Can Neighbourhoods Change the Decisions of Youth on the Margins of University Participation?

Kelly Foley *
Department of Economics
Copenhagen Business School

September 2010

Abstract

This paper uses the Youth in Transition Survey (YITS) to estimate the relationship between neighbourhoods and university participation among Canadian youth. Neighbourhood quality is proxied by the fraction of adults with a Bachelors degree living in close proximity to the youth. The estimated effect is identified using neighbourhood variation within high schools. At the mean, a one percentage point change in neighbourhood BA share is associated with a .16 percentage point change in the probability of attending university. The paper also tests whether the neighbourhood effects differ according to family background and children's reading skills. Neighbourhoods do not affect university participation among youth drawn from either the bottom or the top of the socio-economic distribution. Neighbourhoods have the largest effect on youth drawn from the middle of the socio-economic distribution who also have above median reading skills.

*Correspondence: Department of Economics, Copenhagen Business School, Porclænshaven 16A, 1, Frederiksberg 2000, Denmark, kf.eco@cbs.dk. The statistical analysis in this paper was produced from Statistics Canada microdata. The interpretation and opinions expressed are my own and do not represent those of Statistics Canada. I would like to thank David Green, Thomas Lemieux, Siwan Anderson, Nicole Fortin, Christine Neill, Werner Antweiler, two anonymous referees and participants of the UBC micro-empirical seminar series for helpful comments.

1 Introduction

In the economics literature which focusses on the determinants of university participation the impact of family background plays a dominant role. The family and home environment is not, however, the only arena in which young people develop. Schools and neighbourhoods provide extra-familial contexts where interactions with peers, teachers and other adults might mitigate or reinforce family impacts. The possibility that the local neighbourhood in which a youth resides can affect her chances of participating in university is examined in this paper. Neighbourhood influences are measured by the share of adults living in close proximity who have a Bachelors degree. The neighbourhood BA share proxies the availability of adult role models and peers who might have university aspirations.

An important part of the empirical literature estimating neighbourhood effects has found modest or no impacts on a range of outcomes (Oreopoulos, 2003; Kling et al., 2007). This paper contributes to that literature by demonstrating that there is a connection between neighbourhoods and university attendance particularly among individuals who are near the margins of participation. Taking advantage of a rich Canadian data set, the Youth in Transition Survey (YITS), the paper reports estimates of how neighbourhoods change the marginal probability of university attendance at different points on the distributions of youth's reading skills and their socio-economic background.

The YITS is a longitudinal survey of youth from across Canada, and includes a standardized reading test and a parental survey. A key feature of the data used in this paper is detailed information about where youth attended high school and where they lived at age 15.

Estimating the relationship between neighbourhoods and individual outcomes is particularly difficult because families with similar unobserved characteristics tend to cluster together, making it difficult to disentangle the effects of family characteristics from neighbourhood impacts. To identify the neighbourhood effect, it is assumed that conditional on school choice, family background and a measure of reading skills selection into neighbourhoods is uncorrelated with the unobserved factors determining university participation. Because a neighbourhood in this paper is an area of roughly one to four blocks in size, the variation in neighbourhood BA share is assumed to occur at a

geographic scale that is smaller than the scale over which families make their location decisions. The assumption is grounded in evidence from the literature which suggests that school quality has value for families. If parents focus on schools when choosing their neighbourhood, one would expect considerable sorting across schools. Variation in neighbourhood BA share *conditional on high school* is exogenous because sorting *within* school catchments is constrained by the thinness of markets for houses in specific neighbourhoods.

Estimates suggest that neighbourhoods are strongly related to university participation. Compared to a youth who lives in a community with no university-educated adults, a youth living in a community where all adults hold a Bachelors degree is roughly 16 percentage points more likely to attend university.

The results also imply that marginal effects on university participation differ by family background and reading skills, measured by scores on a standardized test. For youth with university educated parents, neighbourhoods have no impact on participation independent of reading skills. In contrast, for those with high school and college educated parents, neighbourhood effects differ across the distribution of reading skills. The marginal effect is largest for youth whose reading skills are above the median and at least one of their parents have a high school diploma. In contrast, neighbourhoods have no effect on participation on youth from less educated, low-income families.

The remainder of the paper proceeds by first describing evidence from the literature considering the link between neighbourhoods and individual outcomes. The section that follows outlines the identification strategy and empirical specification. The data are described in Section 4. The results from estimating average marginal effects are reported in Section 5. Section 6 includes a discussion of endogenous sorting into neighbourhoods and presents a sensitivity analysis. Finally, the paper reports the estimated marginal effects evaluated at different points along measures of skills and socio-economic background and concludes.

2 Evidence of the Link between Neighbourhoods and Individual Outcomes

Recently, economists have added to the growing psychology literature that explores how individuals are affected by group interactions, behaviour and beliefs.[1] Two broad effects are identified in this literature. The first is peer or conformity effects, which describes the tendency for individuals to behave like group members because they experience some positive utility from group membership, (for example Bernheim, 1994; Akerlof and Kranton, 2000).

In the context of university participation, one can think of how the propensity for university aspirations within a neighbourhood might affect individual behaviour. For example, students who reveal their university aspirations at school, perhaps through active participation, could be targeted for bullying in their neighbourhoods. This argument is similar to a concept described as 'acting white' in the U.S. literature (Roland G. Fryer and Torelli, 2005).

The second broad category of group effects can be described as role model effects, which might arise from incomplete information. Empirical evidence in both Canada and the U.S. suggests that low income youth have less accurate information about the costs and benefits of education (Ludwig, 1999; Junor and Usher, 2004). When information is not fully or easily available, individuals may either gain information from observing those around them or may simply mimic behaviour for lack of information.

For youth, outside the family, schools are probably the most important social arena (Duncan et al., 2001). The model estimated in this paper does not require that local neighbourhoods are the only social forum which affects behaviour. Instead, what matters is that the social influences in a youth's neighbourhood are incrementally important, holding constant school and family effects.

Although theory clearly predicts an important role for neighbourhoods, empirical evidence of a casual link is mixed.[2] Many empirical studies have relied on observational data and multivariate re-

[1] Durlauf (2004) provides a survey of the theoretical literature on neighbourhood and group effects.

[2] Oreopoulos (2008) reviews the Canadian literature, while Vigdor (2006) comprehensively surveys the U.S. literature.

gressions to estimate neighbourhood effects and have found effects on a range of outcomes including income, education (Borjas, 1995; Gibbons, 2002), childhood development and health (Brooks-Gunn et al., 1993; Tremblay et al., 2001; Kohen et al., 2002; Curtis et al., 2004), and social assistance receipt (Corcoran et al., 1992).

In contrast, the branch of literature using experimental and quasi-experimental methods report zero or modest neighbourhood effects. Two prominent examples are the Moving to Opportunities (MTO) project and a natural experiment arising from the Metro Toronto Housing Corporation's allocation policy (Oreopoulos, 2003). MTO, which was evaluated using random assignment, offered low-income residents of public housing a voucher and support to move to higher income neighbourhoods. Although MTO did increase neighbourhood quality on several dimensions, there were no impacts on economic outcomes for adults and the children's schooling outcomes were modest (Orr et al., 2003). In Oreopoulos' study (2003), because of lengthy waiting lists, candidates for public housing were effectively randomly assigned to housing projects of varying quality irrespective of their preferences. Along several dimensions of quality, Oreopoulos (2003) found no significant differences in any economic outcomes.

Another set of studies employs instruments and other types of plausibly exogenous neighbourhood variation. Aaronson (1998), for example, compared sibling outcomes among families who had moved. His estimates suggested that a 10 percentage point increase in the local poverty rate reduces the chance of high school completion by 7 percentage points.

Card and Rothstein (2007) tested whether racial segregation in neighbourhoods and schools can explain the black-white gap in SAT scores. Aggregating across metropolitan areas to eliminate the correlation between neighbours' characteristics and their preferences for integrated neighbourhoods, they find that changing from a fully segregated to a fully integrated city would reduce the test score gap by 25 per cent.

With data from the French Labour Force Survey, Goux and Maurin (2007) examine whether the proportion of youth living in close proximity who have repeated a grade at age 15 will affect the chances that an individual repeats a grade by age 16. The authors use the month of birth

to instrument for neighbours' educational advancement. They find that a one standard deviation increase in the proportion of neighbours who have been held back a grade will increase the chance that a youth is held back by 10 to 15 percentage points.

This paper contributes to this literature by testing whether the impact of neighbourhoods differs along the distributions of skills and family background. The results are consistent with the literature which finds zero impacts in low-income samples. This paper also demonstrates that neighbourhoods can have a large impact for some youth especially if they are drawn from the top half of the reading skills distribution.

3 Estimation and Identification

To disentangle the impact of neighbourhoods from family and individual effects, the probability of university attendance is compared among similar individuals who attend the same high school but live in neighbourhoods of varying quality. This approach is very similar to one employed by Bayer, Ross and Topa (2008), who are interested in whether local neighbourhoods provide effective job referral networks. Using data from the U.S. Census, Bayer et al. (2008) compare the probability that an individual works with a neighbour living on the same city block in metropolitan Boston, to the probability of working with a neighbour from nearby blocks. They argue that while individuals do choose their residential area, defined as a block group consisting of roughly 10 blocks, sorting within block groups is minimal because the market for a specific block is too thin. They find that people living on the same U.S. Census block are 33 per cent more likely to work together when compared to those within the block group.

The variation in neighbourhood BA share used in this paper is at the scale of one to four blocks. The assumption supporting identification is that the geographic scale of this variation is smaller than the scale at which families can choose where they live in practice. In other words, while families may have a preference for a specific block, the chances are small that a house which suits their specific needs and tastes is actually for sale during the family's search period. In practice, to ensure that they can find a house, the typical family would search for houses and would be willing

to live within geographic areas that are much larger than a collection of one to four blocks. Because a large part of why families choose where they live is the quality of the schools their children will attend, a school catchment area is a reasonable way to think about a family's search area.

In the U.S., several studies confirm that families care about school quality when they choose their homes. Specifically, the literature suggests that school quality is capitalized into housing prices (e.g. Barrow, 2002; Bayer et al., 2007). In Canada, while very little research has addressed this issue, there is some evidence that schools are similarly important for Canadian families. One example is work by Ries and Somerville (2010) who exploited a change in school boundaries which meant a group of houses in Vancouver formerly assigned to a low quality school catchment were reassigned to a substantially higher quality school. Their estimates suggest that a 4 per cent change in high school quality is associated with an increase in the median house price of $14,000.

To support the assumption that the availability of houses within a school catchment is random, it is helpful to suggest the possible mechanisms which could generate such exogenous variation. First, there may be heterogeneous preferences for housing characteristics that are uncorrelated with university participation. For example, preferences for ranch verses Victorian style homes. Since heterogenous demand does not necessarily lead to higher prices, families with similar wealth might be exogenously distributed across neighbourhoods according to those preferences.

Another possible source of exogenous variation conditional on schools and income is the process that generates houses on the market. Houses become available in a specific neighbourhood for many reasons that are uncorrelated with the neighbourhood's characteristics. These reasons could include any idiosyncratic family shocks, such as divorce or promotion which requires relocation. Naturally, houses are also available on the market for endogenous reasons, such as declining labour markets or rising crime rates. Such trends in neighbourhood quality might cause sorting across school catchment areas but are unlikely to be so localized as to affect only a few blocks.

Finally, variation will exist within a school catchment area because in any given city or region, the menu of school and housing quality choices will not be continuous. If there were as many schools to choose from as there were family-preference types then families would perfectly sort by

school and there would not be any neighbourhood variation within catchment areas. It is more reasonable to assume that there are many more preference types than there are choices over school and housing quality. Because of the relatively few school catchments within a given city one might, for example, observe variation in neighbourhood BA share because a plumber's family lives beside a school teacher's family.

The empirical specification is a reduced-form linear index representing the net lifetime utility from attending university. The net utility for individual i, living in neighbourhood n, attending school s can be defined as:

$$\omega_{ins} = x_i\beta + x_s\delta + x_n\gamma + \theta_n + \theta_s + \epsilon_{ins} \tag{1}$$

The latent variable is a function of observable individual and family characteristics (x_i), observable (x_s) and unobservable (θ_s) school characteristics, observable (x_n) and unobservable (θ_n) neighbourhood effects and an idiosyncratic error (ϵ_{ins}). The neighbourhood effect in (1) is what Manski (1993) would call a contextual effect. Any endogenous effects which may be present will be reflected in the estimate of γ.[3]

The neighbourhood effect is identified from the assumption that conditional on school selection the process determining a family's neighbourhood is uncorrelated with $\theta_n + \epsilon_{ins}$. This assumption is based on the idea that location choice within a school catchment is restricted by the thinness of housing markets.

The error in (1) is further assumed to be normally distributed so that the probability of university attendance can be estimated in a Probit,[4] with school fixed effects which will absorb both $x_s\delta$ and θ_s. What is of particular interest in this paper is not just whether the neighbourhood effect (γ) is

[3]Manski (1993), Moffitt (2001) and Brock and Durlauf (2007) have described and summarized the issues associated with estimating group interaction effects. In addition to the complications introduced by selection into neighbourhoods and unobservable neighbourhood characteristics, Manski (1993) details the econometric implications of what he termed the 'reflection problem'. The reflection problem arises from the attempt to estimate the impact of average behaviour within a group on the behaviour of the individuals that comprise the group. Manski (1993) calls the relationship between individual behaviour and the behaviour of a defined peer group an endogenous effect. In (1) an endogenous effect would involve an impact of peer preferences on individual preferences.

[4]In general, fixed effects in a Probit do not consistently estimate the parameters. If however, one assumes that the school effects are proportional to the mean of the independent variables then average partial effects are consistently estimated (Wooldridge, 2002). Moreover, the estimates in linear probability models are very similar.

non-zero but also if it varies across family background, reading skills. To examine this possibility the model is also estimated with interactions between individual and neighbourhood characteristics.

4 Data

This paper uses data from the Youth in Transition Survey (YITS). The YITS is a longitudinal survey of Canadian youth. While two cohorts of youth were surveyed in the YITS project, this paper uses data from the younger cohort (YITS-A). The youth were born in 1984, were surveyed first in 2000, and again every two years. Four cycles of data were available for use in this paper. The YITS data is particularly rich because the first cycle for the younger cohort included an internationally comparable reading test, a parental survey and a school administrators survey.

The reading test was administered through the Programme for International Student Assessment (PISA). PISA was an effort, co-ordinated by the Organization for Economic Co-operation and Development (OECD), to generate internationally coherent measures of cognitive skills.

The sample used in this paper is drawn from the 'reading cohort' and includes youth who were respondents to all four waves of the survey and whose parent completed the parental survey in the first cycle.[5] Youth who board at school or commute farther than 45 minutes by car or public transport are excluded. These youth are dropped from the sample because it is less likely that their residential location is linked to their school attendance.

The base sample is 15,034, which represents roughly 45 per cent of the original sample. Most of the sample is lost due to longitudinal attrition. The overall survey response rate is 54.7 per cent. Weights provided by Statistics Canada are used to account for attrition. Further information about attrition is available in an accompanying appendix available on the internet.

The outcome of interest is university participation and is measured when the youth are age 21. Survey respondents who reported ever attending a university prior to December 2005 are defined as university participants. The university participation rate in the sample is roughly 40 per cent.

[5]While the YITS project also includes science and math skills tests, reading scores are used because the whole YITS sample wrote the reading test. The sample within each school was randomly divided so that half the students wrote the math and the other half wrote the science test.

The data also includes a rich set of control variables which are measured when the youth are age 15. Further information about the control variables including the sample means and standard deviations are reported in the accompanying appendix.

Neighbourhood quality is defined as the fraction of adults who have obtained a Bachelors degree. Because of the strong relationship between parental and children's education, this measure of neighbourhood quality might reflect the youth's peer environment. It also describes the potential non-familial adult role models available to youth.

The youth's neighbourhoods are their home residences measured when they are age 15 and are geographically defined by the Enumeration Area (EA). The EA is the smallest geographic unit used by Statistics Canada and is roughly one to four blocks in size. Neighbourhood characteristics are taken from public use files of the 2001 Canadian Census Profiles. The mean BA share weighted by students in the sample is .159 with a standard deviation of .116. Identification in this paper relies on neighbourhood variation *within* schools. There are 1013 schools in the sample; some have very small sample sizes while others have samples over 100 students. The mode sample size is in the order of 25 students.

In a previous version, a number of other variables including income, population, the unemployment rate and the fraction of immigrants were also used to describe neighbourhoods but these variables were substantially less important. The predicted probability of university participation conditional on the BA share alone is highly correlated (.96) with the predicated probability of attendance conditional on BA share and income, population, the unemployment rate and immigrant share.

5 Results

The first column of Table 1 shows the results from a Probit regression of university participation on the proportion with a bachelors degree in an Enumeration Area (EA) and gender. The mean of the marginal effects are reported. In all of the models estimated in the paper, the standard errors are clustered by school. Clustering at the school level accounts for any heterogeneity across individuals

or neighbourhoods in the school effect. This first specification which demonstrates the degree of correlation between neighbourhoods and individual outcomes can be thought of as a 'base' model to be compared with other specifications.

Because the neighbourhood BA share is measured as a proportion, the marginal effect in the first column can be interpreted as percentage point changes. In other words, the marginal effect reported in column 1 implies that a one percentage point increase in the BA share is associated with .93 of a percentage point increase in the probability of university participation.

Fixed effects for high schools are added in the second column of Table 1. After accounting for the school fixed effects, which are jointly significant at the 1 per cent level, the estimated effect of neighbourhood quality drops to .545. For comparison, the model reported in the third column includes parental education but no school controls. Parental education is defined by categories representing the highest level of education in the family: high school, postsecondary education (PSE) below the Bachelors level, one parent has a university degree and both parents have university degrees.

The effects reported in column 3 are relative to a family where both parents were high school dropouts. It is a well established result that parental education is strongly associated with university participation (Carneiro and Heckman, 2002; Black et al., 2005; Drolet, 2005; Oreopoulos et al., 2006). Indeed, a youth whose parents both have a BA is 58 percentage points more likely to attend university when compared to a youth whose parents are high school dropouts.

When compared to column 2 with school fixed-effects, the neighbourhood effect falls by a similar magnitude when parental education is added to the model. This suggests a high degree of correlation between school characteristics and parental education.

The results in column 4, suggests that schools capture more than just parental education. When school fixed effects and parental education are both added to the model, the neighbourhood effect falls to .253.

Adult equivalent family income quartiles are added in the final column of Table 1. This measure of income is derived by dividing total before-tax family income by the square root of the number of

household members. Without other socio-economic variables in the model, income has some impact on the probability of attending university. Youth from families with income in the highest quartile are nearly 7 percentage points more likely to attend university relative to the lowest income families. The addition of these controls has very little impact on the neighbourhood effect.

Table 2 adds to the model a set of background characteristics which characterize families' economic resources more fully than spot measures of income, including family structure, parental labour force status, rural residence, mobility, Aboriginal and immigrant status. The family structure categories include two biological or adoptive parents as the reference group, along with lone parent families, and other two parent.[6] Parental labour force status is measured by two dummy variables, which take on the value one if either the youth's mother or father works. Rural residence is defined in Statistics Canada's Statistical Area Classification as population density less than 400 people per square kilometer. Mobility is measured by the number of household moves the child had experienced by age 15. Variables indicating that the youth was born in another country and if the youth was born in Canada but at least one of her parents were born outside of Canada are included. Also included are a dummy if the youth first learned a non-official language and an indicator for the youth's Aboriginal status.[7]

The first column of Table 2 reports results from a model including the measures of socio-demographic background. The direct effects of these factors are similar to what one would expect based on previous research on university attendance (Zhao et al., 2003; Finnie et al., 2004; Drolet, 2005). Native youth are less likely and second generation youth are more likely to attend university. Youth who first spoke a non-official language are also more likely to attend university. This variable captures to a limited extent differences in the source country of immigration.[8]

Youth living in rural communities are about 5 percentage points less likely to attend university. Compared to youth who live with both of their biological parents, youth from 'other' two parent

[6]Families which combine an adoptive and birth parent are included in the 'other two parent' family category.

[7]Aboriginal status includes North American Indians, Métis and Inuit.

[8]The educational attainment of second generation children varies considerably by the parents' birth country (Boniskowa, 2005; Aydemir and Sweetman, 2006; Abada et al., 2009). Results where controls for country of origin are included are presented in the appendix. The estimated neighbourhood effect changes very little.

families are nearly 11 percentage points less likely to be university participants. After controlling for the demographic characteristics, the estimated neighbourhood effect falls to .207.

The model reported in the second column of Table 2 includes quartiles of the youth's scores on the PISA reading test, which can be thought of as a measure of the stock of reading skills at age 15. Youth who score in the highest quartile are .54 percentage points more likely to attend university compared to those who score in the lowest quartile. Adding a control for reading skills to the model actually increases the neighbourhood effect but only by .007.

Parental aspirations for their children's educational attainment are included in the specification reported in column 3 and are measured by a question asking parents to indicate the highest level of education they 'hope' their child will obtain. The responses were coded into four categories, high school or less, PSE below a BA, which includes college diplomas or trade certificates, Bachelors degree or above, and any level above high school. These variables should capture parents' preferences for education or any knowledge they have about the abilities and motivations of their children. A youth whose parent hopes she will obtain at least a Bachelors degree is about 31 percentage points more likely to attend university relative to a youth whose parents expect her to achieve high school graduation or less. Including parental aspirations in the model once again reduces the neighbourhood effect to .177.

In the final column of Table 2 measures of family behaviour which are correlated with unobserved ability and preferences are included. When the youth were 15 years old, their parents were asked whether they had done anything specific to ensure that their child would have money for further education after high school. As Table 2 column 4 shows, the youth from families who responded yes were about 5 percentage points more likely to have attended university.

The behaviour of siblings is another way to measure the unobserved features of a family's propensity for university. Unfortunately, the YITS does not collect information about sibling participation in PSE; however, in cycle one, parents were asked whether any of the youth's siblings had dropped out of high school. This variable along with a dummy variable indicating that the child had no siblings are included in the regression reported in column 4 of Table 2. Youth with a high school

drop-out sibling are about 7 percentage points less likely to attend university compared to those whose siblings had not dropped out. Including these measures of family behaviour reduces the estimated neighbourhood effect to .163.

In the model with the full set of controls when neighbourhood BA share increases by one standard deviation, the probability of university attendance increases by 1.89 percentage points. This result is considerably smaller than effects found by others. For example, Goux and Maurin (2007), who found that a one standard deviation change in neighbourhood quality resulted in .7 of a standard deviation change in the chances that a youth repeats a grade.

6 Neighbourhood Selection and Sensitivity of Results

Identifying the neighbourhood effect separately from the effects of family and individual characteristics requires a conditional independence assumption. In other words, after controlling for observable characteristics, it is assumed that selection into neighbourhoods is uncorrelated with unobservable determinants of university participation. The key argument used to justify conditional independence in this paper is that while families sort across schools, the thinness of housing markets prevents them from sorting within a school area. If selection on observable characteristics is related to selection on unobserved factors, then one way to gauge the plausibility of this assumption is to examine sorting on observable characteristics after controlling for schools.

Table 3 shows results from a linear regression of neighbourhood BA share on the full set of control from the last column of Table 2. For comparison, the first column of Table 3 does not include school fixed effects, while in the second column the fixed effects are included. Without controlling for schools, virtually all of the variables that affect university participation are also highly correlated with neighbourhood quality. For example, on average, in the neighbourhoods where children whose parents both have a BA live the BA share is 8 percentage points higher than in the neighbourhoods of children whose parents are both high school dropouts.

After controlling for schools, the correlation between neighbourhoods and observable characteristics is considerably weaker. Of the characteristics that are still significantly related to neighbour-

hood quality many of the effects are very small. Notably, there is no sorting based on the children's reading skills. Sorting based on parental education persists to some extent after controlling for schools.

To further examine how sensitive estimates are to the conditional independence assumption, an approach similar to one suggested by Altonji et al. (2004) is adopted. This approach treats the outcome and selection into neighbourhoods as a system of two equations,

$$\omega_{ins} = x_i\beta + x_n\gamma + \alpha_s + \eta_{ins} \tag{2}$$

$$x_n = U(x_i, s) + \upsilon_{ins} \tag{3}$$

If η_{ins} and υ_{ins} are uncorrelated then conditional independence holds. If the two disturbances are correlated then γ is not identified because there is no instrument in the data that determines selection but not university participation. It is possible, however, to impose the degree of correlation between the disturbances and estimate γ using a control function approach,

$$\omega_{ins} = x_i\beta + x_n\gamma + \alpha_s + \rho\hat{\upsilon}_{ins} + e_{ins} \tag{4}$$

where $\rho = \mathbb{E}[\eta_{ins}\upsilon_{ins}] / \mathbb{E}\left[\upsilon_{ins}^2\right]$ and $\hat{\upsilon}_{ins}$ is an estimate of the error in (3). The estimate is the residual obtained by assuming the function U in (3) is linear and regressing neighbourhood BA share on the control variables and the school fixed effects. In other words, it is the residual from the regression reported in column 2 of Table 3. If a value for ρ is assumed then γ can be estimated in a Probit model of (4).

Table 4 presents estimates of the neighbourhood effect assuming different values of the correlation coefficient. As the degree of correlation increases, the estimated value of the neighbourhood effect falls. Assuming a fairly low correlation of .1, the estimated neighbourhood effect is .136. The effect falls to .086 and is statistically significant only at the 10 percent level when $\rho = .3$. The effect is no longer statistically significant when the correlation coefficient is .4 and becomes negative when $\rho = .7$. This suggests that if the true neighbourhood effect was zero the correlation in the

unobserved factors would need to be greater than .6 to entirely explain a positive effect.

7 Marginal Effects Across the Socio-Economic Distribution

Depending on their skills, preferences and financial constraints, some youth may be closer to the margin of university participation and as a result neighbourhoods may have a larger impact on their participation decision. To test for such heterogeneity in the effects, neighbourhood BA share is interacted with gender, parental education, family income, PISA scores and parents' hopes for their children's educational attainment. The estimated coefficients are reported in Table 5. These results suggest that the coefficients on neighbourhood BA share do not differ statistically by PISA scores, income and gender. There are, however, statistically significant differences in the coefficients across parental education categories, and parents' aspirations for their children's education.

Next, the neighbourhood effects are estimated for different socio-economics status (SES) types that are defined according to the youth's family income, their parents' education, their PISA scores and the level of education their parents hope they will achieve. The types are constructed to reflect the correlation between skills and background characteristics that are evident in the data. The first type, called 'Low', is composed of children drawn from the bottom quartile of reading scores, and the bottom quartile of family income, whose parents do not have a high school diploma, and hope their children complete high school or less. Children whose reading scores fell in the second quartile, whose family income is in the second quartile, who have at least one parent with a high school diploma, and who have a parent that hopes they complete college or some level of education above high school are in the 'Medium-High School' group. Children in the 'Medium-PSE' type have reading scores in the third quartile, family income in the third quartile, have at least one parent with PSE below a BA, and have a parent that hopes they complete college or some level of education above high school. The type called 'High' is formed by children from the top of the reading distribution, as well as the top of the income distribution. Their parents' both have a Bachelors degree and the responding parent hopes their child will obtain a university degree.

Table 6 reports the predicted probability of university participation and the marginal effect of a

one percentage point change in the neighbourhood BA share. The strong relationship between SES, skills and university participation is immediately evident from the predicted probabilities in the top panel of Table 6. The probability of attendance for a High type youth is 90 percentage points higher than a Low type. The types have been constructed in such a way that the low type almost never attends university, while the high type almost always does. Not surprisingly neighbourhoods have a larger impact for the two Medium types. The neighbourhood effect is statistically insignificant and actually negative for both the low and high types. One way to interpret the results in the first panel of Table 6 is that the neighbourhood effects are largest for types who are closest to the margin of university participation.

The remaining panels of Table 6 again show the marginal effects and predicted participation rates for each SES type while PISA quartiles are varied. Within lowest and highest socio-economic groups, the neighbourhood effect is statistically insignificant in all four quartiles of the skills distribution. For the two medium socio-economic types, whether neighbourhoods affect university participation depends on the youth's reading skills. Among youth whose reading skills are above the median, neighbourhoods have a stronger relationship with university participation. The marginal effects are largest for the medium-high school types who scored in the third reading quartile. For these youth, a one percentage point change in the neighbourhood BA share is associated with a .35 percentage point change in the probability of attending university. In comparison, the neighbourhood effect for a similar youth drawn from the bottom of the skills distribution is .14 and is statistically insignificant.

Although large neighbourhood effects are found among some groups, it is worth noting that the findings are still consistent with the experimental literature, which found that neighborhood quality had little or no impact on economic outcomes for individuals drawn from the bottom of the socio-economic distribution. For example, the program group in the Moving to Opportunity experiment lived in neighbourhoods where the share of families with a college degree was about 2 percentage points higher relative to control groups' neighbourhoods (Orr et al., 2003). The effect of this change in neighbourhood quality on university participation among participants' children five

years after the experiment began was about .005 and was statistically insignificant. The majority of families in the MTO project were headed by high school dropouts and welfare recipients. The marginal effects estimated in this paper for the children of low-income high school dropouts are also statistically insignificant.

8 Conclusion

This paper has demonstrated that neighbourhoods are strongly related to university participation. At the mean of the data, the estimated effect of living in a neighbourhood where all neighbours have a Bachelors degree relative to a neighbourhood where none have a BA is about 16 percentage points. This size of this effect is about one third of the size of the impact of scoring in the top quartile on the reading skills test relative to the bottom quartile, and roughly two thirds of the impact of growing up with parents who hold a BA relative to parents who are high school dropouts.

The marginal effect of neighbourhoods on the probability of university attendance differs substantially by socio-economic background. For youth whose parents have a Bachelors degree, neighbourhoods have no effect on university participation. Additionally, neighbourhoods have no effect on the marginal probability of attending university among youth whose parents are high school dropouts. This finding is consistent with the literature that also finds small or no neighbourhood impacts within disadvantaged populations. Youth in the middle of the SES distribution are most likely to be affected by neighbourhood quality.

The paper also reveals an interesting interaction between family background, and the pattern of marginal neighbourhood effects along the distribution of reading skills. One of the most striking findings is that the marginal effect of neighbourhoods is largest for skilled youth whose parents have high school diplomas, below median income and who have low educational aspirations for their children.

Overall, the results suggest that parents who have Bachelors degrees provide a family environment that orients youth toward university so that independent of their reading skills and influences in their neighbourhood, these youth attend university with above average probability. This conclu-

sion is consistent with Carneiro and Heckman (2002) and the notion of a missing market for family. In other words, the SES gradient in educational attainment exists in part because children can not 'purchase' the family background that leads to university participation. My results suggest that neighbourhoods may substitute for families but only for a small subset of higher skilled youth.

The result that neighbourhoods have little effect on youth from the lowest socio-economic backgrounds suggests that relocation policies, or desegregation policies may have limited impact on some labour market outcomes where disadvantaged individuals are far from the margin. There may be other outcomes, however, where disadvantaged individuals are sufficiently near the margin so that peer and role model effects do have a meaningful impact. These could include choices about health or parenting styles, and early investments in children's development.

Because of the emphasis on marginal youth, one lesson that can be drawn from the results is the importance of designing policies that can be targeted at specific populations who are most at risk. Rather than broad based relocation policies, it might be the case that smaller scale mentoring type programs are more appropriate. In this context, it might be quite difficult given the types of policy instruments available to governments to define the relevant subgroups who are most likely to benefit from such programs. Effective programs might define target populations by combining observed characteristics such as family background and referral systems involving teachers or community center leaders who have closer contact with youth and are better placed to identify those on the margins.

References

Aaronson, D (1998) 'Using sibling data to estimate the impact of neighborhoods on children's educational outcomes.' *Journal of Human Resources* 33(4), 915–46

Abada, Teresa, Feng Hou, and Bali Ram (2009) 'Ethnic differences in educational attainment among the children of Canadian immigrants.' *Canadian Journal of Sociology* 34(1), 1–28

Akerlof, G., and R. Kranton (2000) 'Economics and identity.' *The Quarterly Journal of Economics* CXV(3), 715–753

Altonji, Joseph, Todd Elder, and Christopher Taber (2004) 'Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools.' *Journal of Political Economy* 113(1), 151–184

Aydemir, Abdurrahman, and Arthur Sweetman (2006) 'First and second generation immigrant educational attainment and labor market outcomes: A comparison of the United States and Canada.' IZA Discussion Papers 2298, Institute for the Study of Labor (IZA), September

Barrow, L. (2002) 'School choice through relocation: Evidence from the Washington, D.C. area.' *Journal of Public Economics* 86(2), 155–189

Bayer, P., F. Ferreira, and R. McMillan (2007) 'A unified framework for measuring preferences for schools and neighborhoods.' *Journal of Political Economy* 115(4), 2001–2052

Bayer, P., S.L. Ross, and G. Topa (2008) 'Place of work and place of residence: Informal hiring networks and labor market outcomes.' *Journal of Political Economy* 116(6), 1150–1196

Bernheim, B.D. (1994) 'A theory of conformity.' *Journal of Political Economy* 102(5), 841–877

Black, S.E., P.J. Devereux, and K.G. Salvanes (2005) 'Why the apple doesnt fall far: Understanding intergenerational transmission of human capital.' *The American Economic Review* 95, 437–449

Boniskowa, A. (2005) 'Educational attainment of three generations of immigrants in Canada: Initial evidence from the Ethnic Diversity Survey.' Discussion Paper 05-21, UBC Department of Economics

Borjas, G. (1995) 'Ethnicity, neighbourhoods, and human-capital externalities.' *American Economic Review* 85(3), 365–390

Brock, W.A., and S.N. Durlauf (2007) 'Identification of binary choice models with social interactions.' *Journal of Econometrics* 140, 52–75

Brooks-Gunn, J, G.J. Duncan, P. Kato-Klebanov, and N. Sealand (1993) 'Do neighbourhoods influence child and adolescent development?' *The American Journal of Sociology* 99(2), 353–395

Card, D., and J. Rothstein (2007) 'Racial segregation and the black-white test score gap.' *Journal of Public Economics* 91, 2158–2184

Carneiro, Pedro, and James J. Heckman (2002) 'The evidence on credit constraints in post–secondary schooling.' *Economic Journal* 112(482), 705–734

Corcoran, M., R Gordon, D. Laren, and G. Solon (1992) 'The association between men's economic status and their family and community origins.' *Journal of Human Resources* 27(4), 575–601

Curtis, Lori J., Martin D. Dooley, and Shelley A. Phipps (2004) 'Child well-being and neighbourhood quality: evidence from the canadian national longitudinal survey of children and youth.' *Social Science & Medicine* 58(10), 1917–1927

Drolet, M. (2005) 'Participation in post-secondary education in Canada: Has the role of parental income and education changed over the 1990s?' Analytical Studies Research Paper Series, Statistics Canada

Duncan, G.J., J. Boisjoly, and K Mullan Harris (2001) 'Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development.' *Demography* 38(3), 427–447

Durlauf, Steven N. (2004) 'Neighborhood effects.' In *Handbook of Regional and Urban Economics,* ed. J. V. Henderson and J. F. Thisse (Elsevier)

Finnie, R., C. Laporte, and E. Lascelles (2004) 'Family background and access to post-secondary education: What happened over the 1990s?' Analytical Studies Branch Research Paper Series 2004226e, Statistics Canada, Analytical Studies Branch

Gibbons, Steve (2002) 'Neighbourhood effects on educational achievement.' CEE Discussion Papers 0018, Centre for the Economics of Education, LSE, February

Goux, D., and E. Maurin (2007) 'Close neighbours matter: Neighbourhood effects on early performance at school.' *The Economic Journal* 117, 1193–1215

Junor, S., and A. Usher (2004) 'The price of knowledge 2004: Access and student finance in Canada.' Technical Report, Millennium Scholarship Foundation, Montreal

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz (2007) 'Experimental analysis of neighborhood effects.' *Econometrica* 75(1), 83–119

Kohen, Dafna E., J. Brooks-Gunn, T. Leventhal, and C. Hertzman (2002) 'Neighbourhood income and physical and social disorder in Canada: Associations with young children's competencies.' *Child Development* 73(6), 1844–1860

Ludwig, J. (1999) 'Information and inner city educational attainment.' *Economics of Education Review* 18, 17–30

Manski, C.F. (1993) 'Identification of endogenous social effects: The reflection problem.' *Review of Economic Studies* 60, 531–542

Moffitt, R. (2001) 'Policy interventions, low-level equilibria, and social interactions.' In *Social Dynamics,* ed. S. Durlauf and P. Young (MIT Press)

Oreopoulos, P. (2003) 'The long run consequences of growing up in a poor neighborhood.' *Quarterly Journal of Economics* 118(4), 1533–1575

_ (2008) 'Neighbourhood effects in Canada: A critique.' *Canadian Public Policy / Analyse de Politiques* 34(2), 237–258

Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens (2006) 'Does human capital transfer from parent to child? The intergenerational effects of compulsory schooling.' *Journal of Labor Economics* 24(4), 729–760

Orr, L., J.D. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, L.F. Katz, J.V. Liebman, and J.R. Kling (2003) 'Moving to Opportunity interim impacts on evaluation.' Technical report, U.S. Department of Housing and Urban Development

Ries, John, and Tsur Somerville (2010) 'School quality and residential property values: Evidence from Vancouver rezoning.' *Review of Economics and Statistics*

Roland G. Fryer, Jr., and Paul Torelli (2005) 'An empirical analysis of acting white.' NBER Working Papers 11334, National Bureau of Economic Research, Inc, May

Tremblay, R.E., B. Boulerice, H. Foster, E. Romano, J. Hagan, and R. Swisher (2001) 'Multi-level effects on behaviour outcomes in Canadian children.' Applied Research Branch of Strategic Policy Working Paper W-01-2E, Human Resources Development Canada

Vigdor, J.L. (2006) *Deviant Peer Influences in Programs for Youth: Problems and Solutions* (New York: Guilford Press)

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data* (Cambridge: MIT Press.)

Zhao, John, Miles Corak, and Garth Lipps (2003) 'Family income and participation in post-secondary education.' Analytical Studies Branch Research Paper Series 2003210e, Statistics Canada, Analytical Studies Branch, October

Table 1: Impact of Neighbourhoods on University Participation

	(1)	(2)	(3)	(4)	(5)
Neighbourhood BA Share	0.929	0.545	0.489	0.253	0.223
	(0.050)***	(0.004)***	(0.049)***	(0.060)***	(0.059)***
Female	0.158	0.167	0.162	0.171	0.172
	(0.008)***	(0.009)***	(0.008)***	(0.013)***	(0.013)***
Parents' highest educational attainment– Reference both have less than high school					
High school			0.123	0.129	0.121
			(0.014)***	(0.018)***	(0.018)***
PSE below BA			0.216	0.226	0.213
			(0.014)***	(0.020)***	(0.020)***
One parent has a BA			0.426	0.428	0.407
			(0.017)***	(0.026)***	(0.027)***
Both parents have a BA			0.584	0.562	0.533
			(0.018)***	(0.029)***	(0.031)***
Family income quartiles– Reference bottom quartile					
Quartile 2					0.006
					(0.012)
Quartile 3					0.040
					(0.013)***
Quartile 4					0.067
					(0.014)***
School fixed effects	N	Y	N	Y	Y
Sample size	15,034	15,034	15,034	15,034	15,034

Marginal effects estimated in a Probit model. Dependent variable = 1 if child ever attended a university
Standard errors clustered by school presented in parenthesis
Stars indicate statistical significance: *** .01, ** .05, * .1

Table 2: Neighbourhoods Effects Controlling for Skills and Family Background

	(1)	(2)	(3)	(4)
Neighbourhood BA Share	0.207	0.214	0.177	0.163
	(0.059)***	(0.060)***	(0.058)***	(0.057)***
Female	0.175	0.103	0.097	0.098
	(0.013)***	(0.010)***	(0.010)***	(0.010)***
Parents' highest educational attainment– Reference both have less than high school				
High school	0.125	0.093	0.080	0.073
	(0.018)***	(0.018)***	(0.018)***	(0.018)***
PSE below BA	0.215	0.144	0.119	0.106
	(0.021)***	(0.019)***	(0.019)***	(0.019)***
One parent has a BA	0.398	0.271	0.231	0.214
	(0.027)***	(0.025)***	(0.025)***	(0.024)***
Both parents have a BA	0.510	0.329	0.284	0.266
	(0.032)***	(0.030)***	(0.029)***	(0.029)***
Family income quartiles– Reference bottom quartile				
Quartile 2	0.011	-0.008	-0.011	-0.015
	(0.012)	(0.012)	(0.012)	(0.012)
Quartile 3	0.055	0.018	0.006	-0.003
	(0.014)***	(0.013)	(0.013)	(0.013)
Quartile 4	0.088	0.051	0.033	0.022
	(0.015)***	(0.014)***	(0.014)**	(0.014)
PISA reading quartiles– Reference bottom quartile				
Quartile 2		0.161	0.136	0.136
		(0.016)***	(0.015)***	(0.015)***
Quartile 3		0.349	0.305	0.305
		(0.021)***	(0.021)***	(0.021)***
Quartile 4		0.542	0.484	0.482
		(0.027)***	(0.026)***	(0.027)***
Parental aspirations for child– Reference high school or less				
PSE below BA			0.063	0.056
			(0.028)**	(0.028)**
BA			0.313	0.298
			(0.042)***	(0.042)***
Any level above high school			0.176	0.168
			(0.037)***	(0.037)***
Sample size	15,034	15,034	15,034	15,034

continued, next page

Table 2 cont'd: Neighbourhoods Effects Controlling for Skills and Family Background

	(1)	(2)	(3)	(4)
Family background				
Aboriginal status	-0.059 (0.023)***	-0.034 (0.024)	-0.047 (0.024)**	-0.043 (0.024)*
Immigrant	0.061 (0.026)**	0.104 (0.028)***	0.088 (0.027)***	0.090 (0.027)***
Second generation	0.056 (0.014)***	0.061 (0.014)***	0.051 (0.013)***	0.051 (0.013)***
Non-official language spoken at home	0.115 (0.023)***	0.116 (0.023)***	0.100 (0.023)***	0.095 (0.023)***
Rural residence	-0.055 (0.021)***	-0.020 (0.019)	-0.013 (0.019)	-0.013 (0.019)
Number of household moves	-0.010 (0.057)	-0.010 (0.058)	-0.010 (0.056)	-0.010 (0.056)
Mother works	-0.003 (0.009)	0.001 (0.009)	-0.001 (0.009)	-0.001 (0.009)
Father works	-0.012 (0.015)	-0.018 (0.014)	-0.018 (0.015)	-0.021 (0.015)
Family structure– Reference two biological/adoptive parents				
Lone parent	-0.019 (0.016)	-0.032 (0.016)**	-0.042 (0.016)**	-0.038 (0.016)**
Other two parent families	-0.107 (0.016)***	-0.087 (0.015)***	-0.084 (0.015)***	-0.078 (0.015)***
Family behaviour				
Parents' prepared financially for education				0.051 (0.009)***
Sibling is high school dropout				-0.069 (0.016)***
Only child				-0.021 (0.016)
Sample size	15,034	15,034	15,034	15,034

Marginal effects estimated in a Probit model. Dependent variable = 1 if child ever attended a university
All specifications include school fixed effects.
Standard errors clustered by school presented in parenthesis
Stars indicate statistical significance: *** .01, ** .05, * .1
Families that combine biological and adoptive parents are included in the 'other two parent family' category.

Table 3: Sorting Across Neighbourhoods

	(1)	(2)
Female	-0.004	-0.001
	(0.003)*	(0.002)
Parents' highest educational attainment		
Reference both have less than high school		
High school	-0.002	-0.004
	(0.005)	(0.004)
PSE below BA	0.007	0.002
	(0.005)	(0.004)
One parent has a BA	0.049	0.019
	(0.006)***	(0.005)***
Both parents have a BA	0.083	0.032
	(0.008)***	(0.006)***
Family income quartiles		
Reference bottom quartile		
Quartile 2	0.006	0.005
	(0.004)	(0.003)*
Quartile 3	0.012	0.006
	(0.004)***	(0.003)**
Quartile 4	0.034	0.019
	(0.004)***	(0.003)***
PISA reading quartiles		
Reference bottom quartile		
Quartile 2	0.008	0.001
	(0.004)**	(0.003)
Quartile 3	0.006	-0.002
	(0.004)	(0.003)
Quartile 4	0.016	-0.002
	(0.005)***	(0.003)
Parental aspirations for child		
Reference high school or less		
PSE below BA	-0.002	0.006
	(0.006)	(0.005)
BA	0.014	0.011
	(0.006)**	(0.005)**
Any level above high school	-0.002	0.003
	(0.007)	(0.006)
School fixed effects	N	Y
Sample size	15,034	15,034

continued, next panel

Table 3 Continued

	(1)	(2)
Family background		
Aboriginal status	-0.015	-0.011
	(0.006)**	(0.004)**
Immigrant	0.032	-0.010
	(0.007)***	(0.006)*
Second generation	0.026	-0.002
	(0.004)***	(0.003)
Rural residence	-0.080	-0.046
	(0.004)***	(0.007)***
Number of household moves	0.000	0.001
	(0.001)	(0.000)
Mother works	-0.005	-0.005
	(0.003)*	(0.002)**
Father works	-0.011	-0.001
	(0.005)**	(0.004)
Family status		
Reference two biological/		
adoptive parents		
Lone parent	0.005	0.000
	(0.006)	(0.004)
Other two parent families	0.000	-0.002
	(0.004)	(0.003)
Family behaviour		
Parents' prepared financially	0.011	0.004
for education	(0.003)***	(0.002)*
Sibling is high school dropout	-0.006	-0.007
	(0.004)	(0.003)*
Only child	0.003	-0.003
	(0.005)	(0.004)
School fixed effects	N	Y
Sample size	15,034	15,034

Dependent variable: Neighbourhood BA share
Standard errors clustered by school presented in parenthesis
Stars indicate statistical significance: *** .01, ** .05, * .1
Families that combine biological and adoptive parents are included in the 'other two parent family' category.

Table 4: Estimates Given Different Assumptions on the Correlation between Selection and University Participation

	Assumed correlation of disturbances			
	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$
Neighbourhood BA Share	0.163 (0.057)***	0.136 (0.052)***	0.111 (0.051)**	0.086 (0.051)*
	$\rho = 0.4$	$\rho = 0.5$	$\rho = 0.6$	$\rho = 0.7$
Neighbourhood BA Share	0.061 (0.051)	0.036 (0.051)	0.011 (0.051)	-0.014 (0.051)
Sample size	15,034			

Marginal effects estimated in a two-step Probit model. The first stage regresses neighbourhood BA share on all the control variables and school fixed effects. The estimated residual from the first stage is included as a regressor in the second stage Probit. The coefficient on the estimated residual is the imposed value for the correlation coefficient.
Standard errors clustered by school presented in parenthesis
Stars indicate statistical significance: *** .01, ** .05, * .1

Table 5: Neighbourhood Heterogeneity

	(1A)		(1B)
Constant	-2.174 (0.212)***	Neighbourhood BA Share	-0.776 (1.768)
Female	0.442 (0.045)***	Neighbourhood BA Share * Female	-0.360 (0.253)
Parents' highest educational attainment– Reference both have less than high school			
High school	0.353 (0.103)***	Neighbourhood BA Share * High school	-0.439 (0.774)
PSE below BA	0.568 (0.098)***	Neighbourhood BA Share * PSE below BA	-1.250 (0.739)*
One parent has a BA	0.937 (0.111)***	Neighbourhood BA Share *One parent has a BA	-1.220 (0.781)
Both parents have a BA	1.369 (0.138)***	Neighbourhood BA Share*Both parents have a BA	-1.836 (0.824)**
Family income quartiles– Reference bottom quartile			
Quartile 2	0.027 (0.066)	Neighbourhood BA Share*Quartile 2	-0.659 (0.444)
Quartile 3	0.030 (0.070)	Neighbourhood BA Share*Quartile 3	-0.659 (0.444)
Quartile 4	0.138 (0.074)*	Neighbourhood BA Share*Quartile 4	-0.659 (0.444)
PISA reading quartiles– Reference bottom quartile			
Quartile 2	0.572 (0.076)***	Neighbourhood BA Share*Quartile 2	-0.034 (0.464)
Quartile 3	1.115 (0.076)***	Neighbourhood BA Share*Quartile 3	-0.011 (0.462)
Quartile 4	1.763 (0.085)***	Neighbourhood BA Share*Quartile 4	-0.302 (0.519)
Parental aspirations for child– Reference high school or less			
PSE below BA	-0.009 (0.192)	Neighbourhood BA Share*PSE below BA	3.312 (1.708)*
BA	0.706 (0.189)***	Neighbourhood BA Share*BA	2.964 (1.731)*
Any level above high school	0.280 (0.197)	Neighbourhood BA Share*Any level above high school	3.553 (1.764)**
Sample size	15,034		

Coefficients estimated in a Probit model. Dependent variable = 1 if child ever attended a university
Includes school fixed effects as well as all family background and behaviour controls
Standard errors clustered by school presented in parenthesis
Stars indicate statistical significance: *** .01, ** .05, * .1

Table 6: Marginal Effects Across the Distributions of Reading Skills and Family Background

	SES Type			
	Low	Medium High School	Medium PSE	High
PISA reading quartile	1	2	3	4
Predicted value	0.027	0.143	0.308	0.930
Marginal effect	-0.056 (0.146)	0.218 (0.118)*	0.242 (0.146)*	-0.060 (0.084)
Sample size	42	120	185	449
PISA reading quartile	1	1	1	1
Predicted value	0.027	0.063	0.082	0.589
Marginal effect	-0.056 (0.146)	0.137 (0.091)	0.112 (0.099)	-0.038 (0.221)
Sample size	42	134	229	36
PISA reading quartile	2	2	2	2
Predicted value	0.078	0.143	0.184	0.695
Marginal effect	-0.131 (0.289)	0.218 (0.118)*	0.173 (0.124)	-0.057 (0.176)
Sample size	18	120	199	88
PISA reading quartile	3 & 4	3	3	3
Predicted value	0.076	0.287	0.308	0.805
Marginal effect	-0.141 (0.273)	0.345 (0.163)**	0.242 (0.146)*	-0.043 (0.147)
Sample size	12	80	185	194
PISA reading quartile		4	4	4
Predicted value		0.439	0.497	0.930
Marginal effect		0.338 (0.215)	0.168 (0.164)	-0.060 (0.084)
Sample size		61	97	449

Marginal effects estimated using coefficients from model reported in Table 5
SES types are defined as follows:
Low–parents' education is below high school, bottom family income quartile, parents hope child achieves high school or less
Medium High School–highest parental education is high school, second family income quartile, parents hope child achieves college or above high school
Medium PSE–parents' highest education is PSE below BA, third family income quartile, parents hope child achieves college or above high school
High–both parents have a BA, top family income quartile, parents hope child achieves a BA
Standard errors clustered by school presented in parenthesis
Stars indicate statistical significance: *** .01, ** .05, * .1