(0.0009)	(0.0009)	(0.0006)
missing: partner's years of education	-0.0094	-0.0098	-0.0119
	(0.0171)	(0.0174)	(0.0162)
missing: partner's childhood violence	-0.0143	-0.0140	-0.0132
	(0.0088)	(0.0092)	(0.0075)
Household variables:			
HH has migrated	-0.0317***	-0.0314***	-0.0235***
	(0.0088)	(0.0092)	(0.0063)
wealth quintile 2	0.0182*	0.0181*	0.0213**

	(0.0085)	(0.0084)	(0.0075)
wealth quintile 3	0.0268**	0.0267**	0.0262**
	(0.0102)	(0.0102)	(0.0084)
wealth quintile 4	0.0320**	0.0317**	0.0275**
	(0.0114)	(0.0116)	(0.0089)
wealth quintile 5 (richest)	0.0309*	0.0306*	0.0218*
	(0.0130)	(0.0133)	(0.0097)
Household composition variables:[§§]			
mother has relatives in HH	-0.0122	-0.0123	-0.0089
	(0.0113)	(0.0112)	(0.0080)
partner has relatives in HH	-0.0099	-0.0098	-0.0066
	(0.0140)	(0.0140)	(0.0097)
# children ages 0-5	-0.0122***	-0.0122***	-0.0121***
	(0.0034)	(0.0035)	(0.0031)
# girls ages 6-11	-0.0060	-0.0059	-0.0054
	(0.0044)	(0.0044)	(0.0036)
# boys ages 6-11	0.0008	0.0009	-0.0003
	(0.0043)	(0.0042)	(0.0034)
# girls ages 12-14	-0.0043	-0.0042	-0.0025
	(0.0060)	(0.0060)	(0.0046)
# boys ages 12-14	-0.0175**	-0.0173**	-0.0139**
	(0.0059)	(0.0061)	(0.0049)
# girls ages 15-17	0.0073	0.0073	0.0064
	(0.0069)	(0.0069)	(0.0052)
# boys ages 15-17	-0.0044	-0.0044	-0.0031
	(0.0062)	(0.0062)	(0.0052)
# women ages 18-64	0.0058	0.0057	0.0050
	(0.0053)	(0.0053)	(0.0038)
# men ages 18-64	0.0076	0.0076	0.0065*
	(0.0044)	(0.0043)	(0.0033)
# women ages 65+	0.0084	0.0082	0.0060
	(0.0111)	(0.0111)	(0.0075)
# men ages 65+	0.0094	0.0092	0.0076
	(0.0118)	(0.0118)	(0.0083)
Geographic variables:			
rural (vs. urban)	-0.0211*	-0.0211*	-0.0183**
	(0.0087)	(0.0086)	(0.0066)
Atlantic region	0.0388***	0.0385***	0.0307***
	(0.0091)	(0.0098)	(0.0074)
Oriental region	0.0465***	0.0462***	0.0416***
	(0.0095)	(0.0100)	(0.0076)
Pacific region	0.0331***	0.0329***	0.0261***

	(0.0091)	(0.0093)	(0.0077)
Bogota region	0.0007	0.0009	0.0045
	(0.0137)	(0.0137)	(0.0105)
Territories region	0.0405***	0.0399***	0.0348***
	(0.0106)	(0.0118)	(0.0087)
Municipality-level variables:			
average wealth factor score	-0.0562*	-0.0559*	-0.0396*
	(0.0255)	(0.0256)	(0.0201)
average years of education, women 25-64	0.0081	0.0081	0.0077
	(0.0069)	(0.0068)	(0.0057)
average years of education, men 25-64	-0.0167**	-0.0166**	-0.0126*
	(0.0062)	(0.0063)	(0.0049)
child-woman ratio (0-4 / f 15-49)	-0.1452**	-0.1438**	-0.1280***
	(0.0442)	(0.0464)	(0.0356)
% of HHs female-headed	0.0003	0.0003	0.0002
	(0.0006)	(0.0007)	(0.0006)
% of population living abroad§§§	0.0051*	0.0050*	0.0037*
	(0.0024)	(0.0024)	(0.0015)
% of employed women in formal work	0.0010	0.0010	0.0007
	(0.0006)	(0.0006)	(0.0005)
% of employed men in formal work	0.0003	0.0003	0.0001
	(0.0006)	(0.0006)	(0.0005)
% HHs with access to piped water	0.0005*	0.0005*	0.0005*
	(0.0003)	(0.0003)	(0.0002)
% HHs with access to sewer	0.0002	0.0002	0.0001
	(0.0002)	(0.0002)	(0.0002)
% HHs cooking with firewood etc.	-0.0006	-0.0006	-0.0004
	(0.0003)	(0.0003)	(0.0003)
constant			0.3363***
			(0.0830)
number of observations	20,429	20,429	20,429

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

§ For Probit and IV Probit reference person has all his/her continuous characteristics set at their mean and their dummy variables set at zero. For IV regress, all characteristics are set at their mean.

§§ includes the index child

§§§ individuals abroad / individuals present in municipality

Table B5. Grade Attended Conditional on Remaining in School, Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	Poisson	IV Poisson	IV Regress
Predicted # of Years Attained by Reference Person	3.236	3.220	4.222
Endogenous regressor:			
mother's intimate partner physical violence	-0.0590***	0.1525	0.1794
	(0.0176)	(0.1162)	(0.1522)
Child variables:			
child is female	0.1865***	0.1830***	0.2486***
	(0.0159)	(0.0158)	(0.0198)
child's age	2.2495***	2.2182***	0.8234***
	(0.0245)	(0.0242)	(0.0261)
child's age-squared/100	-7.5161***	-7.4110***	0.0766
	(0.1158)	(0.1141)	(0.1374)
child is son/daughter of HH head	0.0011	0.0061	0.0265
	(0.0254)	(0.0252)	(0.0332)
Mother variables:			
mother's age when child was age 6	0.0534***	0.0534***	0.0596***
	(0.0128)	(0.0126)	(0.0147)
mother's age squared/100 (when child was 6)	-0.0823***	-0.0801***	-0.0839***
	(0.0197)	(0.0194)	(0.0220)
mother's years of education completed	0.0495***	0.0492***	0.0675***
	(0.0022)	(0.0022)	(0.0030)
mother is married	0.1314***	0.1398***	0.1682***
	(0.0165)	(0.0169)	(0.0227)
mother is widow	-0.0200	-0.0190	-0.0295
	(0.0419)	(0.0413)	(0.0557)
mother is divorced etc.	-0.0617**	-0.0917**	-0.1231**
	(0.0234)	(0.0288)	(0.0397)
Partner-of-mother variables:			
partner's years of education completed	0.0159***	0.0164***	0.0221***
	(0.0020)	(0.0021)	(0.0028)
missing: partner's years of education	-0.0431	-0.0260	-0.0232
	(0.0566)	(0.0573)	(0.0691)
missing: partner's childhood violence	-0.0551*	-0.0632**	-0.0740*
	(0.0241)	(0.0240)	(0.0311)
Household variables:			
HH has migrated	-0.1217***	-0.1234***	-0.1551***
	(0.0196)	(0.0193)	(0.0246)
wealth quintile 2	0.3140***	0.3068***	0.3360***

	(0.0299)	(0.0297)	(0.0344)
wealth quintile 3	0.4388***	0.4298***	0.4935***
	(0.0328)	(0.0327)	(0.0384)
wealth quintile 4	0.4478***	0.4416***	0.5327***
	(0.0356)	(0.0353)	(0.0420)
wealth quintile 5 (richest)	0.3468***	0.3464***	0.4488***
	(0.0388)	(0.0382)	(0.0468)
Household composition variables:[§§]			
mother has relatives in HH	0.0487	0.0545*	0.0576
	(0.0272)	(0.0271)	(0.0355)
partner has relatives in HH	0.0702*	0.0699*	0.0979*
	(0.0353)	(0.0348)	(0.0444)
# children ages 0-5	-0.1087***	-0.1073***	-0.1087***
	(0.0110)	(0.0108)	(0.0126)
# girls ages 6-11	-0.0841***	-0.0852***	-0.1375***
	(0.0122)	(0.0121)	(0.0157)
# boys ages 6-11	-0.0865***	-0.0866***	-0.1051***
	(0.0121)	(0.0120)	(0.0157)
# girls ages 12-14	-0.0882***	-0.0878***	-0.0485*
	(0.0158)	(0.0156)	(0.0208)
# boys ages 12-14	-0.0951***	-0.0962***	-0.1118***
	(0.0159)	(0.0158)	(0.0210)
# girls ages 15-17	-0.0424*	-0.0451*	-0.0637**
	(0.0177)	(0.0175)	(0.0243)
# boys ages 15-17	-0.0735***	-0.0745***	-0.1005***
	(0.0175)	(0.0173)	(0.0236)
# women ages 18-64	0.0288*	0.0304*	0.0332*
	(0.0119)	(0.0118)	(0.0162)
# men ages 18-64	-0.0223*	-0.0232*	-0.0364*
	(0.0111)	(0.0110)	(0.0148)
# women ages 65+	-0.0201	-0.0165	-0.0045
	(0.0258)	(0.0255)	(0.0353)
# men ages 65+	0.0042	0.0089	-0.0016
	(0.0302)	(0.0299)	(0.0391)
Geographic variables:			
rural (vs. urban)	-0.0679**	-0.0621*	-0.0500
	(0.0244)	(0.0241)	(0.0307)
Atlantic region	-0.0969***	-0.0911***	-0.1209***
	(0.0228)	(0.0227)	(0.0305)
Oriental region	0.1291***	0.1221***	0.1454***
	(0.0242)	(0.0242)	(0.0317)
Pacific region	-0.0950***	-0.0956***	-0.1234***

	(0.0240)	(0.0236)	(0.0320)
Bogota region	-0.3350***	-0.3387***	-0.5075***
	(0.0279)	(0.0275)	(0.0390)
Territories region	0.0354	0.0492	0.0572
	(0.0250)	(0.0256)	(0.0349)
Municipality-level variables:			
average wealth factor score	0.0635	0.0696	0.1115
	(0.0685)	(0.0678)	(0.0867)
average years of education, women 25-64	0.0453*	0.0436*	0.0512*
	(0.0190)	(0.0187)	(0.0232)
average years of education, men 25-64	0.0202	0.0204	0.0220
	(0.0162)	(0.0160)	(0.0200)
child-woman ratio (0-4 / f 15-49)	-0.5479***	-0.5550***	-0.6214***
	(0.1237)	(0.1219)	(0.1498)
% of HHs female-headed	-0.0128***	-0.0133***	-0.0148***
	(0.0021)	(0.0021)	(0.0025)
% of population living abroad§§§	0.0127	0.0130	0.0148
	(0.0069)	(0.0068)	(0.0092)
% of employed women in formal work	-0.0028	-0.0029	-0.0044*
	(0.0016)	(0.0016)	(0.0020)
% of employed men in formal work	-0.0045*	-0.0047*	-0.0045
	(0.0019)	(0.0019)	(0.0024)
% HHs with access to piped water	0.0049***	0.0047***	0.0048***
	(0.0008)	(0.0008)	(0.0010)
% HHs with access to sewer	-0.0013*	-0.0012	-0.0017*
	(0.0007)	(0.0007)	(0.0009)
% HHs cooking with firewood etc.	0.0050***	0.0049***	0.0066***
	(0.0010)	(0.0010)	(0.0013)
constant			-5.8540***
			(0.3276)
number of observations	20,370	20,370	20,370

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + p<=0.1

§ For Poisson and IV Poisson reference person has all his/her continuous characteristics set at their mean and their dummy variables set at zero. For IV regress, all characteristics are set at their mean.

§§ includes the index child

§§§ individuals abroad / individuals present in municipality

Table B6. Completed Years of Education Conditional on Entering school, Children Ages 10-14 who ever attended school, Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	Poisson	IV Poisson	IV Regress
Predicted # of Years Attained by Reference Person[§]	4.213	4.099	4.962
Endogenous regressor:			
mother's intimate partner physical violence	-0.0918**	0.3296	0.3489
	(0.0287)	(0.2029)	(0.2479)
Child variables:			
child is female	0.2871***	0.2809***	0.3142***
	(0.0288)	(0.0281)	(0.0292)
child's age	2.4318***	2.3973***	1.2616***
	(0.1385)	(0.1356)	(0.1585)
child's age-squared	-7.2304***	-7.1565***	-1.9424**
	(0.5692)	(0.5567)	(0.6655)
child is son/daughter of HH head	-0.0094	0.0062	0.0034
	(0.0423)	(0.0419)	(0.0500)
Mother variables:			
mother's age when child was age 6	0.0667**	0.0672**	0.0745**
	(0.0238)	(0.0232)	(0.0260)
mother's age squared/100 (when child was 6)	-0.1037**	-0.1001**	-0.1065**
	(0.0367)	(0.0357)	(0.0398)
mother's years of education completed	0.0762***	0.0749***	0.0939***
	(0.0036)	(0.0035)	(0.0044)
mother is married	0.2009***	0.2155***	0.2468***
	(0.0269)	(0.0274)	(0.0332)
mother is widow	-0.1058	-0.1002	-0.1367+
	(0.0675)	(0.0662)	(0.0797)
mother is divorced etc.	-0.1299***	-0.1791***	-0.2290***
	(0.0378)	(0.0431)	(0.0546)
Partner-of-mother variables:			
partner's years of education completed	-0.0086	-0.0321	-0.0385
	(0.0389)	(0.0393)	(0.0456)
missing: partner's years of education	0.0269***	0.0276***	0.0342***
	(0.0033)	(0.0033)	(0.0041)
missing: partner's childhood violence	0.0147	0.0324	0.0555
	(0.0924)	(0.0908)	(0.0973)
Household variables:			
HH has migrated	-0.2179***	-0.2169***	-0.2530***
	(0.0326)	(0.0317)	(0.0377)

wealth quintile 2	0.5063***	0.4909***	0.4647***
	(0.0499)	(0.0492)	(0.0504)
wealth quintile 3	0.7168***	0.6948***	0.6898***
	(0.0553)	(0.0551)	(0.0568)
wealth quintile 4	0.7379***	0.7194***	0.7534***
	(0.0602)	(0.0595)	(0.0627)
wealth quintile 5 (richest)	0.6073***	0.6023***	0.6607***
	(0.0650)	(0.0635)	(0.0687)
Household composition variables:[§§]			
mother has relatives in HH	0.0873*	0.0967*	0.1125*
	(0.0440)	(0.0431)	(0.0509)
partner has relatives in HH	0.1397*	0.1365*	0.1588*
	(0.0562)	(0.0547)	(0.0630)
# children ages 0-5	-0.1659***	-0.1603***	-0.1579***
	(0.0182)	(0.0177)	(0.0186)
# girls ages 6-11	-0.1241***	-0.1245***	-0.1582***
	(0.0196)	(0.0192)	(0.0218)
# boys ages 6-11	-0.1133***	-0.1118***	-0.1186***
	(0.0196)	(0.0191)	(0.0219)
# girls ages 12-14	-0.1160***	-0.1125***	-0.0880**
	(0.0258)	(0.0252)	(0.0283)
# boys ages 12-14	-0.1145***	-0.1130***	-0.1317***
	(0.0259)	(0.0252)	(0.0277)
# girls ages 15-17	-0.0637*	-0.0665*	-0.0804*
	(0.0288)	(0.0282)	(0.0329)
# boys ages 15-17	-0.1106***	-0.1112***	-0.1291***
	(0.0271)	(0.0265)	(0.0303)
# women ages 18-64	0.0447*	0.0443*	0.0457+
	(0.0198)	(0.0193)	(0.0234)
# men ages 18-64	-0.0495**	-0.0490**	-0.0621**
	(0.0179)	(0.0175)	(0.0205)
# women ages 65+	-0.0206	-0.0099	-0.0048
	(0.0402)	(0.0395)	(0.0485)
# men ages 65+	0.0177	0.0328	0.0269
	(0.0500)	(0.0492)	(0.0595)
Geographic variables:			
rural (vs. urban)	-0.0991*	-0.0797*	-0.0765+
	(0.0404)	(0.0400)	(0.0454)
Atlantic region	-0.1867***	-0.1700***	-0.1978***
	(0.0367)	(0.0364)	(0.0438)
Oriental region	0.1180**	0.1111**	0.1155*
	(0.0392)	(0.0382)	(0.0463)

Pacific region	-0.1738***	-0.1680***	-0.2099***
	(0.0387)	(0.0377)	(0.0460)
Bogota region	-0.2895***	-0.2866***	-0.3891***
	(0.0436)	(0.0424)	(0.0560)
Territories region	0.0011	0.0298	0.0252
	(0.0400)	(0.0409)	(0.0508)
Municipality-level variables:			
average wealth factor score	0.1210	0.1259	0.1364
	(0.1124)	(0.1095)	(0.1249)
average years of education, women 25-64	0.0546+	0.0501	0.0480
	(0.0324)	(0.0315)	(0.0346)
average years of education, men 25-64	0.0249	0.0261	0.0299
	(0.0275)	(0.0267)	(0.0301)
child-woman ratio (0-4 / f 15-49)	-0.7531***	-0.7828***	-0.7295**
	(0.2089)	(0.2046)	(0.2270)
% of HHs female-headed	-0.0205***	-0.0215***	-0.0211***
	(0.0035)	(0.0035)	(0.0036)
% of population living abroad§§§	0.0208+	0.0207+	0.0288*
	(0.0122)	(0.0118)	(0.0131)
% of employed women in formal work	-0.0028	-0.0029	-0.0044
	(0.0027)	(0.0026)	(0.0030)
% of employed men in formal work	-0.0089**	-0.0093**	-0.0090**
	(0.0032)	(0.0032)	(0.0035)
% HHs with access to piped water	0.0068***	0.0062***	0.0055***
	(0.0014)	(0.0014)	(0.0014)
% HHs with access to sewer	-0.0021+	-0.0018+	-0.0017
	(0.0011)	(0.0011)	(0.0012)
% HHs cooking with firewood etc.	0.0050**	0.0047**	0.0051**
	(0.0016)	(0.0016)	(0.0018)
constant			-9.4769***
			(1.0745)
number of observations	11,689	11,689	11,689

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

§ For Poisson and IV Poisson reference person has all his/her continuous characteristics set at their mean and their dummy variables set at zero. For IV regress, all characteristics are set at their mean.

§§ includes the index child

§§§ individuals abroad / individuals present in municipality

Table B7. Grades per year, Conditional on being school-age previous year, Colombia, DHS 2005, Parametric Models, Marginal Effects and Standard Errors

Variable	OLS	IV Regress
Predicted # of Grades per Year Attained by Reference Person[§]	0.826	0.826
Endogenous regressor:		
mother's intimate partner physical violence	-0.0166**	0.1015*
	(0.0059)	(0.0458)
Child variables:		
child is female	0.0577***	0.0582***
	(0.0057)	(0.0058)
child's age	0.0726*	0.0821**
	(0.0291)	(0.0298)
child's age-squared	-0.3125**	-0.3493**
	(0.1192)	(0.1221)
child is son/daughter of HH head	-0.0025	0.0020
	(0.0089)	(0.0093)
Mother variables:		
mother's age when child was age 6	0.0103*	0.0109*
	(0.0047)	(0.0048)
mother's age squared/100 (when child was 6)	-0.0157*	-0.0153*
	(0.0073)	(0.0075)
mother's years of education completed	0.0168***	0.0170***
	(0.0008)	(0.0008)
mother is married	0.0416***	0.0470***
	(0.0055)	(0.0061)
mother is widow	-0.0183	-0.0177
	(0.0132)	(0.0137)
mother is divorced etc.	-0.0329***	-0.0486***
	(0.0080)	(0.0101)
Partner-of-mother variables:		
partner's years of education completed	0.0055***	0.0059***
	(0.0007)	(0.0008)
missing: partner's years of education	-0.0016	0.0036
	(0.0183)	(0.0187)
missing: partner's childhood violence	0.0015	-0.0054
	(0.0076)	(0.0082)
Household variables:		
HH has migrated	-0.0398***	-0.0411***
	(0.0067)	(0.0069)
wealth quintile 2	0.0817***	0.0812***

	(0.0091)	(0.0093)
wealth quintile 3	0.1237***	0.1228***
	(0.0101)	(0.0103)
wealth quintile 4	0.1307***	0.1309***
	(0.0112)	(0.0114)
wealth quintile 5 (richest)	0.1153***	0.1179***
	(0.0122)	(0.0125)
Household composition[§§] variables:		
mother has relatives in HH	0.0221*	0.0252**
	(0.0092)	(0.0095)
partner has relatives in HH	0.0220+	0.0223+
	(0.0120)	(0.0121)
# children ages 0-5	-0.0331***	-0.0329***
	(0.0034)	(0.0034)
# girls ages 6-11	-0.0255***	-0.0265***
	(0.0039)	(0.0040)
# boys ages 6-11	-0.0200***	-0.0204***
	(0.0038)	(0.0040)
# girls ages 12-14	-0.0238***	-0.0236***
	(0.0053)	(0.0054)
# boys ages 12-14	-0.0243***	-0.0247***
	(0.0052)	(0.0053)
# girls ages 15-17	-0.0105+	-0.0117*
	(0.0057)	(0.0058)
# boys ages 15-17	-0.0169**	-0.0180**
	(0.0054)	(0.0055)
# women ages 18-64	0.0103*	0.0106*
	(0.0041)	(0.0042)
# men ages 18-64	-0.0108**	-0.0109**
	(0.0036)	(0.0037)
# women ages 65+	-0.0068	-0.0037
	(0.0088)	(0.0091)
# men ages 65+	-0.0039	0.0008
	(0.0110)	(0.0113)
Geographic variables:		
rural (vs. urban)	-0.0169*	-0.0124
	(0.0081)	(0.0084)
Atlantic region	-0.0327***	-0.0290***
	(0.0078)	(0.0081)
Oriental region	0.0273**	0.0266**
	(0.0083)	(0.0085)
Pacific region	-0.0363***	-0.0362***

	(0.0084)	(0.0086)
Bogota region	-0.0685***	-0.0700***
	(0.0105)	(0.0107)
Territories region	0.0059	0.0141
	(0.0089)	(0.0096)
Municipality-level variables:		
average wealth factor score	0.0113	0.0148
	(0.0222)	(0.0228)
average years of education, women 25-64	0.0083	0.0076
	(0.0061)	(0.0062)
average years of education, men 25-64	0.0073	0.0076
	(0.0055)	(0.0057)
child-woman ratio (0-4 / f 15-49)	-0.1416***	-0.1549***
	(0.0396)	(0.0409)
% of HHs female-headed	-0.0045***	-0.0049***
	(0.0006)	(0.0006)
% of population living abroad§§§	0.0049*	0.0051*
	(0.0023)	(0.0024)
% of employed women in formal work	-0.0005	-0.0006
	(0.0005)	(0.0006)
% of employed men in formal work	-0.0017**	-0.0019**
	(0.0006)	(0.0006)
% HHs with access to piped water	0.0014***	0.0013***
	(0.0002)	(0.0003)
% HHs with access to sewer	-0.0004*	-0.0004
	(0.0002)	(0.0002)
% HHs cooking with firewood etc.	0.0007*	0.0006+
	(0.0003)	(0.0003)
constant	0.0884	0.0045
	(0.1976)	(0.2030)
number of observations	11,890	11,890

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, + $p<=0.1$

§ Reference person for OLS and IV-Regress has all his/her characteristics set at their mean value

§§ includes the index child

§§§ individuals abroad / individuals present in municipality