This paper contends that the reverse gender gap in education in Latin America and the Caribbean—girls remaining in school longer than boys—is partly explained by family structure. I first show that living apart from a biological father has a greater negative impact on boys’ education than girls’ education. I then perform a simple standardization exercise that estimates what the gender gap in education in Latin America and the Caribbean would be if children in the region lived with both biological parents as often as children in Asia do. I thus argue that the high rates of single parenthood and union instability in the region help explain gender inequality in education.

Introduction
Latin America and the Caribbean led other developing regions in schooling in the 1960s, but expansion of education has been relatively poor. Now students in Eastern Asian and Pacific countries are more likely to complete secondary school, and the Latin American advantage over poorer developing countries has also dwindled (Barro and Lee 2013). At least 25% of children in the region drop out of school before or during the secondary years (UNICEF 2012). International tests of student achievement provide further evidence of an emergent disadvantage for Latin American students who score near the bottom (Mizala et al. 2002; Park and Sandefur 2006, and references therein).

This poor expansion of education has profound consequences: Slow economic growth in the region can in fact be explained by lower cognitive skills (Hanushek and Woessmann 2012). Given that education contributes to productivity and health at both the individual and national levels, the repercussions of relatively poor educational performance are not surprising.

School quality is undoubtedly an important factor in why many Latin American students reach their late teens functionally illiterate (Hanushek and Woessmann 2012), and has received considerable attention (Fuller and Clark 1994; Woessmann 2010, and references cited therein). Nonetheless, family factors help determine educational outcomes across national contexts with widely varying school quality (Chudgar and Luschei 2009). With Latin America’s history of an early educational advantage over other developing countries and continued governmental commitment to educational goals (e.g., Escobar et al. 2012) including increased per capita expenditures (Ramírez and Téllez 2006), it is appropriate to assess whether children’s living arrangements are one factor that could explain why education has not expanded more rapidly in the region.

It is worth finding out which living arrangements are associated with educational progress for Latin American and Caribbean youth. Literature from Northern countries has documented an educational advantage for children living with their married biological parents (e.g., Amato and Keith 1991; Bernardi and Radl 2014; Hampden-Thompson 2009; McLanahan and Sandefur 1994; Popenoe et al. 1996; Shriner et al. 2010), and non-marital childbearing is common in the region. Even Mexico with low rates of non-marital childbearing by regional standards is on par with Sweden with 55% of births occurring outside of marriage (World Family Map 2014). Further, the both non-marital childbearing and cohabitation are continuing to increase (Esteve et al. 2012). Neither non-marital childbearing nor the “cohabitation boom” automatically increase
the proportion of children reared apart from both biological parents, but they clearly have the potential to do so. Cohabiting unions are more fragile than marriages, even in societies where cohabitation has become common (e.g., de Vos 2001; Ishida 2010; Kiernan 2001), and even in Latin America and the Caribbean where it is accepted as an alternative nuptial system (e.g., Castro Martin 2002; de Vos 2001; Ishida 2010). Recent data show that children are less likely to be living with both biological parents in Latin America and the Caribbean than in any other region besides Africa (World Family Map 2014).

Father absence may be particular detrimental to boys’ continuation of schooling. In earlier work I showed that children living apart from either biological parent were less likely to progress through secondary school, even if their parent was repartnered. Here I test how much stronger this effect is on boys’ schooling than girls’ schooling.

**Data and Methods**

**Data.** Most of my data came from Demographic and Health Surveys (DHS) which are comparable nationally representative surveys fielded in multiple countries.¹ A key advantage that these data have over datasets commonly employed when studying the effects of family background on education was that my data included all youth—not just those attending school. Particularly when seeking to extend knowledge of how living arrangements influence education in poorer contexts, excluding drop-outs is problematic: The measured effect of student background factors may be attenuated when attrition produces greater homogeneity in the remaining student population (Chudgar and Luschei 2009). In other words, living arrangements may be important for continued enrollment, and therefore their effect may be greater in the general population than in the student population.

The data available from the DHS included many of the more populous countries of the region, but Mexico was an important exception. Mexico participated in the DHS in Round I before education information was included in the standard household roster, but not subsequently. I therefore supplemented the DHS data with the 2005 Mexican Family Life Survey (MxFLS-2) which is also a nationally representative survey that identifies whether children have biological parents in the household.² Although these data sources differed in many ways, they were relatively easy to harmonize for my purposes. The treatment of the Mexican data was only different from the other countries in a few places, and those are detailed below. Both sources offered a critical advantage over most census data in that they included whether or not children’s biological parents were in the household.


---

¹ [http://dhsprogram.com/data/available-datasets.cfm](http://dhsprogram.com/data/available-datasets.cfm)
³ The 1996 Brazil DHS is the oldest data in my sample and is of particular concern because the economy has changed so rapidly in recent years. The 2010 Brazilian Census (one of the only censuses that identifies stepchildren of the household head) allows for comparing effects of living arrangements on education among children/stepchildren of the household head (not all children). The results are similar enough to a subsample of children/stepchildren of the household head drawn from the 1996 Brazil DHS to ease concern about using older data.
was great. For example, Guatemala with its highly agrarian economy had a gross secondary enrollment of about 32% at the time of the DHS survey, whereas in more economically diverse Colombia with a more recent DHS, the gross secondary enrollment rate exceeded 95% (World Bank 2014).

Because my data were not from educational surveys and therefore did not include characteristics of schools, I could not determine whether family background mattered more or less than school quality on children’s educational progress. Nonetheless, I was able to assess the effect of living arrangements on educational progress across a wide range of Latin American and Caribbean countries using nationally representative samples, and therefore contributed importantly to what is known about children’s living arrangements and education in the region.

**Dependent variable.** I measured whether or not children aged 11-14 were still progressing on-time through school. Although the countries I studied have various school start ages, this age range covered the important transition from primary school to secondary school everywhere. Children aged 11-14 years who did not attend school in the year before the survey were not on track to complete secondary school, regardless of when they dropped out. Moreover, children behind the expected grade for their age either because of grade repetition or late enrollment are at an elevated risk of dropping out (United Nations 2014). I coded children progressing on-time through school as “0” and those who had either dropped out or fallen behind as “1”.

I recognize that there are many children at an educational disadvantage (e.g., having low test scores or lacking functional literacy) who may have not been identified as disadvantaged by my rough measure, but children out of school or falling behind in school are among the most disadvantaged. I therefore have confidence that my dependent variable measured those least likely to complete secondary school, but also note that relevant differences among relatively advantaged students have not been captured by my analysis.

**Independent variables.**

*Parents and stepparents in the household.* Children were classified as living with both biological parents, just their biological mother, just their biological father, or neither. I identified children as living with a stepparent if they were living with one biological parent and that parent had a partner in the household. This was approximated using relationships of household members to the household head. In most cases the biological parent was either head or spouse of head. I also

---

4 I used country-specific school start ages which ranged from five to seven when calculating whether a child of a given age and completed years of schooling was behind grade for age. Children were only coded as being behind grade-for-age if they were at least two years older than expected grade-for-age. Thus the relative timing of birthdays and surveys did not matter as children only one year older than expected were not counted as behind. For example, an eleven year-old in Honduras where school starts at age six would not be behind if he/she were in his/her fourth year, but would be behind if he/she were only in his/her third year.

5 The DHS included questions in the individual woman’s interviews that identify whether the woman is in union as well as whether her partner is in the household. I did not, however, want to utilize that information since it would have required limiting the sample to children living with interviewed mothers, consequently eliminating children living with only their biological father as well as children living with neither parent. My constructed stepparent variable was of higher quality than the “probable stepparent” variable constructed by IPUMS because I knew whether both biological parents were in the household: the IPUMS methodology only captures stepparents whose age makes them improbable biological parents, while my methodology included biological parents’ partners when the other biological parent was absent.
assumed a spouse was present if the biological parent was a child of the household head and there were as many or more opposite-sex children-in-law of the head in the household (and similarly using numbers of brothers/sisters-in-law when the child’s parent was a sibling of the head). I left ambiguous cases as missing, e.g., when the child’s mother was one of two daughters of the head and there was one son-in-law in the household, or when the child’s parent was unrelated to the head. There were surprisingly few ambiguous cases: 710/83,741 or less than 1% of the entire sample; 710/25,253 or less than 3% of those with one biological parent. There were no missing cases for Mexico as the MxFLS-2 household roster included information on the location of partners of household members in addition to the location of biological parents of children in the household.

If the child lived with one biological parent who did not have a partner in the household, the parent could either be single or in union with a non-coresident partner. Unfortunately, the household roster did not include union status in Bolivia, Brazil, Guatemala, and Honduras. For the other seven countries, parents with an absent partner (i.e., a non-coresident partner) were identified as different from single parents.

Other adults in the household. I counted the number of adult women other than the biological mother or stepmother in the household, but only if they were under 65 and were not currently attending school themselves (because either advanced age or studying would limit their potential contributions). Having no other women in the household was the reference category relative to one, two, three, and four or more women in the household. In the same way, the number of men aged 18-65 and not currently attending school other than the biological father or stepfather was also represented by a vector of dummy variables.

Other children. The presence of other children in the household could compromise the schooling of children aged 11-14 if their labor were needed for income or child care, or more simply because of competition between children for resources like school uniforms, transportation costs and books. Having more siblings has been associated with lower academic achievement (Downey 2001; Parcel and Menaghan 1994), but the sibling effect is smaller in poorer contexts than in richer ones (Chernichovsky 1985; Gomes 1984; Sibanda 2004). I thus included the number of children under age 15 continuously, with all values greater than six considered as equal to six.

Child’s gender and age. Gender was coded 0=female, 1=male. I also included an interaction term between gender and the child’s living arrangements. Age was represented by set of dummy variables because of the sharp drop off in enrollments between primary and secondary school in many countries (see table 1, columns four and five), and also differences between countries in the age at which secondary school should begin (eleven or twelve).

Parental education. I defined parent’s education as the higher of either the mother’s or the father’s education using six categories: no education, incomplete primary, complete primary, incomplete secondary, complete secondary, and higher. For children living with neither parent (and also in the few cases where parent’s education was missing), I used education of the household head. Education of the household head has been shown to be a strong determinant of children’s schooling (Case and Deaton 1999).
Wealth. I constructed an 8-point scale measuring absolute wealth based on housing quality and ownership of consumer durables following Giroux (2008). Paraguay was excluded from this analysis because it did not have information to construct the index. The consumer durable information in the MxFLS-2 was not identical to the DHS and thus the wealth index is not precisely the same, but the substitutions were not absurd. For example, it was assumed that all Mexican households had a radio. Presence of a telephone was substituted for presence of a television, presence of cooking gas was substituted for presence of a refrigerator, and owning an unmortgaged home with at least three bedrooms was substituted for owning a car.

Residence. Residence was represented by a dummy variable (0=rural, 1=urban). Other studies of Latin American education have shown residence to be a persistent factor impacting educational opportunity (Benavides and Mena 2010). All Mexican households were coded as urban in the survey. Although this is clearly in error, much of the rural/urban variation in Mexico was picked up by including Mexico’s federal entities as a set of dummy variables (see Region immediately below).

Region. I included a set of dummy variables to control for fixed effects of regions. Because the geographic distribution of educational opportunities is so uneven in developing countries, and capital cities in Latin America and the Caribbean have a distinct advantage over other areas, particularly in the poorer countries (Escobal et al. 2012; Molinas Vega et al. 2012), regional controls were necessary to account for differences in access to school. The models then estimated the effects of living arrangements and other variables within regions. See Huisman and Smits (2009) who also describe the importance of district controls.

Methods. I first ran logistic regression models separately for each country. The models control for the clustering of observations within communities (sampling clusters for the DHS and municipalities for the MxFLS-2). Having both biological parents in the household was used as the reference category for the other living arrangements variables. For all eleven countries, I divided mother only and father only households into those with stepfathers/stepmothers present and others. For the seven countries where parent’s union status is known, I separated absent partner households from single parent households.

I then pooled the data across countries to provide a clearer summary of how living arrangements were related to educational outcomes in the region. I included all of the observations for Brazil, the largest country in the region, and took a random sample from the other available observations so that each country’s proportion in the pooled sample was the same as in the 2014 population.

---

6 Some countries also have a DHS-provided wealth index that divides households into wealth quintiles within the country (relative wealth). In no case where the two wealth measures could be compared was the statistical significance of living arrangements variables affected by the choice of wealth control. Thus the absolute wealth index was used to retain the maximum number of countries and comparability.

7 2014 population data were from the World Population data sheet (PRB 2014). Brazil had 5434 observations, Mexico 3207, Colombia 1278, Peru 825, Guatemala 426, Haiti 289, Dominican Republic 279, Bolivia 276, Honduras 220, Nicaragua 166, and Guyana 19. Some of the regional variables for Guyana were dropped due to lack of observations.