Essays on Applied Econometrics

By

Ebrahim Azimi

Graduate Diploma, Institute for Advanced Studies, 2008

M.Sc., Institute for Management and Planning Studies, 2001

B.Sc. Amirkabir University of Technology, 1999

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Economics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Spring 2016

Approval

Name:	**Ebrahim Azimi**
Degree:	Doctor of Philosophy (Economics)
Title:	*Essays on Applied Econometrics*

Examining Committee: Chair: **Bertille Antoine**
Associate Professor, Department of Economics

Simon Woodcock
Senior Supervisor
Associate Professor, Department of Economics

Krishna Pendakur
Co-Supervisor
Professor, Department of Economics

Jane Friesen
Supervisor
Associate Professor, Department of Economics

Brian Krauth
Supervisor
Associate Professor, Department of Economics

Alexander Karaivanov
Internal Examiner
Associate Professor, Department of Economics

Guy Lacroix
External Examiner
Professor, Department of Economics
Laval University

Date Defended/Approved: March 7th 2016

Abstract

This thesis is composed of three essays on economics of labour, family, and education.

In the first chapter, I estimate the effect of having children on labour force participation of mothers in urban Iranian areas. I exploit sex composition of children as an exogenous source of variation in family size to account for endogeneity of fertility. Using information from the Iranian Household Income and Expenditure Survey (HIES) over three samples, namely households with one and more, two and more, and three and more children, I find no significant effect of fertility on female labour force participation in Iran.

In the second chapter, I estimate family member's resource shares and investigate gender bias in intra-household resource allocation. I follow Dunbar et al. (2013) in that I estimate the household member's resource shares by observing how budget shares on private assignable goods vary with total expenditure and family size. I extend their methodology to analyze how sex composition of children influences resource shares. Using data from the 2005 Iranian Household Income and Expenditure Survey (HIES), I find that in Iranian rural areas parents assign 1.6 to 1.9% more resources toward their sons. Similarly, I find that mothers in all-boy families get 2.8 to 3.6% less resources than in all-girl families. These effects are more pronounced among farmer families. In contrast, I find no significant role of gender composition on intra-household resource allocation in Iranian urban areas.

In the final chapter I, jointly with Dr. Friesen and Dr. Woodcock, investigate the question of whether schools that charge private tuition deliver higher quality education compared to their public counterparts has proven very challenging. This paper contributes new evidence regarding the quality of private schools relative to public schools. We use a longitudinal student-level data set from British Columbia, Canada that comprises the entire population of students in fourth through seventh grade who enrolled in public or private schools. We apply a procedure developed by Abowd et al. (2002), which allows us to exploit mobility between schools to estimate a full set of both school and student fixed effects.

Keywords: Female labour force participation, children sex composition, son preference, intra-household resource allocation, school quality, private schools, British Columbia

Dedication

To the loves of:

my wife, Maryam

and

my mom, Jannat

Acknowledgements

I am deeply indebted to Professor Simon Woodcock for his continuous support, advice, and encouragements. His wisdom made my Ph.D. studies a valuable journey. I am especially grateful to Professor Krishna Pendakur for his valuable advice and stimulating suggestions. The second chapter of this paper would not have been possible without his precious help. I am also pleased to thank Professor Jane Friesen and Professor Brian Krauth for their valuable time and support on my research and wonderful advice on my teaching.

I would like to thank all faculties and staffs in Economics department at Simon Fraser University especially Chris Muris, Andrew McGee, Hitoshi Shigeoka, Fernando Aragon, Ramazan Gencay, and Gregory K. Dow for all wonderful discussions and their comments.

I thank Mehdi Majbouri, Hossein Abbasi, Nick Dadson, Ideen Riahi, Natt Hongdilokkul, all the participants of the labour and family session at the Canadian Economic Association (2010) conference at University of Ottawa, participants at the 2011 HAND forum at Massachusetts Institute of Technology, all participants of "International Evidence on Empirical Microeconomics" session in 2014 CEA conference for their helpful comments and discussions. All remaining mistakes are mine.

The administrative data used in the third chapter was extracted from the British Columbia Ministry of Education's student records by Maria Trache at Edudata Canada. Klaus Edenhoffer created the digital maps used to link student postal codes to school catchment areas using information provided by school district personnel. The programs used to create the school choice variables used in the analysis were adapted from codes written by Mohsen Javdani and Benjamin Harris. I appreciate their contribution to this project.

I acknowledge the financial support provided by Simon Fraser University's Community Trust Endowment Fund and the Social Sciences and Humanities Research Council of Canada along this project.

Table of Contents

List of Tables

List of Figures

List of Acronyms

FLFP	Female Labour Force Participation
IV	Instrumental Variable
HIES	Household Income and Expenditure Survey

Chapter 1.

The Effect of Children on Female Labour Force Participation in Urban Iran

In economic literature, children are often considered a barrier to female labour force participation (FLFP). In the last three decades, fertility in Iran has dropped sharply from an average of seven births per woman in 1984 to less than two births in 2005. Although fertility in Iran has experienced one of the fastest declines in modern human history, no considerable rise in FLFP in Iran is documented (Aghajanian 1995; Abbasi-Shavazi et al. 2009; Majbouri 2010). Using household-level information, this paper investigates the effect of children on FLFP of mothers in urban Iran.

The association between fertility and FLFP is extensively documented in theoretical models of work and family. While it is difficult to empirically estimate the endogenous effect of fertility on FLFP (Schultz 1981; Goldine 1995), several studies estimate its causal effect by exploiting an exogenous variation in family size. For example, Rozenzweig and Wolpin (1980) and Bronars and Grogger (1994) use twinning at the first birth. Angrist and Evans (1998) use an instrumental variables (IV) strategy based on sibling sex composition in families with two or more children. Agüero and Marks (2008) exploit random assignment of infertility as an exogenous variation in family size. Using the Iranian Household Income and Expenditure Survey (HIES), this paper contributes new evidence on the effect of fertility on FLFP by using an IV strategy.

I follow Angrist and Evans (1998) strategy to construct IV estimates of the effect of fertility on FLFP based on sex composition of children. While in the US, parents are more likely to have a third child if their first two children are of the same sex, in Iran, as parents prefer sons to daughters, the presence of daughters in their previous children acts as a positive shock to fertility. To show this relationship and investigate the effect of children on FLFP, I construct three samples: one with families with one and more children (1^+), another with two and more children (2^+), and the third with three and more children (3^+). In all these samples, families with more daughters than sons are more likely to have another child. In other words, presence of more girls relative to boys results in an increased likelihood of having another child. Considering this relationship, I construct IV based on the sex composition of previous children to investigate the effect of fertility on FLFP by using dummies for the gender of the first child, the first two children, and the first three children in the samples of 1^+, 2^+, and 3^+, respectively. As sex-selective abortion and infanticide are rare in Iran, I consider sex composition among children as essentially random. To support this claim, I follow Almond and Edlund (2008) by observing that sex ratio does not vary significantly with birth order parity and sex composition of the previous children.

To the best of my knowledge, this is the first estimation of the effect of fertility on FLFP in Iran. While most empirical estimations of the effect of fertility on FLFP find a negative impact, which in most cases is less negative than its ordinary least squares (OLS) counterparts, I find no significant effect of children on the labour force participation of Iranian mothers in urban areas. This result is similar to Agüero and Marks (2008), who find an insignificant effect of fertility on FLFP in six Latin American countries.

The rest of this chapter is organized as follows. In the next section, I present the data. In section 3, I describe the methodology and explain how fertility in Iran is influenced by the sex composition of previous children. Section 4 presents the results, and section 5 concludes.

1.1. Data and Descriptive Statistics

I use data from the HIES (1994–2003). This survey is conducted annually by the Statistical Center of Iran. For each member of a family, this survey contains information on demographic characteristics such as geographic location, age, gender, education, relationship with the householder, marriage status, employment status, occupation, and income. It also contains information on family expenditures, housing characteristics, and ownership of assets and amenities.

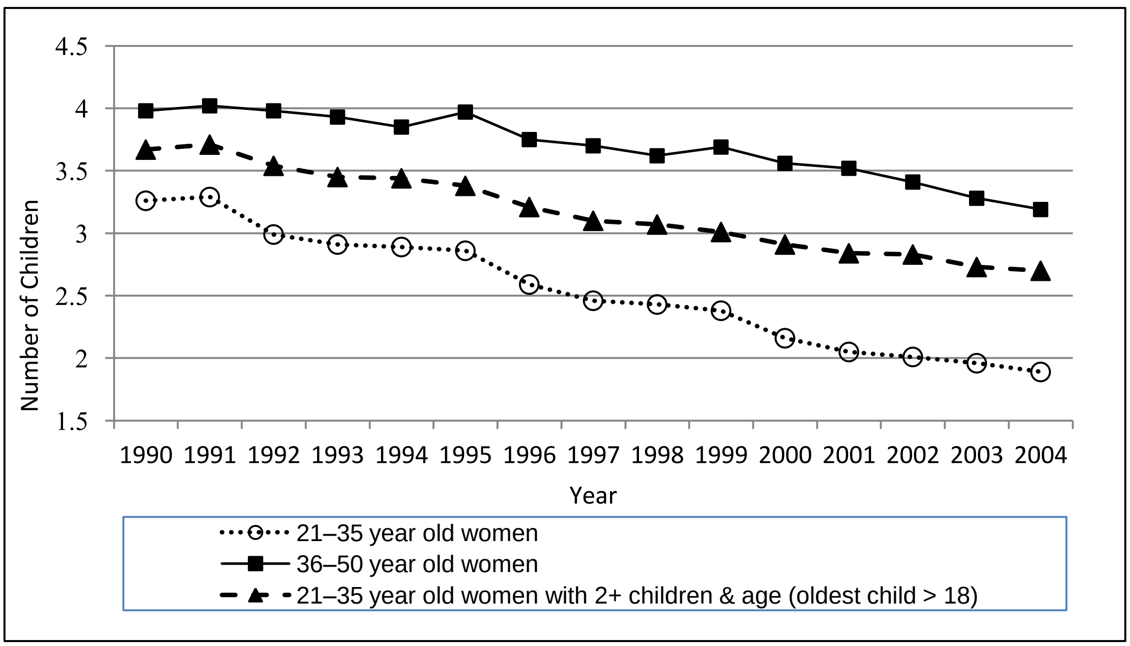
The number of households in the HIES ranged from 17,500 in 1998 to 36,500 in 1995. To ensure that the insignificant effect of fertility on FLFP is not a result of insufficient data, I use 10 rounds of the HIES data (1994–2003), all of which follow the same definition for labour force participation.

The following restrictions are applied to the sample. 1) Polygamous families are excluded from the sample; otherwise, it would have been impossible to match the children to the women. 2) I exclude all cases wherein two or more families share a common residence. Given that household identification, which is based on residential address, is the only way to distinguish families, I am unable to distinguish children of families that share the same residence. 3) Similar to most household surveys, the HIES does not track children according to their households. To match the children with their respective mothers, I restrict the sample to women aged between 20–35 years whose oldest child is younger than 18 years. Few women younger than 19 have two or more children, and women older than 35 are likely to have children who have already migrated from the family. By restricting the sample, I ensure that the family's oldest child is still living with the parents and has not migrated from the family on account of marriage, higher education, or work.

Table A1 in the appendix compares the selected sample with the overall sample of women for some measures of fertility and FLFP. The three samples of women include: (1) women aged between 20–35 years; (2) women aged between 36–50 years; and (3) women aged between 20–35 years with two or more children and whose oldest child is younger than 18. I highlight three features of fertility and FLFP in Iran from 1990–2004 in this table. First, we observe a declining trend in the number of children; second, depending on the year of the interview, a low FLFP rate of 10–14 percent in urban areas, for the selected sample, is observed. The rigidity of FLFP in this period is an important feature that has been addressed in the literature (Majbouri 2010). Third, fertility and FLFP are comparable for all three categories.

Trends in fertility and FLFP are depicted in Figures 1 and 2, respectively, for three subsamples of women from the HIES data.

Figure 1.1: Trend in Number of Children by Age Group and Family Size: 1990-2004



While fertility shows a sharp decline during this period, based on economic theory, we expect an increase in FLFP. However, we observe no such increase over the period in figure 2. The objective of this paper to investigate whether and to what extent fertility impacts FLFP in urban Iran. I continue this discussion in section 4, where I present the results.

Figure 1.2: Trend in Female Labour Force Participation by Age Group and Family Size: 1990-2004

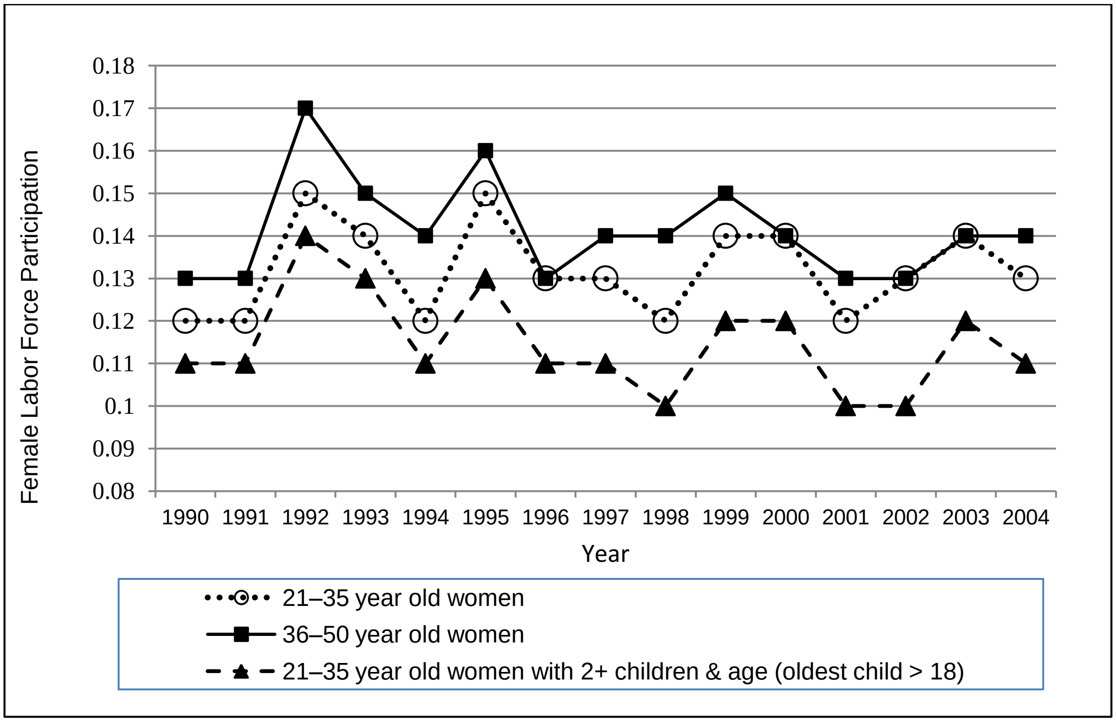


Table 1.1 reports the summary statistics for each sample of the 20–35 year-old mothers in the three samples of 1^+, 2^+, and 3^+.

1.2. Children's Sex Composition and Fertility

Similar to Angrist and Evans (1998), I use the following two-stage least squares (2SLS) regression model of FLFP:

$$x_i = \alpha_0 . w_i + \alpha_1 . z_i + \varepsilon_i \quad (1.1)$$

$$y_i = \beta_0 . w_i + \beta_1 . \hat{x}_i + \eta_i \quad (1.2)$$

Here, x_i is a measure of fertility for woman i; y_i is an indicator of FLFP; w_i includes socioeconomic characteristics of i; and z_i denotes the instrumental variable based on the children's sex composition. The instrumental variable is an indicator for the children's sex composition. The theoretical framework underlying the effect of son preference on fertility is captured by the quality–quantity model of fertility developed by Becker and Lewis (1973). Based on this model, in the current study, parents derive utility from quantity and quality of children. Children sex composition in this model is viewed as a source of utility related to the quality of children. If parents are not satisfied with the sex composition of children, they will be more likely to expand their family to draw utility from quantity of children or from a more desired composition of children. Therefore, having more children is a response to dissatisfaction from children's sex composition.

Table 1.1: Summary statistics: women aged 21–35 years

	1^+ sample		2^+ sample		3^+ sample	
	mean	St. dev.	mean	St. dev.	mean	St. dev.
FLFP	0.129	(0.335)	0.116	(0.320)	0.098	(0.297)
number of children	2.627	(1.431)	3.110	(1.279)	3.917	(1.131)
three-and-more-children indicator	0.447	(0.497)	0.579	(0.494)	1	(0)
four-and-more-children indicator	0.235	(0.424)	0.305	(0.460)	0.526	(0.499)
first-born son indicator	0.514	(0.500)	0.509	(0.500)	0.497	(0.500)
second-born son indicator	0.510	(0.500)	0.510	(0.500)	0.502	(0.500)
two-son indicator	0.260	(0.439)	0.260	(0.439)	0.255	(0.436)
two-daughter indicator	0.241	(0.428)	0.241	(0.428)	0.257	(0.437)
age	28.88	(3.967)	29.69	(3.685)	30.53	(3.376)
age at first birth	19.98	(3.462)	19.25	(3.096)	18.49	(2.745)
years of schooling	6.715	(4.267)	6.023	(4.076)	4.810	(3.727)
presence of relatives in the family	0.089	(0.285)	0.090	(0.286)	0.096	(0.295)
incidence of zero in non labour income	0.812	(0.391)	0.809	(0.393)	0.807	(0.394)
logarithm of non labour income	2.581	(5.426)	2.616	(5.449)	2.622	(5.427)
Observations	56845		43868		25399	

Note: I report the mean of each variable with the standard deviation in parentheses. The variable "age at first birth" is measured by assuming that matching of the woman with their respected children is perfect. Therefore, this variable is measured by error.

Based on the above framework and similar to Ben-Porath and Welch (1976), I analyze the effect of child sex composition on fertility among Iranian families. Table 1.2 reports how the firstborn's sex influences the number of children. In the 1^+ sample, the difference by the first child's sex suggests that families with a firstborn daughter are one percentage point more likely to have a second child. This is consistent with the preference for a son among Iranian parents. Similarly, in the 2^+ and 3^+ samples, the presence of a firstborn son reduces the likelihood of a third and a fourth child. Thus, the firstborn's sex is a plausible instrumental variable for number of children.

Similarly, among families with two or more children, the difference by the firstborn's sex suggests that those with a firstborn daughter are 2.4 percentage points more likely to have a third child (table 1.3). This finding is also consistent with the fact that Iranian parents have a marked son preference, especially for the first birth.

Table 1.2: Fraction of families who had a second child depending on the sex of the first child

	1^+ sample		2^+ sample		3^+ sample	
	fraction of sample	fraction that had a 2nd child	fraction of sample	fraction that had a third child	fraction of sample	fraction that had a 4th child
(1) first-born son	51.3%	0.790	51.0%	0.603	50.6%	0.559
		(0.002)		(0.003)		(0.004)
(2)first-born daughter	48.7%	0.799	49.0%	0.627	49.4%	0.583
		(0.002)		(0.003)		(0.004)
Difference (1)-(2)		-0.009***		-0.024***		-0.024***
		(0.003)		(0.004)		(0.005)

Note: Standard Errors are reported in parenthesis. ***: significant at 1% level.

Similar to the case of the firstborn daughter, the second daughter also increases the likelihood of having a third child. The effect, however, is smaller than that of the firstborn girl. Parents of a second daughter are 1.7 percentage points more likely to have a third child. Further, while parents with two sons are 1.2 percentage points less likely to have a third child, parents with two girls are 4.5 percentage points more likely to have a third child. This also shows that parents of mixed sex children are 2.4 percentage points less likely to have a third child relative to parents of same-sex children.

There are three points worth mentioning about table 1.3. First, all the evidence demonstrates a marked preference for sons among Iranian families; while sons reduce the likelihood of a third child, daughters increase it. Second, the effect of two daughters is larger compared to other combinations. Third, this table shows that child sex composition has a strong explanatory power on fertility decisions in Iranian families. Based on these results, I use an indicator of two daughters as an instrumental variable to estimate the effect of a third child on FLFP in the 2^+ sample.

Similarly, table 1.4 reports the relation between sex composition of children and probability of having a fourth child in families with at least three children. Based on this table, I use the indicator of having at least two sons as an instrument to investigate the effect of a fourth child on FLFP in the 3^+ sample.

Random assignment of sex composition makes it very likely that these IV estimates of the effect of fertility on FLFP have a causal interpretation. Selective abortion and infanticide are quite rare in Iran because both have been illegal since the 1979 Islamic Revolution, except in very specific cases where the mother is in serious danger or the baby is expected to born with a severe disease (Hoodfar 1996; Mehryar et al. 2007). Thus, we treat the gender composition of children as essentially random.

Table 1.3: Fraction of families who had another child depending on the sex composition of previous children: 2^{+} sample

	fraction of sample	fraction that had a third child
(1) first-born son	51.0%	0.603 (0.003)
(2) first-born daughter	49.0%	0.627 (0.003)
Difference: (1)-(2)		-0.024*** (0.004)
(1) second-born son	51.1%	0.607 (0.003)
(2) second-born daughter	48.9%	0.624 (0.003)
Difference: (1)-(2)		-0.017*** (0.004)
(1): two sons	26.2%	0.607 (0.004)
(2):not(two sons)	73.8%	0.618 (0.002)
Difference: (1)-(2)		-0.012*** (0.005)
(1): two daughters	24.0%	0.649 (0.004)
(2): not (two daughters)	76.0%	0.604 (0.002)
Difference: (1)-(2)		0.045*** (0.005)
(1): mixed sex	50.2%	0.627 (0.003)
(2): same sex	49.8%	0.603 (0.003)
Difference: (1)-(2)		0.024*** (0.004)

Note: Standard Errors are reported in parenthesis.
***: significant at 1% level.

Table 1.4: Fraction of families who had a fourth child depending on Parity and sex composition of previous children

	fraction of sample	fraction that had a 4th child
(1): 3 sons	13.6%	0.560
		(0.007)
(2): other mixes	86.4%	0.573
		(0.003)
Difference: (1)-(2)		-0.013**
		(0.007)
(1): 3 daughters	12.7%	0.557
		(0.007)
(2): other mixes	87.3%	0.573
		(0.003)
Difference: (1)-(2)		-0.015**
		(0.008)
(1): same sex	25.8%	0.594
		(0.005)
(2): mixed sex	74.2%	0.563
		(0.003)
Difference: (1)-(2)		0.031***
		(0.006)
(1): 2 sons, 1 daughter	38.0%	0.548
		(0.004)
(2): other children compositions	62.0%	0.585
		(0.003)
Difference: (1)-(2)		-0.036***
		(0.005)
(1): 1 son, 2 daughters	36.2%	0.578
		(0.004)
(2): other children compositions	63.8%	0.567
		(0.003)
Difference: (1)-(2)		0.011**
		(0.005)

Note: Standard Errors are reported in parenthesis. ***: significant at 1% level, **: significant at 5% level.

One empirical test for random assignment of child sex composition is to compare the sex ratio by birth order and sex of previous children (Almond and Edlund 2008). For the general human population, sex ratio, defined as the proportion of males to females, is about 1.05 with the exception of during and after war times, where it is documented to be slightly higher. Otherwise, a higher-than-normal sex ratio is a sign of sex-selective abortion. In such cases, the sex ratio will depend on previous children's sex composition. For example, due to the availability of prenatal sex determination, a two-child family that already has a daughter and prefers sons is more likely to abort a girl fetus relative to a boy fetus. Table 1.5 reports sex ratio by birth order and previous children's sex composition. No significant inflated sex ratio is evident from this table. Although the sex ratio is slightly larger than 1.05 for the third child, there is no significant difference between the sex ratios of families with two sons and families with two daughters. Thus, child sex composition can be treated as random. The results of the 1996 and 2006 Iranian Censuses in table A2 in the appendix support the randomness of sex composition in Iran.

Table 1.5: Sex ratio by parity and sex composition of previous children

birth order	previous children	observations	sex ratio	95% confidence interval	
				lower bound	upper bound
first		132,599	1.055	1.044	1.066
second	girl	51,689	1.032	1.015	1.050
	boy	53,920	1.049	1.031	1.066
third	girl, girl	17,356	1.082	1.051	1.115
	girl, boy	33,703	1.067	1.044	1.090
	boy, boy	17,685	1.080	1.048	1.112

An alternative test for random assignment of child sex composition is to compare demographic characteristics of families by the sex composition of children (Angrist et al. 1998). If child sex composition is random, there is no significant difference between demographic characteristics of families by this composition, as the insignificant differences in table 1.6 suggest.

Table 1.6: Difference in mean for demographics by sex composition of children

	a first-born son	two daughters	two or more sons
age	-0.012	0.030	-0.020
	(0.026)	(0.034)	(0.036)
literacy	-0.002	-0.004	0.006
	(0.003)	(0.005)	(0.006)
education	-0.013	-0.050	0.051
	(0.035)	(0.044)	(0.044)
husband's years of education	-0.004	-0.042	0.081
	(0.038)	(0.049)	(0.053)
Sample	1^+ sample	2^+ sample	3^+ sample

Note: Standard errors are reported in parentheses.

The random assignment of child sex composition makes it very likely that the reduced form regressions of fertility and FLFP have a causal interpretation. Section 4 reports this estimation, and the results confirm that the three dummies of firstborn son, two daughters, and two sons and more are plausible instrumental variables for investigating the effect of fertility on FLFP in the 1^+, 2^+, and 3^+ samples, respectively. I use number of children and indicators of having more than two and three children as the measures of fertility for the 1^+, 2^+, and 3^+ samples, respectively.

1.3. Estimation Results

In this section, I estimate the effect of having more children on FLFP, using the 2SLS regression model of FLFP shown in equations (1.1) and (1.2). As explained earlier, I estimate the model for three subsamples of women. I use number of children, an indicator denoting that a woman has more than 2 children, and indicator denoting that a woman has more than three children as measures of fertility for the samples 1^+, 2^+, and 3^+, respectively. The respective instrumental variables are indicators of a firstborn son, two daughters, and two or more sons.

All the specifications include indicators of age, namely, I(25 ≤ age ≤ 29), I(30 ≤ age ≤ 35); indicators of schooling, namely, I(1 ≤ schooling ≤ 5), I(6 ≤ schooling ≤ 8), I(9 ≤ schooling ≤ 12), I(13 ≤ schooling); age at first birth, log(nonlabour income), and year effects.

Table 1.7 reports the results of the OLS and IV estimates of the effect of children on FLFP for all three samples as well as the results of the first-stage estimates.

All three samples confirm the strong association between fertility and child sex composition. The F-statistics of a test for weak instrument hypothesis based on Stock and Yogo (2005) strongly rejects the null hypothesis of a weak instrument. The IV estimates suggest no significant effect of fertility on FLFP, while the OLS estimates report negative effects. Although the OLS estimates are small, they are significant. Most studies of the causal effect of fertility on FLFP find a negative effect, which is usually smaller than the OLS estimates but is still significant. For example, using the firstborn's sex as an instrumental variable; Chun et al. (2002) find that an additional child reduces the labour force participation of Korean mothers by 27%. Angrist et al. (1988) estimate the effect of a third child on women's income to be -0.12. On the other hand, Agüero and Marks (2008) find no evidence that fertility has a causal effect on FLFP.

Although the insignificant effect of fertility on FLFP is consistent with the aggregate trend of FLFP for urban Iranian families (see Figures 1 and 2), the result is nevertheless surprising.

As explained in section 2, I restrict the sample according to the ages of the women and their oldest child, to match the women with their respective children. If the match is not close to perfect, the estimate of the variable "age at first birth" will be erroneous resulting in bias in estimates. In my result, however, excluding this variable does not change the main finding of the paper; that is, fertility continues to have an insignificant effect on FLFP.

Table 1.7: OLS and 2SLS estimates of the effect of fertility on FLFP in urban Iran

	1^+ sample		2^+ sample		3^+ sample	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
First-stage results: fertility equation						
a first-born son		-0.106***				
		(0.009)				
two daughters				0.042***		
				(0.004)		
two or more sons						-0.044***
						(0.005)
Estimation Results: FLFP equation						
number of children	-0.008***	0.005				
	(0.001)	(0.026)				
more than 2 children			-0.011***	-0.033		
			(0.004)	(0.060)		
more than 3 children					-0.009**	-0.019
					(0.004)	(0.068)
Observations	56,845	56,845	43,868	43,868	25,399	25,399
1st stage R-squared		0.562		0.344		0.226
2nd stage R-squared	0.183	0.183	0.152	0.151	0.084	0.083
IV F-statistics		149.253		149.430		89.917

Note: All the specifications include indicators of age: I(25≤age≤29), I(30≤age≤35); indicators of schooling: I(1≤ schooling≤ 5),I(6≤ schooling ≤8),I(9≤ schooling≤ 12), I(schooling≤ 13); age at first birth, log(non-labour income), incidence of zero in non-labour income and year effects. Standard errors are reported in parentheses. ***: significant at 1% level, **: significant at 5% level.

One potential concern with the estimation of FLFP equation is endogeneity of nonlabour income as it highly depends on previous labour income (Lise and Seitz, 2011). The endogeneity of income can potentially create a bias to the estimated effect of children on FLFP. To control for this endogeneity problem, one potential instrument is changes in liquidity such as changes in local housing market prices (Hurst and Lusardi, 2004). In this study, however, using similar instrument is not feasible with data restriction. Therefore, as a robustness check, I estimate the FLFP equation by excluding the non-labour income. No effect of children on FLFP is robust to the omission of nonlabour income which reduces concerns about the endogeneity of nonlabour income in this study.

My finding suggests that children at the extensive margin are not a barrier for the female labour supply. However, one possibility is that fertility influence labour supply of women at the intensive margin. That is, women may change their work hours in response to having a larger family. Unfortunately, information on hours of work are not available in the HIES. Therefore, it is not feasible to estimate the effect of children on female labour supply at the intensive margin. Even if this is a plausible explanation, the rigidity of FLFP at low levels is still surprising.

The low rate of FLFP in Iran and its rigidity over the last three decades have been referred to as a puzzle in development literature (Majbouri 2010). It is more surprising to know that this rigidity was concurrent with a sharp decline in fertility, as mentioned earlier, and a considerable increase in the education of women. For example, according to the Statistical Center of Iran, the female-to-male student ratio in Iranian public and private colleges has increased from less than 0.4 in 1990 to about 1.2 in 2005.

With the sharp decline in fertility and the considerable increase in women's education in Iran, we might expect an increase in FLFP. However, FLFP has remained low at 10–14 percent. While investigating the reasons behind the low rate and rigidity of FLFP is outside the scope of this paper, I state a few potential reasons. One possible explanation is the hidden employment resulting from the definition of employment in governmental surveys that measure FLFP including HIES. These surveys consider someone employed only if they work at least two days per week. This may preclude people employed on a part-time basis. Another possible explanation is that traditional norms stemming from religion have a strong role in impeding women from seeking employment. In addition, employment of women can increase their autonomy within the family and it reduces the autonomy of their husband. The fact that women employment influences the intra-household bargaining power provides further explanation on barriers to FLFP in Iran.

The recent development literature provides some additional possible explanations for low FLFP in the region and its rigidity in the last few decades. Many observers believe that religion is the main reason for the low representation of women in the labour market (Sharabi 1988). They emphasize the common problem of low FLFP throughout the Middle East and attribute it to the influence of traditional and religious norms. The suggested mechanism is that traditional norms result in discrimination against women in the labour market, by restricting both labour supply and labour demand for women. This explanation, however, does not seem plausible as countries like Bangladesh and Indonesia are predominantly Muslim but have high rates of FLFP.

Using cross-country data, Ross (2008) shows that the effect of Islam disappears as he controls for oil and gas income, and he concludes that oil, not Islam, is responsible for low rates of FLFP in Iran and other Middle Eastern countries. Majbouri (2015) challenges this argument by proposing a mechanism through which oil and gas income along with traditional institutions account for the rigidity of FLFP. He explains that oil and gas income acts as rent and strengthens traditional norms and the religion's influence among Muslim countries with access to oil income.

1.4. Conclusion

Rigidity of FLFP in urban Iran in the last three decades has been a consensus in development literature. It is of more surprise to know that it has been simultaneous with the period in which fertility has sharply declined and women's education has considerably increased (Majbouri 2010). I investigate the causal effect of fertility on FLFP to shed light on a part of this puzzle.

Following Angrist et al. (1988), I exploit the random assignment of child sex composition as an instrument to investigate the causal effect of fertility on FLFP among Iranian families in urban areas. As Iranian parents prefer to have sons relative to daughters, I show that the presence of sons reduces the likelihood of having more children in Iranian families. Based on this information, I exploit children's sex composition to investigate the effect of fertility on FLFP.

While most estimates of the causal effect of fertility on FLFP report negative effects, I find no evidence that the presence of more children is a barrier for mothers to work. This finding is similar to that of Agüero and Marks (2008) and consistent with the aggregate trend in FLFP in urban Iran. Oil and gas income along with traditional institutions in Iran are considered to account for the rigidity of FLFP (Majbouri, 2015).

Chapter 2.

Intra-household Resource Allocation and Gender Bias in Iran

Intra-household resource allocation has been widely studied in analyzes of gender bias, poverty, and standards of living. The gender gap observed in adult outcomes is generally thought to be a consequence of gender bias in intra-household resource allocation (Deaton, 1997). While preference for sons as well as the gender gap in a number of socioeconomic has been reported in Iran, intra-household resource allocation and its gender bias consequences remain open questions. Given the foregoing, I estimate the resource share allocated to men, women, and children among Iranian families in rural and urban areas.

In particular, I investigate two potential consequences of son preference[1] on intra-household resource allocation in Iran. First, I examine whether and to what extent son preference among Iranian parents results in the allocation of family resources in favor of boys relative to girls. Second, I investigate the effect of family gender composition on the resource share of parents. If parents prefer boys relative to girls, the trade-off between their own consumption and their children's consumption is expected to differ for boys and girls; specifically, there would be relatively less consumption for parents when they have sons. Furthermore, I expect the extent of this trade-off to differ between parents. As Iranian culture is strongly patriarchal (Moghadam, 1992), the effect of more sons in a family on the mother's resource share is of great scholarly importance.

[1] Son preference refers to the attitude based on which a female child is less valued than a male child by her parents and the society.

The effect of patriarchy on women's resource share is ambiguous because at least two opposing forces exist. First, since giving birth to a son is a source of pride and privilege for a woman, the presence of more boys in the family leads to more autonomy for the mother, resulting in her higher resource share. In this case, the father makes the largest sacrifice by allocating more of his own resources to his son. Second, the presence of boys in the family may create competition between the mother and her sons, resulting in a relatively low resource share for the mother. While both views are similarly plausible, I investigate which one is consistent with household behavior in Iran.

This study provides the first estimates of intra-household resource allocation in Iran, and quantifies the effect of gender composition on resources allocated to each family member among Iranian families. This has two principle implications in welfare economics. First, it sheds light on a major source of economic inefficiency, especially in developing countries. Gender bias in intra-household resource allocation is a source of discrimination that could lead to inequality in adulthood by negatively influencing the well-being of female offspring and promoting a gap in the educational and labour market outcomes of girls compared with boys. Second, the estimation of the resource shares of household members helps measure well-being at an individual level. If households unevenly allocate resources among family members, then a household-level analysis of poverty, for example, might mask the effect of adverse economic conditions on individual family members.

Unitary models of households are unable to examine the within-household distribution of consumption (i.e., the resource shares of household members). By contrast, collective household models (e.g., Chiappori, 1988, 1992), wherein the household is modeled as a collection of individual people with distinctly defined utility functions, are more suitable for assessing distribution among household members. The main contribution of this study is thus using a collective household model to investigate empirically gender bias in intra-household resource allocation.

In general, the effect of gender composition on the resource allocation of Iranian families is difficult to study because household surveys usually provide expenditure information at the household level and not at the individual level. This study uses the collective household model of Dunbar et al. (2013) as well as information on assignable goods in household-level expenditure data plus a structural model of demand to estimate the share of household resources allocated to each family member. The approach of Dunbar et al. (2013) identifies resource shares by observing how family expenditure on private assignable goods such as clothing varies by household type and total expenditure. I use the Iranian Household Income and Expenditure Survey (HIES) 2005, and, like Dunbar et al., use clothing as the private assignable good. But, in contrast to Dunbar et al., I focus on how resource shares depend on gender composition of the family.

The estimated results, the first on intra-household resource allocation in Iran to my best knowledge, suggest that Iranian families in rural areas allocate more resources toward their sons. Depending on family size, Iranian parents in rural areas devote 15–17% of their resources toward boys. The share of girls, however, is on average 1.6–1.9 percentage points less than that of boys. In addition, the presence of boys in the family diverts resources from their mothers. In all-boy families, mothers obtain fewer resources than those in all-girl families, by 2.8–3.6 percentage points. By contrast, the results from families in urban areas suggest no strong effect of gender composition on household members' resource shares.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 summarizes the econometric methodology. Section 4 describes the data set. Section 5 presents and discusses the empirical results. Section 6 concludes.

2.1. Iranian Context and Related Literature

Studies of economic well-being are of particular interest for Iranian households because the country has experienced three decades of adverse economic conditions. The revolution of 1979, the 1980–1988 war with Iraq, and international economic sanctions have all negatively influenced Iranian families' standards of living. In addition, considering the patriarchal culture among Iranian families, it is important to account for within-family inequality when measuring individual well-being, especially in developing countries such as Iran where family decisions play a dominant role in children's future options. In Iran, for example, families take on the roles assigned to other social institutions in more developed countries. In this regard, Iranian households decide whether and to what extent a child should be educated. By contrast, because the law in developed countries enforces minimum education levels, there is less potential for discrimination against girls.

If resource allocation in Iranian families is discriminatory against female offspring, then it may have wide-ranging influences on many dimensions of their lives. Indeed, gender bias in intra-household resource allocation could explain the differences in many socioeconomic outcomes, such as the higher mortality rate of girls relative to boys (UNFPA, 2010), their lower secondary school enrollment rate (UNICEF, 2007), and poor female labour market outcomes (Abbasi and Farjadi, 1999). Indeed, Azimi (2015) finds that the gender composition of children even influences Iranian parents' decisions regarding fertility.

Detecting gender bias in intra-household resource allocation is a challenging task. If consumption were observed at the individual level, one could directly observe whether some members received a greater share of household resources than others. Unfortunately, household consumption surveys usually provide information on various types of commodity expenditures at the household level not at the individual level. Therefore, drawing conclusions from these data about gender bias at the household member level is not straightforward.

Some of the stream of research on intra-household resource allocation uses information on aggregate consumption (e.g., using expenditure on food, education, or healthcare as a proxy) and family composition to investigate how household expenditure needs depend on the age and gender composition of families. For example, Deaton (1997) uses a regression of food budget share on the natural logarithm of family size, the natural logarithm of per capita total expenditure, the proportion of family members in each age–sex group, and household demographics. The share of food in family total expenditure is considered as an inverse measure of household welfare. The coefficient of each age–sex group identifies the increase in expenditure share of the analyzed good as a result of adding one more member in that group. Although researchers have used this method to analyze a number of countries including India, Bangladesh, Pakistan, and China (Subramanian and Deaton, 1991; Bhalotra and Attfield, 1998; Gong et al., 2000), establishing that household consumption needs are related to the gender composition of families; this finding does not imply that consumption is shared unequally within the household.

Although the analysis of household-level food, health, or educational consumption is not in general informative about discrimination in the intra-household resource allocation, the household-level consumption of exclusive goods is informative in this regard. An exclusive good is one consumed by a single identifiable household member. Such goods are thus useful because the household-level consumption of the good equals the individual-level consumption. The analysis of exclusive goods is a common approach for investigating gender bias in intra-household resource allocation. In this context, however, we need goods exclusively consumed by different genders (and potentially ages).

Rothbarth (1943) first applied this concept to observe how the slope of the Engel curve for an adult good such as alcohol, tobacco, or adult clothing varies with the gender of children. In this case, the consumption of adult goods is a proxy for parents' welfare. The slope of the Engel curve shows the extent to which adults are willing to trade off their own welfare as well as that of their children. If the extent of this trade-off is different for boys and girls, it is a sign of gender bias. While this approach originally estimated the cost of children, researchers have since used it extensively to estimate intra-household resource allocation. Gronau (1988) shows that the theoretical basis for the Rothbarth methodology is the assumption of separability and constant preferences. Alcohol, tobacco, clothing, and footwear are among the exclusive goods used in the literature. This method and its extensions have been widely used for intra-household resource allocation studies in India and China (Lancaster et al. 2003; Gong et al. 2000; Kingdon 2005; Zimmerman 2012).

Two main criticisms arise from the assumptions of previous studies on the household-level consumption of shared or exclusive goods. First, earlier research has typically used a unitary model of households wherein the household optimizes a single utility function rather than an amalgam of each member's utility function. Investigating intra-household resource allocation and gender bias based on a model that comprises one utility function (i.e., only one person) seems to be unsuitable. Second, these methods do not account for returns to scale in consumption. Adding the same resources to families of different sizes can result in different welfare. Since large families usually benefit from economies of scale in consumption, research on intra-household resource allocation should overcome this discrepancy.

This study avoids these drawbacks by following Browning et al. (2013) and Dunbar et al. (2013), who propose collective households (i.e., they model the household as a collection of members with individual utility functions) that allow for scale economies in consumption. In particular, I adapt those models to address the possibility that the intra-household distribution of resources depends on the gender composition of children.

2.2. Econometric Methodology

The collective model of Dunbar et al. (2013), based on that of Browning et al. (2013), proceeds as follows. Households consist of three types of individuals t={f,m,c}, denoting the father, mother, and children, respectively. The number of children s={0,1,2,3} is used to index household type. Let y denote the household's total expenditure and p denote the market price vector.

Household consumption decisions take into account the utilities of each person; hence, I assume that such decisions reach the Pareto frontier in the sense that trade across household members does not lead to unexploited gains. This assumption allows me to decentralize the household problem into a separate optimization problem for each member. Separability assumption also allows for separability of consumption and production.

The personal optimization problem aims to maximize utility against a budget constraint characterized by a shadow price vector *r* (which is the same for all household members) and a shadow budget $\eta_{ts}y$.

The shadow price vector *r* does not equal the market price vector *p* because shareable goods have shadow prices below their market prices. These shadow prices capture what Browning et al. (2013) call the "consumption technology," while the differences between the shadow and market prices account for scale economies in household consumption. Importantly, the shadow price of an exclusive good is equal to its market price, because exclusive goods are by definition not shareable.

The scalar η_{ts} is the resource share of person t in a household with s children. The resource share is a scalar-valued measure of an individual budget constraint within the household, and it measures the amount of consumption flowing to that person (i.e., the object of interest in this paper).

Define $W_{ts}(y)$ as the share of a household's total expenditure spent by member t on his or her private assignable good in a household of type s. Let $w_t(y)$ be the hypothetical share of y that t would spend on his or her private good when maximizing his or her own utility function subject to shadow price r and budget y. In general, Dunbar et al. (2013) show that if the data have no single-member households, further structure is necessary to identify the resource shares and thus they propose the restriction that resource shares are independent of y and a pair of preference restrictions, which allows them to identify resource shares from the variation in the Engel curve.

These authors also specify the solution to this household optimization problem in the form of an Engel curve relating budget shares $W_{ts}(y)$ to budgets y, as follows:

$$
\begin{aligned}
W_{fs}(y) &= \eta_{fs} \, . \, w_f(\eta_{fs}\, y) \\
W_{ms}(y) &= \eta_{ms} \, . \, w_m(\eta_{ms}\, y) \\
W_{cs}(y) &= s \, . \, \eta_{cs} \, . \, w_c(\eta_{cs}\, y)
\end{aligned}
\tag{2.1}
$$

In Eq. (2.1), $W_{ts}(y)$ is observable from the information available on household expenditure information. Here, I assume that η_{ts} does not depend on y. The goal is to identify the resource share in Eq. (2.1). However, the challenge in identifying the resource share in this equation is that for every observable $W_{ts}(y)$ on the left-hand side, two unknown functions of resource share exist on the right-hand side, namely η_{ts} and $w_t(\eta_{ts} y)$. This is where the preference restrictions come in.

Dunbar et al. (2013) impose that the functions $w_t(\eta_{ts} y)$ have similar shapes (essentially fixed curvatures) for either variation in household size s or variation in person t. Under this structure, no further restriction on the shape of the preference functions $w_t(\eta_{ts} y)$ is necessary to identify the resource shares.

For estimation simplicity, let preferences be price-independent generalized logarithmic (PIGLOG; Muellbauer, 1976). In this case, demands are given by

$$
\begin{aligned}
W_{fs}(y) &= \eta_{fs}\left(\delta_{fs} + \beta_{fs}\ln\eta_{fs}\right) + \eta_{fs}\,\beta_{fs}\ln y \\
W_{ms}(y) &= \eta_{ms}\left(\delta_{ms} + \beta_{ms}\ln\eta_{ms}\right) + \eta_{ms}\,\beta_{ms}\ln y \\
W_{cs}(y) &= s\eta_{cs}\left(\delta_{cs} + \beta_{cs}\ln\eta_{cs}\right) + s\eta_{cs}\,\beta_{cs}\ln y
\end{aligned}
\tag{2.2}
$$

These two restrictions are when preferences are similar across people (SAP) or when preferences are similar across types (SAT). The SAP condition for a PIGLOG indirect utility function is equivalent to $\beta_{ts}=\beta_s$ for all individual types. Under SAP, the linear regression of the private good's household budget shares on ln(y) allow us to identify the slopes of the abovementioned household Engel curves. Considering that resource shares sum to one for each household type, we can identify three resource shares and β_s. Alternatively, the SAT condition for a PIGLOG indirect utility function is equivalent to $\beta_{ts}=\beta_t$ for all household and individual types. Considering four household types, Eq. (2.3) provides 12 equations in addition to the three sets of resource shares equal to one. This enables us to identify the three resource shares for each household type plus the three preference parameters, β_t. With more than four household types, the model is overidentified.

We can use additional information to test the model or improve the precision of the estimates. Dunbar et al. (2013) combine both restrictions by imposing $\beta_{ts}=\beta$. In this study, I use information on Iranian families with zero to three children. The presence of childless families makes the SAT restriction a strong assumption. Therefore, I estimate the parameters by assuming that a combination of SAT and SAP holds for families with one to three children (β is allowed to differ for childless families). This assumption can be stated as $\beta_{ts}=\beta_1$ for s=1,2,3 and $\beta_{ts}=\beta_0$ for s=0. The corresponding household Engel curves for the private assignable goods under the assumption of a combination of SAP and SAT are

$$
\begin{aligned}
W_{fs}(y) &= \eta_{fs}\left(\delta_{fs} + \beta \ln \eta_{fs}\right) + \eta_{fs}\,\beta \ln y \\
W_{ms}(y) &= \eta_{ms}\left(\delta_{ms} + \beta \ln \eta_{ms}\right) + \eta_{ms}\,\beta \ln y \qquad (2.3) \\
W_{cs}(y) &= s\eta_{cs}\left(\delta_{cs} + \beta \ln \eta_{cs}\right) + s\eta_{cs}\,\beta \ln y
\end{aligned}
$$

where $\beta=(\beta_1 * I_{s>0}+\beta_0 * I_{s=0})$.

I use the following parametric specification that allows resource shares to depend on household size and the gender composition of children:

$$\eta_{ms} = (\alpha_{0m} + \alpha_{1m} * s + \alpha_{2m} * s^2)(1 + \gamma_m * \frac{s_g}{s})$$

$$\eta_{cs} = (\alpha_{1c} * s + \alpha_{2c} * s^2)(1 + \gamma_c * \frac{s_g}{s}) \qquad (2.4)$$

$$\eta_{fs} = 1 - \eta_{ms} - s\eta_{cs}$$

This specification consists of a quadratic function of family size multiplied by a linear function of the proportion of girls in the family. $\alpha_{0t}, \alpha_{1t}, \alpha_{2t}$ are coefficients of the quadratic term, while s_g is the number of girls and γ_t the coefficient of the proportion of girls for each family member, t. This specification has a number of useful features. First, the quadratic functional form allows returns to scale in consumption. Second, the intercept in the female and male equations identifies the resource share in a childless family. Third, the children equation does not include an intercept because children's resource share in childless families is zero. Finally, the extent to which each member is influenced by the presence of female children in the family is proportional to γ_t.

To estimate this model, I add an error term to each equation and use nonlinear seemingly unrelated regression techniques to estimate the resource shares on a sample of Iranian households with zero to three children.

2.3. Data

The HIES is an annual survey conducted by the Statistical Center of Iran. For each family member, the HIES reports demographic information such as geographical location, age, gender, education, relationship with the family head, marital status, employment status, occupation, income, and sources of income. The focus of this survey is, however, on household expenditure, for which the HIES reports information on about 600 items. Expenditure on clothing and footwear is reported separately for male adults, female adults, and children in different age groups. I combine the children age groups to construct clothing and footwear expenditure for children younger than 14 years. The recall period for consumption expenditure is one month. To capture seasonal effects, households are interviewed in different seasons (the interview season is reported). I use information on interview season to test the robustness of results when I account for any price variation over a year.

The HIES included about 24,000 families in 2005. I restrict the sample to childless couples and families composed of married couples with one to three children. Furthermore, I exclude polygamous families, households with any member older than 65 years, two or more families sharing a common residence[2], and families with children older than 14 years. Finally, I limit the sample to families comprising a father, mother, and zero to three children (i.e., I omit families that include other relatives such as grandparents). The final sample includes 7,496 households. While it is technically difficult to measure the effect of the above sample restrictions, it is unlikely that these restrictions change the main results. In section 5, however, I briefly explain a potential consequence of the sample restrictions.

I use clothing as the exclusive good for adults. The exclusion of families with children 14 or over allows us to assign clothing expenditure to either adults or children, and since the data separate clothing expenditure for men and women, to husbands and wives[3]. For children, I create a children's good that includes both clothing and other children's private goods such as toys, dolls, and children's clothing. Table 2.1 presents the means of the exclusive goods for different family members, total household expenditure, and demographic characteristics for the sample by family size.

Further, a rich set of demographic variables from the HIES is used in this study: an indicator of urban–rural residential area, age of father, age of mother, average age of children, age of the youngest child, father's education, mother's education, indicator of parents' occupation in farms, and indicators of the interview season.

The inclusion of demographics, although not required to identify resource shares, helps to identify the share more precisely. Hence, I estimate the resource shares for a reference family (i.e., a family for which all the demographics have zero values). For all

[2] Given that household identification (based on residential address) is the only way in which to distinguish families, I am unable to decompose those families that share the same residence.

[3] Because four-child families in the sample were rare, I restrict the sample to couples with zero to three children.

age and educational variables where zero values are rare or impossible, I include the deviation from the modal values. Specifically, I measure the ages of the father and mother as deviations from 35 and 30 years, respectively and the average age of children and age of the youngest child as deviations from five. Similarly, I measure parents' education levels as deviations from five. Five years of schooling is equivalent to completing primary school, which is the most common level of education among parents in the sample. Using deviation from the mode compared with deviation from the mean ensures a sufficient proportion of reference families.

Table 2.1: Data Means

	couples with				
	no child	1 child	2 children	3 children	0-3 children
men's clothing and footwear	0.028	0.022	0.019	0.018	0.021
women's clothing and footwear	0.023	0.018	0.014	0.012	0.016
children's private good	0	0.035	0.042	0.047	0.033
number of children	0	1	2	3	1.45
number of daughters	0	0.48	0.98	1.46	0.71
log(total expenditure)	14.5	14.6	14.7	14.6	14.6
men's age(demean)	1.58	-2.81	-0.45	0.73	-0.67
women's age(demean)	1.42	-2.61	-0.24	1.05	-0.51
men's schooling (demode)	2.98	3.44	2.61	1.19	2.75
women's schooling (demode)	2.25	3.10	1.75	-0.38	1.98
average age of children	-5	-0.47	1.21	2.04	-0.35
children age difference	0	0	3.98	6.90	2.38
first born girl indicator	-	0.48	0.51	0.49	-
proportion of girls	-	0.48	0.49	0.49	-
Observations	1377	2379	2718	1022	7496

All demographics, including age of father, age of mother, average age of children, children's age difference, father's education, and mother's education, enter the model through linear shifters to the resource share functions η_{ts} and through linear shifters of each member's preferences through δ_{fs}. I use all other information for robustness checks.

2.4. Results

I first present the coefficient estimates in Eq. (2.3) and then use these estimates to recover the estimates of resource shares for members of Iranian families by size. Thereafter, I report the effect of family gender composition on each individual member's resource share. Finally, I report the estimates of the effect of demographics on overall resource shares. Table 2.2 shows the coefficient estimates in Eq. (2.4) and their standard errors for families in rural and urban areas, separately. The left panel reports the parameters of equations for rural areas. The parameters of the quadratic equation in family size for women, reported in the top panel, are 0.59, −0.25, and 0.049, respectively. These parameters suggest that 59.2% of resources in childless families are devoted to women. The coefficients of −0.25 and 0.049 for s and s^2 suggest that a woman's share decreases sharply as family size grows. The estimate of γ_m, which is the effect of the proportion of girls in the family on the woman's resource share, is 0.098, which suggests that as the proportion of girls in the family increases, more resources are devoted to the mother.

The second panel from the top in Table 2.2 reports the parameters of the children's equation. As mentioned in Section 3, this equation does not contain a constant, as children's resource share in childless families is zero. The coefficients of family size and its square in the children's equation are 0.182 and −0.01, respectively. This result suggests that in one-child families, 17.1% of resources are devoted to that child. The small coefficient of the squared term suggests that the share of each child does not decrease considerably with the size of the family. The estimate of γ_c is −0.111, supporting the existence of gender bias in resource allocation among children in rural areas. In other words, the presence of girls in a family diverts resources away from the children.

The bottom panel in Table 2.2 reports the resulting coefficients of resource share for the father's equation. The coefficients of the quadratic equation in s are 40.8, 0.07, and −0.038, suggesting that households assign 40.8% of resources to the man in childless families. In contrast to women, men retain or lose their resources smoothly as family size grows.

Table 2.2: Estimates of Parameters

	Rural Areas		Urban Areas	
	Coefficient	S.E.	Coefficient	S.E.
mothers				
α_{0m}	0.592***	0.053	0.455***	0.045
α_{1m}	-0.252***	0.066	-0.133**	0.064
α_{2m}	0.049***	0.018	0.027	0.019
γ_m	0.098***	0.037	0.043	0.030
children				
α_{1c}	0.182***	0.038	0.239***	0.029
α_{2c}	-0.011	0.017	-0.037***	0.012
γ_c	-0.111***	0.037	0.001	0.032
fathers				
α_{0f}	0.408***	0.053	0.545***	0.045
α_{1f}	0.070	0.071	-0.105	0.065
α_{2f}	-0.038*	0.022	0.011	0.020
N	3343		4153	

*: significant at 10%, **: significant at 5%, ***: significant at 1%.

The right panel in Table 2.2 reports similar results for families in urban areas. The parameters of the quadratic expression in family size for women are 0.455, −0.133, and 0.027, respectively. These estimates state that among childless families in urban areas, the resource share devoted to women is 45.5%, which is smaller than the estimate of women's resource share among childless couples in rural areas. In addition, it states that this resource share decreases sharply as families increase in size. The decrease in women's resource share by family size, however, is smaller than that seen in rural areas. The estimate of γ_m is roughly 0.043 and its sign is similar to that of the rural areas analysis. Finally, the insignificant estimate of γ_m means that children's gender composition does not influence the resources allocated to women in urban areas.

The second panel from the top reports the estimates of the quadratic function of family size for children. The estimates of 0.239 and −0.037 suggest that resource share devoted to children is 20.2%, which is larger than the estimated share for the child in one-child families in rural areas. The larger coefficient of the quadratic term relative to that of rural areas suggests that the share of each child in two- and three-child families decreases faster in urban areas.

The last panel reports the coefficients of the quadratic function in s for the man as 0.545, −0.105, and 0.011. These estimates suggest that men's resource share in urban childless families is 54.5%, which is considerably larger than that in rural areas. However, men's resource share in urban areas declines at a faster rate as the number of children rise.

2.4.1. Estimates of Resource Shares

Table 2.3 reports the estimates of resource shares by family size for families in rural and urban areas, which derive from the coefficient estimates presented in Table 2.2. These estimates report the resource share for a reference family (see Section 4). All the estimates in Table 2.3 are significant at the 1% significance level.

The top panel reports the estimates of resource shares for childless couples, showing that 40.8% (54.5%) of family resources in rural (urban) areas are devoted to men. The next panel (one-child families) highlights that the resource shares allocated to men, women, and children in rural and urban areas are 43.9%, 39%, and 17.1% and 45.1%, 34.8%, and 20.1%, respectively. These results are important in two aspects. First, unlike childless families, the estimated resource shares are similar between urban and rural areas, except that families in urban areas allocate more resources toward their children. Second, the resources allocated to children are mostly diverted from their mother's share because in one-child families, women's resource share declines considerably relative to that of the women in childless families. The man−woman gap in resource shares is 4.9% in rural areas and 10.3% in urban areas, both in favor of men.

The next panel for two-child families reports that the man, woman, and children in rural areas receive 39.4%, 28.7%, and 32% of family resources, respectively. The corresponding estimates for urban areas are similar (37.7%, 29.5%, and 32.8%). Despite a gap of 3% in the child's resource share between one-child families in rural and urban areas, this gap decreases to only 0.4% for each child in two-child families. The gap between the resource share of men and women is 10.7% in rural areas compared with 8.2% in urban areas.

Table 2.3: Estimates of Resource Shares by Household Size

	Rural Areas		Urban Areas	
	Coef.	S.E.	Coef.	S.E.
childless families				
Man	0.408	0.053	0.545	0.045
Woman	0.592	0.053	0.455	0.045
one-child families				
Man	0.439	0.029	0.451	0.028
Woman	0.390	0.027	0.348	0.027
Children	0.171	0.025	0.201	0.022
two-child families				
Man	0.394	0.043	0.377	0.046
Woman	0.287	0.039	0.295	0.044
Children	0.320	0.037	0.328	0.040
three-child families				
Man	0.272	0.076	0.325	0.082
Woman	0.282	0.062	0.295	0.077
Children	0.446	0.079	0.381	0.075
N	3343		4153	

All estimated coefficients are significant at the 1% significance level.

The last panel reports the resource allocation among three-child families. Resources allocated to men, women, and children in rural areas are 27.2%, 28.2%, and 44.6%, and in urban areas are 32.5%, 29.5%, and 38.1%, respectively. The first important aspect of these results is that the man−woman gap in resource share shrinks considerably compared with families that have one or two children. Second, the resources allocated to each child decline to 14.8% in rural areas and 12.7% in urban areas. This finding suggests that consumption among Iranian households follows economies of scale, which distinguishes the models underlying this study from the literature that follows Working’s (1943) Engel curve method.

Two points are worth emphasizing. First, women's resource share is larger than the estimates of previous studies of developing countries. For example, Dunbar et al. (2013) estimate women's resource share to be about 37%, 22%, and 17% for Malawian families with one, two, and three children, respectively. It is thus surprising to see that among childless families in Iranian rural areas, women's resource share exceeds that of men's. Two potential explanations worth mentioning in this regard: First, in rural areas marriage is sometimes considered as a strategic bind between the groom's and the bride's families. In this regard, the bride, in the husband's family, is treated with extra respect as a sign of respect for her families. The respect of husband family for the bride in this case results in higher resource share for her. An alternative explanation is that the results might be a consequence of sample restriction. In rural areas it is common for newly married couples to live with groom's family. The sample, however, excludes the couples who live in an extended family; therefore, the remaining couples are those with higher autonomy of women. It is because woman's autonomy can enforce a separate residence right after the marriage. This autonomy often comes from woman's family reputation relative to that of man.

The second point regarding the results is that as families grow parents' resource share decreases to compensate for the increase in children's resource share. However, the decline in women's resource share by family size is more pronounced than that in men's resource share. This result differs from the finding of Dunbar et al. (2013) that men's resource share among Malawian families is almost unaffected by family size. For example, these authors estimate that men's resource share is about 46% in three-child families. Therefore, women's and children's resource shares decline in response to a growing family.

Various economic and cultural factors influence resource share in families (Das Gupta et al., 2003), including occupation. In this study, I categorize families by using a variable (*farmer*) that indicates whether either parent earns any income from a farming job, including gardening, ranching, and other related tasks. Table 2.4 reports the estimates of the resource shares for urban and rural areas by farmer indicator.

Table 2.4: Estimates of Resource Shares by Household Size and Farmer Status

	rural/ nonfarmer		rural/ farmer		urban/ nonfarmer		urban/ farmer	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
childless families								
Father	0.331	0.072	0.449	0.077	0.537	0.045	0.680	0.136
Mother	0.669	0.072	0.551	0.077	0.463	0.045	0.320	0.136
one-child families								
Father	0.405	0.043	0.495	0.041	0.448	0.029	0.499	0.124
Mother	0.468	0.040	0.354	0.040	0.356	0.028	0.309	0.112
Children	0.127	0.037	0.151	0.026	0.197	0.021	0.192	0.118
two-child families								
Father	0.385	0.059	0.430	0.060	0.384	0.047	0.409	0.181
Mother	0.314	0.055	0.257	0.056	0.282	0.045	0.301	0.136
Children	0.301	0.051	0.313	0.049	0.335	0.040	0.289	0.181
three-child families								
Father	0.269	0.104	0.254	0.096	0.345	0.081	0.411	0.321
Mother	0.208	0.089	0.260	0.085	0.241	0.081	0.297	0.161
Children	0.523	0.092	0.486	0.100	0.414	0.079	0.292	0.317
N	1870		1473		3913		240	

An important feature of this table is that the farmer–urban sample consists of only 240 observations. The small size of this group results in insignificant estimates of resource share for this sample. Table 2.4 highlights the fact that relative to nonfarmers, farmer families in both rural and urban areas direct more resources toward the man and fewer resources toward the woman compared with nonfarmer families. This finding suggests that women in Iranian families that work in agriculture have less autonomy than women in nonagricultural sectors. The resources allocated to children do not differ by farmer indicator in rural areas, except in one-child families in which families devote fewer resources to the child. In urban areas, however, farmer families allocate a smaller resource share to children relative to nonfarmer families.

The fact that a large proportion of families in rural areas are involved in agricultural production may create some concerns regarding the validity of separability of production and consumption. However, the collective household model of BCL and DLP does not require the separability of production and consumption to hold for all the goods and services that an agricultural family consumes. It only requires the separability assumption to hold for the private assignable good. That is, I assume that families purchase their clothing and footwear. Given the fact that Iranian families in rural areas are mostly involved in farming and carpet production, this assumption of the model is consistent with Iranian context.

2.4.2. Children’s Gender Composition

As stated in Section 4, the results presented thus far are estimates of resource share for a reference family whose children are boys. In this section, I thus investigate whether and to what extent the presence of girls in the family changes the resources allocated to each family member. Table 2.5 presents the results separately for families with one to three children. The top panel shows the effect of female children on resource allocation among one-child families, highlighting that one-girl families in rural areas allocate 3.8 percentage points more resources to women and 1.9 percentage points fewer resources to children compared with one-boy families. The effect on fathers’ resource share is negative and insignificant.

A similar pattern exists among two and three-child families in rural areas. Two-child families with two daughters devote 2.8 percentage points more resources to women and 3.5 percentage points fewer resources to children relative to families with two boys. Among three-child families, in families with three daughters, 2.8 (4.9) percentage points more (fewer) resources are devoted to women (children) compared with the reference family. Again, no significant effect of gender composition on men’s resource share is evident.

The effects of children’s gender composition on intra-household resource allocation among Iranian families in rural areas are thus threefold. First, 1.6–1.9 percentage points fewer resources are devoted to each girl relative to each boy. To contextualize the extent of this effect, recall that the resource share of boys is 14.8–

17.1% depending on family size[4]. Second, the presence of girls in the family increases women's resource share by between 2.8 and 3.8 percentage points depending on family size. Third, we observe no significant effect of family gender composition on men's resource share in rural areas. On the contrary, gender composition plays no significant role in intra-household resource allocation in urban areas.

Table 2.5: The Effect of Female Children on Family Members' Resource Shares

	Rural Areas		Urban Areas	
	Coef.	S.E.	Coef.	S.E.
one-child families				
Man	-0.019	0.013	-0.015	0.010
Woman	0.038***	0.014	0.015	0.010
Children	-0.019***	0.007	0.000	0.006
two-child families				
Man	0.007	0.013	-0.013	0.011
Woman	0.028***	0.010	0.013	0.009
Children	-0.035***	0.012	0.000	0.011
three-child families				
Man	0.022	0.018	-0.013	0.013
Woman	0.028**	0.011	0.013	0.009
Children	-0.049***	0.018	0.000	0.012
N	3343		4153	

: significant at 5%, *: significant at 1%.

The fact that presence of boys in a family is associated with lower income share for the mother supports a competition scenario as discussed earlier. It shows that fathers in rural areas may sacrifice their own resource shares for their male children. This result is similar to that of Rose (1999) and DLP. DLP found that in all-girl families, the combined share of children is about 6 percent lower than that of all-boy families and that the extra resource share almost fully diverts to the women.

[4] These results follow from assuming a linear effect by modeling the effect of family gender composition using the proportion of girls.

The fact that such gender bias is evident only in rural areas reflects the differences in economic and cultural factors. For example, gender differences in the family's contribution to the economy and old-age support for parents are among the strongest economic motives for allocating resources in favor of boys. Allowing for production in the model provides an alternative explanation to the results. Given the possibility of working on a farm in rural areas, the difference in resource allocation between boys and girls in urban vs. rural areas might reflect the difference between boys and girls in their productivity in the two environments. In rural areas, boys may be more productive than girls because they are more likely to work in farms and they may be more productive in physical activities in the agricultural sector. Therefore, it may be efficient for households to allocate higher resources to boys than girls. However, in urban areas boys and girls similarly play more similar roles in the household. Therefore, it may be efficient for households to allocate resources equally between girls and boys. The view that boys and girls have different production function in rural areas and similar production function in urban areas is supported by theoretical and empirical literature. (Michael 1974; Polachek and Polachek 1989; Gugl and Welling 2012).

Similarly, the kinship system and construction of gender in society are among the cultural norms that influence the roles of family members (Das Gupta et al., 2003). The distinction between the driving forces behind the gender bias in rural areas lies in the scope of their influence as well as policies that can alleviate such behavior. Contrary to economic forces, cultural forces are wider in influence and usually require more time to abate.

While distinguishing between the influence of economic and cultural factors is challenging, one simple test is to produce the estimates by farmer indicator similar to those in Section 5.1. The estimates of the influence of gender composition on resource share by farmer status in Table 2.6 report the effects of female children on intra-household resource allocation for four subsamples: rural–nonfarmer, rural–farmer, urban–nonfarmer, and urban–farmer. These results show that the role of gender composition in the nonfarmer sample in rural areas is considerably smaller than that in the farmer subsample.

Table 2.6: The Effect of Female Children on Family Members' Resource Shares

	rural/ nonfarmer		rural/ farmer		urban/ nonfarmer		Farmer/ urban	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
one-child families								
Father	-0.014	0.019	-0.037*	0.021	-0.019*	0.011	0.022	0.039
Mother	0.024	0.019	0.057***	0.022	0.020*	0.011	-0.036	0.040
Children	-0.010*	0.006	-0.020**	0.009	-0.000	0.006	0.014	0.028
two-child families								
Father	0.007	0.017	-0.001	0.019	-0.015	0.011	0.014	0.046
Mother	0.016	0.013	0.041***	0.015	0.016*	0.009	-0.035	0.041
Children	-0.023	0.014	-0.041**	0.017	-0.001	0.011	0.022	0.042
three-child families								
Father	0.029	0.025	0.021	0.029	-0.013	0.014	0.013	0.051
Mother	0.011	0.009	0.042**	0.018	0.013	0.008	-0.035	0.042
Children	-0.040	0.025	-0.063**	0.027	-0.001	0.013	0.022	0.046
N	1870		1473		3913		240	

*: significant at 10%, **: significant at 5%, ***: significant at 1%.

In nonfarmer families, the effect is significant only for children's resource share in one-child families, at −1 percentage point. Among farmer families, while the effect on men in the farmer sample is still insignificant, except for one-child families, the effect on women and children is more pronounced than the estimates presented in Table 2.5 for aggregate rural areas. All-girl farmer families with one to three children in rural areas devote 5.7, 4.1, and 4.2 percentage points more resources to women, respectively, each of which is larger than the corresponding effect in aggregate rural areas (3.8, 2.8, and 2.8 percentage points). The resource share of children in all-girl farmer families with one to three children in rural areas is 2, 4.1, and 6.3 percentage points smaller than that in all-boy families, respectively. These effects are larger than those reported for all rural areas (1.9, 3.5, and 4.5 percentage points).

Regarding the effect of gender composition on intra-household resource allocation by farmer status for urban areas, the results presented in Table 2.6 show no considerable difference between the nonfarmer sample and previous results on aggregate urban areas[5]. These estimates support the important role of economic factors in the presence of gender bias in intra-household resource allocation in rural areas. Because the majority of families in Iranian rural areas are involved in agricultural employment, boys are considered as a source of greater economic contribution to the family relative to girls. Furthermore, family structure in rural areas is such that girls leave the family after marriage and only boys provide for their parents in old age. These are the principal reasons behind the fact that parents in rural areas treat boys more favorably than girls in terms of intra-household resource allocation. By contrast, in urban areas, the family system is more flexible and parents' jobs are usually independent of family members' contributions. In addition, the more developed social infrastructures in urban areas suggest a channel from which parents obtain old-age support. For example, the development of pension funds plays a similar role to that of old-age support in rural areas.

The fact that there are larger differences between boys' and girls' resource allocation in rural farmer families than in non-farmer families is consistent with the idea that there may be productivity differences between boys and girls in rural agriculture that result in the allocation of more resources to boys.

Other studies have reported on the difference between rural and urban areas in terms of son preference. For example, Long (1989) finds that the sex ratio in Shanghai (a major city) has become fairly normal at about 1.05 in sharp contrast to the inflated sex ratios in rural areas and smaller cities, which reflect the higher mortality rate of girls and is associated with son preference. In the same vein, Rose (1999) reports the difference in parents' son preference in India by state, reporting the sex ratio at birth to be as low as 0.99 and 1.06 in Assam and Kerala but higher than 1.4 in West Bengal, Punjab, and Uttar Pradesh.

Iranian families devote more resources to sons than to daughters for one of three reasons. First, Iranian culture is strongly patriarchal (Moghadam 1992; Rasul-Ronning 2013). Fathers typically make the final decision on important family matters. In some instances, grandfathers even wield power over large extended families. In such a cultural environment, it is natural for parents to value the son as the next generation's father. In addition, children in Iran get their names from their father's side. Therefore, son

[5] As mentioned in Section 5.1, the farmer sample in urban areas includes only 240 observations, which results in imprecise estimated coefficients.

preference can be viewed as a way of retaining the family name. In this context, the first-born has a more important role in the family and sons are even more valuable if born first. Indeed, in more traditional areas, a mother is more respected if she gives birth to a boy first (Moghaddam, 1992).

Second, a woman usually leaves the family after marriage and moves in with the husband (or husband's family). The son, however, typically supports the family, especially as his parents age. In multi-son families, this role is assigned to the oldest son. Third, Iranian women have low labour market participation (Mehryar et al., 2002; Majbouri, 2010). Over the past three decades, female labour force participation has been below 20%, suggesting that men are the main economic providers for the majority of families. Although some of the abovementioned mechanisms, including the kinship system, are evolving, driven by urbanization and economic development, it is reasonable to expect that cultural effects will continue to influence the behavior of people in Iran.

Given the infrequency of reported purchases in HIES, total expenditure can be subject to measurement error. Dunbar et al. (2013) address this via a generalized Method of Moments estimator using household income as an instrument for total expenditure. However, they found that their results regarding the intrahousehold resource allocation and gender bias are invariant to instrumenting for total expenditure.

I address this concern by instrumenting for total expenditure in a control function framework (Heckman and Robb, 1985). The instruments are several measures of wealth, including family income from different sources (wage and salary, self-employment, and non-labour income), home ownership and car ownership.

I apply the control function by regressing the logarithm of total expenditure on instruments and demographics. Then, I use the residual from the first stage as a linear shifter in preference in demand equations for each family member. I should note that a Generalized Method of Moments approach is preferred as it is consistent even when the system of equation cannot be specified as a triangular form. For further discussion regarding the control functions and systems of equations please refer to Blundell et. al (2013).

Table 2.7 shows the effect of family gender composition on resource shares. This table supports two main findings of this study. First, fewer shares of family resources are devoted to female children in rural areas. Second, the lower share of girls in rural areas is associated with higher shares for their mother. Third, family gender composition does not play any role in intra-household resource allocation in urban areas.

Table 2.7: The Effect of Female Children on Family Members' Resource Shares

	Rural Areas		Urban Areas	
	Coef.	S.E.	Coef.	S.E.
one-child families				
Man	-0.017	0.016	-0.016	0.011
Woman	0.035***	0.016	0.016	0.011
Children	-0.018***	0.007	-0.001	0.007
two-child families				
Man	0.012	0.016	-0.013	0.012
Woman	0.024***	0.011	0.014	0.009
Children	-0.035***	0.014	-0.001	0.011
three-child families				
Man	0.026	0.023	-0.014	0.014
Woman	0.027**	0.013	0.015	0.011
Children	-0.052***	0.021	-0.001	0.012
N	3343		4153	

: significant at 5%, *: significant at 1%.

2.4.3. Effect of Demographics

The effects of demographics on resource share reported in Table 2.8 show that paternal education in cities diverts resources from mothers to fathers. Each extra year of the husband's education increases his resource share by about 0.67 percentage points, all of which is diverted from the mother's resource share. Maternal education in rural areas diverts the father's resource share to children (0.4 percentage points). In urban areas, it diverts resources from fathers (0.65 percentage points) to mothers (0.61). For example, if a woman in an urban area moves from the median education level (grade 5) to a high school diploma level (grade 12), her resource share increases by 4.2 percentage points, her husband's resource share decreases by 4.5 percentage points, and her children's resource share remains almost unchanged. This result is consistent with a bargaining mechanism among couples in the allocation of resources, in which education provides the educated party with higher bargaining power. Education thus acts as a distributional factor that influences resource share. Another interesting finding regarding maternal education in rural areas is that it diverts resources from fathers to children and results in a significant positive impact on children's resource share. In this regard, Dunbar et al. (2013) find that maternal education diverts resources from fathers to mothers (60%) and to children (40%). However, paternal education in the Dunbar et al. (2013) study does not have a substantial effect on family members' resource shares.

Table 2.8: The Effect of Demographics on Intra-household Resource Allocation

	Rural Areas		Urban Areas	
	Coef.	S.E.	Coef.	S.E.
man education				
Man	0.0012	0.0036	0.0067**	0.0032
Woman	-0.0005	0.0035	-0.0064**	0.0031
Children	-0.0007	0.0011	-0.0003	0.0007
woman education				
Man	-0.0042	0.0042	-0.0065*	0.0034
Woman	0.0004	0.0041	0.0061*	0.0034
Children	0.0038**	0.0013	0.0004	0.0007
man age				
Man	0.0003	0.0025	-0.0022	0.0024
Woman	-0.0002	0.0024	0.0021	0.0024
Children	-0.0001	0.0006	0.0001	0.0005
woman age				
Man	0.0038	0.0025	0.0037	0.0025
Woman	-0.0043*	0.0025	-0.0036	0.0025
Children	0.0005	0.0007	-0.0000	0.0005
children's average age				
Man	-0.0097*	0.0053	-0.0031	0.0055
Woman	0.0098*	0.0052	0.0034	0.0054
Children	-0.0001	0.0008	-0.0003	0.0006
children age difference				
Man	-0.0010	0.0091	0.0070	0.0118
Woman	-0.0142*	0.0077	-0.0058	0.0089
Children	0.0152**	0.0075	-0.0012	0.0115
N	3343		4153	

*: significant at 10%, **: significant at 5%.

The substantial positive effect of maternal education on children's resources is important for formulating policies on reducing child poverty. Previous studies in developing countries also support that improvements in women's status can contribute to their children's well-being and outcomes (Hill and King, 1995).

A woman's age in rural areas has a negative impact on her resource share. As women age, they lose their productive power and, consequently, their bargaining power in the family. For each extra year of age, women's resources are reduced by 0.43 percentage points. For instance, a 35-year-old woman in a rural area loses 2.1 percentage points of her resource share as she approaches 40 years. This finding contrasts with that of Das Gupta et al. (2003), who report that women's power dramatically rises over their lifespan as power and autonomy shift from fathers to mothers in older age.

Older children in rural areas divert resources toward their mothers. For example, one extra year of the average age of children diverts about 1 percentage point of resources from fathers to mothers. Raising the average age of children from five to 10 years in rural areas also diverts resources from fathers (5 percentage points) to mothers. This finding, that older children divert resources from fathers to mothers, is consistent with the results of Dunbar et al. (2013) from Malawian rural areas. The higher diversity of children's ages diverts resources mostly from mothers to children in Iranian rural areas. One extra year of children's age diversity increases their resource share by 1.5 percentage points, which is diverted mostly from their mother's resources. In rural Malawi, Dunbar et al. (2013) find that a higher variance in children's ages diverts resources from fathers to mothers.

A comparison of the effect of covariates in rural and urban areas suggests that in urban areas, the most influential variable in intra-household resource allocation is parents' education. In rural areas, however, education plays less of a role than children and women's age. This finding supports the notion that in rural areas, inter-family bargaining power is governed by traditional social norms, which values family members according to their age and productivity. On the contrary, educational attainment has a small role in bargaining power in rural areas compared with urban areas, in which education plays a more important role than age and physical productivity.

2.5. Conclusions

This study estimated family members' resource shares and investigated gender bias in intra-household resource allocation among Iranian families. Methodologically, it extended the approach taken by Dunbar et al. (2013) to allow for resource share to depend on family gender composition as well as on total expenditure and family size.

The estimates of resource share presented herein suggest that among families with one or two children, there is a gap between the resources allocated to men and women in favor of men. However, this gap shrinks in three-child families. In childless families in urban areas, the gap is small, while women's resource share outweighs that of men in rural areas. These results suggest that women, especially among families with one and two children, trade off more between their own and their children's consumption. Children's resource share increases as family size grows, from 17.1% (20.1%) in one-child families to 44.6% (38.1%) in three-child families in rural (urban) areas.

In addition, the study found that intra-household resource allocation in rural areas is influenced by gender composition. Parents assign 1.6–1.9 percentage points more resources toward their sons relative to their daughters, an effect more pronounced among farmer families. Such strong gender bias among farmer families in rural areas suggests that economic factors are more influential in the presence of parents' discriminatory behavior toward their children. Male children in farmer families are a source of greater contribution relative to female children and, therefore, families treat boys relatively more favorably.

Similarly, this study found that women in all-boy families in rural areas are assigned 2.8–3.6 percentage points fewer resources than those in all-girl families. As before, these effects are higher among farmer families. The fact that the presence of male children has a negative effect on their mother's resource share suggests either that the reward argument is not plausible for explaining the effect of gender composition on women's resource share or that the effect is influential in the short run (i.e., it does not change women's lifelong autonomy). In contrast to the rural sample, the study found no significant role of gender composition on intra-household resource allocation in urban areas.

Different effect of gender composition on intra-household resource allocation between rural and urban areas might suggest that family member's production functions are different in rural and urban areas. Likewise, the difference between farmer and nonfarmer families in production function might be an explanation for more pronounced effect of gender composition on intra-household resource allocation in farmer families. Boys in rural areas considered to be more productive than girls in farm production and later as old age support for parents. Therefore, an efficient resource allocation will allocate more resources toward boys. In such environments, the extra resource share of boys comes at the expense of mothers as they are considered less productive than fathers. In urban areas, however, there is no difference in production function of boys and girls. Consequently, we do not observe any evidence of gender bias in intra-household resource allocation in urban areas.

Finally, the fact that maternal education has a substantial effect on children's share has important policy implications for the alleviation of child poverty.

Chapter 3.

Private Schools and Student Achievement [6]

Critics frequently argue that public schools are hamstrung by a variety of constraints that undermine their effectiveness, including enrolment rules that restrict competition, top-down management structures that limit their discretion to innovate, and strong union rules that prevent them from disciplining or dismissing poor teachers. Private schools, in contrast, are forced to compete in order to attract tuition-paying students, can dismiss teachers who underperform, and may apply selective enrolment rules in order to generate complementarities among students. With greater access to resources, private schools may be able to hire better or more experienced teachers, and offer smaller classes and programs that are specially designed to serve their students.

Given these advantages, economists would expect private schools to provide higher quality education than their public counterparts. Surprisingly, however, the empirical evidence consistently finds that private schools in advanced economies do not do a better job of teaching core reading and numeracy skills than public schools, and may in fact do worse on average[7]. Highly credible random assignment studies of the effects of tuition fee voucher programs in the United States find little if any effect of

[6] This chapter is a joint work with Dr. Jane Friesen and Dr. Simon Woodcock.

[7] Evidence of the relative performance of private versus public schools in developing economies is more favorable. See, for example, estimates of private school effects for Columbia (Angrist 2002, 2006), India (Muralidharan and Sundararaman 2015; Singh 2015), Pakistan (Andrabi et al. 2011; Alderman et al. 2001).

private schools on students' test scores, with the possible exception of African-American students (see Epple et al. (2015a) for a review). A second literature that uses instruments based on geography and/or religiosity to address selection into private Catholic high schools finds very little evidence of any effect on test scores, except among urban minorities (e.g. Figlio and Stone 1997; Grogger and Neal 2000).[8] Along with others, Altonji et al. (2005) question the validity of these instruments. They instead control for as many observable student characteristics as possible in a value-added model of achievement, and then generate bounds on the true causal effect under plausible assumptions about the relationship between selection on observable and unobservable student characteristics. Like the previous IV studies, they find no effect of private Catholic high schools on test scores in the U.S. Applying Altonji et al.'s method, Elder and Jepsen (2014) find negative effects of private Catholic primary schools on numeracy scores in the U.S., and Nghiem et al. (2015) find negative effects of private Catholic primary schools on both reading and literacy scores in Australia. Earlier studies that use the ECLS-K data to study the effect of private Catholic schools on primary test scores find small negative and insignificant results (Carbonaro 2006; Lubienski et al. 2008; Reardon et al. 2009). Jepsen (2003) finds small positive effects of Catholic primary schools on high school test scores in data from the Prospects project.

Recent evidence of the effects of secular private schools is more limited but somewhat more promising. Figlio and Stone (1997) find positive effects of private secular high schools in the U.S. on mathematics and science scores. Nghiem et al. (2015) find generally positive but statistically insignificant effects on numeracy and literacy skills in Australian private secular primary schools. Lefebvre et al. (2011) control for individual heterogeneity via student fixed effects rather than lagged test scores; they find positive effects on private secular high school numeracy scores in Quebec, Canada.

[8] A number of papers find that private schools do, however, confer an advantage with respect to high school completion (e.g. Evans and Schwab 1995; Grogger and Neal 2000; Neal 1997).

Our paper advances this literature by providing the first evidence of the effects of private schools on student achievement based on longitudinal population data. This data affords several advantages over previous studies of private school effects on achievement. First, it provides a much larger sample,[9] allowing us to contribute new and precise estimates to the literature on the effectiveness of Catholic and secular private schools, and to extend this literature to include other types of faith schools. Second, because it includes both a substantial number of observations concentrated within fairly small geographic areas and detailed information about students' residential locations, we can control for neighborhood characteristics or neighborhood fixed effects in our specifications. Doing so allows us to control for an important potential source of unobserved heterogeneity in family background and student ability that is correlated with neighborhood choice. Moreover, these neighborhood controls ensure that private school effects are identified either from comparisons of schools attended by students who live in similar types of neighborhoods (when we include neighborhood characteristics), or who live in the same neighborhoods (when we include neighborhood fixed effects). This feature is important: some previous studies report positive effects of U.S. Catholic schools when they restrict their attention to urban populations, although sample sizes become very small. A common explanation offered for this pattern of results is that the quality of public schools is lower in urban areas than elsewhere, so that local private schools can be superior to these schools while not outperforming public schools in general (see Martinez-Mora, 2006, for a theoretical treatment of this issue). By identifying private school effects from comparisons between schools attended by students who live in nearby or similar neighborhoods, our approach avoids this potentially confounding factor. A third advantage of using multiple cohorts of population data is that it provides us with a sufficiently large number of observations per school to

[9] Specifically, our data include more than seven times the number of students enrolled in private Catholic schools compared to the data used by Nghiem et al. (2015) and Elder and Jepsen (2014), and more than nine times the number of students enrolled in private secular schools compared to the data used by Nghiem et al. (2015) and Lefebvre et al. (2011). McKewan (2001) has a larger sample of private Catholic school students in his data from Chile, but only one test score per student. Other jurisdictions that provide researchers with access to population-based longitudinal student-level data typically do not include records for students enrolled in private schools (e.g. Florida, England, North Carolina and Texas).

estimate individual school effects, allowing us to characterize heterogeneity both within and across school sectors. Finally, the density of our data within one contiguous geographic area allows us to estimate a full set of school and student fixed effects using methods developed by Abowd et al. (2002). We use these estimates to characterize the entire distribution of student as well as school quality both within and across school sectors.

Our student-level administrative data from fourteen school districts in British Columbia, Canada encompasses the population of students enrolled in virtually all private as well as public schools. Our longitudinal records follow five cohorts of grade 4 students (aged 8-9) for a total of four years, when most of these students are in grade 7. British Columbia (B.C.) administers standardized tests in reading and numeracy to students in grades 4 and 7. Along with information about test scores and basic demographic characteristics, individual records include information about the school attended and the student's residential postal code in each year.

B.C.'s public education system has many standard features and offers considerable choice within the public sector to a highly diverse student population. B.C. students consistently rank very well on international comparisons of math, science and reading achievement (Brochu et al. 2012). Almost 12% of B.C. primary school students were enrolled in private schools in seventh grade in 2010/11 (British Columbia Ministry of Education 2011), roughly comparable to national private school enrolment rates in the United States (Rouse and Barrow 2009). Most private schools have been operating for many years, and have received public funding since 1977. This universal voucher, typically worth half of the basic allocation provided for each student attending a public school, is available to both religious and secular private schools. In order to be eligible for the voucher, private schools are required to hire qualified teachers, offer the approved core curriculum and meet minimum requirements for instructional time, and they are subject to provincial inspection and evaluation. They are fully autonomous with respect to personnel decisions, pay, school calendar, and admissions criteria.

We estimate several different specifications of the education production function, using both a lagged value-added model and a student fixed effects model, and conduct a number of robustness checks based on sub-samples of the data. We obtain a number of important results that are consistent across specifications and samples. We find that private schools on average outperform public schools by at least 0.10 standard deviations in both reading and numeracy, and these results cannot be accounted for by reasonable assumptions about the degree selection on unobservable factors relative to selection on observable student and neighborhood characteristics.

Unlike most previous studies, we find that Catholic private schools outperform their public counterparts in both math and reading, with an effect size of at least 0.10 standard deviations. Non-Christian faith private schools perform even better, with an effect size of at least 0.20 standard deviations. Other Christian (i.e. non-Catholic) private schools, in contrast, do not outperform public schools; estimated effects are small and statistically insignificant. We are also able to obtain precise and robust estimates of the effects of a group of secular private schools that we refer to as “prep” schools; these are schools that emphasize academic achievement. We find that prep private schools on average perform very well – somewhat better than Catholic private schools and as well as non-Christian faith private schools.

3.1. Related literature

Several related literatures yield results that can inform our expectations of private school performance relative to public schools. Clark (2009) investigates the performance of “grant maintained” public schools in England. These schools have full authority to hire and fire teaching staff and can apply some criteria for selecting students, but they are not permitted to charge tuition fees or use admissions tests. He finds that converting to grant-maintained status increased achievement gains by 0.25 standard deviations. Changes in student characteristics can account for less than half of this improvement; the remaining gains appear to be associated with turnover in teaching staff and an increase in the number of teachers.

In the U.S., the charter school movement has grown as an alternative to regular public schools. Charter schools are granted the freedom to innovate with respect to curriculum and pedagogy. Compared to regular public schools, they are less likely to be unionized and exhibit higher rates of teacher turnover (Epple et al. 2015a). Unlike private schools, charter schools cannot charge tuition and must use lotteries to admit students if they are oversubscribed. Lottery-based studies of oversubscribed charter schools tend to find fairly large positive effects. Estimates of a broader set of charter school effects, including those that are not oversubscribed, find no overall effect, and sometimes negative effects in the early years of newly established charter schools (see Epple et al. (2015b) for a review).

3.2. Institutional Context

3.2.1. Public school choice and funding

The majority of public primary schools in British Columbia offer Kindergarten through grade 7, with high schools offering grades 8 through 12. There are many exceptions, however, with some schools offering Kindergarten through grades 3, 4, 5 or 6, some middle schools that begin in grade 6 and some "junior" high schools that begin in grade 7.

Students in B.C. are guaranteed access to a "catchment" public school based on their residential address. They may also choose to enroll in a regular public school other than their catchment area school. Before July 2002, the provincial education authority (the Ministry of Education) mandated that out-of-catchment enrollment in a regular (non-magnet) public school required permission of the principals of both the catchment area school and the preferred school. Since July 2002, students have been free to enroll in any public school in the province that has space and facilities available after students who reside in the catchment area have enrolled. Transportation to non-catchment schools is not provided. When catchment schools are over-subscribed, provincial legislation requires that school boards give priority to students who reside within the district. Boards may elect to give priority to siblings of children who are already enrolled. Within these enrolment categories, principals of regular public schools have discretion over which students to enroll.

Parents in B.C. may also choose to enroll their children in a public magnet program. The most popular form of magnet program is French Immersion, which enrolls about 10 percent of Kindergarten students in the province (BC Ministry of Education 2011). Entry into French Immersion programs is restricted to students entering Kindergarten or grade 1, and space is often allocated by lottery.

The B.C. Ministry of Education provides operating and capital funding directly to public districts. Operating funds are provided in proportion to total district enrolment, with supplementary funding for each student who is Aboriginal, gifted or disabled, or who qualifies for English as a Second Language (ESL)[10] instruction. Public districts and schools are not authorized to raise any additional revenue, and are required to offer the provincial curriculum. Hiring, firing and remuneration of teachers is governed by strict rules specified in a collective agreement between the Province and the powerful union that represents B.C. teachers (the British Columbia Teachers Federation).

3.2.2. Private school choice and funding

Since 1977, British Columbia has provided public grants to private schools that conform to provincial curriculum standards and meet various provincial administrative requirements (B.C. Federation of Independent Schools[11] Associations 2015). The value of the grant varies with school operating costs: since 1989, schools whose operating costs are no higher than in the public system (Group 1 schools) receive an amount per

[10] The ESL program was renamed to English Language Learning (ELL) in January 2012.

[11] The “independent school” has been widely used in BC province to refer private schools. We use both terms interchangeably.

student equal to 50 percent of the value of the per student grant to public schools; those with higher operating costs (Group 2 schools) receive 35 percent of the public school grant (B.C. Ministry of Education 2005). The Ministry of Education does not limit the total number of funded private school spaces, and private schools are not constrained in their selection of students. The formula for supplementary funding for special education students in private schools changed in 2005. Private schools had historically received half as much per student special education funding as public schools; since 2005, private schools have received the full value of the public school special education supplement.

In order to be eligible for funding, private schools must hire qualified B.C. teachers and offer the provincial curriculum. Private schools may charge tuition, apply any admissions criteria that do not violate the Canadian Charter of Rights and Freedoms or the provincial Human Rights Code, and can hire, fire and remunerate teachers subject only to provincial labour standards legislation.

3.2.3. Testing and accountability

All public and provincially funded private schools in British Columbia are required to administer standardized tests to students in grades 4 and 7 in reading and numeracy each year. These scores do not contribute to students' academic records and play no role in grade completion, and there are no financial incentives for teachers or schools related to student performance. The Ministry of Education began posting school-average test scores on their website in 2001 (B.C. Ministry of Education 2001). The Fraser Institute, an independent research and educational organization (Fraser Institute 2008), began issuing annual "report cards" on B.C.'s elementary schools in June 2003 (Cowley and Easton 2003). These reports include school scores and rankings based on test scores. From the outset, the school report cards have received widespread media coverage in the province's print, radio and television media.

3.3. Data

Our estimates are based on extracts from two administrative databases collected and maintained by the B.C. Ministry of Education. The first is an enrolment database that records the school at which each student is enrolled on September 30 of each year. Our extract includes five cohorts of grade 4 students who were enrolled in a public or private school located within the geographic boundaries of the fourteen school districts in the Lower Mainland of B.C.[12] in 1999/2000 through 2003/2004, and follows them for the following four years. Students remain in our data so long as they remain within the provincial public or private school system in any of these districts. The individual records include indicators for the language spoken in the student's home (English, Chinese, Punjabi, and other), whether the student self-identified as Aboriginal in any year, whether the student was registered in ESL or special education (i.e. a gifted or disabled program), whether the student was enrolled in French Immersion, whether the school is public or private, and the student's gender. In addition, the extract provides the student's residential postal code and unique student, school and district identifiers. We attach average family income, proportion of immigrant families, and proportion of people with different levels of education in the student's Census neighborhood (enumeration area), based on a postal code match. An enumeration area is the smallest geographic area for which public-use Census data are produced, and typically comprises several hundred households. A detailed description of our procedures for locating residential postal codes within enumeration areas is provided in a data appendix.

The second database provides student-level data on participation and scores on standardized tests administered in grades 4 and 7 for the 1999/2000-2006/2007 school

[12] The Lower Mainland consists of the city of Vancouver and its suburbs. It is geographically isolated by the Canada/U.S. border to the south, rugged mountains to the east and north, and the Salish Sea to the west.

years. We merge students' test scores with the enrolment database via the unique student identifier provided in both files.

3.4. Methodology

3.4.1. The value-added model

The literature on specification issues related to education production functions is highly developed (e.g. Todd and Wolpin 2003; Andrabi et al. 2011). The most commonly used specification in the education literature is the lagged value added model, which includes a lagged test score as a sufficient statistic for the entire history of inputs in a model of current outcomes. We write our basic version of this model as:

$$y_{i,7} = \beta C_{i,7} + \lambda y_{i,4} + \alpha \mathrm{Private}_{s(i,7)} + \mathrm{year}_{t(i,7)} + \varepsilon_{i,7} \qquad (3.1)$$

where $y_{i,7}$ is an outcome measure observed when student i is in grade 7, $C_{i,7}$ is a vector of student characteristics observed in grade 7, $\mathrm{Private}_{s(i,7)}$ is an indicator that student i attends a private school in grade 7, $\mathrm{year}_{t(i,7)}$ is a fixed effect associated with the year that student *i* is in grade 7 and $\varepsilon_{i,7}$ is a stochastic error. The parameter of interest, α, captures the average contribution of private versus public schools to student test score growth between grades 4 and 7.

The key identifying assumption is that unobserved factors affecting grade 7 test scores are mean-independence with respect to enrolment in a private versus public school, conditional on grade 4 test scores and other observable student characteristics:

$$E\left(\varepsilon_{i7} \middle| \mathrm{Private}_{s(i,7)}, y_{i,4}, \mathrm{year}_{t(i,7)}, C_{i,7}\right) = 0 \qquad (3.2)$$

This condition requires that there are no unobserved factors that affect test score growth that are systematically related to the decision to attend a private versus public school in grade 7, conditional on observable student and neighborhood characteristics. A further threat to identification comes from the inclusion of the lagged dependent variable in this model, since test scores are prone to measurement error. However, despite their theoretical limitations, value-added models that include an adequate set of controls have been shown to deliver estimates of the effects of various inputs with a fairly small degree of bias in a range of contexts.[13] We include a number of individual student characteristics, including home language, Aboriginal identity and gender. We use two alternative sets of controls based on the student's residential neighborhood. In one case we add mean neighborhood socioeconomic characteristics, including family income, parents' education and proportion of immigrants. In another case we include neighborhood (postal code) fixed effects. Gibbons and Silva (2011) find that including home postal code fixed effects as controls in their value-added model accounts for a substantial amount of the estimated differences in quality between faith-based and secular public schools in England.

We begin by estimating this basic model, which provides an estimate of the difference between the average test score of students attending private versus public schools. We then extend this model to include a full set of school fixed effects. The condition for identification in this case can be stated analogously to (3.2), with the private school indicator replaced by a full set of school fixed effects. In this case, identification requires that there are no unobserved factors that affect test score growth that are

[13] Several recent studies have compared the results from value-added estimators to experimental results from the same data set. Andrabi et al. (2011) find biases from measurement error and unobserved heterogeneity are offsetting in their study of private schools in Pakistan, such that the aggregate bias on the private school coefficient is not significant. Deming et al. (2014) find no significant differences between experimental estimates of school effects based on lottery data and estimates from a value-added model that controls for a previous test score. A number of studies find similar results when comparing experimental estimates to value-added estimates of teacher effects (Angrist et al. 2013; Kane and Staiger 2008; Kane et al. 2013).

systematically related to the decision to attend each private versus public school in grade 7, conditional on observable student and neighborhood characteristics.

3.4.2. The student fixed effects model

We next estimate an alternative model of the education production function, in which the coefficient on the lagged score, λ, is constrained to be zero. This restriction is supported by growing evidence that the effects of lagged inputs dissipate rapidly (see for example Andrabi et al. 2011, Jacob et al. 2010, and Kane and Staiger 2008 in the context of teacher effects). Since we observe test scores three years apart, any bias introduced by this restriction is likely to be very small. We control for individual heterogeneity by including student fixed effects. The student fixed effects specification has been used widely in the school quality literature, including several recent papers on the effects of charter schools (e.g. Imberman 2011). We write this model as:

$$y_{i,g} = \beta_1 C_{i,g} + \alpha \text{Private}_{s(i,g)} + \text{grdyr}_{t(i,g)} + \text{student}_i + \varepsilon_{i,g} \quad g=4,7 \qquad (3.3)$$

where $\text{grdyr}_{t(i,g)}$ is a grade-by-year fixed effect.

The key identifying assumption in the student fixed effects model is that, among the subset of students who switch between private and public schools between grades 4 and 7, unobserved factors that affect a student’s test score in a given grade are uncorrelated with private school status in that grade, conditional on grade-level transitory shocks, student fixed effects and observable time-varying student characteristics:

$$E\left(\varepsilon_{ig} \middle| \text{Private}_{s(i,7)}, \text{grdyr}_{t(i,7)}, C_{i,7}, \text{grdyr}_{t(i,4)}, C_{i,4}, \text{Private}_{s(i,4)}, \text{student}_i\right) = 0$$

$$\text{for } g = 4, 7 \qquad (3.4)$$

At least two plausible scenarios could threaten the validity of this assumption. First, students may be heterogeneous with respect to an unobserved effect in test score growth between grades 4 and 7. The student fixed effects estimator will be biased if this effect is correlated with patterns of student mobility. Suppose, for example, that students who switch from a public school in grade 4 to a private school in grade 7 on average have characteristics associated with higher rates of test score growth, all else equal, than those who switch from private to public. In this case, the estimator will attribute the higher rate of test score growth that is caused by this pattern of unobserved heterogeneity to the effect of the private school.

Second, transitory shocks that cause students to change schools may be correlated with unobserved factors that affect test scores. Suppose, for example, that some students change schools at the end of grade 4 following a family break-up or job loss, and this event also adversely affects student achievement in grade 4. If the student's grades recover by grade 7, the estimated quality of the grade 4 school will be biased downwards relative to the grade 7 school. If students enrolled in private schools in grade 4 experience this scenario with the same frequency and degree as those enrolled in public schools in grade 4, the estimated difference between public and private school quality will not be affected. If, for example, these types of shocks more frequently result in student moving from private school to public schools, rather than vice versa, the estimated quality of private schools will be biased downwards.

We address this second selection problem in two ways. First, we investigate the sensitivity of our results to several sample restrictions designed to eliminate bias associated with different patterns of correlation between mobility patterns and transitory shocks to student achievement. By excluding students who changed schools immediately after grade 4, we eliminate the threat of bias from shocks that precipitate a move at the end of grade 4 and affect grade 4 test scores. By excluding students who changed schools immediately after grade 6, we eliminate the threat of bias from shocks that precipitate a move at the end of grade 6 and affect grade 7 test scores. Second, we estimate our model for a sample of students who are required by the grade configuration of schools to move between grades 4 and 7. The types of transitory shocks that lead to the dynamic selections problems described above are likely to be less frequent among compulsory movers than among students who elect to change schools.

Again, we begin by estimating this basic model, and then extend the model to include a full set of school fixed effects. The condition for identification in this case can be stated analogously to (3.4), with the private school indicators replaced by a full set of school fixed effects for each grade. In this case, identification requires that, among the subset of students who switch between private and public schools between grades 4 and 7, unobserved factors that affect a student's test score in a given grade are uncorrelated with the choice of school (rather than school type) in that grade, conditional on grade-level transitory shocks, student fixed effects and observable time-varying student characteristics. We estimate this two-way fixed effects model using a procedure developed by Abowd et al. (2002).

3.4.3. Bounding the effects of selection on unobservables

None of these approaches fully addresses the threat from unobserved individual fixed effects in test score growth. Following some of the recent literature (e.g. Nghiem et al. 2015; Elder and Jepsen 2014), we generate bounds on some of main estimates under assumptions about the relationship between selection on unobservable and observable characteristics, using the procedure introduced by Altonji et al. (2005) and extended by Oster (2015).

Defining the effects of observables as $O_{i7} = \beta C_{i,7} + \lambda y_{i,4} + year_{t(i,7)}$, we can write the value-added model with a simple private school indicator as:

$$y_{i,7} = O_{i,7} + \alpha Private_{i,7} + \varepsilon_{i,7} \qquad (3.5)$$

Call the estimate of α obtained from this "long" version of the regression model $\tilde{\alpha}$, and the corresponding R^2, $\tilde{R}$. Now consider a "short" regression that includes only the private school indicator:

$$y_{i,7} = \alpha \text{Private}_{i,7} + \varepsilon_{i,7} \tag{3.6}$$

Call the estimate of α obtained from this model α_0, and the corresponding R^2, R_0. Oster (2015) shows that

$$\frac{\tilde{\alpha} - \alpha}{\alpha_0 - \tilde{\alpha}} = \delta \frac{R_{max} - \tilde{R}}{\tilde{R} - R_0} \tag{3.7}$$

where δ is the degree of selection on unobservables relative to the selection on observables and R_{max} is the value of R^2 we would obtain if we included all of the observable and unobservable factors that affect the dependent variable. We can obtain estimates of $\tilde{\alpha}$ and α_0 and the corresponding values of $\tilde{R}$ and R_0 by estimating models (3.6) and (3.7). Oster notes that the value of R_{max} will be less than one if there is measurement error in the dependent variable. For an assumed value of R_{max}, we can calculate the value of δ, the ratio of the degree of selection on unobservables to selection on observables, under the null hypothesis that α=0. Oster (2015) shows that this method of calculating δ is equivalent to the method proposed by Altonji et al. (2005) under the assumption that $R_{max} = 1$. Both Oster (2015) and Altonji et al. (2005) argue that it is unreasonable to expect the degree of selection on unobservables to exceed the degree of selection on observables. A value of $\delta > 1$ therefore would lead to the conclusion that the estimated treatment effect is too large to be solely a consequence of bias due to selection on unobservables. We use this method to calculate values of δ using $R_{max} = 1.3\tilde{R}$, as recommended by Oster (2015).

3.5. RESULTS

Our estimation sample consists of all students in the population of interest who are observed in our data in grades 4, 5, 6 and 7. We restrict our attention to students attending an English language[14] public or private school that enrolls at least five students in the relevant grade and year, and who have non-missing values for all relevant variables.

3.5.1. Descriptive statistics

Table 3.1 presents selected school characteristics by school type. Our estimation sample includes a total of 649 schools, of which 554 are public and 95 are private. Of the private schools, 33 are Catholic, 39 are other Christian, 7 are other faith, and 15 are secular prep schools. Three-quarters of private schools are in funding group 1, receiving a per student subsidy equal to 50% of the public school subsidy; the remaining private schools are in funding group 2, receiving a 35% subsidy. A small number of other (non-prep) secular private schools offer Montessori or Waldorf programs, or specialized education for students with special learning needs. Given their small number, typically small size, and the diversity of programming, we exclude these schools from our analysis. Almost all private schools in our sample offer both grades 4 and 7. In contrast, 101 public schools offer grade 4 but do not offer grade 7, and 35 schools offer grade 7 but not grade 4.

Table 3.2 presents descriptive statistics for the grade 7 students in our estimation sample who have non-missing test scores. Private school students on average earn higher test scores than public schools students, but this difference varies across private

[14] French language instruction is offered to English speaking students via public French Immersion programs. Attrition from these programs is very high (see Shack 2015). Francophone students may choose to attend one of a small number of schools operated by the public francophone school board.

school types. Students enrolled in private prep schools excel; their average grade 7 test scores are 0.86 and 0.93 standard deviations above the mean in reading and numeracy. Students enrolled in faith private schools also score relatively well, averaging 0.43 standard deviations above the mean in reading and 0.54 standard deviation in numeracy at private Catholic schools, 0.27 standard deviations in reading and 0.28 standard deviations in numeracy at private other Christian schools, and 0.32 standard deviations in reading and 0.52 standard deviations in numeracy at non-Christian private faith schools, compared to 0.02 and 0.12 standard deviations among public school students.

Table 3.1: Selected school characteristics, by school type

	(1)	(2)	(3)	(4)	(5)	(6)
	Public			Private		
		All	Catholic	Christian	Other faith	Prep
No. of schools	554	95	33	39	7	15
Funding group 1	.	76	31	30	5	9
Funding group 2	.	19	2	9	2	6
Grade 4 only	101	1	0	0	0	1
Grade 7 only	35	4	0	2	1	1
Grades 4 and 7	418	85	33	37	6	13

Notes: see text and Data Appendix for details of sample selection and construction, and for variable definitions.

The remaining rows of Table 3.2 demonstrate the extent to which sorting across schools produces differences in the observables characteristics of students attending each school type. These patterns demonstrate the substantial potential for differences in student characteristics to account for the observed differences in mean achievement across school types. Students enrolled in private prep schools are the most positively selected with respect to several observable characteristics that have a known association with test scores. These students on average live in neighborhoods with very high mean family income and highly educated household heads. Relatively few prep schools students are Aboriginal or speak Punjabi at home, characteristics that in British Columbia are associated with low mean test scores on average (see Friesen and Krauth

2011). Students enrolled in private non-Christian faith schools also live in relatively high SES neighborhoods. However, these students are far less likely to speak English at home than any other group. Notably, over one-third of these students are Punjabi-speakers who attend Sikh faith schools. Students enrolled in Catholic schools are also positively if slightly less strongly selected with respect to neighborhood SES. They are more likely to speak English at home than public schools students and are less likely to be Aboriginal. In contrast, neighborhood SES is very similar for students enrolled at other Christian private schools compared to public school students. Punjabi speakers and Aboriginal students are underrepresented at these schools, but the proportion that speaks another non-English home language (e.g. Korean, Vietnamese or Tagalog) is higher than among public school students.

Table 3.2: Selected student characteristics, grade 7 students with non-missing test scores, by school type

	(1)	(2)	(3)	(4)	(5)	(6)
	Public	Private				
		All	Catholic	Christian	Other faith	Prep
No. of students	87113	11924	4565	4448	532	2379
% of sample	88.00	12.00	4.60	4.50	0.50	2.40
Reading score	0.02	0.45	0.43	0.27	0.33	0.86
Numeracy score	0.12	0.52	0.54	0.28	0.52	0.93
Home language						
English	0.67	0.74	0.81	0.72	0.36	0.71
Chinese	0.12	0.08	0.05	0.09	0.00	0.12
Punjabi	0.08	0.03	0.00	0.01	0.36	0.03
Other	0.13	0.16	0.14	0.18	0.29	0.15
Aboriginal	0.05	0.01	0.01	0.01	0.00	0.00
Female	0.48	0.49	0.52	0.46	0.54	0.48
Neighborhood mean						
Immigrant	0.07	0.06	0.07	0.05	0.10	0.07
High school	0.25	0.25	0.25	0.26	0.24	0.23
Some college	0.30	0.28	0.29	0.30	0.24	0.24
Bachelor’s	0.17	0.22	0.20	0.15	0.23	0.36
Family income	6.71	7.89	7.06	6.67	7.07	12.00

Notes: see text and Data Appendix for details of sample selection and construction, and for variable definitions.

Table 3.3: Grade 7 student characteristics, by school type and mover status

ALL MOVERS	(1)	(2)	(3)	(4)
	Public/public	Public/private	Private/public	Private/private
No. of students	32370	1691	1410	862
% of full sample	32.68	1.71	1.42	0.87
Reading score	-0.07	0.41	0.02	0.44
Numeracy score	-0.01	0.53	0.07	0.45
Home language				
English	0.69	0.59	0.61	0.69
Chinese	0.09	0.17	0.05	0.09
Punjabi	0.07	0.06	0.17	0.03
Other	0.15	0.20	0.16	0.20
Aboriginal	0.08	0.02	0.04	0.00
Female	0.48	0.44	0.47	0.47
Neighborhood mean				
Immigrant	0.07	0.07	0.07	0.06
High school	0.26	0.24	0.25	0.24
Some college	0.31	0.28	0.29	0.28
Bachelor’s	0.15	0.23	0.18	0.24
Family income	6.39	8.21	6.98	8.43
COMPULSORY MOVERS ONLY				
	Public/public	Public/private	Private/public	Private/private
No. of students	14946	179	109	63
% of full sample	15.09	0.18	0.11	0.06
Reading score	-0.07	0.33	0.06	0.54
Numeracy score	-0.04	0.42	-0.03	0.60
Home language				
English	0.77	0.61	0.75	0.90
Chinese	0.08	0.15	0.07	0.02
Punjabi	0.04	0.07	0.04	0.02
Other	0.12	0.20	0.14	0.06
Aboriginal	0.07	0.03	0.04	0.02
Female	0.47	0.48	0.45	0.49
Neighborhood mean				
Immigrant	0.06	0.08	0.05	0.04
High school	0.26	0.25	0.25	0.23
Some college	0.32	0.29	0.32	0.28
Bachelor’s	0.15	0.18	0.14	0.30
Family income	6.43	6.74	6.48	11.49

Notes: see text and Data Appendix for details of sample selection and construction, and for variable definitions.

The variation across school types in gender composition is also worth noting. Private prep school students are 48 percent female, the same proportion as in public schools. Catholic schools and other (non-Christian) faith private schools are disproportionately female (52 percent and 54 percent respectively), while other (non-Catholic) Christian private schools are disproportionately male (46 percent female).

Table 3.3 presents characteristics of students who change schools between grades 4 and 7 by type of move. This mobility is used to identify school and student fixed effects in the student fixed effects specification of our model. As described above, identification of school effects requires that mobility patterns between private and public schools are uncorrelated with unobserved factors that affect test score growth between grades 4 and 7, conditional on time-varying student characteristics. The second and third columns of Table 3.3 present characteristics of students who move from a public school in grade 4 to a private school in grade 7 and vice versa. Among all movers, shown in the top panel, the characteristics of students who move from public to private schools differ on average from those who move from private to public schools in a number of dimensions. To the extent that unobservable characteristics that affect test score growth are correlated with these observables, this evidence suggests that the student FE estimator may be subject to bias from this source.

The lower panel of Table 3.3 presents characteristics of students who change schools, by type of move, within the sub-sample of students who were forced to change schools between grades 4 and 7 because of the grade configuration of their grade 4 schools. Relatively few movers in this group move between the private and public sectors. Of those who do, students who move from public to private schools reside in neighborhoods with very similar mean family income to those who move from private schools to public schools, and their test scores are more similar to one another. This pattern of observables suggests that estimates from this sample may be less prone to bias from this source.

3.5.2. Results from the private school indicator model

The main results from the simple private school indicator version of our model are reported in Table 3.4. The top panel reports results for reading and the bottom panel for numeracy. The first two columns report results from the value-added model. Along with individual covariates, the specification reported in the first column includes a set of neighborhood characteristics, while the second column includes postal code fixed effects. The results show that the estimates of the private school effect is not sensitive to these alternative methods of controlling for unobserved factors that are correlated with neighborhood characteristics. These estimates are positive and statistically significant for both reading and numeracy, with effect sizes between 0.16 and 0.18 standard deviations.

The remaining columns of table 3.4 present results from the corresponding student fixed effects model. The estimated effect for the full sample, presented in the third column, is positive and statistically significant, but substantially smaller than in the value-added specification, only 0.10 standard deviations in reading and 0.08 standard deviations in math. The relatively small magnitude of these estimates compared to the value-added estimates could reflect several factors. First, they incorporate the effects of middle schools that don't offer grade 7, while the value-added estimates do not. Second, the effects of private schools on students who don't change school sectors may be greater than for those who do; since stayers do not contribute to the identification of school fixed effects, this could result in lower estimates.

Table 3.4: Estimates of private school effect on reading and numeracy scores, value-added and student fixed effects models

	(1)	(2)	(3)	(4)	(5)	(6)
READING						
	Value-added[a]		Student Fixed Effects[b]			
				No movers between		
	All students		All students	G4/G5	G6/G7	Compulsory movers
Private school	0.17***	0.16***	0.10***	0.09***	0.12***	0.12**
	[0.022]	[0.018]	[0.017]	[0.020]	[0.019]	[0.036]
R-squared	0.484	0.502	0.028	0.027	0.027	0.047
# of students	84774	84774	169419	151950	149071	26450
Census EA characteristics	x		x	x	x	x
Postal code fixed effects		x				
NUMERACY						
	Value-added[a]		Student Fixed Effects[b]			
				No movers between		
	All students		All students	G4/G5	G4/G5	Compulsory movers
Private school	0.18***	0.17***	0.08***	0.06**	0.10***	0.09*
	[0.035]	[0.028]	[0.021]	[0.023]	[0.023]	[0.045]
R-squared	0.471	0.494	0.056	0.054	0.056	0.089
# of students	83375	83375	166613	149573	146799	25843
Census EA characteristics	x		x	x	x	x
Postal code fixed effects		x				

Notes: Standard errors clustered at the school level. Additional control variables in all specifications include indicators for English as a Second Language and Aboriginal identity. Census neighborhood controls consist of mean family income in the student's Census EA/DA, proportion of household heads in the student's Census EA/DA who are immigrants, whose highest level of education is high school, a trade certificate, some college and a bachelor's degree, and. [a]Dependent variable is the student's grade 7 FSA test score. Additional control variables include gender, home language (Chinese, Punjabi, other non-English) and year fixed effects. [b]Dependent variable is the student's FSA test score in grade *g*. Additional control variables include grade-by-year fixed effects.

As described earlier, it is also possible that the student fixed effects estimates are subject to bias from dynamic selection into mover status. We address this threat by estimating our model for samples that exclude students who move between school sectors after different grades. The results in the fourth column of Table 3.4 exclude students who change school sectors immediately after grade 4; this exclusion will eliminate bias from transitory shocks that systematically affect the school sector choice of these movers and are correlated with grade 4 test scores. The results in the fifth column exclude students who change school sectors immediately after grade 6; this exclusion will eliminate bias from transitory shocks that systematically affect the school sector choice of these movers and are correlated with grade 7 test scores. Excluding post-grade 4 movers produces a smaller point estimate while excluding post-grade 6 movers produces a larger point estimate. This pattern of results is consistent with a scenario where negative shocks to test scores are more highly correlated with transitory shocks that cause public school students to move to private schools than with transitory shocks that cause private school students to move to public schools. The results in final column of table 4 correspond to the sub-sample of students who are required to change schools between grades 4 and 7 because of the grade configuration of their grade 4 schools. The point estimates in both reading and numeracy are very close to those from the full sample, despite the fact that they reflect effects for a sub-set of schools.

3.5.3. Results from the individual school fixed effects model

We next present estimates from specifications that include a full set of school fixed effects rather than a simple private school indicator variable. These estimates allow us to investigate heterogeneity within the public and private school sectors.[15] We begin this investigation by computing the differences between the average of the estimated school fixed effects for each type of private school and the average for public schools. These enrolment-weighted mean differences are presented in table 3. 5. As in table 3.4, the

[15] Recent work on charter school quality uses population data on public schools to investigate within-sector heterogeneity and characterize the evolution of the distribution of charter school effects over time in Texas (Baude et al. 2015) and North Carolina (Ladd et al. 2015). This dynamic perspective is of less interest in the context of B.C.'s mature private school sector.

Table 3.5: Student-weighted means of estimated private school fixed effects on reading and numeracy scores, by type of private school, value-added and student fixed effects models

	(1)	(2)	(3)	(4)	(5)	(6)
			READING			
				Student Fixed Effects[b]		
	Value-added[a]			No movers between		Compulsory movers
	All students		All students	G4/G5	G6/G7	
Prep	0.26***	0.25***	0.17***	0.20***	0.14**	0.34***
	[0.064]	[0.046]	[0.036]	[0.034]	[0.044]	[0.083]
Catholic	0.21***	0.21***	0.11***	0.09***	0.15***	0.04
	[0.030]	[0.022]	[0.022]	[0.026]	[0.026]	[0.079]
Other Christian	0.08***	0.07***	0.03	-0.01	0.05*	0.07
	[0.021]	[0.020]	[0.022]	[0.025]	[0.027]	[0.049]
Other faith	0.26**	0.25**	0.20***	0.20**	0.23***	0.59***
	[0.086]	[0.080]	[0.060]	[0.064]	[0.067]	[0.149]
R-squared	0.485	0.504	0.028	0.027	0.027	0.048
# of students	84774	84774	169419	151950	149071	26450
Census char.	x		x	x	x	x
Postal code FE		x				
			NUMERACY			
				Student Fixed Effects[b]		
	Value-added[a]			No movers between		Compulsory movers
	All students		All students	G4/G5	G6/G7	
Prep	0.26**	0.26**	0.18***	0.18***	0.16**	0.35***
	[0.087]	[0.061]	[0.047]	[0.043]	[0.055]	[0.090]
Catholic	0.29***	0.29***	0.14***	0.13***	0.16***	0.18*
	[0.046]	[0.035]	[0.027]	[0.031]	[0.031]	[0.075]
Other Christian	0.04	0.00	-0.05	-0.11***	-0.02	0.02
	[0.032]	[0.032]	[0.025]	[0.026]	[0.027]	[0.059]
Other faith	0.24*	0.25*	0.18*	0.20**	0.24**	-0.17
	[0.119]	[0.089]	[0.071]	[0.064]	[0.074]	[0.178]
R-squared	0.473	0.497	0.056	0.055	0.056	0.090
# of students	83375	83375	166613	149573	146799	25843
Census EA char.	x		x	x	x	x
Postal code FE		x				

Notes: Standard errors clustered at the school level. Reported effects are differences between student-weighted averages of estimated private school and public school fixed effects, by private school type. Control variables in all specifications include indicators for English as a Second Language and Aboriginal identity. Census neighborhood controls consist of mean family income in the student's Census EA/DA, proportion of household heads in the student's Census EA/DA who are immigrants, whose highest level of education is high school, a trade certificate, some college and a bachelor's degree. [a]Dependent variable is the student's grade 7 FSA test score. Additional control variables include grade 4 test score, gender, home language (Chinese, Punjabi, other non-English) and year fixed effects. [b]Dependent variable is the student's FSA test score in grade *g*. Additional control variables include grade-by-year fixed effects and a full set of individual fixed effects.

first two columns present results from the value-added model, with neighborhood characteristics and postal code fixed effects respectively, and the remaining columns present results from the student fixed effects model estimated for several different samples.

The mean differences for the value-added school effects estimates are positive and statistically significant for all private school types. The effect sizes are similar for prep schools, Catholic schools and non-Christian faith schools, ranging between 0.21 and 0.29 standard deviations. The effect size for Other (non-Catholic) Christian private schools is smaller, 0.08 standard deviations in reading and 0.04 standard deviations and statistically insignificant in numeracy. The corresponding mean differences for the student fixed effects model are smaller for non-Christian faith schools and prep schools, but are still positive, substantial and statistically significant, ranging between 0.17 and 0.20 standard deviations. The point estimates for Catholic schools are smaller, 0.11 standard deviations in reading and 0.14 standard deviations in numeracy, but remain statistically significant. The point estimates for other Christian schools are small and statistically insignificant. In most cases, excluding post-grade 4 movers produces a smaller point estimate while excluding post-grade 6 movers produces a larger point estimate. Again, this pattern of results is consistent with a scenario where negative shocks to test scores are more highly correlated with factors that cause public school students to move to private schools than with factors that cause private school students to move to public schools. In the case of prep schools, this pattern is reversed. When we restrict the sample to compulsory movers only, the effects for other Christian schools remain small and statistically significant. Among Catholic schools, the effect for reading is now smaller and statistically insignificant; among other faith schools the effect for numeracy is now negative and statistically insignificant. The estimated effects for the subset of prep schools included in this sample remain positive, substantial and statistically significant.

We further characterize heterogeneity within and across school types in table 3.6, using estimates of individual school and student fixed effects from the student fixed effects model. The first column shows the differences between school-weighted averages of estimated private school and public school fixed effects, by private school type. These mean estimates differ somewhat from those presented in the third column of table 3.5 because they are school-weighted, rather than student-weighted, averages of estimated school fixed effects. The second column of table 3.6 shows differences between average estimated student fixed effects in private and public schools, by private school type. Prep schools enroll the highest achieving students; these students on average would score about half a standard deviation above the public school mean in both reading and numeracy if they were enrolled in public schools. Students enrolled in Catholic and other Christian private schools are also positively selected on ability, but to a lesser degree. These students would score between 0.14 and 0.22 standard deviations above the public school mean if they were enrolled in public schools. Students attending other faith private schools are only slightly positively selected on ability, by 0.07 to 0.06 standard deviations.

The last two columns of table 3.6 characterize the standard deviations of estimated school and student effects by type of school. The standard deviation of school effects is approximately half again as large among private schools compared to public schools. The prep, Catholic and especially other faith private school sectors are all substantially more heterogeneous with respect to school effects than public schools, while other Christian private schools are no more heterogeneous than public schools. The within-school variation in student effects is moderately lower on average among prep schools and other faith private schools compared to public schools. These schools may be attracting a more homogeneous group of applicants, or applying more selective admission procedures. Students at Catholic and other Christian private schools exhibit slightly higher degrees of heterogeneity on average, but these schools still enroll relatively homogeneous students with respect to reading ability compared to public schools.

Figure 1 presents kernel density estimates of the distribution of school fixed effects for private and public schools estimated from the student FE model. The range of estimated school fixed effects is wider for private than for public schools because of fatter tails at both ends of the distribution. A fairly small proportion of private schools are less effective than the average public school in reading. This proportion is higher in the case of numeracy. Figure 2 presents analogous kernel density estimates of the student-weighted distribution of school fixed effects, which leads to a similar conclusion. Figure 3 presents estimates of the distribution of student fixed effects from the same set of estimates. The entire distribution of student fixed effects is shifted slightly to the right in private schools compared to public schools.

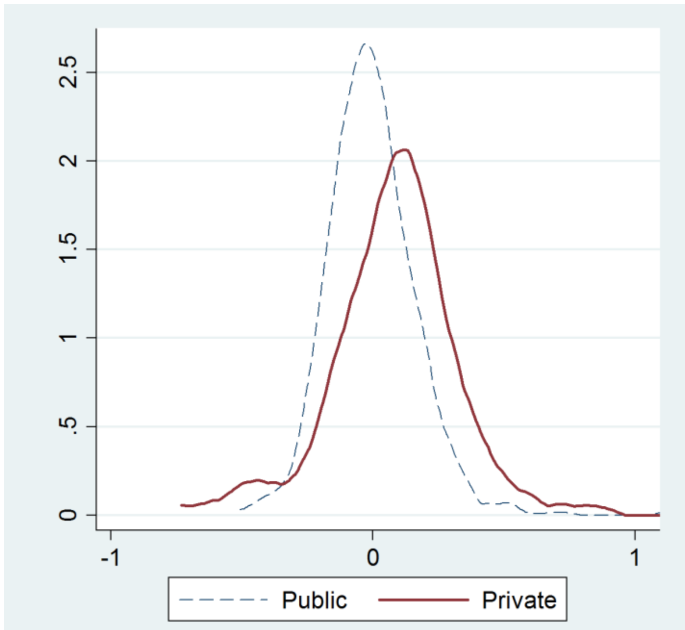
Table 3.6: School-weighted means and standard deviations of estimated private school and student fixed effects on reading and numeracy scores, by type of school, student fixed effects model

	(1)	(2)	(3)	(4)
	READING			
	Mean difference from public school mean[a]		Standard deviation	
	School effect/se	Student effect/se	School effect[b]	Student effect (within school)[c]
All public			0.17	0.79
All private	0.10***	0.27***	0.30	0.75
	0.021	0.023		
Prep	0.22***	0.57***	0.23	0.71
	0.037	0.044		
Catholic	0.11***	0.22***	0.21	0.76
	0.033	0.035		
Other Christian	0.03	0.21***	0.23	0.77
	0.034	0.033		
Other faith	0.14**	0.07	0.76	0.74
	0.054	0.060		
	NUMERACY			
	Mean difference from public school mean[a]		Standard deviation	
	School effect/se	Student effect/se	School effect[b]	Student effect (within school)[c]
All public			0.23	0.81
All private	0.09***	0.22***	0.31	0.80
	0.019	0.020		
Prep	0.21***	0.51***	0.36	0.77
	0.036	0.040		
Catholic	0.15***	0.14***	0.26	0.80
	0.035	0.039		
Other Christian	-0.06*	0.18***	0.27	0.81
	0.030	0.034		
Other faith	0.21***	0.06	0.45	0.74
	0.066	0.065		

Notes: Standard errors are bootstrapped. Dependent variable is the student's FSA test score in grade *g*. Control variables include indicators for English as a Second Language, mean family income in the student's Census EA/DA, proportion of household heads in the student's Census EA/DA who are immigrants, whose highest level of education is high school, a trade certificate, some college and a bachelor's degree, grade-by-year fixed effects. Reported effects are: [a]differences between school-weighted averages of estimated private school and public school fixed effects, by private school type; [b]standard deviations of estimated individual school fixed effects, by school type; and [c]within-school standard deviations of estimated individual student fixed effects, by school type.

Figure 3.1: Kernel density estimates of the distribution of estimated school effects, student FE model, full sample, by school sector

Reading



Numeracy

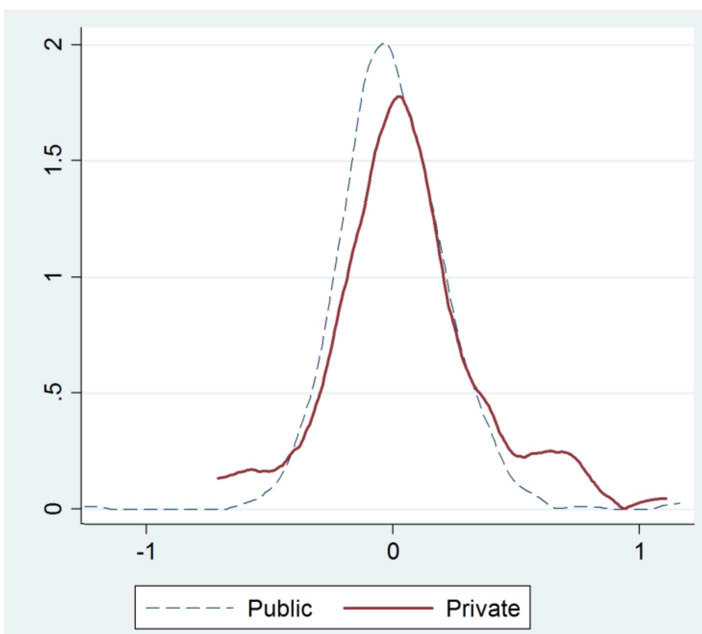
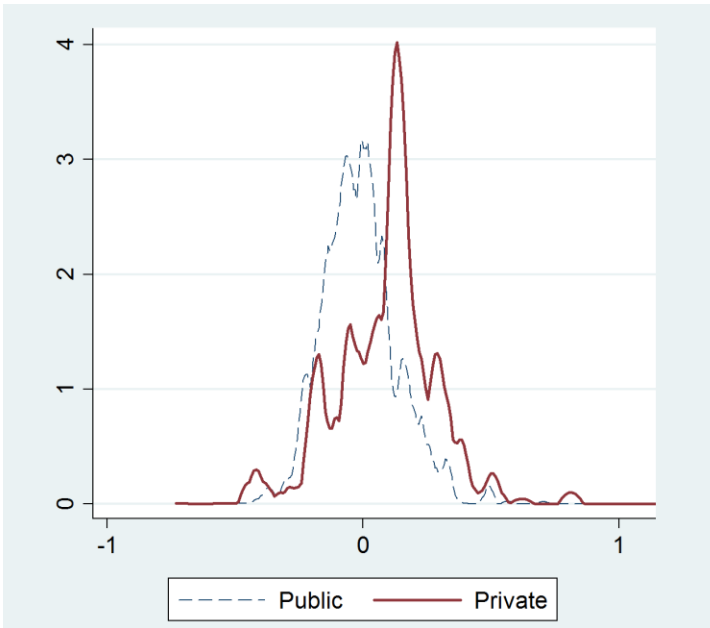


Figure 3.2: Kernel density estimates of the student-weighted distribution of estimated school effects, student FE model, full sample, by school sector

Reading



Numeracy

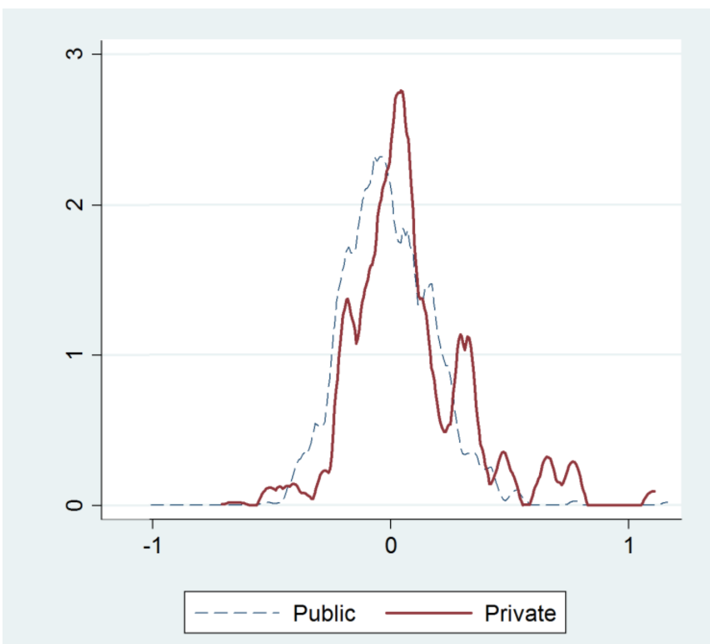
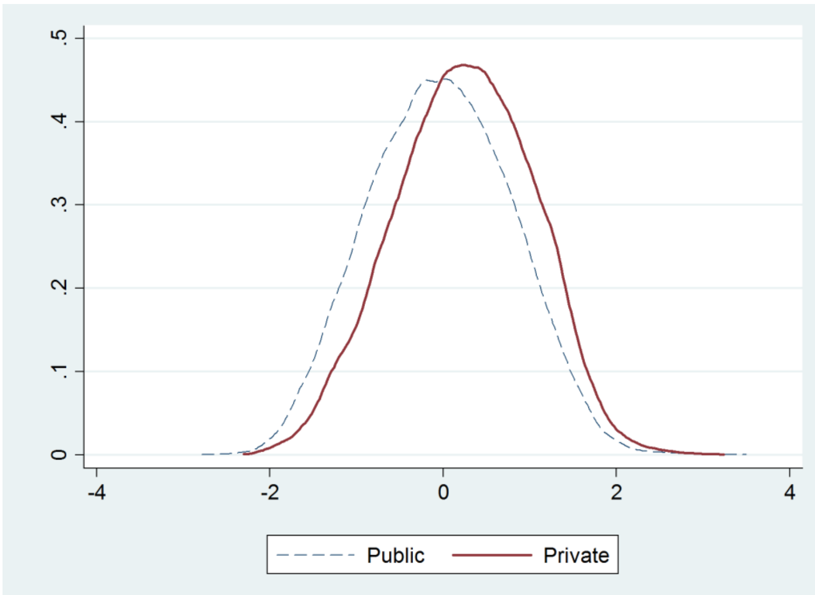
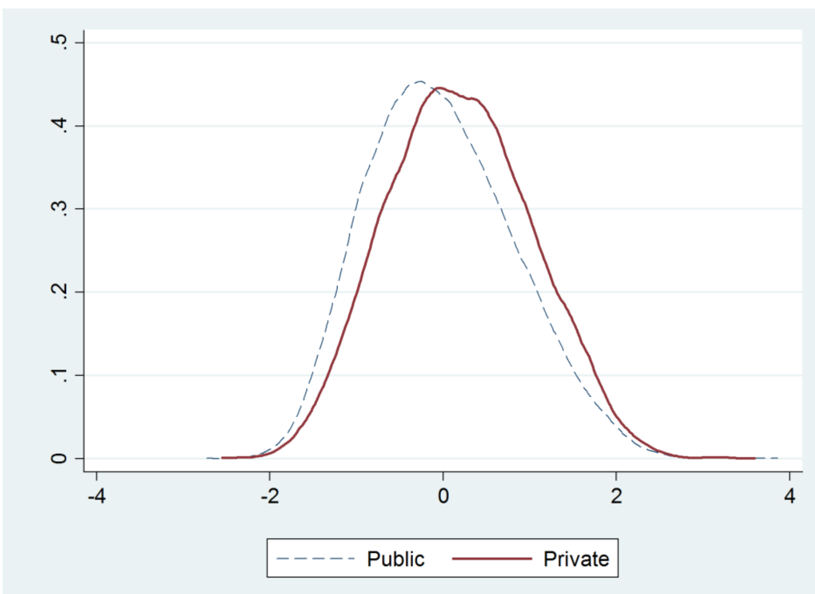


Figure 3.3: Kernel density estimates of the distribution of estimated student effects, student FE model, full sample, by school sector

Reading



Numeracy



3.6. CONCLUSION

All of the results in this paper point to the relative success of the average private primary school in British Columbia with respect to literacy and numeracy skill development, compared to its public counterpart. This evidence contradicts the results of most previous studies of the effects of private schools on test scores in other advanced economies; Catholic private schools in particular have been associated with no better or even relatively poor test scores in both the U.S. and Australia (e.g. Altonji et al. 2005; Elder and Jepsen 2014; Nghiem et al. 2015). This qualitative difference in our conclusions about the relative effectiveness of private schools may reflect differences in institutions. Like private schools in many jurisdictions, B.C.'s private schools enjoy substantial autonomy with respect to teacher hiring, firing, discipline and remuneration, are free to apply a wide range of admissions criteria and can charge tuition. Unlike private schools in some jurisdictions, B.C. private schools are required to hire teachers who are provincially certified, and they receive partial funding from the provincial government. Published school-level test score results include private schools, and these receive a substantial amount of media and public attention.

Nevertheless, the range of quality among private schools is substantial, and weaker private schools appear to be less effective than many public schools. The survival of these schools, many of which have been in operation for decades, invites several potential explanations. One possibility is that parents may value private schools for a variety of reasons beyond their contributions to cognitive skill development; religious or moral instruction, peer quality, enriched supervision and monitoring of student behavior are examples of school qualities that parents are willing to pay for that may not substantially affect cognitive skills. A second possibility is that some private schools may be located in neighborhoods with relatively weak public schools. If it is cheaper for some parents to send their children to a local private school than to relocate to a more expensive neighborhood with better public schools, these private schools need only outperform their local public counterparts. To the extent that the equilibrium distribution of school quality reflects these sorts of constraints on individual school choice, it also suggests a possible explanation for why our results differ qualitatively for the average private school than those of previous studies. Unlike research based on national survey data, we are able to include detailed geographic controls in our population-based study. Our estimates therefore are based to a greater degree on comparisons of public and private schools that enroll students who live in similar neighborhoods.

While we cannot claim to have fully addressed the issue of non-random selection into private schools, we also have several reasons to believe that any positive bias from this source may be offset at least to some degree by other sources of negative bias. In the value-added estimates, measurement error in the lagged test score may be substantial, leading to attenuation bias in our estimates of private school effects. Moreover, our approach is more likely to produce estimates that capture general equilibrium effects of private school competition on public school quality compared to previous studies, since these effects will be more pronounced in comparisons among schools that are geographically proximate. Since, these spillover effects onto public school quality via competition will reduce the quality difference between public and private schools, our estimates will capture only part of the total effect of private schools on student achievement (see Epple et al. 2015 for a review of relevant evidence of the effects of private school competition).

Our methodology provides no direct evidence of the mechanisms that may be driving our results. However, it is of some interest to note that our analysis by private school type finds positive effects for some but not all groups of faith private schools as well as positive effects for secular prep private schools. This pattern suggests that it is not faith-based education per se that is contributing to the relative effectiveness of some private schools. This interpretation is consistent with existing direct evidence that faith public schools are no more effective than secular public schools (Gibbons and Silva 2011). It is also worth noting that, while private schools on average enroll higher quality students and deliver higher quality outcomes, this is not true in all private school sectors.

In particular, other Christian schools enroll relatively high quality students but do not get better results on average than public schools, while other faith schools enroll fairly average students but get better than average results. This pattern is not consistent with a simple story of peer effects driving variation in school performance across public versus private school types.

References

Abbasi, H. Farjadi, Gh. (1999) Labour Force Participation in Iran. (In Farsi) In: ed. M. Tabibian. Iran's Economy 2. Institute for Research on Planning and Development Publication, Tehran, Iran.

Abbasi-Shavazi, M. J., McDonald, P., & Hosseini-Chavoshi, M. (2009). *The fertility transition in Iran:* Revolution and reproduction. Springer, Dordrecht

Abowd, J., R. Creecy and F. Kramarz (2002). Computing person and firm effects using linked longitudinal employer-employee dataset, Technical report, US Census Bureau.

Aghajanian, A. (1995). A new direction in population policy and family planning in the Islamic Republic of Iran. *Asia-Pacific Population Journal / United Nations*, *10*(1), 3–20.

Agüero, J. M., & Marks, M. S. (2008). Motherhood and female labour force participation: evidence from infertility shocks. *The American Economic Review*, 500–504.

Alderman, Harold, Peter F. Orazem and Elizabeth M. Paterno (2001). School quality, school cost and the public/private school choices of low-income households in Pakistan. Journal of Human Resources 36(2): 304-26.

Almond, D., & Edlund, L. (2008). Son-biased sex ratios in the 2000 United States Census. *Proceedings of the National Academy of Sciences*, *105*(15), 5681–5682.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005a). An evaluation of instrumental variable strategies for estimating the effects of Catholic schooling,Journal of Human Resources 40(4): 791-821.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005b). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools, Journal of Political Economy 113: 151–184.

Andrabi, T., J. Das, A.I. Khwwja and T. Zajonc (2011). Do value-added estimates add value? Accounting for learning dynamics. American Economic Journal: Applied Economics 3(3): 29-54.

Angrist J, Bettinger E, Kremer M. (2006). Long-term educational consequences of secondary school vouchers: evidence from administrative records in Colombia, American Economic Review 96(3):847-62.

Angrist Joshua, Eric Bettinger E, Bloom E, King E, Kremer M. (2002). Vouchers for private schooling in Columbia: randomized natural experiment. American Economic Review 92(5):1535-58.

Angrist, J. D., & Evans, W. N. (1998). Children and Their Parents' Labour Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, *88*(3), 450–477.

Angrist, Joshua, P.A. Pathak and C.R. Walters (2013). Explaining charter school effectiveness. American Economic Journal: Applied Economics 5(4): 1-27.

Azimi, E. "The effect of children on female labour force participation in urban Iran." IZA Journal of Labour & Development 4.1 (2015): 1-11.

Baude, P.L., M. Casey, E.A. Hanushek and S.G. Rivkin (2014). The evolution of charter school quality. NBER Working Paper Number 20645, National Bureau of Economic Research.

Becker GS, Lewis HG (1974) Interaction between quantity and quality of children. In Economics of the family: Marriage, children, and human capital. UMI, pp 81–90

Ben-Porath, Y., & Welch, F. (1976). Do sex preferences really matter? *The Quarterly Journal of Economics*, 285–307.

Bhalotra, S., & Attfield, C. (1998). Intrahousehold resource allocation in rural Pakistan: a semiparametric analysis. Journal of Applied Econometrics, 13(5), 463–480.

Blundell, R., Kristensen, D., & Matzkin, R. L. (2013). Control functions and simultaneous equations methods. The American Economic Review, 103(3), 563-569.

British Columbia Federation of Independent Schools Association (2015). Who are we? http://www.fisabc.ca/who-are-we/history

British Columbia Ministry of Education (2001) Interpreting and communicating Foundation Skills Assessment Results 2001 (Victoria: Government of British Columbia).

British Columbia Ministry of Education (2005). Overview of Independent Schools in British Columbia. http://www.bced.gov.bc.ca/independentschools/geninfo_05.pdf

British Columbia Ministry of Education (2011). Provincial Report: Student Statistics. http://www.bced.gov.bc.ca/reports/pdfs/student_stats/prov.pdf

Brochu, Pierre, Marie-Anne Deussing, Koffi Houme and Maria Chuy (2012). Measuring up: Canadian results of the OECD PISA study. Council of Ministers of Education, Canada.

Bronars, S. G., & Grogger, J. (1994). The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment. *American Economic Review*, *84*(5), 1141–1156.

Browning, M., Chiappori, P.-A., & Lewbel, A. (2013). Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power. Review of Economic Studies, 80(4), 1267–1303.

Buddin, Richard J., Joseph J. Cordes and Sheila Nataraj Kirby (1998). School choice in California: Who chooses private schools? Journal of Urban Economics 44(1): 110-134

Carbonaro, William (2006). Public-private differences in achievement among kindergarten students: differences in learning opportunities and student outcomes. American Journal of Education 113(1): 31-66.

Chiappori, P.-A. (1988). Rational Household Labour Supply. Econometrica, 56(1), 63.

Chiappori, P.-A. (1992). Collective Labour Supply and Welfare. Journal of Political Economy, 100(3), 437–67.

Chun H, Oh J (2002) An instrumental variable estimate of the effect of fertility on the labour force participation of married women. Applied Economics Letters 9(10):631-634

Clark, D. (2009). The performance and competitive effects of school autonomy.Journal of Political Economy, 117(4):745-783.

Cowley, Peter and Stephen T. Easton (2003). Report Card on British Columbia's Elementary Schools: 2003 Edition (Vancouver, B.C.: Fraser Institute).

Das Gupta, M. Zhenghua, J. Bohua, L. Zhenming, X. Chung, W. & Hwa-Ok, B. Why is Son Preference so Persistent in East and South Asia? A Cross-Country Study of China, India, and the Republic of Korea. (2003). The Journal of Development Studies, 40(2), 153-187.

Deaton, A. (1997). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. World Bank Publications.

Deming, Dave J., Justine S. Hastings, Thomas J. Kane and Douglas O. Staiger (2014). School choice, school quality and academic achievement. American Economic Review 104(3): 991-1013.

Dunbar, G. R., Lewbel, A., & Pendakur, K. (2013). Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi. American Economic Review, 103(1), 438–471.

Elder, Todd and Christopher Jepsen (2014). Are Catholic primary schools more effective than public primary schools? Journal of Urban Economics 80:28-38.

Epple, Dennis, Richard E. Romano and Miguel Urquiola (2015a). School vouchers: A survey of the economics literature. NBER Working Paper Number 21523, National Bureau of Economic Research.

Evans, William N. and Robert M. Schwab (1995). Finishing high school and starting college: Do Catholic schools make a difference? Quarterly Journal of Economics 110 (4): 941–974.

Figlio, David N. and Joe A. Stone (1997). School choice and student performance: Are private schools really better? Discussion Paper 1141-97, Institute for Research on Poverty.

Figlio, David N., and Cassandra M. D. Hart (2014). Competitive effects of means-tested school vouchers. American Economic Journal: Applied Economics 6(1): 133-56.

Fraser Institute (2008). Who We Are. http://www.fraserinstitute.org/aboutus/whoweare/

Friesen, Jane and Brian V. Krauth (2011). Ethnic enclaves in the classroom. Labour Economics 18: 656-663.

Gibbons, Stephen and Olmo Silva (2011). Faith primary schools: Better schools or better pupils? Journal of Labour Economics 29(3): 589-635.

Goldin, C. (1995). *Career and family: College women look to the past*. (No. w5188) National Bureau of Economic Research.

Gong, X., van Soest, A., & Zhang, P. (2000). Sexual Bias and Household Consumption: A Semiparametric Analysis of Engel Curves in Rural China (IZA Discussion Paper No. 212). Institute for the Study of Labour (IZA).

Grogger, Jeffrey and Derek Neal (2000). Further evidence on the effects of Catholic secondary schooling. In: Gale,William G.,Pack, Janet Rothenberg (Eds.), Brookings-Wharton Papers On Urban Affairs, 2000. Brookings Institution Press,Washington, DC: 151–193.

Gronau, R. (1988). Consumption Technology and the Intrafamily Distribution of Resources. Journal of Political Economy, 96(6), 1183–1205.

Gugl, E., & Welling, L. (2012). Time with sons and daughters. Review of Economics of the Household, 10(2), 277-298.

Hanushek, Eric A., Kain, John F., Rivkin, Steve G.(2004). Disruption versus Tiebout improvement: the costs and benefits of switching schools. Journal of Public Economics 88(9–10): 1721–1746.

Heckman, J. J. and R. Robb (1985), "Alternative Methods for Evaluating the Impact of Interventions," in Longitudinal Analysis of Labour Market Data, (eds J. Heckman and B. Singer), CUP.

Hill, M. A., & King, E. (1995). Women's education and economic well-being. Feminist Economics, 1(2), 21–46.

Hoodfar, H. (1996). Bargaining with fundamentalism: women and the politics of population control in Iran. *Reproductive Health Matters*, *4*(8), 30–40.

Hurst, E. and A. Lusardi (2004): "Liquidity Constraints, Household Wealth, and Entrepreneurship," Journal of Political Economy, 112, 319–347.

Imberman, Scott A. (2011). Achievement and behavior in charter schools: Drawing a more complete picture. Review of Economics and Statistic 93(2): 416-435.

Jacob, Brian A., Lars Lefgren and David P. Sims (2010). The persistence of teacher-induced learning. Journal of Human Resources 45(5): 915-943.

Jepsen, Christopher (2003). The effectiveness of Catholic primary schooling. Journal of Human Resources 38(4): 928-941.

Kane, Thomas J. and Douglas O. Staiger (2008). Estimating teacher impacts on student achievement: an experimental evaluation. Working Paper 14607, National Bureau of Economic Research.

Kane, Thomas J., Daniel F. McCaffrey, Trey Miller and Douglas O. Staiger (2013). Have we identified effective teachers? Validating measures of effective teaching using random assignment. Research Paper. MET Project. Technical report. Bill and Melinda Gates Foundation.

Kingdon, G. G. (2005). Where has all the bias gone? Detecting gender bias in the intrahousehold allocation of educational expenditure. Economic Development and Cultural Change, 53(2), 409–451.

Kramarz, Francis, Stephen Machin and Amine Ouazad (2014). Using compulsory mobility to identify school quality and peer effects. Oxford Bulletin of Economics and Statistics :1-22.

Ladd, Helen F., Charles T. Clotfelter and John B. Holbein (2015). The growing segmentation of the charter school sector in North Carolina. NBER Working Paper No. 21078. National Bureau of Economic Research.

Lancaster, G., Maitra, P., & Ray, R. (2003). Endogenous power, household expenditure patterns and new tests of gender bias: evidence from India. In Preliminary Version of paper presented at the NEUDC conference.

Lefebvre, Pierre, Philip Merrigan and Matthieu Verstraete (2011). Public subsidies to private schools do make a difference for achievement in mathematics: Longitudinal evidence from Canada. Economics of Education Review 30: 79-98.

Lise, J., & Seitz, S. (2011). Consumption inequality and intra-household allocations. The Review of Economic Studies, 78(1), 328-355.

Long, X. (1989). The higher sex ratio in the modern population of China. Wei Ding Gao, 8, 53-59.

Lubienski, Christopher, Corinna Crane and Sarah Thule Lubienski (2008) What do we know about school effectiveness? Academic gains in public and private schools. Phi Delta Kappan 89(9): 689-695.

Majbouri, M. (2010). Against the Wind: Labour Force Participation of Women and Economic Instability in Iran. *Available at SSRN 2419323*.

Majbouri, M. (2015). Oil, Patriarchy, and Female Labour Force Participation. *Working paper*

Martinez-Mora, Francisco (2006). The existence of non-elite private schools. Journal of Public Economics 90: 1505-1518.

McEwan, Patrick J. (2001). The effectiveness of public, Catholic and non-religious private schools in Chile’s voucher system. Education Economics 9(2): 103-128.

Mehryar, A. H., Ahmad-Nia, S., & Kazemipour, S. (2007). Reproductive health in Iran: pragmatic achievements, unmet needs, and ethical challenges in a theocratic system. *Studies in Family Planning*, *38*(4), 352–361.

Mehryar, A. H., Farjadi, G., Tabibian, M., & Population Studies and Research Center for Asia and the Pacific (Iran). (2002). Labour force participation of women in Iran: trends and covariates. Tehran, Iran: Ministry of Science, Research & Technology, Population Studies and Research Center for Asia and the Pacific.

Michael R. T., 1974. "Education and the Derived Demand for Children," NBER Chapters, in: Economics of the Family: Marriage, Children, and Human Capital, pages 120-159 National Bureau of Economic Research, Inc.

Moghadam, V. M. (1992). Patriarchy and the Politics of Gender in Modernising Societies: Iran, Pakistan and Afghanistan. International Sociology, 7(1), 35–53.

Muellbauer, J. (1976). Community Preferences and the Representative Consumer. Econometrica, 44(5), 979–99.

Muralidharan, Karthik and Venkatesh Sundararaman (2015). The aggregate effect of school choice: Evidence from a two-stage experiment in India. Quarterly Journal of Economics, forthcoming.

Neal, Derek (1997). The effect of Catholic secondary schooling on educational attainment, Journal of Labour Economics 15(1 Part 1): 98–123.

Nghiem, Hong Son, Ha Trong Nguyen, Rasheda Khanam and Luke B. Connelly (2015). Does school type affect cognitive and non-cognitive development in children? Evidence from Australian primary schools. Labour Economics 33: 55-65.

Oster, Emily (2015). Unobservable selection and coefficient stability: Theory and evidence. Working Paper.

Polachek, D. E. and Polachek, S. W. (1989), An Indirect Test of Children's Influence on Efficiencies in Parental Consumer Behavior. Journal of Consumer Affairs, 23: 91–110.

Rasul-Ronning Z. (2013). Women in Iraq and Iran: Leadership and the Right to Equal Participation. In: Winkler, P. Confronting the International Patriarchy: Iran, Iraq and the United States of America. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Reardon, Sean F., Jacob F. Cheadle and Joseph P. Robinson (2009). The effects of Catholic schooling on math and reading development in kindergarten through fifth grade. Journal of Research on Educational Effectiveness 2(1): 45-87

Rivkin, Steve G., Eric A. Hanushek and John F. Kain (2005). Teachers, schools, and academic achievement. Econometrica 73(2): 417–458.

Rose, E. (1999). Consumption Smoothing and Excess Female Mortality in Rural India. Review of Economics and Statistics, 81(1), 41–49.

Rosenzweig, M. R., & Wolpin, K. I. (1980). Testing the quantity-quality fertility model: the use of twins as a natural experiment. *Econometrica: Journal of the Econometric Society*, *48*(1), 227–240.

Rothbarth, E. (1943), "Note on a Method of Determining Equivalent Income for Families of Different Composition," in: Madge, C. (Ed.), Wartime Pattern of Saving and Spending, London: Faber and Faber Limited, 123-130.

Rouse, Cecilia and Lisa Barrow (2009). School vouchers and student achievement: Recent evidence, remaining questions, Annual Review of Economics 17-42.

Sander, William (1996). Catholic grade schools and academic achievement. Journal of Human Resources 31(2): 540-548.

Schultz, TP (1981) Economics of population. Perspectives on Economics Series

Shack, Daniel (2015). The impact of parents' uncertainty and learning on dynamic educational choices and student achievement. Presented at the CEA meetings.

Sharabi, Hisham. 1988. Neopatriarchy: A Theory of Distorted Change in Arab Society. New York: *Oxford University Press.*

Singh, Abhijeet (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. Journal of Development Economics 113: 16-32.

Stock JH, Yogo M (2005) Testing for weak instruments in linear IV regression. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg

Subramanian, S., & Deaton, A. (1991). Gender effects in Indian consumption patterns. Sarvekshana, 14(4), 1–12.

Todd, Petra E. and Kenneth I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. Economic Journal 113(485): F3–F33.

UNICEF (2007) State of the World's Children : the Double Dividend of Gender Equality, New York: UNICEF

United Nations Population Fund (UNFPA) (2010) State of the World's Population 2010. From conflict and crisis to renewal: generations of change, UNFPA, New York, p.96

Working, H. (1943). Statistical Laws of Family Expenditure. Journal of the American Statistical Association, 38(221), 43–56.

Zimmermann, L. (2012). Reconsidering Gender Bias in Intrahousehold Allocation in India. Journal of Development Studies, 48(1), 151–163.

Appendix A.
Further Statistics

Table A1: Fertility and FLFP in Iran during 1990–2004

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
21–35 year old women															
number of children	3.26	3.29	2.99	2.91	2.89	2.86	2.59	2.46	2.43	2.38	2.16	2.05	2.01	1.96	1.89
more than 2 children	0.64	0.64	0.56	0.54	0.54	0.53	0.46	0.42	0.41	0.40	0.33	0.30	0.30	0.27	0.25
FLFP	0.12	0.12	0.15	0.14	0.12	0.15	0.13	0.13	0.12	0.14	0.14	0.12	0.13	0.14	0.13
observation	3802	3656	2564	2730	4816	7742	4183	4182	2983	4385	4664	4454	5584	4060	4297
36–50 year old women															
number of children	3.98	4.02	3.98	3.93	3.85	3.97	3.75	3.70	3.62	3.69	3.56	3.52	3.41	3.28	3.19
more than 2 children	0.75	0.76	0.75	0.75	0.74	0.77	0.74	0.75	0.73	0.73	0.72	0.71	0.69	0.67	0.64
FLFP	0.13	0.13	0.17	0.15	0.14	0.16	0.13	0.14	0.14	0.15	0.14	0.13	0.13	0.14	0.14
observation	2689	2835	1829	2086	3995	6926	3850	3832	3065	4806	4344	4474	5366	3946	4184
21–35 year old women with two or more children where the oldest child is younger than 18 years															
number of children	3.67	3.71	3.54	3.45	3.44	3.38	3.21	3.10	3.07	3.01	2.91	2.84	2.83	2.73	2.70
more than 2 children	0.74	0.74	0.70	0.68	0.68	0.66	0.62	0.58	0.57	0.56	0.51	0.47	0.49	0.45	0.43
FLFP	0.11	0.11	0.14	0.13	0.11	0.13	0.11	0.11	0.10	0.12	0.12	0.10	0.10	0.12	0.11
observation	3280	3150	2059	2182	3836	6201	3141	3032	2162	3161	3050	2789	3408	2477	2517

Source: Iranian Household Expenditure and Income Surveys

Table A2: Under-18 Population by Gender and the Sex Ratio in Iran

	1996 Census	2006 Census
male	14,465,424	11,712,684
female	13,889,253	11,142,509
sex ratio	1.041	1.051

Source: Statistical Yearbook, Statistical Center of Iran

Appendix B.

Merging Student information to Census characteristics

We incorporate students' residential information to our data set by merging it to Canadian Census. This enables us to create a rich data set by including variables such as mean family income and the percentage of household heads in the student's Census Enumeration/Dissemination area who are immigrants, who are without high school diploma, who have a high school diploma, a post-secondary certificate or a bachelor's degree.

We merge students in 2006 school year to 2006 census, students in 2001-05 school years to 2001 Census, and students in 1999-2000 school years into 1996 Census. We use Postal Code Conversion Files (PCCF) to associate each postal code in student data set to Enumeration Area (EA)/ Dissemination Area (DA) codes in Census.

Education variables are reported differently in different Census years. The above categorization provides us with variables that are consistent and comparable for different years of census.

Please refer to Statistics Canada website for further information regarding the Canadian Census, area codes, and PCCF.