# Neighbourhood Measures: Quantifying the Effects of Neighbourhood Externalities*

**Ben Jensen**
**Melbourne Institute of Applied Economic and Social Research**
**The University of Melbourne**
**and**
**Mark N. Harris**
**Monash University**

**Melbourne Institute Working Paper No. 4/03**

**ISSN 1328-4991 (Print)**

**ISSN 1447-5863 (Online)**

**ISBN 0 7340 3116 5**

**February 2003**

*The authors would like to thank Lisa Farrell, Carli Cochi Ficano, Tim Fry, and Robert Haveman for their helpful comments. The authors would also like to thank the Ronald Henderson Foundation for their support in this research.

**Melbourne Institute of Applied Economic and Social Research**
**The University of Melbourne**
**Victoria 3010 Australia**
***Telephone* (03) 8344 3701**
***Fax* (03) 8344 5630**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

**Abstract**

In recent years, analyses of neighbourhood externalities have grown with the perceived importance of their influence upon outcomes. Despite this growth, a clear understanding of the role of neighbourhoods in determining outcomes remains elusive. Various attempts have been made to quantify the role of neighbourhoods and limit problems of misspecification that have plagued this literature. Recent research suggests that neighbourhood proxies that measure characteristics similar to the dependent variable may better capture neighbourhood externalities. We explore variation in estimations including distinct neighbourhood proxies by estimating the influence of neighbourhood externalities upon youths' education expectations. Misspecification tests for normality and heteroscedasticity show particular neighbourhood proxies are more susceptible to misspecification. Monte-Carlo experiments show these neighbourhood proxies are also more likely to produce biased estimates if particular family characteristics are not fully captured. We find estimations including neighbourhood proxies measuring characteristics proximate to youths' education are less likely to suffer misspecifications. We also find that different geographic definitions of neighbourhoods can lead to erroneous findings, particularly considering variation in school quality.

## 1. Introduction

In recent years, analyses of neighbourhood externalities have grown with the perceived importance of their influence upon socio-economic outcomes (Wilson 1987; Corcoran and et al. 1990; Case and Katz 1991; Crane 1991; Borjas 1995; Mayer 1996; O'Regan and Quigley 1998). Despite this growth, a clear understanding of the role of neighbourhoods in determining outcomes remains elusive (Ginther, Haveman and Wolfe 2000). As yet, there exists no consistent findings on the impact of neighbourhood externalities upon youths' outcomes. While some research has found neighbourhoods to play a significantly large role in determining outcomes, other research has found the influence to be insignificant (Brooks-Gunn, Duncan, Klebanov and Sealand 1993; Plotnick and Hoffman 1999). Variation in estimation method, the types of neighbourhood proxies and diagnostic testing have hindered progress towards commonly accepted heuristics in neighbourhood research.

There is some evidence that misspecification in the estimation of neighbourhood externalities varies with particular neighbourhood proxies (Manski 1993; Duncan, Connell and Klebanov 1997; Ginther, Haveman and Wolfe 2000). This paper builds upon this research by including different neighbourhood proxies in estimations of probit models of youths' education expectations. We employ several methods in an attempt to accurately quantify the importance of neighbourhoods and thereby explain variation in previous research. First, standard probit models provide initial estimations of the significance of 11 separate neighbourhood characteristics. Second, diagnostic testing for normality and heteroscadicity reveal variation in heteroscedasticity across estimations including distinct neighbourhood characteristics. Third, a heteroscedastic-probit model provides more accurate estimations of neighbourhood effects. Fourth, implied probabilities calculated from both the standard probit and the heteroscedastic-probit identify variation among findings with different neighbourhood characteristics. Fifth, two sets of Monte-Carlo experiments are presented to illustrate the role of omitted variables in estimations including distinct neighbourhood characteristics. Variation in estimations of neighbourhood characteristics is further analysed through the use of two sets of neighbourhood characteristics. Neighbourhood characteristics of students' home neighbourhoods are compared with characteristics of the neighbourhood where their school is located. This provides a form of sensitivity analysis and offers some indication of the effects of unobserved school characteristics.

The plan of this paper is as follows. Section 2 includes a brief summary of the previous literature analysing neighbourhood externalities. Differences in these findings highlight the need for a more consistent approach to the analysis of neighbourhood externalities. In Section 3, the model and data utilised for the analysis is detailed. Attention is paid to the use of youths' expectations as the dependent variable. We consider expectations to be central to human capital investments. Previous research has shown that youths' expectations should be considered rational judgements upon which they form decisions of their appropriate human capital investments (Dominitz and Manski 1996). The empirical results are presented in Section 4. Section 5 concludes by examining the findings here in the context of previous research and how these findings can influence future neighbourhood research.

## 2. Neighbourhood Research and Problems of Misspecification

Economists' analyses of neighbourhood externalities have primarily followed the epidemic and collective socialisation models developed chiefly by Sociologists (Gans 1968; Moynihan 1968; Wilson 1980). Epidemic models emphasise the effects of peer interaction. The models estimate the likelihood of an individual undertaking a given action as a positive function of the proportion of that individual's peers who undertake that action (Crane 1991; Evans, Oates and Schwab 1992). Collective socialisation models emphasise the importance of the actions and characteristics of adults from within a given neighbourhood. Adults act as role models who both create and enforce social norms that affect youths' behaviour. They are also the prime source of information flows that can aid both employment and education outcomes. The likelihood of a youth undertaking a given activity is influenced by the actions and norms of the adults within that neighbourhood (Wilson 1980; Wilson 1987; Wilson 1993; Wilson 1996).

Empirically, the characteristics of peers and neighbourhoods are utilised as proxies for neighbourhood externalities. Characteristics such as high-status employment are used to identify a 'good' neighbourhood where positive externalities would imply higher achieving outcomes upon youths. Conversely, low-status socio-economic characteristics that typify 'poor' neighbourhoods may influence youths in adverse ways and restrict their opportunities for social advancement (Borjas 1995). In essence, neighbourhood characteristics act as proxies for neighbourhood externalities that potentially affect socio-economic outcomes.

Substantial variation exists in findings of the influence of neighbourhood externalities. Neighbourhood externalities have been found to significantly influence youths' education outcomes (Crane 1991; Borjas 1995). Yet, the magnitude of the effect of Neighbourhoods has often been small and/or significant among only particular demographic groups (Datcher 1982; O'Regan and Quigley 1998). Several studies have found no significant neighbourhood effect (Brooks-Gunn, Duncan, Klebanov and Sealand 1993; Plotnick and Hoffman 1999). Other studies have found the influence of youths' peers to play a more important role than neighbourhood effects upon youths' outcomes (Case and Katz 1991; Evans, Oates and Scwab 1992).

Inconsistency in the findings of previous research stems from problems in correctly specifying neighbourhood externalities. Previous research has been unable to identify the most appropriate proxy of these externalities (Jencks and Mayer 1990). Numerous studies have found only particular neighbourhood proxies to significantly influence socio-economic outcomes (Corcoran and et al. 1990; Duncan, Connell and Klebanov 1997; Plotnick and Hoffman 1999). Difficulties in estimating neighbourhood externalities can occur if neighbourhood proxies are incomplete or endogenous to family characteristics, or if unmeasured family characteristics affect youths' outcomes, their residential location or choice of school. Recent research suggests that neighbourhood proxies that measure characteristics proximate to the dependant variable may better capture neighbourhood externalities. For example, a neighbourhood proxy that measures levels of education within a neighbourhood may provide a more accurate estimation of the influence of neighbourhood externalities upon the education of youths living in that neighbourhood (Ginther et al., 2000; Manki, 1995). Manski (1993) explored difficulties in identifying neighbourhood externalities in his analysis of what he termed the 'reflection problem'. This occurs with analysis of the 'whether the average behaviour in some group influences the behaviour of individuals that comprise the group' (Manski 1993, p. 532). Critical to overcoming this problem is the relationship between sets of independent variables and the outcome variable. Neighbourhood characteristics that are more related, or what Ginther et al. term 'proximate', to the outcome variable may overcome potential bias incurred with the reflection problem. However, this may differ with the extent that family background characteristics are included in the model estimation (Ginther, Haveman and Wolfe 2000). To explore differences in neighbourhood proxies, the model presented in section three includes several different neighbourhood proxies and identifies their appropriateness in various specifications.

## 3. Model and Data

A survey of 1207 secondary school students from forty metropolitan schools was utilised to estimate the influence of neighbourhoods upon youths' education expectations[1]. The survey was conducted at students' schools and obtained information concerning the individual, family, and neighbourhood. The response rate of individual students in schools was very high. Over 98 per cent of students who were asked to participate, completed the survey[2]. Participating students were still yet to complete the final three years of their secondary education. It is during this academic year that students choose the subjects they undertake in the final years of their secondary education that determine their entry to a post-secondary education institution.

Relative to other social sciences, analyses of youths' expectations are rare within the economics discipline (Dominitz and Manski 1996). Analyses of expectations can be greeted with scepticism if they conflict with actual outcomes. Yet, expectations are central to the human capital model as youths' invest in their human capital in the expectation of increased future earnings (Becker 1964). Analysis of education expectations requires that youths' expectations are rational and reliable indicators of youths' perceptions of their opportunities. A number of analyses have tested both the reliability and determinants of youths' expectations (Manski 1990; McClelland 1990; Dominitz and Manski 1996; Dominitz and Manski 1997; Wolter 2000; Reynolds and Pemberton 2001). These expectations have centred on youths' earnings expectations. Reliability of expectations of future earnings has been gauged with comparison to actual earnings within the economy. The earnings expectations of secondary school and college aged youths have been found to be consistent with actual wages offered in the labour market (Dominitz and Manski 1996; Dominitz and Manski 1997; Wolter 2000; Reynolds and Pemberton 2001). Expectations were considered rational as they reflected the available information of wages in the labor market. However, expectations will not perfectly match outcomes. This does not imply that the expectations are not rational but

---

[1] Nineteen of these schools were government schools, 12 were Catholic, and nine were independently operated. It was considered that this was largely representative of Schools in Melbourne (ABS, 1998).

[2] To ensure an appropriate cross-section of students were sampled from each school, students were surveyed in classes that were compulsory subjects. This negates the risk that only high or low performing students were surveyed at particular schools.

that unforseen events can alter expectations and outcomes. Conclusions drawn from analyses of expectations should therefore differ with analyses of outcomes. We consider youths' post-secondary education expectations an appropriate measure to identify the importance of neighbourhood externalities. They are central to youths' human capital investments that, in turn, possess clear implications for future inequality and social mobility.

A dichotomous variable (POSTSEC) identifying whether a student expects to undertake post-secondary education was utilised as the dependant variable (equal to 1 of the student expects to undertake post-secondary education and 0 otherwise). Just over 68 per cent of students expected to pursue post-secondary education, and twenty-two per cent of students expected to complete their secondary education but not undertake post-secondary education. The remaining 10 per cent expected to 'drop-out' before completing secondary school.

As is common in the literature (Fry, Brooks, Comley and Zhang 1993), an underlying random variable model was employed

$$y_i^* = x_i'\beta + u_i \tag{3.1}$$

where $y_i^*$ was the underlying index of the desirability of post schooling expectations, $x_i$ a vector of individual, family and neighbourhood characteristics with unknown weights $\beta$ and $u_i$ a random disturbance term assumed to follow a standard normal distribution (Maddala 1983). However, this latent variable is not observed. The observed realisation of this is

$$y_i = \begin{cases} 1 \text{ if } y_i^* > 0 \\ 0 \text{ if } y_i^* \le 0 \end{cases}. \tag{3.2}$$

The variables chosen to enter $x_i$ were a set of individual, family, and neighbourhood characteristics.[3] The variables include students' gender (MALE); a variable identifying recent migrant status of student's family (GENERATIONS/MIGRANT STATUS[4]); the

[3] The variables included in the estimation are detailed in Table A.1 in the Appendix.

[4] This is a dichotomous variable equal to 1 if the student is a first or second generation migrant. Unfortunately, the distribution of the migrant status of students is not representative of all of Melbourne. Caution must therefore be taken in drawing conclusions of findings of this variable.

number of books in the household (BOOKS); the sector of the student's school (NON-GOVERNMENT SCHOOL); a home ownership variable (OWN HOME); and the employment status and education level of students' parents. The data was largely representative of the Melbourne population (ABS 1996). Information could not be obtained concerning family income. Previous surveys have found difficulties in obtaining accurate measures of family income from secondary school students (Jensen and Seltzer 2000). Other family variables were included that should capture the influence of differences in family wealth and income upon youths' education expectations. We expect variables indicating family wealth and income to possess positive coefficients. Yet, we do not have definitive theoretical prior expectations for the sign of the gender variable. On the one hand, we would expect males to more heavily invest in education due to the greater rewards received as male incomes remain greater than female incomes in the labour market (Teese 2000). Conversely, female participation in higher education has exceeded the participation of males for over a decade (ABS 1998; ABS 1998).

Due to the multicollinearity of neighbourhood proxies, only one neighbourhood variable was included in each estimation. Therefore, a number of estimations were conducted to estimate the appropriateness of distinct neighbourhood proxies. Neighbourhood proxies were obtained from the 1996 census (ABS 1996). Two sets of neighbourhood proxies were utilised and are presented in Appendix Table 2 and Appendix Table 3. The first included measures from youths' home neighbourhoods and the second included measures from the neighbourhoods in which students' schools were located. While the differences between these two sets of measures is not large (the correlation between each measure of the students' home neighbourhoods and their school's neighbourhoods was statistically significant at the 1% level), differences may indicate sources of potential bias. Comparison of findings with these two sets of neighbourhood measures further extends the analysis of inconsistency of findings of neighbourhood externalities. It explores two issues that may lead to misspecification of estimations. Previous research has found that variation in school quality could influence expectations and possibly bias the estimation (Dearden, Ferri and Meghir 1997; Card and Krueger 1998; Hoxby 2000). The Victorian education system is funded at the state level rather than the local level as in the United States (Teese, 1999). Variability in the quality of government schools is therefore likely to be lower across the sample[5]. Private sector schools may have greater resources and therefore provide a better quality education. A variable identifying whether a student attended a government or a non-government school was included in the estimation to capture these effects. The geographic measurement of neighbourhoods may also influence findings of neighbourhood externalities. The sociological and psychological research underpinning much empirical analysis illustrates, among other issues, youths' interaction with their peers and adults within their neighbourhood. Geographic measures of these neighbourhoods have not been precisely detailed. Emphasis is placed upon issues of social isolation and exclusion from networks that can increase education and employment opportunities (Wilson 1987). Empirical estimations of the influence of neighbourhood externalities have not tackled this issue. Data constraints have yielded estimations that have included postcode, census collectors district, and local government area proxies of neighbourhoods (Datcher 1982; Wilson 1987; Case and Katz 1991; Crane 1991; Corcoran and et al. 1992; Borjas 1995; Mayer 1996; O'Regan and Quigley 1998). Incorporating two sets of neighbourhood measures does not attempt to solve the problem of imprecise geographic measures of neighbourhoods that influence youths. It is possible to argue that youths may be equally influenced by their home neighbourhood and the neighbourhood in which they spend their time at, and travelling to and from school. However, the use of two sets of neighbourhood measures facilitates analysis of potential sources of bias in estimations of neighbourhood externalities.

## 4. Results

Eleven estimations were conducted, each with a separate measure of students' home neighbourhoods. For clarity, we first present the results of four estimations in Table 1. This shows the consistency of results for non-neighbourhood variables across four estimations. These estimations include a neighbourhood unemployment measure (model 1); an occupation status measure (model 2); an income measure (model 3); and an education measure (model 4).

[5] We do not discount the possibility of school quality influencing students' expectations. Unfortunately, greater information of variability in school quality was not available.

**Table 1: probit estimation of education expectations**

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Male | -0.510*** | -0.513*** | -0.505*** | -0.508*** |
| | (0.096) | (0.096) | (0.095) | (0.096) |
| Generations/Migrant status | 0.227** | 0.272*** | 0.215** | 0.248** |
| | (0.100) | (0.100) | (0.100) | (0.099) |
| Books | 0.186*** | 0.172*** | 0.187*** | 0.175*** |
| | (0.043) | (0.043) | (0.043) | (0.043) |
| Non – government school | 0.647*** | 0.525*** | 0.652*** | 0.525*** |
| | (0.098) | (0.103) | (0.098) | (0.103) |
| Own home | 0.15 | 0.163 | 0.152 | 0.158 |
| | (0.136) | (0.137) | (0.136) | (0.137) |
| Father employed | 0.154 | 0.129 | 0.164 | 0.14 |
| | (0.158) | (0.157) | (0.158) | (0.157) |
| Mother employed | 0.239* | 0.217* | 0.245** | 0.228* |
| | (0.123) | (0.124) | (0.123) | (0.123) |
| Father – university educated | 0.254** | 0.211 | 0.264** | 0.214** |
| | (0.128) | (0.129) | (0.128) | (0.129) |
| Father – less than secondary | 0.129 | 0.142 | 0.126 | 0.136 |
| | (0.137) | (0.137) | (0.137) | (0.137) |
| Mother – university educated | 0.303** | 0.231 | 0.305** | 0.232 |
| | (0.140) | (0.143) | (0.139) | (0.142) |
| Mother – less than secondary | 0.16 | 0.181 | 0.157 | 0.178 |
| | (0.142) | (0.143) | (0.142) | (0.143) |
| Neighbourhood variable | -0.007 | 0.027*** | -0.001 | 0.018*** |
| | (0.014) | (0.007) | (0.008) | (0.005) |
| Constant | -1.160*** | -1.50*** | -1.224*** | -1.520*** |
| | (0.340) | (0.305) | (0.370) | (0.307) |
| | | | | |
| Chi - Squared | 144.87 | 161.22 | 144.61 | 159.47 |
| P > Chi2 | 0.000 | 0.000 | 0.000 | 0.000 |
| R-Squared | 0.131 | 0.146 | 0.131 | 0.144 |
| Log Likelihood | -480.930 | -472.753 | -481.057 | -473.628 |
| N | 894 | 894 | 894 | 894 |

Standard errors in parentheses with p<0.10 = *, p<0.05 = **, p<0.01 = ***.
Note: Separate neighbourhood characteristics were included in each estimation. Neighbourhood unemployment was included in the first estimation; the percentage of employed individuals over the age of 15 years in professional employment in the second estimation ; median individual income in the third; and the percentage of individuals over the age of 15 years with an education of at least a bachelor degree in the fourth.

Individual, family, and neighbourhood variables were found to significantly influence youths' post-secondary education expectations. These variables were significant across estimations. Female students were more likely to possess higher education expectations. Students who were first or second generation migrants also possessed higher expectations. The variable BOOKS, representing the number of books and frequency of reading in youths' homes, was also found to be significant. Of the family characteristics, maternal employment exhibited a significant and positive influence upon education expectations. University educated parents also possessed a positive influence. Importantly, there was little variation in the coefficients of these variables across estimations including distinct neighbourhood proxies. Similar findings existed for estimations including student's school postcode. While some family variables were insignificant, we do not omit them from the estimation as we have a priori expectations of their importance in determining youths' education, and we are concerned it may overestimate the importance of the neighbourhood characteristics.

Variability in the significance of neighbourhood proxies was evident across the four estimations. Significant neighbourhood effects were found with neighbourhood measures of education and occupation status. As expected, the direction of the coefficient indicated neighbourhoods with adults possessing higher levels of education and occupation status exhibited a positive influence on youths' education expectations. Yet, neighbourhood unemployment and income measures were not significant. Moreover, the coefficient of the neighbourhood income was small and negative when the expected outcome had been a positive coefficient. As detailed in section two, this variation in findings of neighbourhood effects is not uncommon. Further estimations were conducted to explore differences in neighbourhood proxies and problems of misspecification. We present results of the significance of neighbourhood proxies from the two sets of neighbourhood measures and the diagnostic testing of these estimations.

### *4.1. Specification Tests*

Misspecification problems have plagued analyses of neighbourhood externalities (Mayer 1996). An important assumption underlying maximum likelihood estimation of the probit model, is that the functional form for the disturbance term is an *i.i.d.* standard normal variate. Unlike the normal regression setting, if the disturbances are non-normal, or are heteroscedastic, maximum likelihood estimators are inconsistent and the covariance matrix is inappropriate (Yatchew and Grilliches 1985). Thus, it is imperative to test for these.

Following Pagan and Vella (1989) for example, tests were conducted for both normality and heteroscedasticity (Pagan and Vella 1989). A RESET-type test for normality is given by a joint test of significance of the variables $\left(x_i'\hat{\beta}\right)^2$ and $\left(x_i'\hat{\beta}\right)^3$ included in addition to $x_i$ in an auxiliary probit regression. A test for heteroscedasticity can be based on the joint significance of the variable used in an auxiliary regression given by a test of significance of the additional variables $\left(x_i'\hat{\beta}\right)z_i$, where the variables $z_i$ are those thought to influence the heteroscedasticity. To test variation among neighbourhood proxies, 11 neighbourhood measures were chosen to enter into $z_i$. The 11 neighbourhood measures were: the neighbourhood unemployment rate; the unemployment rate for youths' aged 15-19 years; the unemployment rate for youths aged 20–24 years; the median individual income; the median household income; the percentage of adults in the neighbourhood with an individual income in the bottom 10 per cent of income earners in Metropolitan Melbourne; the percentage of adults in the neighbourhood with an individual income in the top 10 per cent of income earners in Metropolitan Melbourne; the percentage of professionally employed adults; the percentage of adults with an education level of at least a Bachelor degree; the percentage of adults with an education level of a higher degree (above Bachelor level); and the percentage of adults who ceased their education at, or below the age of 16 years. Estimations of these neighbourhood measures allow comparison of findings with distinct neighbourhood measures. Previous research has shown greater statistical significance may be found in measures proximate to the dependent variable (Manski 1993; Ginther, Haveman and Wolfe 2000). We consider neighbourhood measures of education and professional occupation status to be proximate to the dependent variable of post-secondary education expectations.

The significance levels and diagnostic testing results for each neighbourhood measure of students' home neighbourhoods are presented in Table 2[6]. The findings of estimations including students' school neighbourhoods are displayed in Table 3. In Table 2, all neighbourhood measures of education and occupation status were found to be highly significant. Conversely, income and unemployment measures were not significant. The one caveat to this was the proportion of high-income earners in students' neighbourhoods. This finding is similar to previous research that explored the significance of high-socio-economic status adults in a neighbourhood (Brooks-Gunn, Duncan, Klebanov and Sealand 1993). All but two of the neighbourhood proxies from students' school neighbourhoods (presented in Table 3) were found to be significant. The increased significance of these variables may be due to variation in school quality. Although the variation in socio-economic characteristics between students' home and school neighbourhoods is not large, variation in school quality

[6] The complete results of each estimation are available from the authors.

may cause misspecification of particular neighbourhood variables and lead to erroneous conclusions of the importance of neighbourhoods in youths' education.

**Table 2: Normality and Heteroscedasticity Test Results: Students' home neighbourhoods**

| | Probit coefficients | Normality | | | Heteroscedasticity | | |
|---|---|---|---|---|---|---|---|
| | | Chi-Square | df | p-value | Chi-Square | df | p-value |
| **Unemployment** | | | | | | | |
| Unemployment | -.0073 | 0.891 | 2 | 0.641 | 3.15 | 1 | 0.0759 |
| Unemployment: 15-19 y.o. | -.0031 | 0.941 | 2 | 0.625 | 2.87 | 1 | 0.0903 |
| Unemployment: 20-24 y.o. | -.0036 | 0.949 | 2 | 0.622 | 3.21 | 1 | 0.0732 |
| **Income** | | | | | | | |
| Individual Income | -.0000 | 1.06 | 2 | 0.588 | 1.90 | 1 | 0.168 |
| Household Income | -.0003 | 1.27 | 2 | 0.531 | 2.66 | 1 | 0.103 |
| Low-income earners (%) | -.0415 | 0.945 | 2 | 0.624 | 0.0618 | 1 | 0.804 |
| High-income earners (%) | .0236** | 0.448 | 2 | 0.799 | 1.47 | 1 | 0.225 |
| **Education** | | | | | | | |
| Bachelor degree + | .0178*** | 0.515 | 2 | 0.773 | 0.00263 | 1 | 0.959 |
| Higher degree | .0824*** | 0.725 | 2 | 0.696 | 0.0615 | 1 | 0.804 |
| Education < 16 y.o. | -.0511*** | 1.09 | 2 | 0.579 | 0.0860 | 1 | 0.769 |
| **Occupation** | | | | | | | |
| Professionals | .0265*** | 0.260 | 2 | 0.878 | 0.823 | 1 | 0.364 |

Note: p<0.10 = *, p<0.05 = **, p<0.01 = ***.

**Table 3 Normality and Heteroscedasticity Test Results: Students' school neighbourhoods**

| | Probit coefficients | Normality | | | Heteroscedasticity | | |
|---|---|---|---|---|---|---|---|
| | | Chi-Square | df | p-value | Chi-Square | df | p-value |
| **Unemployment** | | | | | | | |
| Unemployment | -.0319** | 0.84 | 2 | 0.657 | 5.41 | 1 | 0.020 |
| Unemployment: 15-19 y.o. | -.021** | 1.20 | 2 | 0.548 | 4.59 | 1 | 0.032 |
| Unemployment: 20-24 y.o. | -.0186 | 1.02 | 2 | 0.599 | 2.61 | 1 | 0.106 |
| **Income** | | | | | | | |
| Individual Income | .0017** | 1.90 | 2 | 0.387 | 7.06 | 1 | 0.008 |
| Household Income | .0001* | 1.05 | 2 | 0.592 | 5.94 | 1 | 0.015 |
| Low-income earners (%) | -.0335 | 0.728 | 2 | 0.695 | .0005 | 1 | 0.982 |
| High-income earners (%) | .031*** | 3.47 | 2 | 0.176 | 3.99 | 1 | 0.046 |
| **Education** | | | | | | | |
| Bachelor degree + | .0182*** | 2.74 | 2 | 0.255 | 0.903 | 1 | 0.342 |
| Higher degree | .0835*** | 1.84 | 2 | 0.399 | 0.291 | 1 | 0.590 |
| Education < 16 y.o | -.0437*** | 1.0 | 2 | 0.605 | 1.77 | 1 | 0.183 |
| **Occupation** | | | | | | | |
| Professionals | .0247*** | 1.54 | 2 | 0.462 | 1.04 | 1 | 0.308 |

Note: p<0.10 = *, p<0.05 = **, p<0.01 = ***.

Variation was evident in results of diagnostic testing of estimations including distinct neighbourhood proxies. In estimations including each of the 11 proxies of students' home neighbourhoods the null-hypothesis of normality cannot be rejected. However, the results for heteroscedasticity differ across each neighbourhood measure. Estimations including neighbourhood unemployment measures appear to be susceptible to heteroscedasticity. While they are significant at the 10 per cent level, the Chi-square statistic for the neighbourhood unemployment measures are above the 5 per cent critical value. Estimations including neighbourhood education, occupation, and income measures appear unaffected by heteroscedasticity (although the chi-square statistic for the neighbourhood household income measure is very close to the 5 per cent critical value). Diagnostic testing of estimations with students' school neighbourhoods indicate greater misspecifications (Table 3). With the exception of the second youth unemployment measure (Unemployment: 20-24 y.o.) and the measure of low-income earners (Low-income earners (%)), the neighbourhood unemployment and income measures were found to suffer from heteroscedasticity.

In light of the heteroscedasticity results, heteroscedastic-probit estimations were conducted for models that failed the diagnostic test of heteroscedasticity. Estimations included unemployment measures of student's home neighbourhoods and four of the six unemployment and income measures of students' school neighbourhoods. Following (Harvey 1976), the general form of heteroscedasticity was assumed to be

$$Var(\varepsilon) = \left[\exp(z_i'\gamma)\right]^2. \tag{4.1}$$

Again, the variables chosen to enter $z_i$ were the respective neighbourhood measures. The probability that $Y = 1$ is now accordingly

$$\Pr(Y = 1) = \Phi\left[x_i'\beta / \exp(z_i'\gamma)\right] \tag{4.2}$$

where $\Phi(.)$ and $\phi(.)$ - used later - are the distribution and density functions, respectively, of the standard normal distribution. The log-likelihood is

$$\begin{aligned} \ln L &= \sum_i \ln \Phi(a_i) \\ a_i &= (2y_i - 1)x_i'\beta \times \exp(-z_i'\gamma). \end{aligned} \tag{4.3}$$

The marginal effects, that is the change in the probability for a small change in one of the explanatory variables, are given by

$$\frac{\partial \Pr(Y=1)}{\partial x_k} = \phi\left[\frac{x_i'\beta}{\exp(z_i'\gamma)}\right] \times \frac{\beta_k}{\exp(z_i'\gamma)} \tag{4.4}$$

which are evaluated at the MLE estimates of $\beta$ and $\gamma$. However, if $z_i$ contains $x_i$, as is the case here, (for example, the neighbourhood unemployment rate influence both the mean and the variance of the probit regression) equation (4.4) is more complex

$$\frac{\partial \Pr(y_i=1)}{\partial x_k} = \phi\left[\frac{x_i'\beta}{\exp(z_i'\gamma)}\right] \times \left[\frac{\beta_k - (x_i'\beta)\gamma_k}{\exp(z_i'\gamma)}\right].k\ . \tag{4.5}$$

Table 4 presents the results for the heteroscedastic-probit estimation of youths' post-secondary education expectations. Three estimations are presented to provide a more accurate estimation of characteristics influencing education expectations. The three estimations include the three unemployment measures of students' home neighbourhoods as the neighbourhood variable. These are the neighbourhood unemployment rate (model 1), the neighbourhood youth unemployment rate of 15-19 y.o. (model 2), the neighbourhood youth unemployment rate of 20-14 y.o. (model 3). These estimations provide comparison with standard probit estimations presented in Table 1.

**Table 4: Heteroscedastic probit estimation results**

| | 1 | 2 | 3 |
|---|---|---|---|
| Male | -0.84*** | -0.811*** | -0.808 |
| | (0.271) | (0.278) | (0.294) |
| Generations/ Migrant status | 0.375* | 0.347* | 0.352* |
| | (0.192) | (0.191) | (0.198) |
| Books | 0.294*** | 0.28** | 0.288** |
| | (0.106) | (0.110) | (0.113) |
| Non – government school | 0.937*** | 0.941*** | 0.942*** |
| | (0.288) | (0.310) | (0.320) |
| Own home | 0.244 | 0.254 | 0.243 |
| | (0.231) | (0.224) | (0.228) |
| Father employed | 0.205 | 0.216 | 0.2 |
| | (0.267) | (0.253) | (0.261) |
| Mother employed | 0.356* | 0.354* | 0.366* |
| | (0.214) | (0.212) | (0.218) |
| Father – university educated | 0.361* | 0.36* | 0.373* |
| | (0.211) | (0.217) | (0.217) |
| Father – less than secondary | 0.236 | 0.198 | 0.236 |
| | (0.233) | (0.226) | (0.233) |
| Mother – university educated | 0.417* | 0.418* | 0.433* |
| | (0.228) | (0.225) | (0.242) |
| Mother – less than secondary | 0.204 | 0.216 | 0.212 |
| | (0.224) | (0.223) | (0.225) |
| Neighbourhood – mean | 0.0148 | 0.00551 | 0.00796 |
| | (0.039) | (0.020) | (0.026) |
| Neighbourhood - variance | 0.046 | 0.0214 | 0.0307 |
| | (0.031) | (0.016) | (0.024) |
| Constant | -1.99 | -1.92 | -1.95 |
| | (0.956) | (0.990) | (0.985) |

Standard errors in parentheses with $p<0.10$ = *, $p<0.05$ = **, $p<0.01$ = ***.
Note: A separate characteristic of students' home neighbourhood was included in each estimation. Neighbourhood unemployment was included in the first estimation; youth unemployment (15-19 y.o.) in the second; and a second measure of youth unemployment (20-24y.o.) in the third.

The structural coefficients of the heteroscedastic-probit model are interpreted as before, while those in $z_i$ are interpreted as affecting the variance of the equation. However, in calculating the total marginal effects of variables that are both in the mean and variance equations, it is important to realise these have a joint effect (although their effect on the mean is quantitatively small). The findings of non-neighbourhood variables largely follow the probit estimation with a variety of individual and family characteristics found to be significant.

### *4.2. Implied Probabilities*

The implied probabilities calculated from each neighbourhood measure illustrate variation in findings between neighbourhood measures and highlights the importance of comprehensive diagnostic testing. The implied probabilities of both the standard and heteroscedastic-probit model are data dependent. These are evaluated for the mean student, with all characteristics held at the sample means. Implied probabilities were calculated for five levels of each neighbourhood measure. These are the minimum, maximum, mean, and one standard deviation above and below the mean value for each neighbourhood measure (Table A.2). The implied probabilities of the mean student expecting to attend post-secondary education at five values of each neighbourhood measure are presented in Table 5. These are calculated with measures of students' home neighbourhoods. Implied probabilities calculated with students' school neighbourhood are presented in Table 6. The probabilities calculated for estimations including neighbourhood measures found to possess heteroscedasticity in the standard probit model were then calculated using the coefficients from the heteroscedastic-probit model. These probabilities are calculated following equations (4.4) and (4.5).

**Table 5: Probabilities of mean students' Expectations: Student's home neighbourhoods**

| | Min | Mean - SD | Mean | Mean + SD | Max |
|---|---|---|---|---|---|
| *Standard probit* | | | | | |
| Bachelor degree + | 0.594 | 0.634 | 0.712 | 0.780 | 0.905 |
| Education < 16y.o. | 0.827 | 0.767 | 0.711 | 0.670 | 0.568 |
| Higher degree | 0.633 | 0.646 | 0.714 | 0.775 | 0.918 |
| Professionals | 0.601 | 0.632 | 0.713 | 0.785 | 0.896 |
| Low-income earners (%) | 0.786 | 0.734 | 0.709 | 0.683 | 0.628 |
| High-income earners (%) | 0.626 | 0.661 | 0.708 | 0.751 | 0.851 |
| Individual income (median) | 0.713 | 0.712 | 0.710 | 0.708 | 0.704 |
| Household income (median) | 0.743 | 0.723 | 0.708 | 0.693 | 0.652 |
| *Heteroscedastic-probit* | | | | | |
| Unemployment rate | 0.733 | 0.724 | 0.705 | 0.687 | 0.621 |
| Unemployment: 15-19 y.o. | 0.759 | 0.721 | 0.705 | 0.690 | 0.644 |
| Unemployment: 20-24 y.o. | 0.764 | 0.724 | 0.707 | 0.690 | 0.641 |

Note: Probabilities for the first eight characteristics were calculated using coefficients obtained in the standard probit estimation. The other three neighbourhood measures utilised coefficients form the heteroscedastic-probit estimation.

**Table 6: Probabilities of mean students' Expectations: Student's school neighbourhoods**

| | Min | Mean - SD | Mean | Mean + SD | Max |
|---|---|---|---|---|---|
| *Standard probit* | | | | | |
| Bachelor degree + | 0.593 | 0.624 | 0.712 | 0.788 | 0.857 |
| Education < 16y.o. | 0.780 | 0.755 | 0.709 | 0.670 | 0.624 |
| Higher degree | 0.635 | 0.631 | 0.715 | 0.788 | 0.886 |
| Professionals | 0.589 | 0.631 | 0.712 | 0.784 | 0.836 |
| Low-income earners (%) | 0.771 | 0.730 | 0.709 | 0.688 | 0.672 |
| Unemployment: 20-24 y.o. | 0.748 | 0.732 | 0.708 | 0.684 | 0.576 |
| *Heteroscedastic-probit:* | | | | | |
| Unemployment rate | 0.776 | 0.753 | 0.699 | 0.655 | 0.543 |
| Unemployment: 15-19 y.o. | 0.794 | 0.752 | 0.704 | 0.663 | 0.587 |
| Individual income (median) | 0.615 | 0.652 | 0.699 | 0.756 | 0.874 |
| Household income (median) | 0.623 | 0.660 | 0.704 | 0.757 | 0.858 |
| High-income earners (%) | 0.614 | 0.641 | 0.697 | 0.767 | 0.911 |

Note: Probabilities for the first six characteristics were calculated using coefficients obtained in the standard probit estimation. The other five neighbourhood measures utilised coefficients form the heteroscedastic-probit estimation.

Considerable variation was found between implied probabilities calculated with each neighbourhood measure. Differences were greater with neighbourhood measures of students' home neighbourhoods than their school neighbourhoods. The one caveat to this was the measure of high-income earners in students' home neighbourhoods. The ranges of probabilities evident at each level of the neighbourhood measures included in the heteroscedastic-probit estimations were consistently smaller than those evident in the probit estimations. For example, probabilities obtained with the neighbourhood education measure (Bachelor degree +) ranged from 0.594 to 0.905, while the neighbourhood unemployment measure had a substantially smaller probability range (0.621 - 0.733). Yet, without diagnostic testing the range of reported implied probabilities would have been smaller. Standard probit estimations found to possess heteroscedasticity yielded a smaller range of implied probabilities than those calculated with the more accurate heteroscedastic-probit estimations. Heteroscedastic-probit estimations provide a more accurate measure of neighbourhood externalities and reduce the variation in the quantitative importance of various neighbourhood characteristics. This further illustrates the need for comprehensive diagnostic testing to identify the influence of neighbourhood externalities.

Support was found for the importance of neighbourhood measures proximate to the dependent variable. Larger ranges of implied probabilities were found for neighbourhood education and professional occupation status measures. The education levels of individuals in a neighbourhood were found to have a large effect upon the probability of students expecting to attend post-secondary education. The mean student living in a neighbourhood where 25.02 per cent of the population have at least a bachelor degree (mean level) was 8 per cent more likely to expect to undertake post-secondary education than the mean student who lived in a neighbourhood where the proportion of bachelor degree holders dropped to 12.9 per cent (mean minus standard deviation). This disparity clearly grows for students from neighbourhoods with greater disparities of education levels. Similar probabilities were found with the proportion of professionals in the sample neighbourhoods. Neighbourhood measures of low and high levels of education possess extremes in the variation in the probabilities associated with the mean student's education expectations. Neighbourhoods with a high proportion of individuals with low levels of education impacted upon youths such that the probability of the mean student expecting to attend post-secondary education falls substantially. Conversely, neighbourhoods with a high proportion of individuals who had obtained higher degrees in education had a large positive effect upon the probability of students expecting to attend post-secondary education.

### 4.3. Analysing the Likely Extent of Omitted Variable Bias

To analyse the likely impact of omitted variable bias a limited set of Monte Carlo experiments was undertaken. These experiments estimate the likely impact of omitting variables upon remaining coefficients. For example, omitting specific family characteristics in a Monte Carlo experiments allows an estimation of the extent of upward bias incurred by various neighbourhood characteristics. Specifically, the estimated $\beta$'s from the full set of explanatory variables are taken as the "true" values, and $y_i^*$ and $y_i$ generated according to equations (3.1) and (3.2), and where the $u_i$ are random draws from the standard normal distribution. In subsequent estimations, particular independent variables are purposely excluded and the remaining estimated coefficients are compared with the "true" ones. This process was repeated 500 times and average parameter values collected. In this way, it is possible we can ascertain the extent of likely omitted variable bias and how this bias varies among various neighbourhood characteristics.

Six models were used that estimated the effect of omitted variable bias upon neighbourhood measures. The six models follow previous research illustrating the degree of upward bias of neighbourhood externalities when excluding particular groups of variables (Ginther et al., 2000). The variables excluded in each model are described in Table 7. The first model contains all of the independent variables in the probit estimation presented in Table 1. Model 2 omits parental employment characteristics. Model 3 omits these characteristics and parental education characteristics. Model 4 provides contrast by omitting characteristics other than parental employment and education, and the neighbourhood measure. Model 5 omits all characteristics but gender and neighbourhood measure. Model 6 omits all characteristics but the neighbourhood measure. Only estimations including students' home neighbourhoods are presented. Little difference was found between estimations including students' home and school neighbourhoods in the Monte Carlo experiments.

**Table 7: Variables Included in Model Specification**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Male | X | X | X | | X | |
| Generations/Migrant Status | X | X | X | | | |
| Books | X | X | X | | | |
| Non-government | X | X | X | | | |
| Own home | X | X | X | | | |
| Father employed | X | | | X | | |
| Mother employed | X | | | X | | |
| Father – university educated | X | X | | X | | |
| Father – less than secondary education | X | X | | X | | |
| Mother – university educated | X | X | | X | | |
| Mother – less than secondary education | X | X | | X | | |
| Neighbourhood measure | X | X | X | X | X | X |

Note: The variables included in each model are denoted by X.

The six model estimations were simulated for each of the 11 neighbourhood characteristics. The parameter values of the neighbourhood measures in each of the six models are presented in Table 8. This details the extent of bias in different neighbourhood measures due to the exclusion of particular independent characteristics.

**Table 8: Average Bias in measures of students' home neighbourhoods**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Unemployment | | | | | | |
| Unemployment | -.007 | -.012 | -.018 | -.003 | -.025 | -.021 |
| Unemployment: 15-19 y.o. | -.003 | -.005 | -.009 | -.002 | -.011 | -.011 |
| Unemployment: 20-24 y.o. | -.004 | -.006 | -.010 | -.001 | -.016 | -.013 |
| Income | | | | | | |
| Individual Income | -.001 | .002 | .006 | -.010 | -.155 | -.129 |
| Household Income | -.003 | -.003 | -.002 | -.004 | .000 | -.001 |
| < 10% Individual Income | -.042 | -.040 | -.034 | -.048 | -.031 | -.034 |
| > 88% Individual Income | .024 | .027 | .032 | .032 | .047 | .043 |
| Education | | | | | | |
| Bachelor degree + | .018 | .019 | .021 | .025 | .031 | .029 |
| Higher degree | .082 | .087 | .100 | .125 | .152 | .147 |
| Education < 16 y.o. | -.051 | -.045 | -.048 | -.081 | -.075 | -.077 |
| Occupation | | | | | | |
| Professionals | .027 | .028 | .031 | .035 | .043 | .041 |

Omitting independent variables upwardly biased findings of most neighbourhood characteristics. As expected, extreme levels of bias were found in Models 5 and 6. Comparing the average parameter values in model 6 to the standard probit coefficients in model 1, greater bias was found for estimations including neighbourhood measures of unemployment that suffered from heteroscedasticity than with neighbourhood education measures (although the greatest variation was found for neighbourhood median income, both median income measures had an unexpected sign). The average parameter value in model 6 for the neighbourhood unemployment rate was nearly three times larger than the standard probit coefficient. The two youth unemployment rate measures were 3.5 times larger in model 6 than model 1. This confirms previous findings that estimates of neighbourhood externalities are susceptible to omitted variable bias (Ginther et al., 2000; Mayer, 1996).

Variation was also evident among changes in the average parameter values of neighbourhood characteristics across each model. Neighbourhood unemployment measures exhibited greater bias with the omission of parental employment and education characteristics in Model 3. This contrasted with the neighbourhood education and occupation characteristics. Greater bias was found in these estimations in Model 4 that omitted individual, school, and other family variables. Relatively little bias was found with these characteristics in Models 2 and 3 that omitted family education and employment characteristics. This may indicate that estimations of neighbourhood measures of education and professional occupation status, measures considered proximate to the dependent variable, are less susceptible to unobserved family characteristics.

An implicit belief appears to exist in discussions of omitted variable bias in estimations of neighbourhood externalities that neighbourhood measures are particularly susceptible to omitted variable bias. Certainly, variation in findings of previous research supports this belief. However, all independent variables can be susceptible to omitted variable bias. The range of average parameter values for each of the independent variables found in each of the six models are presented in Table 9. For clarity, we presented the ranges for only three neighbourhood measures: neighbourhood education, unemployment, and income. For each neighbourhood measure, the ‘true’ betas for each variable are presented. These are contrasted with the range of the average biases of the estimated coefficient across the different estimations.

**Table 9: Range of Average Bias in Estimations with Omitted Variables**

| | Education | | | Unemployment | | | Income | | |
|---|---|---|---|---|---|---|---|---|---|
| | *Probit* | Min | Max | *Probit* | Min | Max | *Probit* | Min | Max |
| Male | -0.508 | -0.522 | -0.456 | -0.510 | -0.524 | -0.420 | -0.505 | -1.670 | 0.154 |
| Generations/Migrant Status | 0.248 | 0.216 | 0.260 | 0.227 | 0.201 | 0.258 | 0.215 | -0.962 | 0.802 |
| Books | 0.175 | 0.187 | 0.192 | 0.186 | 0.191 | 0.211 | 0.187 | -0.960 | 0.771 |
| Non-government | 0.525 | 0.525 | 0.656 | 0.647 | 0.648 | 0.686 | 0.652 | -0.497 | 1.247 |
| Own home | 0.158 | 0.148 | 0.187 | 0.150 | 0.151 | 0.178 | 0.152 | -0.991 | 0.740 |
| Father employed | 0.140 | 0.167 | 0.186 | 0.154 | 0.169 | 0.227 | 0.164 | -0.983 | 0.794 |
| Mother employed | 0.228 | 0.208 | 0.256 | 0.239 | 0.229 | 0.252 | 0.245 | -0.910 | 0.806 |
| Father – university educated | 0.214 | 0.233 | 0.315 | 0.254 | 0.269 | 0.407 | 0.264 | -0.870 | 0.984 |
| Father – less than secondary | 0.136 | 0.126 | 0.143 | 0.129 | 0.126 | 0.134 | 0.126 | -1.026 | 0.687 |
| Mother – university educated | 0.232 | 0.234 | 0.304 | 0.303 | 0.304 | 0.341 | 0.305 | -0.847 | 0.900 |
| Mother – less than secondary | 0.178 | 0.170 | 0.186 | 0.160 | 0.149 | 0.167 | 0.157 | -1.006 | 0.728 |
| Neighbourhood measure | 0.018 | 0.019 | 0.031 | -0.007 | -0.025 | -0.003 | -0.001 | -1.154 | 0.565 |

Note: The neighbourhood measures included were the neighbourhood education measure of individuals over the age of 15 years with at least a bachelor degree, the neighbourhood unemployment rate, and the neighbourhood median individual income.

The Monte Carlo experiment shows that omitting particular variables biases the parameter values of independent variables[7]. Upward bias in the average parameter values of various family variables illustrates that omitted variables bias the coefficients of all variables not just neighbourhood measures. For example, the largest average parameter value of the neighbourhood education measure was 0.03 compared to the standard probit coefficient of 0.02. Bias of similar magnitude was evident for both paternal and maternal university education (0.32 compared to 0.21, and 0.30 compared to 0.23 respectively). A similar pattern emerges for the neighbourhood unemployment and income characteristics although the

[7] The complete results of each estimation are available from the authors.

variation in the average parameter values of these neighbourhood measures is larger. The potential bias among non-neighbourhood variables illustrates an aspect of analyses of youths' attainment that appears neglected. While omitted variable bias often distorts measures of the effect of neighbourhood externalities, this bias also distorts the coefficients of all characteristics that influence education attainment. Moreover, if neighbourhood externalities are important, findings of the influence of individual and family characteristics may be upwardly biased if neighbourhood measures have been omitted.

## 5. Conclusion

A lack of consistent findings of neighbourhood externalities has led to questions of their influence upon youths' socio-economic outcomes. This paper illustrated how findings of the significance of neighbourhood externalities can differ with different neighbourhood proxies. Neighbourhood measures of education levels and professional occupation status were considered proximate to the dependent variable. Neighbourhood proxies measuring characteristics proximate to the dependent variable were found to provide more consistent measures of the influence of neighbourhood externalities upon youths' education expectations. Larger ranges of implied probabilities were also found for neighbourhood education and professional occupation status measures. Extensive diagnostic testing showed that estimations including these neighbourhood measures were less susceptible to misspecification than other neighbourhood measures, particularly neighbourhood unemployment measures. Heteroscedastic-probit estimations showed that inference drawn from standard probit estimations could bias the reported effect of neighbourhood externalities.

Monte Carlo experiments showed that neighbourhood education and professional occupation measures were less susceptible to omitted variable bias, particularly with omitted variables identifying family characteristics. Variation in the effects of omitted variable bias across neighbourhood proxies may partly explain variation in previous findings. Yet, neighbourhood characteristics, particularly education and occupational status, were no less susceptible to omitted variable bias than particular family characteristics.

Differences were found between estimations including measures of students' home neighbourhoods and their school neighbourhoods. Standard probit estimations yielded significant findings among a greater number of students' school neighbourhood characteristics. Yet, these estimations were found to be more susceptible to misspecification following extensive diagnostic testing. If this difference is due to unobserved school characteristics, then it may be an important finding that provides a rationale for variation in previous research. Particularly, given difficulties in quantifying variation in school quality. Importantly, significant and quantitatively important neighbourhood externalities were evident with neighbourhood measures of characteristics proximate to the dependent variable in students' home neighbourhood.

## Appendix

**Appendix Table 1: Individual and Family Variables – Mean, Standard Deviation & Expected Sign**

| | Mean (Standard Deviation) | Expected Sign |
|---|---|---|
| Male | 0.497 | - |
| | (0.500) | |
| Generations/ Migrant Status | 0.465 | ? |
| | (0.499) | |
| Books in household | 5.261 | + |
| | (1.124) | |
| Non-government school | 0.404 | + |
| | (0.490) | |
| Own home | 0.864 | + |
| | (0.343) | |
| Father employed | 0.907 | + |
| | (0.290) | |
| Mother employed | 0.805 | + |
| | (0.396) | |
| Mother – University educated | 0.194 | + |
| | (0.396) | |
| Mother - Less than secondary education | 0.170 | - |
| | (0.376) | |
| Father - University educated | 0.260 | + |
| | (0.439) | |
| Father - Less than secondary education | 0.187 | - |
| | (0.390) | |

Note: Standard deviations are shown in parentheses.

**Appendix Table 2: Descriptive Statistics of student's home Neighbourhoods**

| | Minimum | Maximum | Mean | Standard Deviation | Mean - SD | Mean + SD |
|---|---|---|---|---|---|---|
| Unemployment rate | 4.21 | 27.76 | 9.34 | 3.56 | 5.78 | 12.90 |
| Unemployment rate - 15-19 y.o. | 0.00 | 44.12 | 18.93 | 5.75 | 13.18 | 24.69 |
| Unemployment rate - 20-24 y.o. | 0.00 | 32.52 | 13.57 | 4.33 | 9.24 | 17.89 |
| Median individual income | 17.95 | 54.95 | 31.88 | 6.63 | 25.24 | 38.51 |
| Median household income | 39.95 | 119.95 | 71.66 | 13.52 | 58.15 | 85.18 |
| Less than 10% of income | 4.29 | 15.50 | 10.10 | 1.79 | 8.32 | 11.89 |
| At least 90% of income | 1.30 | 31.85 | 10.87 | 5.50 | 5.37 | 16.37 |
| Professionals (%) | 5.35 | 43.24 | 16.97 | 8.56 | 8.42 | 25.53 |
| Bachelor degree or higher | 6.97 | 67.28 | 25.02 | 12.13 | 12.90 | 37.15 |
| Higher degree | 0.00 | 12.78 | 2.75 | 2.31 | 0.43 | 5.06 |
| Dropped out at or below 16 yo | 8.52 | 23.66 | 16.09 | 3.37 | 12.72 | 19.46 |

Note: The 11 neighbourhood measures include the neighbourhood unemployment rate; the unemployment rate for youths' aged 15-19 years; the unemployment rate for youths aged 20–24 years; the median individual income (calculated as the median level for all individuals in the labour force over the age of 15 years); the median household income (calculated for all households regardless of labour force participation); the percentage of individuals in the neighbourhood with an individual income in the bottom 10 per cent of income earners in Metropolitan Melbourne (calculated by identifying the bottom 10 per cent of weekly income earners of all labour force participants); the percentage of individuals in the neighbourhood with an individual income in the top 10 per cent of income earners in Metropolitan Melbourne (calculated by identifying the top 10 per cent of weekly income earners of all labour force participants); the percentage of professionally employed individuals (calculated from the percentage of all labour force participants over the age of 15 years); the percentage of individuals with an education level of at least a Bachelor degree (calculated as a percentage of all individuals over the age of 15 years); the percentage of individuals with an education level of a higher degree (calculated as a percentage of all individuals over the age of 15 years); and the percentage of individuals who ceased their education at, or below the age of 16 years (calculated as a percentage of all individuals over the age of 15 years).

**Appendix Table 3: Descriptive Statistics of student's school Neighbourhoods**

| | Minimum | Maximum | Mean | Standard Deviation | Mean - SD | Mean + SD |
|---|---|---|---|---|---|---|
| Unemployment rate | 4.63 | 27.76 | 9.06 | 3.20 | 5.86 | 12.26 |
| Unemployment rate - 15-19 y.o. | 9.33 | 38.03 | 18.68 | 5.25 | 13.43 | 23.93 |
| Unemployment rate - 20-24 y.o. | 7.01 | 32.52 | 13.38 | 3.78 | 9.60 | 17.16 |
| Median individual income | 179.50 | 549.50 | 336.09 | 77.07 | 259.02 | 413.17 |
| Median household income | 399.50 | 1249.50 | 768.39 | 181.65 | 586.74 | 950.04 |
| Less than 10% of income | 4.29 | 13.10 | 9.95 | 1.85 | 8.10 | 11.80 |
| At least 90% of income | 1.30 | 31.85 | 12.47 | 6.90 | 5.57 | 19.37 |
| Professionals (%) | 5.54 | 36.08 | 19.12 | 9.15 | 9.96 | 28.27 |
| Bachelor degree or higher | 10.02 | 55.79 | 27.89 | 13.29 | 14.59 | 41.18 |
| Higher degree | 0.69 | 10.97 | 3.35 | 2.78 | 0.57 | 6.13 |
| Dropped out at or below 16 yo | 24.41 | 68.87 | 45.56 | 9.86 | 35.71 | 55.42 |

**References**


ABS 1996. CDATA 96. Canberra, Australian Bureau of Statistics.

ABS (1998a). *Education and Training in Australia*. Canberra, Australian Bureau of Statistics.

ABS (1998b). *Schools Australia*. Canberra, Australian Bureau of Statistics.

Becker, Gary S. (1964). *Human Capital: A theoretical and empirical analysis with special reference to education*. New York, Columbia University Press.

Borjas, George 1995. "Ethnicity, neighbourhoods, and human-capital externalities." *American Economic Review* 85: 365-90.

Brooks-Gunn, Jeanne, Greg. J. Duncan, Pamela Klebanov and Naomi Sealand 1993. "Do Neighbourhoods Influence Child and Adolescent Development." *American Journal of Sociology* 99: 353-93.

Card, David and Alan B. Krueger 1998. "School Resources and Student Outcomes." *Annals of the American Academy of Political and Social Science* 559(0): 39-53.

Case, Anne C. and Lawrence F. Katz 1991. "The Company You Keep: The Effects of family and neighbourhood on disadvantaged youths." *Working Paper 3705, National Bureau of Economic Research, Cambridge, Massachusetts*.

Corcoran, Mary, Roger Gordon, Deborah Laren and Gary Solon 1990. "Effects of Family and Community Background on Economic Status." *American Economic Review* 80(2): 362-66.

Corcoran, Mary, Roger Gordon, Deborah Laren and Gary Solon 1992. "The Association between Men's Economic Status and Their Family and." *Journal of Human Resources* 27(4): 575-601.

Crane, Jonathan (1991a). Effects of Neighbourhoods on Dropping Out of School and Teenage Childbearing. *The Urban Underclass*. C. Jencks and P. E. Peterson. Washington, D.C., The Brookings Institute.

Crane, Jonathan 1991b. "The Epidemic Theory of Ghettos and Neighbourhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology* 96: 1226-59.

Datcher, Linda 1982. "Effects of Community and Family Background on Achievement." *Review of Economics and Statistics* 64(1): 32-41.

Dearden, Lorraine, Javier Ferri and Costas Meghir 1997. "The Effect of School Quality on Educational Attainment and Wages." *Institute for Fiscal Studies Working Paper* W98/03.

Dominitz, J. and C.F Manski 1996. "Eliciting Student Expectations of the Returns to Schooling." *Journal of Human Resources* 31(1): 1-26.

Dominitz, J. and C.F Manski 1997. "Perceptions of Economic Insecurity: Evidence from the Survey of Economic Expectations." *Public Opinion Quarterly* 61(2): 261-87.

Duncan, Greg J., James R. Connell and Pamela Klebanov (1997). Conceptual and Methodological Issues in Estimating Causal Effects of Neighbourhood and Family Conditions on Individual Development. *Neighbourhood Poverty: Context and Consequences for Children*. J. Brooks-Gunn, G.J. Duncan and L. Aber. New York, Russell Sage Foundation. 1.

Evans, William, Wallace Oates and Robert Schwab 1992. "Measuring Peer Group Effects: A Study of Teenage Behaviour." *Journal of Political Economy* 100(5): 966-991.

Fry, Timothy R.L., R.D. Brooks, B.R. Comley and J Zhang 1993. "Economic motivations for Limited Dependent and Qualitative Variable Models." *Economic Record* 69: 193-205.

Gans, Herbert (1968). Culture and Class in the Study of Poverty: An Approach to Antipoverty Research. *On Understanding Poverty: Perspectives from the Social Sciences*. Daniel P. Moynihan. New York, Basic Books.

Ginther, Donna, Robert Haveman and Barbara Wolfe 2000. "Neighbourhood attributes as determinants of children's outcomes: How robust are the relationships?" *Journal of Human Resources* 35(4): 603-42.

Harvey, Andrew C. 1976. "Estimating Regression Models with Multiplicative Heteroscedasticity." *Econometrica* 44(3): 461-65.

Hoxby, Caroline 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115(4): 1239-85.

Jensen, Ben and Andrew Seltzer 2000. "Neighbourhood and Family Effects in Educational Progress." *Australian Economic Review* 33(1): 17-31.

Maddala, G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, Cambridge University Press.

Manski, Charles F. 1990. "The Use of Intentions Data to Predict Behavior: A Best Case Analysis." *Journal of American Statistical Association* 85(412): 934-40.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60: 531-42.

Mayer, Christopher J. 1996. "Does Location Matter?" *New England Economic Review* May/June: 26-40.

McClelland, Katherine 1990. "Cumulative Disadvantage among the Highly Ambitious." *Sociology of Education* 63(2): 102-21.

Moynihan, Daniel P. (1968). *On Understanding Poverty: Perspectives from the Social Sciences*. New York, Basic Books.

O'Regan, Katherine M. and John M. Quigley 1998. “Spatial Effects upon Employment Outcomes: The Case of New Jersey Teenagers.” *Working Paper, University of California, Berkeley*.

Pagan, Adrian and Frank Vella 1989. “Tests for Models Based on Individual Data: A Survey.” *Journal of Applied Econometrics* 4(Supplement): S29-59.

Plotnick, Robert and Saul Hoffman 1999. “The Effect of Neighbourhood Characteristics on Young Adult Outcomes: Alternative Estimates.” *Social Science Quarterly* 80(1): 1-18.

Reynolds, John R. and Jennifer Pemberton 2001. “Rising College Expectations Among Youth in the United States.” *Journal of Human Resources* 36(4): 703-26.

Teese, Richard (2000). *Academic success and social power : examinations and inequality*. Carlton, Vic. :, Melbourne University Press,.

Wilson, William J. (1980). *The declining significance of race : Blacks and changing American institutions*. Chicago :, University of Chicago Press,.

Wilson, William J. (1987). *The truly disadvantaged : the inner city, the underclass, and public policy*. Chicago :, University of Chicago Press,.

Wilson, William J. (1993). *Sociology and the public agenda*. Newbury Park, Calif. :, Sage Publications,.

Wilson, William J. (1996). *When work disappears : the world of the new urban poor*. New York :, Knopf : Distributed by Random House, Inc.,.

Wolter, Stefan C. 2000. “Wage Expectations: A Comparison of Swiss and US Students.” *Kyklos* 53(1): 51-69.

Yatchew, Adonis J. and Zvi Grilliches 1985. “Specification Error in Probit Models.” *Review of Economics and Statistics* 67(1): 134-39.