Family Impacts on Cognitive Development of Young Children: Evidence from Australia

by

Jessica Meredith, Frank Neri and Joan Rodgers

School of Economics,
University of Wollongong,
Wollongong N.S.W.,
Australia

August 2013

Abstract

This paper investigates the manner and extent to which family structure impacts upon the cognitive development of young Australian children. Our methodology draws on the standard household production model of Becker but also includes control variables emphasised by parental investment and good-parent theories of child development. We use data from the Longitudinal Study of Australian Children (LSAC) and from the National Assessment Program – Literacy and Numeracy (NAPLAN) in cross sectional, panel, instrumental variables and fixed-effects analyses. Our results suggest that the large negative effects initially associated with single parent families disappear when child characteristics and parental preferences for education are controlled for. On the other hand parental completion of Year 12 education, 'warm' parent-child interactions, a stress-free home environment and positive parental aspirations for their children are persistently strong determinants of the educational success of young children.

Corresponding author:
Frank Neri
Email: frank_neri@uow.edu.au
Tel: 61 02-42214671
Fax: 61 02-42213725

JEL classifications: I21, I24, I28

Family Impacts on Cognitive Development of Young Children: Evidence from Australia

I. Introduction

In 2011, single-parent families accounted for nearly 21 per cent of all Australian families containing at least one child under the age of 15 years, a near four percentage point increase since 1992. More than four out of five of these families are headed by a single mother. Australian social policy acknowledges that single-parent families are at a higher risk of disadvantage. For example, 34 per cent of single-parent families with dependent children are classified as low-income compared with only 19 per cent of two-parent families with dependent children. Similarly, 44 per cent of children in single-parent families have an unemployed resident parent, whilst only five per cent of children in two-parent families have unemployed resident parents (ABS 2011, 4102.0). These statistics suggest that children in single-parent families enjoy fewer resources. The quality of such resources may also be lower in single-parent families due to tighter time constraints and increased stress. These and other disadvantages suggest that children in single-parent families are likely to be more vulnerable to social inequality. One way that this vulnerability may become manifest is through diminished cognitive development.

The objective of this paper is to analyse the impact of family structure and income on the cognitive development of young Australian children. If family income has significant impacts on cognitive development, the arguments for increased income redistribution or for additional income supplements to single-parent and other low-income families may be strengthened. Similarly if home environment factors are influential then there may be scope for additional programs such as parental mentoring and relief services. Of course our results may also be salient to individuals considering parenthood where the composition of their family is a choice variable.

The remainder of this paper is structured as follows. Section II sets out the theoretical foundations of our study. Section III then summarises the previous literature in the area. Section IV discusses the data used in this study. Sections V and VI present the results of our analysis using alternate measures of cognitive achievement. Finally, section VII summarises our findings and makes some concluding comments.

II. Theoretical Foundations

Economic Theory

The household production model of Becker (1965, 1981) assumes that parents are utility-maximising agents who care about the welfare of their children. In Becker's model, the cognitive development (and other outcomes) of children are commodities produced by the household using available resources.[1] Parents combine market goods and services such as childcare, tuition and books, with their own human capital and leisure time to invest in the development, cognitive and otherwise, of their children. This in turn impacts on the utility of the household.

Other things equal, a two-parent household will achieve a higher level of utility since two parents can invest more time and other resources into their children's development, and can do so in a more complementary manner. For example, in a two-parent household each parent likely allocates time according to the principal of comparative advantage. The parent who is relatively more efficient at labour market activities dedicates relatively more time to paid work whilst the second parent dedicates relatively more time to childcare. Clearly a single parent will be relatively more time constrained in one or both activities and so his or her household will enjoy fewer market and/or non-market resources.

[1] Becker and Tomes (1986) posit that child outcomes are also dependant on the transmission of genetic and cultural endowments.

Furthermore, due to imperfect capital markets, parents face borrowing constraints. So to invest in their children, parents must forego current consumption and/or increase their labour force activity. Since higher income households face fewer capital constraints they have greater potential to invest in their children. Hence this model predicts, other things equal, that children in higher income households likely achieve higher levels of cognitive development, which typically contribute to higher incomes in adulthood.[2]

Non-Economic Theory

Parental stress theory (Mayer 1997, Violato *et al.* 2011) proposes that low family income makes good parenting less likely because it magnifies the financial and emotional stress of parents onto the children. Parental-child interactions in such families may be less frequent and/or of poorer quality and so are likely to be less conducive to the cognitive and social development of children. This theory thus emphasises income redistribution policy to reduce stress and thereby improve child outcomes.

In a related vein, role-model theory focuses on the long-term changes in parental behaviour and values after a permanent adverse income shock (Mayer 1997). If family income diminishes permanently, parents adjust their attitudes and behaviours accordingly to cope with their permanently diminished material standard of living. This theory predicts that the cognitive and social development of children in such families is likely to be reduced due to deterioration in the behaviours, values and aspirations of their parent(s) following the shock. Interestingly, the role model theory focuses not on the impact of lower family income *per se*, but rather on the impact of unexpected transitions to a permanently lower family income following an adverse shock such as parental separation or illness.

[2] Mayer (1997) even argues that this success becomes further manifest via the availability of better marriage partners.

III. Previous Literature

A large body of empirical literature has examined the impact of family structure on child cognitive development. Studies can be classified according to the nature of the data, the estimation methods used, and the variables held constant when measuring the impact of income and family structure on child outcomes. In general, the results suggest that family structure loses its impact on child cognitive development once (permanent) family income is controlled for, and that family income loses much, though not all, of its impact once other family factors are taken into account.

An early example of these studies is Thomson *et al.* (1994) who find that the cognitive disadvantage of children in one-parent families largely disappears when family income is controlled for. More recently, Ginther and Pollack (2004) find that children from one-parent families display lower cognitive development than their peers from two-parent families, but that this effect disappears when family income is included as a control.[3] A recent Australian example, but with a focus on emotional rather than cognitive development, is Baxter et al. (2011) who, like us, use data from the Longitudinal Study of Australian Children (LSAC). Prior to controlling for child background variables, children living in two-parent families had better emotional outcomes than children living in one-parent families. After controlling for demographic, socioeconomic and personal characteristics, children from single-parent families continued to display lower emotional development.

Early examples of studies with more sophisticated empirical strategies include Mayer (1997) and Blau (1999) who supplement OLS models with fixed effects models to control for

---

[3] The one exception to this finding is for children from so-called 'blended' families, those headed by a parent and a step-parent, who do worst of all even controlling for family income. Adding further weight to the notion that not all non-traditional two-parent family structures are equal, Wadsworth and McLean (1986) and Biblarz and Gottainer (2000) both find that children from widowed single-parent families suffer no cognitive disadvantage, whereas children from divorced single-parent families do, when compared with their peers from two-parent families. However these latter results are based on unconditional means and so should be treated with caution.

the likely endogeneity of family income. They find that the coefficients on permanent income remain statistically significant but are so small in magnitude as to preclude income transfers as a feasible approach to achieving substantial cognitive improvements in low-income children. Aughinbaugh and Gittleman (2003) supplement their OLS models with random effects models and again find little support for income transfers to low-income families. Taylor *et al* (2004) use a similar empirical approach and conclude that a $10,000 annual increase in permanent family income would result in an increase of only 3.6 per cent of a standard deviation in child assessment scores. Studies have also addressed non-linear effects of income by using quadratic functional forms (Løken *et al.* 2010) and spline regression functions (Duncan *et al.* 1998; Aughinbaugh and Gittleman 2003). The common finding is that changes in permanent family income exhibit diminishing marginal returns to cognitive outcomes.

Finally, some studies have also controlled for differences in non-pecuniary family factors. For example, Aughinbaugh and Gittleman (2003) find that, overall, a one standard deviation reduction in emotional strain improves assessment scores of US (UK) children by 13-28 (14-23) per cent of a standard deviation. Similarly, Violato *et al.* (2011) find that the positive and significant impacts of family income on child cognitive and behavioural development fall by approximately 60 per cent when controls are added for parental stress, parental investments in their children, and other relational variables. Variables exhibiting the largest partial effects were parent-child time in cognitively stimulating activities, the quality of the childcare arrangements, the quality and safety of the neighbourhood, and the nature and extent of home appliances.

IV. Data

The Longitudinal Study of Australian Children[4]

This paper uses unit record data from the first three waves of Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC). The LSAC follows two cohorts of children: the child (K) cohort, which consists of children born between March 1999 and February 2000 and the baby (B) cohort which consists of children born between March 2003 and February 2004. Our study is based on the 4,983 children in the K cohort, the B cohort children being too young to have taken one of the cognitive tests used in our study. By Wave 3 the K cohort had been observed at ages 4-5, 6-7 and 8-9 years. (Soloff *et al.*, 2005, 2006).

The Dependent Variables

We use two measures of cognitive achievement as dependent variables. The first is the child's score on the National Assessment Program – Literacy and Numeracy (NAPLAN) test, which was introduced to Australian schools in 2008 and is taken by children in Years 3, 5, 7 and 9. The NAPLAN test assesses students' ability in the domains of reading, writing, spelling, grammar and punctuation, and numeracy (National Assessment Program 2011a, 2011b). Year 3, 5, 7 and 9 NAPLAN scores are measured on a common assessment scale ranging from zero to 1000. The scale has ten bands in total with each year recording student performance over six bands. Students in Year 3 are expected to achieve NAPLAN scores between bands one and six, with band two assigned as the national minimum standard.

The second measure of cognitive achievement is the child's score on the Peabody Picture Vocabulary Test (PPVT), which is taken by children at ages 4-5, 6-7 and 8-9 years. A

---

[4] The Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) funds and manages the LSAC on behalf of the Australian Government. The Australian Bureau of Statistics collects and processes the data. The Australian Institute of Family Studies (AIFS) is responsible for the design and content of the study, and the preparation of research and statistical reports. The findings and views reported in this paper, however, are those of the authors and should not be attributed to either FaHCSIA, the ABS or AIFS.

shortened version of the PPVT Third Edition was included in the LSAC as a means of testing children's receptive competence and understanding of the Standard American English vocabulary (AIFS 2011a, 2011b). The PPVT is adjusted at each wave to account for the child's age and acts as a measure of scholastic aptitude and verbal ability that is comparable across waves (Blau 1999; Rothman 2005).

One NAPLAN test score (Year 3) and three PPVT test scores are recorded for each child in the K cohort in Waves 1-3 of LSAC data. Consequently, our sample of the NAPLAN scores is cross-sectional and that of the PPVT scores is longitudinal.

Explanatory Variables

Family structure is represented by three continuous variables in our NAPLAN cross-sectional analysis. They indicate the proportion of waves spent in a two-biological-parent family (*twoparent)*, a single-biological-parent family (*singleparent*) and in a family with a biological mother and a stepfather or with a biological father and a stepmother (*stepparent*).[5] Three dummy variables represent family structure in a particular year of the PPVT panel data analysis: *2biopar* equals one for children living in a family with their two biological parents, *singpar* equals one for children living in a single-biological-parent family, and *biostep* equals one for children living in a one-biological-parent, one-stepparent family.

Income is an important control variable because it is correlated with family structure. Our income measure is the aggregate, real, equivalised, usual, weekly income of the main adult(s) in the family.[6] Usual weekly income is defined as the amount of income earned from wages and salaries, dividends, interest, business income, superannuation, child support,

[5] Only a small number of children in the data set were living solely with their biological father so they were combined with children living solely with their biological mother.

[6] Family income is constructed by adding the usual weekly income of 'Parent 1' and 'Parent 2', which is the most comprehensive measure of family income available in all waves of LSAC data. Nominal usual weekly income is converted to 2003-2004 dollars using the Consumer Price Index (CPI) (ABS Cat. No. 6401.0, 2012) and then adjusted using the 'modified OECD' equivalence scale to account for variation in family size and composition. This scale assigns one point for the first adult in the family, 0.5 points for each additional adult and 0.3 points for each child.

worker's compensation, property rentals, government benefits and allowances in an average week, before income tax (AIFS 2011b). In accordance with other studies (Blau 1999; Taylor *et al*. 2004; Violato *et al.* 2011) income is "smoothed" in the NAPLAN analysis by taking its arithmetic average over all available LSAC waves. In the PPVT analysis, income is not smoothed due to the longitudinal nature of the analysis. Both the NAPLAN and PPVT analyses use the natural logarithm of income to model non-linear income effects.

Characteristics of the children and their parents are also used as controls in our study. Child characteristics are Aboriginal or Torres Strait Islander status, number of siblings, number of older siblings, gender, age, birth weight, health and an indicator of whether the child lives in a metropolitan area. Parental characteristics are age, education, employment status, measures of health and depression and whether the mother smoke or drank alcohol during pregnancy. For each parental characteristic, a variable is constructed that is suitable for all family structures regardless of whether the family has one or two parents.

Three additional sets of control variables are included to mirror the theoretical foundations of our analysis. The variables used to capture parental investment indicate whether the child has access to a computer and whether he or she lives in a safe and secure home. There is also a variable measuring the 'liveability' of the child's neighbourhood. Parental stress theory is reflected in a series of 'coping with life' variables that control for any effects that the emotional well-being of parents may have upon the educational attainment of their children. Parental well-being is captured by three dummy variables indicating whether the parent(s) have more than a few problems, are coping and feel rushed. There is also a variable for 'parental warmth', which is the mean of the responses by the parent(s) to several questions about their interactions with the child.[7] Role model theory is

[7] The questions are as follows. "How often do you express affection by hugging, kissing and holding this child?", "How often do you hug or hold this child for no particular reason?", "How often do you tell this child

taken into account by two dummy variables that capture parental aspirations for the child's education: whether the parent(s) believe that the child will leave school prior to the completion of year 12, and whether the parent(s) believe their child will undertake post-secondary studies.

The NAPLAN Sample

Table 1 shows the number and proportion of children in Wave 3 of LSAC with Year 3 NAPLAN test scores in each of the five domains.

Table 1 Family Structure Distributions by NAPLAN Sample

	Reading	*Writing*	*Spelling*	*Grammar and Punctuation*	*Numeracy*
Two biological parents	1,793 (82.9%)	1,793 (83%)	1,796 (83%)	1,796 (83.1%)	1,793 (83%)
Single biological parent	301 (13.9%)	299 (13.8%)	299 (13.8%)	297 (13.7%)	299 (13.8%)
One biological parent, one stepparent	68 (3.1%)	68 (3.1%)	68 (3.1%)	68 (3.1%)	68 (3.1%)
Total	**2,162 (100%)**	**2,160 (100%)**	**2,163 (100%)**	**2,161 (100%)**	**2,160 (100%)**

Source: LSAC, Waves 1 to 3.

The PPVT Sample

Not all children have scores on all three PPVT tests so the analysis utilises an unbalanced panel. There are 10,992 child-wave observations, corresponding to K-cohort children with at least two consecutive PPVT test scores. Table 2 presents the distribution of family structures for the children in the PPVT sample. The numbers of children of various ages appear in the last row. For example, 3,850 children were observed at age 6-7, which is approximately 77.3 per cent of the 4,983 K-cohort children in Wave 1.

how happy he/she makes you?", "How often do you have warm, close times together with this child?", "How often do you feel close to this child when he/she is happy and when he/she is upset?"

Table 2 Family Structure Distribution by PPVT Sample

Children with	*PPVT (age 4-5)*	*PPVT (age 6-7)*	*PPVT (age 8-9)*
Two biological parents	3,043 (86.1%)	3,187 (82.8%)	2,949 (81.8%)
Single biological parent	430 (12.2%)	568 (14.8%)	528 (14.6%)
One biological parent, one stepparent	62 (1.8%)	95 (2.5%)	130 (3.6%)
Total	**3,535 (100%)**	**3,850 (100%)**	**3,607 (100%)**

Source: LSAC, Waves 1 to 3.

V. Results of the NAPLAN Analysis

Six specifications of the cross-sectional NAPLAN model were estimated (see Table 3). The first five specifications state the relationship between children's NAPLAN scores and family structure with varying numbers of controls; they were estimated with ordinary least squares (OLS). The sixth model was estimated with instrumental variables. Model I expresses NAPLAN scores as a function of family structure with no controls; it acts as the baseline model. Model II also includes family income. The inclusion of more and more control variables in Models III, IV and V helps to remove bias caused by endogeneity of family structure and income.

Table 3 Model Specifications

Model	*Explanatory Variables*
Model I	Family structure
Model II	Model I + income
Model III	Model II + child characteristics
Model IV	Model III + parental characteristics
Model V	Model IV + parental investment, parental stress and role model variables
Model VI	Model V + instruments for family structure and family income

Unobserved Heterogeneity

Family structure and income are likely to remain endogenous even after a large number of explanatory variables are included in the model. The unobservable characteristics of parents in traditionally structured families are feasibly different from those of parents who are divorced or who engaged in non-marital childbearing (Ginther & Pollak 2004). Furthermore, unobservable characteristics of the child's parent(s), such as ability, motivation and ambition, are likely to be correlated with family income. Unobservable parental preferences are likely to influence the allocation of family income to their children (Blau 1999). With these variables omitted from the model, the empirical results will be biased (Wooldridge 2009, p.94; Blau 1999; Taylor et al. 2004; Violato et al. 2011).

We use two-stage least squares (2SLS) to address the problem of unobserved heterogeneity in our cross-sectional data. Model VI was estimated by 2SLS with instruments for family structure and income. The instrument for family structure is a dummy variable equal to one if Parent 1's own parents (that is, the child's grandparents) were divorced or permanently separated prior to when Parent 1 was aged ten. In accordance with Shea (2000), the instrument for family income is a dummy variable equal to one if Parent 1 is involved, in an unpaid capacity, in an industrial union. To be a valid instrument the parent's union status must be correlated with income but have no direct effect on the child's NAPLAN score. It is argued that "union and industry wage premiums reflect rents rather than unobserved ability differences" (Shea 2000, p.157), which is required for union status to be exogenous.

Sample Selection Bias

NAPLAN test scores are only observed for children who are in the appropriate year of school to take the NAPLAN test. If the children who are enrolled in a different year of school have intrinsically higher or lower cognitive ability then omitting these children from the sample will introduce selection bias into an empirical analysis of the factors affecting educational outcomes (Verbeek 2008, p.249; Wooldridge 2009, p.606)[8]. Children in the K

[8] Selection bias does not occur if selection into the sample is solely related to the explanatory variables in the outcome equation (Verbeek 2008, p.250). Hence selection bias does not arise if children from two-parent

cohort, being 8-9 years old in Wave 3 of the LSAC, are expected to be in Year 3 of school and approximately 70 per cent are. But 24 per cent of children are in the year ahead and six per cent are in the year behind (Sipthorp & Taylor 2012)[9]. NAPLAN test scores of children who sat the NAPLAN examinations in 2008 or 2009 have been linked to Wave 3 of the LSAC data, so children enrolled in Year 2 or Year 3 in 2008 have Year 3 NAPLAN scores recorded in the data, [10] but children enrolled in Year 4 in 2008 do not.

Models I through V were re-estimated using a Heckman-correction procedure that accounts for possible sample selection bias resulting from there being no Year 3-NAPLAN scores for children who were in Year 4 in 2008. The child's month of birth (*MOB*) is employed as an instrumental variable to identify the child-outcome equation. All children in the data set were born between March 1999 and February 2000 but given State and Territory laws about the age at which a child must start primary school, the oldest children have a greater likelihood of being in Year 4 in 2008 and therefore missing a Year 3 NAPLAN score. Accordingly month of birth (*MOB*) is a binary variable equal to one for children who are likely to be in Year 3 or Year 2 in 2008 and zero for children who are likely to be enrolled in Year 4 based on their month of birth and state of residence.[11] No Heckman correction for sample selection is applied to Model VI. It is estimated only in the case of those NAPLAN domains that exhibit no evidence of sample selection bias in Model V.

---

families, or high-income families, have a greater propensity than others to be included in the NAPLAN sample, since family structure and income are both included as explanatory variables.

[9] Based on Wave 1 data.

[10] A small number of these students do not have scores because their parents did not give permission for them to be linked to the LSAC data.

[11] All South Australian and Tasmanian children as well as children in the ACT, NSW, Queensland, Victoria, Northern Territory and Western Australia who were born between July 1999 and February 2000 have *MOB* equal to one; for other children *MOB* equals zero.

Table 4 Probit Model Selection Equation Estimates for the NAPLAN Samples

	Model I	Model II	Model III	Model IV	Model V
Reading					
singleparent	-0.27***	-0.15**	0.19**	0.21*	0.28**
	(0.07)	(0.08)	(0.08)	(0.107)	(0.13)
stepparent	-0.45***	-0.29*	-0.29*	-0.17	-0.16
	(0.15)	(0.15)	(0.16)	(0.20)	(0.21)
log(income)	-	-0.07***	0.29***	0.26***	0.27***
		(0.01)	(0.05)	(0.06)	(0.06)
MOB	0.38***	0.81***	0.70***	0.73***	0.73***
	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)
N	4,039	4,001	3,947	3,102	2,853
ρ	0.508	0.295	0.360	0.244	0.329
Wald χ^2 *test*	8.78	5.48	3.21	1.01	2.17
Prob > χ^2	0.000	0.019	0.0731	0.315	0.141
	Model I	Model II	Model III	Model IV	Model V
Writing					
singleparent	-0.28***	-0.16**	0.19**	0.19*	0.25**
	(0.07)	(0.08)	(0.08)	(0.11)	(0.13)
stepparent	-0.44***	-0.29*	-0.28*	-0.16	-0.16
	(0.15)	(0.15)	(0.16)	(0.20)	(0.21)
log(income)	-	-0.07***	0.29***	0.26***	0.28***
		(0.01)	(0.05)	(0.06)	(0.06)
MOB	0.38***	0.82***	0.71***	0.74***	0.74***
	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)
N	4,039	4,001	3,947	3,102	2,853
ρ	0.190	0.157	0.199	0.144	0.127
Wald χ^2 *test*	6.47	3.59	5.38	1.62	1.10
Prob > χ^2	0.011	0.058	0.020	0.203	0.295
	Model I	Model II	Model III	Model IV	Model V
Spelling					
singleparent	-0.29***	-0.17**	0.19**	0.21*	0.27**
	(0.07)	(0.08)	(0.08)	(0.11)	(0.13)
stepparent	-0.44***	-0.29*	-0.28*	-0.16	-0.16
	(0.15)	(0.15)	(0.16)	(0.20)	(0.21)
log(income)	-	-0.07***	0.31***	0.27***	0.29***
		(0.01)	(0.05)	(0.06)	(0.06)
MOB	0.38***	0.81***	0.70***	0.74***	0.74***
	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)
N	4,039	4,001	3.947	3.102	2,853
ρ	0.206	0.128	0.169	0.032	0.019
Wald χ^2 *test*	2.24	1.37	1.31	0.03	0.01
Prob > χ^2	0.135	0.243	0.252	0.859	0.905
	Model I	Model II	Model III	Model IV	Model V
Grammar and Punctuation					
singleparent	-0.30***	-0.19**	0.17**	0.20*	0.26**
	(0.07)	(0.07)	(0.09)	(0.11)	(0.13)
stepparent	-0.44***	-0.29*	-0.29*	-0.16	-0.18
	(0.15)	(0.15)	(0.16)	(0.20)	(0.21)
log(income)	-	-0.07***	0.30***	0.27***	0.28***
		(0.01)	(0.05)	(0.06)	(0.06)
MOB	0.38***	0.81***	0.70***	0.74***	0.74***
	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)
N	4,039	4,001	3,947	3,102	2,853
ρ	0.514	0.285	0.373	0.216	0.337
χ^2 *test*	6.59	4.51	4.08	0.78	1.86
Prob > χ^2	0.010	0.034	0.043	0.376	0.172

	Model I	Model II	Model III	Model IV	Model V
Numeracy					
singleparent	-0.28***	-0.16**	0.19**	0.20*	0.25**
	(0.07)	(0.07)	(0.08)	(0.11)	(0.13)
stepparent	-0.43***	-0.29*	-0.29*	-0.14	-0.14
	(0.15)	(0.15)	(0.16)	(0.20)	(0.21)
log(income)	-	-0.07***	0.31***	0.28***	0.28***
		(0.01)	(0.05)	(0.06)	(0.06)
MOB	0.37***	0.80***	0.66***	0.67***	0.67***
	(0.03)	(0.06)	(0.06)	(0.07)	(0.08)
N	4,039	4,001	3,947	3,102	2,853
ρ	0.756	0.462	0.667	0.661	0.659
χ^2 *test*	100.99	3.65	22.27	11.87	8.85
Prob > χ^2	0.000	0.056	0.000	0.001	0.003

Source: LSAC, Waves 1 to 3.
Note: The NAPLAN estimates are weighted using the Wave 3 cross-sectional sampling weights.
Robust standard errors appear in parentheses. *, ** and *** denote that the variables are statistically significant at the 10%, 5% and 1% levels of significance respectively.

The Heckman Selection Equations

The probability of selection into our NAPLAN sample is taken to be a function of the explanatory variables in Models I through V, plus the month of birth instrument. A separate probit function was estimated for each of the five domains, the dependent variable being the probability of having a NAPLAN score in the appropriate domain. Key results of the 25 estimated models are given in Table 4. More detailed results appear in Appendix A.

The month of birth instrument is nontrivial in size and highly statistically significant (p=0.000) in all 25 specifications, suggesting that *MOB* is a strong instrument for identifying the selection process. For example, in Model I the probability of selection into the NAPLAN reading sample for children living in a single-biological-parent family in all three waves is given by $\Phi(-0.27(1)+0.38(1)) = \Phi(0.11) = 0.54$ if *MOB*=1 and by $\Phi(-0.27(1)+0.38(0)) = \Phi(-0.27) = 0.39$ if *MOB*=0, indicating that children in the *MOB*=1 category have a fifteen percentage-point higher probability of being in the NAPLAN reading sample.[12]

The ρ statistic in Table 4 measures the correlation between the error terms in the selection and outcomes equations whilst the Wald statistic is used to test the null hypothesis

[12] Φ is the standard, normal cumulative density function.

that ρ equals zero. In the domains of NAPLAN reading, writing and grammar and punctuation, the Wald test indicates sample selection bias in Models I, II and III but not in Models IV and V. Accordingly the Heckman-corrected cross-sectional results will be discussed for Models I-III and the standard OLS results will be discussed for Models IV and V. There is no evidence that the NAPLAN spelling sample suffers from sample selection bias in any of Models I-V and hence only OLS results will be discussed. On the other hand, all five specifications of the NAPLAN numeracy model suffer from sample selection bias so only the Heckman-corrected cross-sectional results will be discussed. Key results appear in Table 5 and more detailed results are in Appendix A.

The Outcome Equations

Family Structure and Smoothed Income

The negative and statistically significant coefficients of *singleparent* and *stepparent* in Model I of Table 5 are evidence that children living with both their biological parents achieve significantly higher NAPLAN test scores in all domains than children living in single-parent families or in families with one biological parent and one step parent.[13] For example, children who spent all three waves in two-biological-parent families achieved NAPLAN reading scores that were 50.45 and 28.81 marks above those of children who spent all three waves in single-biological-parent families and one-biological-parent, one-stepparent families, respectively[14].

Introducing income into the model reduces the negative effect on NAPLAN scores of living in a non-traditional family. In most cases the statistical significance of the family

---

[13] There is one exception only: the coefficient of stepparent in the NAPLAN spelling domain is negative but not significantly different from zero.

[14] These increments are substantial. The national minimum standard is a score in band two, between 101 and 200 on a 1000 point scale. So these increments represent, respectively, 25%-50% and 14%-28% improvements for a child with a national minimum standard score.

Table 5 NAPLAN Results, Heckman-Corrected and Standard OLS Estimates, Models I-V

	Model I		Model II		Model III		Model IV		Model V	
	OLS	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*
Reading										
singleparent	-42.69***	-50.45***	-13.50*	-15.96**	-16.47**	-12.68	-9.81	-7.32	-10.44	-6.18
	(7.23)	(7.96)	(7.83)	(8.03)	(7.71)	(7.90)	(9.14)	(9.30)	(10.65)	(11.01)
stepparent	-17.36	-28.81**	-9.68	-12.78	-11.14	-14.61	3.84	2.91	-2.79	-4.01
	(13.11)	(14.45)	(14.21)	(14.45)	(14.02)	(14.32)	(16.88)	(16.88)	(17.62)	(17.70)
log(income)	-	-	36.66***	35.43***	31.02***	36.19***	30.47***	33.31***	24.30***	28.27***
			(4.07)	(4.14)	(4.46)	(5.23)	(5.33)	(6.00)	(5.33)	(5.96)
N	2,162	2,162	2,162	2,162	2,142	2,142	1,738	1,739	1,617	1,618
Wald χ^2	-	43.53	-	114.34	-	191.95	-	274.00	-	319.58
F	19.12	-	39.47	-	17.61	-	13.73	-	11.54	-
Writing										
singleparent	-29.54***	-31.73***	-9.45	-10.54	-12.34*	-10.83*	-4.14	-3.13	-3.44	-2.31
	(5.97)	(6.08)	(6.50)	(6.56)	(6.47)	(6.46)	(7.72)	(7.72)	(8.64)	(8.71)
stepparent	-26.76**	-29.95***	-21.49*	-22.78**	-20.56*	-22.02*	-5.60	-5.99	-8.08	-8.41
	(10.60)	(10.70)	(11.15)	(11.19)	(11.32)	(11.28)	(13.36)	(13.31)	(13.78)	(13.70)
log(income)	-	-	25.16***	24.62***	21.68***	23.85***	21.07***	22.37***	17.21***	18.39***
			(3.09)	(3.12)	(3.37)	(3.41)	(3.85)	(3.96)	(3.95)	(4.03)
N	2,160	2,160	2,160	2,160	2,140	2,140	1,737	1,737	1,616	1,616
Wald χ^2	-	36.70	-	99.55	-	205.78	-	265.95	-	269.92
F	16.33	-	33.94	-	18.14	-	12.93	-	9.23	-

	Model I		Model II		Model III		Model IV		Model V	
	OLS	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*
Spelling										
singleparent	-38.63***	-41.36***	-20.43***	-21.48***	-20.52***	-19.08***	-13.94	-13.67	-8.86	-8.66
	(6.60)	(6.94)	(7.07)	(7.17)	(7.20)	(7.14)	(9.35)	(9.34)	(9.85)	(9.99)
stepparent	-14.87	-18.72	-10.00	-11.16	-9.68	-11.09	9.53	9.42	1.49	1.42
	(14.11)	(14.36)	(14.83)	(14.82)	(14.82)	(14.77)	(17.18)	(17.08)	(17.23)	(17.10)
log(income)	-	-	22.94***	22.45***	18.38***	20.54***	20.06***	20.41***	15.56***	15.77***
			(3.44)	(3.47)	(3.75)	(4.09)	(4.48)	(4.79)	(4.46)	(4.67)
N	2,163	2,163	2,163	2,163	2,143	2,143	1,741	1,741	1,620	1,620
Wald χ^2	-	36.61	-	79.98	-	146.27	-	157.44	-	270.04
F	17.81	-	26.78	-	13.24	-	7.91	-	9.19	-
Grammar and Punctuation										
singleparent	-35.07***	-44.06***	-6.66	-9.64	-7.37	-4.12	1.48	3.48	-4.93	-0.99
	(7.61)	(8.39)	(8.20)	(8.44)	(7.91)	(8.13)	(9.75)	(9.89)	(11.26)	(11.58)
stepparent	-30.02**	-41.16***	-22.29	-25.10*	-23.40*	-26.85*	-4.32	-5.07	-8.87	-9.97
	(12.69)	(13.33)	(13.63)	(13.61)	(13.96)	(14.05)	(13.76)	(13.88)	(13.26)	(13.60)
log(income)	-	-	36.20***	35.00***	32.70***	38.06***	30.86***	33.41***	26.37***	30.48***
			(3.80)	(3.90)	(4.18)	(4.69)	(5.08)	(5.66)	(5.12)	(5.73)
N	2,161	2,161	2,161	2,161	2,142	2,142	1,740	1,740	1,619	1,619
Wald χ^2	-	37.40	-	118.40	-	223.74	-	266.05	-	343.41
F	13.93	-	41.06	-	19.93	-	13.03	-	12.23	-

	Model I		Model II		Model III		Model IV		Model V	
	OLS	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*	*OLS*	*Heckman*
Numeracy										
singleparent	-33.18***	-43.60***	-7.53	-10.85*	-6.29	1.05	-4.54	3.12	-10.30	-1.34
	(5.30)	(6.20)	(5.83)	(6.30)	(5.97)	(6.65)	(7.99)	(8.75)	(9.03)	(10.10)
stepparent	-21.78*	-37.79***	-15.04	-19.16	-16.66	-22.63*	-0.00	-1.84	-1.71	-3.13
	(11.27)	(13.12)	(12.10)	(12.58)	(11.62)	(12.90)	(14.25)	(15.70)	(13.71)	(15.16)
log(income)	-	-	32.18***	30.49***	29.75***	39.38***	29.00***	37.14***	23.91***	32.00***
			(3.33)	(3.43)	(3.70)	(4.57)	(4.33)	(5.32)	(4.28)	(5.30)
N	2,160	2,160	2,160	2,160	2,140	2,140	1,735	1,735	1,614	1,614
Wald χ^2	-	57.86	-	130.01	-	173.15	-	178.93	-	239.33
F	21.48	-	46.13	-	17.62	-	10.90	-	10.14	-

Source: LSAC, Waves 1 to 3.
Note: The NAPLAN estimates are weighted using the Wave 3 cross-sectional sampling weights.
Robust standard errors appear in parentheses. *, ** and *** denote that the variables are statistically significant at the 10%, 5% and 1% levels of significance respectively.

structure coefficient is reduced as well, suggesting that the large negative effect in Model 1 is largely attributable to the tendency of these families to be income-poor. The effect of income itself on all domains of NAPLAN performance is positive and highly statistically significant in Model II. For example, a $100 increase in weekly smoothed income improves NAPLAN Writing outcomes of the average child by (100x24.62/686.39 =) 3.59 marks.[15]

In Models III, IV and V the effect of family structure on NAPLAN scores is not statistically significant and it is either about the same size or even smaller than in Model II. This suggests that once income and other controls are taken into account, growing up in a non-traditionally structured family is not significantly detrimental to NAPLAN scores. However, the effect of income on all domains of NAPLAN performance is statistically significant in all models. The behaviour of the income coefficient in Models III, IV and V is similar in the domains of reading, writing, spelling and grammar and punctuation. In these domains the introduction of child characteristics has little effect on the income coefficient (see Model III) but the inclusion of parental characteristics reduces the effect of income (Model IV), reflecting the fact that parental education and employment are correlated with income. The coefficient of income in Model V, which has the full set of controls, is even smaller than in Model II. For example, an increase in smoothed weekly income of $100 improves NAPLAN reading results by (100x24.30/686.29 =) 3.54 marks in Model V.

On the other hand, the effect of income on NAPLAN numeracy scores does not decrease in size as additional controls are added. In fact, the coefficients of income in Models II and V are approximately the same size. For example, in Model V an increase in smoothed weekly income of $100 improves NAPLAN numeracy results by (100x32.00/686.54 =) 4.66 marks. The larger income effect obtained in the NAPLAN Numeracy sample suggests that

[15] Average smoothed income is equal to $686.29, $686.39, 687.72, $688.12 and $686.54 for children in the NAPLAN reading, writing, spelling, grammar and punctuation and numeracy samples, respectively.

mathematical ability may be more sensitive than general language and comprehension skills to the availability of monetary resources and external learning environments.

The Control Variables

The non-significance of the family structure effects and the relatively small, albeit statistically significant, effects of family income in Model V draws attention to the other statistically significant explanatory variables across the NAPLAN domains. In Model V the child characteristics, parental characteristics and parental role model vectors are each jointly statistically significant (based on *F* tests of joint significance) suggesting that the basic attributes of family members are important for the educational success of young children. However the results do not support Becker and Tomes' (1986) theory of parental investment as the joint effect of home facilities are not statistically significant. The parental stress vector exhibits joint statistical significance in the NAPLAN reading, grammar and punctuation, and numeracy samples suggesting that a stress-free home environment is more important to a child's NAPLAN scores than are home facilities.

Endogeneity of the Key Explanatory Variables

Whilst the Heckman-correction procedure addresses sample selection bias, it does not ensure that family structure and income are exogenous. Model VI of the NAPLAN analysis introduces Parent 1's childhood family structure and union status as instruments for family structure and smoothed income respectively. The explanatory variables in the OLS and 2SLS regressions are the same as those in Model V of the preceding NAPLAN analysis except that *singleparent* and *stepparent* are combined to form one '*non-traditional family*' dummy variable, which mirrors the family structure instrument. 2SLS was run only on the NAPLAN samples exhibiting no evidence of sample selection bias in Model V (see Table 6 for key results and Appendix B for detailed results). Accordingly, the NAPLAN numeracy scores were not re-estimated using 2SLS.

The F statistics obtained in the first-stage are in excess of ten in all NAPLAN domains, signalling that the instruments are not weak. Since the structural equation is exactly identified, the exogeneity of the instruments cannot be examined by a Hansen J test of over-identifying restrictions. This, and the fact that union status *may* still be correlated with parental ability and preferences, suggests that the income estimator may be inconsistent.

Table 6 Instrumental Variables, Model VI, 2SLS and OLS Estimates

	Reading		*Writing*		*Spelling*		*Grammar and Punctuation*	
	OLS	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*
Non-traditional family	-11.02	19.59	-6.03	-8.45	-3.87	-2.06	-12.43	9.62
	(8.05)	(34.07)	(6.15)	(26.73)	(7.15)	(29.89)	(7.98)	(33.54)
log(income)	24.16***	22.92	16.84***	64.65	16.22***	1.73	25.27***	81.62
	(5.19)	(69.76)	(3.93)	(46.47)	(4.42)	(61.94)	(4.93)	(76.54)
N	1,618	1,584	1,616	1,582	1,620	1,586	1,619	1,585
Wald χ^2	-	288.25	-	232.09	-	224.67	-	285.73
Prob > χ^2	-	0.000	-	0.000	-	0.000	-	0.000
F	12.10	-	9.75	-	9.39	-	12.77	-
Prob > F	0.000	-	0.000	-	0.000	-	0.00	-
First stage F(1)	-	47.38	-	46.66	-	46.89	-	45.45
First stage F(2)	-	30.87	-	30.94	-	30.91	-	30.35

Source: LSAC, Wave 3. Note: The estimates are weighted using the Wave 3 cross-sectional sampling weights. Robust standard errors appear in parentheses. *, ** and *** denote that the variables are statistically significant at the 10%, 5% and 1% levels of significance respectively.

The 2SLS family structure coefficients exhibit great variability over the four NAPLAN domains and are close to their corresponding OLS estimates only in the NAPLAN domains of writing and spelling. Furthermore the 2SLS standard errors are four to five times larger than in the OLS specification indicating that the family structure effect is not significantly different from zero in any NAPLAN domain. We are unable to reject the null hypothesis that the true income coefficient is zero in all four NAPLAN domains, which suggests that in Models I-V income is correlated with unobservable characteristics that were confounding its true effect. The insignificance of the income effect in Model VI implies that the level of family income does not determine the educational success of children. In summary, assuming that Parent 1's union status and childhood family structure are valid instruments for smoothed income and the child's family structure, the 2SLS estimation suggests that neither of the key explanatory variables explain discrepancies in children's NAPLAN scores.

VI. PPVT Results

The PPVT models were estimated using fixed-effects to take account of correlation between time-invariant unobservables and the explanatory variables. Random-effects models were also estimated and a Hausman specification test was conducted to assess which of these competing models has consistent estimators. The model specification for the fixed and random-effects estimations mirror Model V of the NAPLAN specification and includes a control for the assessment year[16]. The Hausman chi-square specification test strongly rejects the null hypothesis that family structure and income are exogenous[17]. In other words, the fixed-effects estimators of the effect of family structure and income on PPVT results are

Table 7 PPVT Results: Random-and Fixed-Effects Estimates

	Model V	
	RE	*FE*
singpar	0.15	0.65
	(0.28)	(0.51)
biostep	-0.71	-0.36
	(0.45)	(0.72)
log(income)	0.73***	0.08
	(0.14)	(0.18)
R^2	0.5279	0.5808
N	7,536	8,585
F	-	585.21
Prob>F	-	0.000
χ^2	13,695.79	-
Prob> χ^2	0.000	-
Median λ	0.40	-

Source: LSAC, Waves 1 to 3.
Note: The PPVT estimates are unweighted. Cluster robust standard errors are in parentheses. *, ** and *** denote that the variables are statistically significant at the 10%, 5% and 1% levels of significance respectively.

16 All explanatory variables are included in the random-effects model but only time-varying explanatory variables are included in the fixed-effects model.
17 χ^2=248.29 and Prob> χ^2=0.000.

consistent but the random-effects estimators are not, and so emphasis is placed on the fixed-effects estimates in the discussion below. (See Table 7 and Appendix C for results.)

The results indicate that living in a non-traditionally structured family is not detrimental to PPVT scores once observed and unobserved heterogeneity have been taken into account. The effect of income is also small and non-significant in the fixed-effects model (but much larger and significant in the random-effects model). This result runs parallel to international findings where income effects are close to zero and statistically non-significant after eliminating the fixed-error component while specifications ignoring unobserved heterogeneity typically report larger, statistically significant income effects (Blau 1999; Løken *et al.* 2010; Violato *et al*. 2011).

The Control Variables

Child characteristics (p=0.000) and parental characteristics (p=0.0128) both display a high level of joint statistical significance. The parental stress variables are jointly statistically significant at the one per cent level, suggesting that children are susceptible to the emotional and psychological well-being of their parents, which may itself be affected by family income. However, contrary to Violato *et al*. (2011), who found that parental investment variables exhibit large statistically significant effects on children's cognitive outcomes, we found parental investment to be the least important vector in explaining PPVT outcomes (the *F*-test of joint significance had p=0.2078).

VII. Summary and Conclusions

The results of our empirical analysis indicate that after controlling for a comprehensive set of covariates and accounting for selection bias using a Heckman-correction procedure, family structure was unable to explain variations in children's NAPLAN scores. Family income on the other hand remained statistically significant after controlling for all confounding factors, although the magnitude of the income effect indicates that a *very* large increase in weekly income would be required to improve NAPLAN scores substantially. NAPLAN Numeracy scores exhibit the most sensitivity to changes in income, suggesting that children's mathematic skills are more responsive to monetary resources than are general language and comprehension skills.

When instrumental variables and 2SLS estimation were used to remove the endogeneity of family structure and income, the effects of both family structure and income became statistically non-significant. Therefore, we conclude that family structure is irrelevant to children's academic achievement and that income re-distribution or income supplements to income-poor families are not likely to improve children's NAPLAN results.

The Hausman specification test signalled that the fixed-effects methodology was preferred to random effects for the purpose of estimating the PPVT model. The fixed-effects results suggest that once the correlation between the key regressors and unobservable characteristics is removed, living in single-biological-parent or a one-biological-parent, one-stepparent family is inconsequential for PPVT scores, as is family income.

A question arising from these results is why income plays such a minor and non-significant role in determining the educational outcomes of young Australian children. The first reason, as suggested by Shea (2000), points to the possibility that government expenditures in the schooling system neutralises income inequalities within the home. A second explanation is that parents may not act altruistically; changes in parental income do not necessarily flow to the benefit of children. Finally, income may indirectly facilitate familial relationships which in turn promote the educational achievement of young children.

Whilst income transfers may not be a practical solution to improving children's NAPLAN scores, the large negative effects of low parental education in several NAPLAN domains provides support for increasing the minimum education level in Australia to Year

12. The role-model theory suggests that increasing the qualifications of low-skilled parents through minimum schooling legislation, on-the-job training or government incentive schemes may increase parental preferences for human capital accumulation and, via a familial transmission effect, that of their children. The effect of having a highly educated parent was not statistically different from zero, suggesting that parental education to year 12 level would be sufficient to positively affect children's NAPLAN scores. As for the PPVT, basic child attributes accounted for a significant portion of the variation in PPVT outcomes, providing minimal scope for public policy in improving children's educational and cognitive outcomes.

The results obtained in the NAPLAN and PPVT analyses provide little support for Becker's household production model and theory of parental investment since (i) home facilities and family income were found to be unimportant in determining children's scores, and (ii) children from two-parent families did not perform significantly better than children in single-biological-parent families. On the other hand the parental stress and parental role model variables were important in both the NAPLAN and PPVT analyses. These results suggest that 'warm' parent-child interactions, a stress-free home environment and positive parental aspirations for their children are strong determinants of the educational success of young children. Accordingly, factors that promote parental well-being such as affordable access to childcare and parental mentoring services may improve the home environment and thereby lead to better educational outcomes for children. Whilst Baxter *et al*. (2011) concluded that family structure impacts significantly on the emotional wellbeing of Australian children, this study finds no support for the hypothesis that family structure has a similar impact on their educational success in the early primary school years.

Appendix A: Probit, OLS, and Heckman-Corrected Coefficients (Model V, smoothed income)

	Reading			Writing			Spelling		
	OLS	*Probit*	*Heckman*	*OLS*	*Probit*	*Heckman*	*OLS*	*Probit*	*Heckman*
singleparent	-10.44	0.28**	-6.18	-3.44	0.25**	-2.31	-8.86	0.27**	-8.66
stepparent	-2.79	-0.16	-4.01	-8.08	-0.16	-8.41	1.49	-0.16	1.42
log(income)	24.30***	0.73***	28.27***	17.21***	0.74***	18.39***	15.56***	0.74***	15.77***
atsi	-23.61	0.27***	-19.97	-8.39	0.28***	-7.29	-34.46**	0.29***	-34.27**
ensib	2.18	0.13	2.96	2.19	0.13	2.40	-0.33	0.13	-0.30
enoldsib	-9.50***	0.06*	-10.12***	-5.00**	0.06	-5.17**	-4.10	0.06	-4.12
female	18.32***	-0.06	16.93***	25.20***	-0.05	24.77***	14.78***	-0.05	14.70***
echildage	0.19	-0.11**	-1.01	0.14	-0.11**	-0.23	0.14	-0.12**	0.08
zbirthwgt	0.00	-0.03***	0.00	0.00	-0.03***	0.00	0.00	-0.03***	0.00
ehealth	19.21	0.00	17.66	13.87	0.00	13.25	3.39	0.00	3.29
emetropolitan	-1.72	-0.08	-2.51	4.70	-0.14	4.51	8.64**	-0.14	8.60**
eavparage	1.34**	-0.05	1.33**	0.92**	-0.04	0.93**	0.94*	-0.04	0.94*
epgyear12	0.08	0.00	0.62	-1.60	0.00	-1.51	-12.68	0.00	-12.66
eplyear12	-32.25**	0.04	-32.45**	-26.82**	0.02	-26.95**	-22.20*	0.03	-22.22*
eftemploy	-21.49	0.05	-23.94*	-19.43*	0.04	-20.14*	-25.57**	0.04	-25.69**
eemploy	-8.66	-0.16	-9.10	-8.81	-0.12	-9.00	-10.16	-0.12	-10.19
ephealth	-8.10	-0.02	-6.11	1.69	0.00	2.30	-1.58	0.00	-1.48
ek6depression	4.44	0.10	4.78	4.57	0.10	4.64	0.70	0.11	0.71
zsmoke	-10.99*	0.05	-10.99*	-12.01**	0.05	-11.89**	-6.86	0.05	-6.84
zalcohol	6.26	0.02	7.79	5.48	0.04	5.95	6.87*	0.04	6.94*
ecomputer	1.23	0.13**	4.13	0.04	0.13**	0.86	8.05	0.13**	8.19
esecurehome	-2.59	0.20	-2.39	4.20	0.18	4.25	0.62	0.19	0.62
enliveb	-6.58*	0.03	-5.71	0.49	0.02	0.72	3.05	0.01	3.09
epprob	13.18*	0.05	13.07**	6.92*	0.03	6.87*	1.23	0.04	1.22
eprush	8.31*	-0.03	7.73	4.50	-0.02	4.33	5.53	-0.02	5.50
epncope	-26.89	-0.03	-23.59	-11.10	-0.04	-10.16	-21.65	-0.04	-21.49
epwarm	-4.29	0.36	-4.02	-3.90	0.36	-3.80	-5.20	0.35	-5.19
explyr12	-11.33	0.00	-16.30	-28.55	0.00	-30.07	-12.94	0.00	-13.20

expgyr12	51.05***	-0.23	50.62***	30.21***	-0.23	30.06***	49.24***	-0.24	49.22***
constant	154.60*	-0.05	224.28**	221.91***	-0.05	242.94***	265.35***	-0.05	268.97***
N	1,617	2,698	1,618	1,616	2,698	1,616	1,620	2,698	1,620
F	11.54	-	-	9.23	-	-	9.19	-	-
Prob>F	0.00	-	-	0.00	-	-	0.00	-	-
R^2	0.18	-	-	0.17	-	-	0.15	-	-
ρ	-	0.301	-	-	0.118	-	-	0.015	-
λ	-	24.18	-	-	61.76	-	-	1.02	-
Wald χ^2 *test*	-	2.17	-	-	0.91	-	-	0.01	-
Prob $> \chi^2$	-	0.141	-	-	0.326	-	-	0.926	-
χ^2	-	-	319.58	-	-	269.92	-	-	270.04
Log-PL	-	-	-10,029.09	-	-	-9,667.32	-	-	-9,869.23

Appendix A (continued)

	Grammar and Punctuation			Numeracy		
	OLS	*Probit*	*Heckman*	*OLS*	*Probit*	*Heckman*
singleparent	-4.93	0.26**	-0.99	-10.30	0.25**	-1.34
stepparent	-8.87	-0.18	-9.97	-1.71	-0.14	-3.13
log(income)	26.37***	0.74***	30.48***	23.91***	0.67***	32.00***
atsi	-58.26***	0.28***	-54.37***	-28.64**	0.28***	-21.58
ensib	2.86	0.13	3.56	0.85	0.13	2.51
enoldsib	-8.86**	0.06**	-9.39***	-5.33*	0.06*	-6.63**
female	22.32***	-0.05	20.78***	-7.86**	-0.05	-10.21***
echildage	0.66	-0.12**	-0.58	0.56	-0.12**	-1.78**
zbirthwgt	0.00	-0.03***	0.00	0.01	-0.03***	0.00
ehealth	7.69	0.00	5.46	5.81	0.00	1.37
emetropolitan	2.33	-0.13	1.61	1.02	-0.10	-0.57
eavparage	0.92*	-0.04	0.93*	0.97**	-0.04	0.99**
epgyear12	0.19	0.00	0.47	1.51	0.00	1.54
eplyear12	-22.07	0.02	-22.50	3.58	0.04	1.59
eftemploy	-27.11*	0.05	-29.00*	-7.77	0.03	-13.02
eemploy	-10.34	-0.10	-10.41	-2.14	-0.19	-3.40
ephealth	-2.68	0.03	-0.68	-5.75	-0.05	-0.87
ek6depression	5.29	0.10	5.50	4.00	0.11	5.13
zsmoke	-9.55	0.05	-9.31	-7.55	0.06	-7.73
zalcohol	8.70*	0.03	10.28**	9.58**	0.02	12.41***
ecomputer	-6.60	0.13**	-3.85	6.50	0.11*	11.83
esecurehome	-7.56	0.17	-7.48	-5.78	0.19	-5.60
enliveb	-4.33	0.01	-3.57	-1.69	0.02	0.11
epprob	10.59**	0.04	10.38**	9.31**	0.06	9.07*
eprush	10.01**	-0.03	9.35**	12.71***	-0.04	11.78***
epncope	-29.18	-0.04	-25.78	-23.28	-0.02	-19.10
epwarm	-8.58*	0.36	-8.34*	-8.82**	0.35	-7.98**

explyr12	-28.13	-0.01	-33.15*	-6.21	0.00	-15.02
expgyr12	49.57***	-0.22	49.23***	41.23***	-0.20	40.71***
constant	150.36*	-0.04	223.22**	146.80**	-0.01	276.94***
N	1,619	2,698	1,619	1,614	2,698	1,614
F	12.23	-	-	10.14	-	-
Prob>F	0.00	-	-	0.00	-	-
R^2	0.18	-	-	0.15	-	-
ρ	-	0.305	-	-	0.575	-
λ	-	24.37	-	-	75.06	-
Wald χ^2 *test*	-	1.97	-	-	1.37	-
Prob $> \chi^2$	-	0.161	-	-	0.241	-
χ^2	-	-	343.41	-	-	239.33
Log-PL	-	-	-10,017.76	-	-	-9,791.14

Appendix B: NAPLAN Instrumental Variables Regression (Model VI)

	Reading		Writing		Spelling		Grammar and Punctuation	
	OLS	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*
childnontrad	-11.02	19.59	-6.03	-8.45	-3.87	-2.06	-12.43	9.62
logsmoothinc	24.16***	22.92	16.84***	64.65	16.22***	1.73	25.27***	81.62
atsi	-23.72	-22.53	-8.73	-4.96	-34.25**	-35.37**	-58.78***	-52.56***
ensib	2.09	3.58	1.97	6.15	-0.01	-1.24	2.29	8.42
enoldsib	-9.43***	-11.52***	-4.87*	-2.49	-4.27	-5.38	-8.53**	-6.61
female	18.28***	18.19***	25.24***	25.53***	14.70***	14.90***	22.38***	23.17***
echildage	0.18	0.18	0.13	-0.09	0.14	0.13	0.65	0.57
zbirthwgt	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ehealth	19.16	19.99	13.57	12.17	3.57	4.02	7.28	5.92
emetropolitan	-1.77	-1.65	4.73	-2.46	8.57**	10.64	2.34	-5.60
eavparage	1.29**	1.67**	0.92**	0.53	0.94*	1.05	0.86*	0.72
epgyear12	0.04	-1.22	-1.59	-5.03	-12.89	-12.22	0.40	-2.97
eplyear12	-31.94**	-34.99**	-26.65**	-19.33	-22.44*	-23.79*	-21.31	-17.03
eftemploy	-22.47*	-10.42	-20.72*	-44.12*	-24.77**	-14.65	-29.75*	-48.77
eemploy	-9.73	1.74	-10.28	-27.27	-9.15	-1.41	-13.40	-26.04
ephealth	-8.34	-2.67	1.52	4.75	-1.05	2.09	-3.68	2.35
ek6depression	4.71	3.45	4.86	3.29	0.50	1.13	6.01*	2.62
zsmoke	-10.66*	-16.03*	-11.96**	-5.46	-6.92	-8.00	-9.02	-5.93
zalcohol	6.26	6.73	5.52	1.28	6.85*	7.66	8.74**	4.40
ecomputer	1.00	4.98	0.04	-8.99	8.23	10.61	-7.09	-14.99
esecurehome	-3.23	2.99	3.86	-5.45	0.80	4.84	-8.74	-15.22
enliveb	-6.83*	-6.06	0.42	5.31	2.96	0.82	-4.60	1.74
epprob	13.74***	8.77	7.41*	9.17*	1.04	0.03	11.74**	9.71
eprush	8.53*	6.22	4.89	3.16	5.13	5.31	10.91**	6.84
epncope	-26.95	-30.75	-10.53	-3.86	-22.73	-18.69	-27.74	-22.08
epwarm	-4.22	-4.27	-3.95	-6.28	-5.01	-4.09	-8.74*	-11.71**
explyr12	-10.83	-14.02	-28.31	-27.63	-12.91	-16.99	-27.41	-29.08*

expgyr12	51.24***	51.37***	30.20***	24.90***	49.51***	51.01***	49.50***	43.51***
constant	160.01**	129.21	226.16***	0.41	259.88***	328.74	164.98***	-143.50
N	1,618	1,584	1,616	1,582	1,620	1,586	1,619	1,585
Wald χ^2	-	288.25	-	232.09	-	224.67	-	285.73
Prob > χ^2	-	0.000	-	0.000	-	0.000	-	0.000
F	12.10	-	9.75	-	9.39	-	12.77	-
Prob > F	0.000	-	0.000	-	0.000	-	0.00	-
First stage F(1)	-	47.38	-	46.66	-	46.89	-	45.45
First stage F(2)	-	30.87	-	30.94	-	30.91	-	30.35
χ^2 *test of exogeneity*	-	1.55	-	1.17	-	0.09	-	0.66
Prob > χ^2	-	0.461	-	0.558	-	0.955	-	0.719

Appendix C: PPVT Fixed-and Random-Effects Estimates (Models I-V)

	Model I		Model II		Model III		Model IV		Model V	
	FE	*RE*	*FE*	*RE*	*FE*	*RE*	*FE*	*RE*	*FE*	*RE*
year	3.51***	3.51***	3.47***	3.43***	1.54***	1.90***	1.49***	1.87***	1.58***	1.84***
singpar	-0.27	-1.09***	-0.25	-0.30	-0.23	-0.45**	0.41	0.49*	0.65	0.15
biostep	-0.98*	-1.45***	-0.70	-1.07***	-0.84	-1.06***	-0.89	-0.83*	-0.36	-0.71
log(income)			-0.01	1.21***	0.00	1.02***	0.06	0.81***	0.08	0.73***
atsi					-	-1.06***	-	-0.75*	-	-0.63
ensib					0.55***	-0.13	0.61**	0.03	0.68***	0.01
enoldsib					-0.78**	-0.52***	-0.53	-0.76***	-0.73*	-0.70***
female					-	0.09	-	0.01	-	-0.05
echildage					0.15***	0.13***	0.14***	0.12***	0.13***	0.12***
zbirthwgt					-	0.00***	-	0.00***	-	0.00***
ehealth					0.83**	1.18***	0.74	1.09***	0.76	1.25***
emetropolitan					0.08	-0.23*	-0.09	-0.37***	-0.02	-0.44***
eavparage							0.27*	0.12***	0.29**	0.10***
epgyear12							-0.55	0.31	-0.45	0.28
eplyear12							-0.54	-0.65	-0.23	-0.42
eftemploy							0.05	0.13	-0.02	-0.26
eemploy							0.47	0.85**	0.40	0.48
ephealth							0.67	0.64*	0.54	0.17
ek6depression							0.20	0.39***	0.24*	0.49***
zsmoke							-	-0.61***	-	-0.12
zalcohol							-	0.55***	-	0.48***
ecomputer									0.36	0.79***
esecurehome									0.38	0.07
enliveb									-0.04	-0.23**
epprob									0.09	0.39***
eprush									0.72***	0.81***
epncope									-0.69	0.29
epwarm									1.15***	0.35**

explyr12									-	-1.98***
expgyr12									-	1.39***
constant	$-7.0e^{3}$***	$-7.0e^{3}$***	$-6.9e^{3}$***	$-6.8e^{3}$***	$-3.0e^{3}$***	$-3.8e^{3}$***	$-2.9e^{3}$***	$-3.7e^{3}$***	$-3.1e^{3}$***	$-3.7e^{3}$***
N	10,992	10,992	10,379	10,379	10,372	10,250	8,859	8,176	8,585	7.536
F	6,402.87	-	4,141.47	-	1,854.58	-	875.32	-	585.21	
Prob>F	0.000	-	0.000	-	0.000	-	0.000	-	0.000	
χ^2	-	19,099.48	-	16,620.73	-	16,913.19	-	14,497.10	-	13,695.79
Prob> χ^2	-	0.000	-	0.000	-	0.000	-	0.000	-	0.000
R^2	0.52	0.52	0.52	0.51	0.54	0.51	0.57	0.52	0.58	0.53
Median λ	-	0.43	-	0.41	-	0.41	-	0.43	-	0.40

References

Aughinbaugh, A. & Gittleman, M. (2003). Does Money Matter? A Comparison of Income on Child Development in the United States and Great Britain. *The Journal of Human Resources*. 38(2): 416-440.

Australian Bureau of Statistics (2011). *Australian Social Trends Data Cube – Family and Community: National Summary 1998-2011*. ABS Cat. No. 4102.0. Canberra.

Australian Bureau of Statistics (2012). *Consumer Price Index Australia, June 2012.* ABS Cat. No. 6401.0. Canberra.

Australian Institute of Family Studies (2011a). *Longitudinal Study of Australian Children Data User Guide – August 2011*. Melbourne.

Australian Institute of Family Studies (2011b). *Longitudinal Study of Australian Children Data Dictionary*. Melbourne.

Baxter, J., Weston, R. & Qu, L. (2011). Family Structure, Co-parental Relationship Quality, Post-Separation Paternal Involvement and Children's Emotional Wellbeing. *Journal of Family Studies*. 17(2): 86-109.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*. 75(299): 493-517.

Becker, G. S. (1981). *A Treatise on the Family*. Harvard University Press, Cambridge, Massachusetts.

Becker, G. S. & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labour Economics*. 4(3): S1-S39.

Biblarz, T. J. & Gottainer, G. (2000). Family Structure and Children's Success: A Comparison of Widowed and Divorced Single-Mother Families. *Journal of Marriage and the Family*. 62(2): 533-548.

Blau, D. M. (1999). The Effect of Income on Child Development. *The Review of Economics and Statistics*. 81(2): 216-276.

Duncan, G. J., Yeung, J. W., Brooks-Gunn, J. & Smith, J. R. (1998). How Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review*. 63(3): 406-423.

Ginther, D. K. & Pollak, R. A. (2004). Family Structure and Children's Educational Outcomes: Blended Families, Stylized Facts, and Descriptive Regressions. *Demography*. 41(4): 671-696.

Løken, K. V., Mogstad, M. & Wiswall, M. (2010). What Linear Estimators Miss: Re-Examining the Effects of Family Income on Child Outcomes. Discussion Paper

No.4971. Institute for the Study of Labor.

Mayer, S. E. (1997). *What Money Can't Buy: Family Income and Children's Life Chances*. Harvard University Press, Cambridge, Massachusetts.

National Assessment Program (2011a). *NAPLAN*. Viewed 6 February 2012, Available: <http://www.nap.edu.au>

National Assessment Program (2011b). *Assessment Scales - NAPLAN*. Viewed 6 February 2012, Available: < http://www.nap.edu.au/results-and-reports/how-to-interpret/scales>

Rothman, S. (2005). Report on Adapted PPVT-III and Who Am I? LSAC Data Issue Paper No.2. Australian Council of Economic Research.

Shea, J. (2000). Does Parents' Money Matter? *Journal of Public Economics*. 77(2): 154-184.

Sipthorp, M. & Taylor, M. (2012). Data User Workshop. Presented at Australian Institute of Family Studies, Melbourne, 24 July 2012.

Soloff, C., Lawrence, D. & Johnstone, R. (2005). Sample Design. LSAC Technical Paper No. 1. Australian Institute of Family Studies, Melbourne.

Soloff, C., Lawrence, D., Misson, S. & Johnstone, R. (2006). Wave 1 Weighting and Non-Response. LSAC Technical Paper No. 3. Australian Institute of Family Studies, Melbourne.

Taylor, B. A., Dearing, E. & McCartney, K. (2004). Incomes and Outcomes in Early Childhood. *The Journal of Human Resources*. 39(4): 980-1007.

Thomson, E., Hanson, T. L. & McLanahan, S. S. (1994). Family Structure and Child Well-Being: Economic Resources vs. Parental Behaviour. *Social Forces*. 73(1): 221-242.

Verbeek, M. (2008). *A Guide to Modern Econometrics*, 3rd Edition. John Wiley & Sons Ltd, West Sussex, England.

Violato, M., Petrou, S., Gray, R. & Redshaw, M. (2011). Family Income and Child Cognitive and Behavioural Development in the United Kingdom: Does Money Matter? *Health Economics*. 20(10): 1201-1225.

Wadsworth, M. E. J. & MacLean, M. (1986). Parents' Divorce and Children's Life Chances. *Children and Youth Services Review*. 8(2): 145-159.

Wooldridge, J. M. (2009). *Introductory Econometrics A Modern Approach*, 4^{th} Edition. South-Western Cengage Learning, OH, USA.