NBER WORKING PAPER SERIES

NEW EVIDENCE ON THE IMPACTS OF ACCESS TO AND ATTENDING UNIVERSAL CHILDCARE IN CANADA

Michael J. Kottelenberg
Steven F. Lehrer

Working Paper 18785
http://www.nber.org/papers/w18785

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2013

We would like to thank Habiba Djebbari and seminar participants at the University of Toronto, 2010 CEA Annual meetings, John Deutsch Institute's conference on Economic Relations Between Children and Parents, 2012 RES annual meetings, and the CEPS/INSTEAD's conference on the Theory and Practice of Program Evaluation for helpful comments and suggestions. We would like to thank Kevin Milligan for generously answering a number of questions regarding earlier analysis of the data used in the study. This paper is a revised version of a portion of Kottelenberg's Queen's University 2009 Master's research paper. Lehrer wishes to thank SSHRC for research support. We are responsible for all errors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

New Evidence on the Impacts of Access to and Attending Universal Childcare in Canada
Michael J. Kottelenberg and Steven F. Lehrer
NBER Working Paper No. 18785
February 2013
JEL No. C31,H2,I28,J2

ABSTRACT

In Canada, advocates of universal child care often point to policies implemented in Quebec as providing a model for early education and care policies in other provinces. While these policies have proven to be incredibly popular among citizens, initial evaluations of access to these programs indicated they led to a multitude of undesirable child developmental, health and family outcomes. These research findings ignited substantial controversy and criticism. In this study, we show the robustness of the initial analyses to i) concerns over whether negative outcomes would vanish over time as suppliers gained experience providing child care, ii) concerns regarding multiple testing, and iii) concerns that the original test measured the causal impact of childcare availability and not child care attendance. A notable exception is that despite estimated effects stemming from the policy indicating declines in motor-social development scores in Quebec relative to the rest of Canada, our analyses imply that on average attending childcare in Canada leads to a significant increase in this test score. However, our analysis reveals substantial heterogeneity in program impacts that occur in response to the Quebec policies and indicates that most of the negative impacts reported in earlier research are driven by children from families who only attended childcare in response to the implementation of this policy.

Michael J. Kottelenberg
Department of Economics
Queen's University
Kingston, Ontario K7L 3N6
Canada
kottelen@econ.queensu.ca

Steven F. Lehrer
School of Policy Studies
and Department of Economics
Queen's University
Kingston, ON K7L 3N6
CANADA
and NBER
lehrers@post.queensu.ca

1 Introduction

While many parents, teachers, and politicians have concluded that there is strong evidence to support increased public sector support for the delivery of childcare services, the opinion is not supported by the entire body of evidence from the research community. Proponents of early childhood education often point to a large body of evidence summarized in Heckman (2006, 2011) which indicates that investment and intervention aimed at young children will yield greater societal and individual benefits relative to those aimed at older children or adults. Much of this evidence is drawn from evaluations of targeted early childhood education programs.[1] Unfortunately, less is known about the effects of universal care on child outcomes. The few studies that have examined the effects of introducing universal childcare have yielded at best what Baker (2011) described as mixed evidence.[2] Despite the paucity of evidence indicating its effectiveness, there has been a growing focus—in both public and academic spheres—on the need for governmental participation in the provision of childcare.

Ongoing interest in childcare has been stimulated in part by trends associated with an increase in female labour force participation. In Canada, the employment rate of mothers with children under the age of six has risen from 31 percent in 1976 to 71 percent in 2008.[3] Both the supply and cost of childcare have also added impetus to demands for publicly funded childcare. In 2008 CBC News reported growing complaints concerning the limited availability of space in existing day-care facilities. The childcare Advocacy Association of Canada suggested that, "the federal government has simply failed to meet the childcare

[1]Findings that investing in children at early ages are not new and a wide body of research reports that pre-kindergarten programs can boost subsequent outcomes for disadvantaged children. See Currie (2001) or the wide-body of evaluation research that examined either the Perry Preschool Project or the Abecedarian Project.

[2]For example, Baker et al.(2008) and Datta Gupta and Simonsen (2010) report significant declines in a number of developmental and behavioural outcomes, whereas Havnes and Mogstad (2011) report strong positive effects on children's long-run outcomes as adult.

[3]The statistics for Canada are derived by the authors from the Canadian Labour Force Survey. Note, a less pronounced trend exists in the United States where data from the CPS indicates that the employment rate of mothers with children under the age of six has risen from 34 percent in 1976 to 56 percent in 2001.

needs of Canadian families."[4]

Regarding costs of childcare, data from the Survey of Income and Program Participation indicate that the average American family using childcare pays a whopping $6,708(US) annually for this service. In this climate, it is unsurprising that universal, publicly funded child care occupies a growing place on political agendas. During the 2011 federal election in Canada, the Liberal Party's campaign included promises to provide the provinces with $500 million a year to create child-care spaces. Throughout North America whenever discussions on reforming early childhood education are debated, Quebec's heavily subsidized $7-a-day childcare system is presented as a template. Quebec's child care system was introduced as a component of the Quebec Family Policy, which remains one of the most comprehensive policy measures taken by any North American government in response to childcare trends and concerns. In 1997, the Quebec government implemented a bold set of policies in hopes of encouraging higher birth rates, primarily by strengthening governmental support for parents. In large part, this support came in the form of an expansion of the childcare system. Under the Quebec Family Policy, parents with children aged zero to four were granted access to childcare at a rate of five-dollar-per-day (increased to seven-dollar-per-day in 2004).[5] This program was implemented gradually; access was extended to children aged four in 1997, aged three in 1998, aged two in 1999 and aged zero to two in 2000. Additionally, full-day kindergarten was introduced (for children age five) and, extraneous to the aforementioned policy, more childcare spaces for school aged children.[6]

The Quebec Family Policy's extension of highly subsidized universally available childcare to children aged zero to four provides a unique opportunity to examine the impact of a switch to a comprehensive system of childcare support. In an influential paper, Baker, Gruber and

[4]Examining data from Canada, one can observe that the province of Quebec significantly increased the number of licensed childcare places available to children living in the province over the past 15 years. In fact, the OECD (2007) reports that by itself, Quebec accounts for almost all the increase in regulated early childhood education and care (ECEC) places in Canada since 1998.

[5]The Quebec Family Policy also increased parental leave benefits and provided families with a standard child allowance based on income, family type (single parent, two parent), and number of children.

[6]See Tougas (2002) for more details.

Milligan (2008) (henceforth referred to as BGM) provided the first formal analyses of the impacts of the Quebec Family Policy. The analyses in BGM provided evidence that the introduction of universal childcare led to statistically significant reductions in a variety of child health, developmental, and behavioural measures. Outcomes include increased hyperactivity, inattention, physical aggression, and decreased motor social development scores. In addition, the authors' analyses indicated that parenting practices and family functioning were negatively affected in Quebec.[7] Their results ignited substantial controversy and advocacy groups such as the United Early Childhood Employees and Child Care Advocacy Association of Canada provided critiques.[8] Several of these critiques including Cleveland (2007) highlighted the fact that the BGM study involved examination of the impact of access to childcare, not the impact of the utilization of childcare. As such, the BGM study does not provide evidence regarding the outcomes associated with child care attendance and use, which, according to the above mentioned advocacy groups, should be the true parameters of policy interest.

This study extends Baker, Gruber, and Milligan (2008) in three ways. First, BGM evaluated the Quebec Family Policy at a time when the program was newly implemented and childcare centers were attempting to expand their services accordingly. The implementation of large scale social programs is rarely frictionless, since new programs require time for recruitment of high quality employees and integration in society. Therefore, early evaluations of child care programs may capture short-term changes that differ from results that may be found when child care centers are better established. Since the BGM study, two cycles of data have been made available from the National Longitudinal Study of Children and Youth, enabling the present study to revisit the original results to determine whether BGM results

[7]The authors as well as Lefebvre and Merrigan (2009) showed that the Quebec Family Policy significantly increase maternal labour supply. Calculations in Fortin (2011) suggest that the increases in labour supply generate sufficient additional tax revenue to fully finance this program.

[8]See testimony to The Standing Senate Committee on Social Affairs, Science and Technology on Thursday, February, 14, 2008 by the late Fraser Mustard who founded the Council for Early Child Development for a heated critique of this research (Mustard, 2008) as well as a report by Cleveland et al. (2006) that was prepared for The Canadian Council on Learning.

are robust to the inclusion of data from a later time period.[9]

Second, we extend past research, which focused primarily on intent to treat estimates of childcare in Canada. Using an instrumental variables estimator, we will first attempt to identify the causal impact of attending childcare due to the Quebec policy on both child and parent outcomes. This empirical strategy allows us to report whether the negative impacts found by BGM are driven by childcare itself, as many readers of BGM hypothesized. From a policy perspective, this parameter is quite important since it provides information on all individuals whose childcare decisions were affected at the margin by the policy. Second, we will use an inverse propensity score reweighting estimator to recover the average impact of childcare attendance in Canada under the assumption that factors leading to childcare attendance are observed by the researcher. By comparing the estimates across empirical strategies we are attempting to uncover whether the subgroup whose childcare attendance decision is affected by the policy has differential impacts from the rest of the population.

The use of different estimators allows recovery of alternative causal parameters, permitting us to begin an exploration into possible treatment effect heterogeneity. In applications of childcare at the universal level, as in the Quebec Family Policy, significant treatment effect heterogeneity is plausible. For example, children who receive high quality one-on-one parental care at home may experience negative outcomes when they shift to public childcare that includes higher adult-child ratios, while other children may benefit from the transition to public childcare. Further, since the universal childcare policy led to an increase in maternal employment, newly employed parents may have higher levels of stress than they did prior to the implementation of the childcare policy.

Third, we account for a statistical issue that affects the methods by which inference is conducted when researchers examine the effectiveness of a single policy on a multitude of child and family outcomes. Specifically we utilize techniques that incorporate the dependence in child and parent outcomes across multiple domains for the same individual. A failure

[9]Note that variants of this analysis also appear in Kottelenberg (2009) and Lefebvre et al. (2011).

to account for multiple outcomes from the same treatment(s) may falsely yield significant results. For example, if the effectiveness of access to universal childcare is assessed on six outcomes, each at a significance level of five percent (two-sided tests), the chance of finding at least one false positive statistically significant test increases to 15.9 percent. Accounting for multiple outcomes can have a substantial influence on the rate of false positive conclusions. In the early education policy setting, where many outcomes are present, researchers may be tempted to select the most favourable results from an analyses, reporting significance where there is none. To this end, we adopted a multiple testing procedure that controls for the probability of at least one rejection of a true null hypothesis and allows the number of false rejections one is willing to tolerate to vary with the total number of rejections. In this way we present a more detailed analysis of the effectiveness of the policy.[10] To the best of our knowledge, this issue has been largely ignored in the evaluation of results from natural experiments in Canada.

Overall, our results are in line with BGM's findings, namely that the introduction of the Quebec Family Policy led to a significant decline in child, parent and family outcomes. The initial estimates in BGM are robust to inclusion of additional years of data and concerns regarding multiple testing. Estimates of the causal impact of childcare attendance are also negative with the notable exception of the motor-social development score. On average, attending subsidized childcare leads to a significant increase in this score. Further, our results suggest that the negative impacts reported in BGM are driven by children in families who decided to attend childcare in response to the implementation of the policy. This heterogeneity in program impacts suggests an important avenue for further research.[11]

This paper is structured as follows. Section 2 provides a description of the data used for the analyses. Our empirical strategy to meet the three aims described above is presented

[10] In contrast, earlier research has examined each of these outcomes independently.

[11] In Kottelenberg and Lehrer (2012a, 2012b) we have explored heterogeneity across gender and the unconditional outcome distribution respectively. In both papers, we present significant evidence of policy relevant treatment effect heterogeneity. See also Kottelenberg (2009) and Lefebvre et al. (2011) for an exploration into heterogeneity in program impacts based on age of child's attendance to childcare.

in section 3. Empirical results are presented and discussed in section 4. Finally, in the concluding section we summarize our findings from this paper and suggest that, given the substantial treatment effect heterogeneity from attending childcare, to boost child outcomes policy-makers should consider targeting childcare rather than developing policies that would introduce universal coverage.

2 Data

To undertake our proposed analyses, we use the National Longitudinal Study of Children and Youth (NLSCY), a nationally representative longitudinal study that tracks cohorts of Canadian children from early childhood. This study's first sample was drawn in 1994-95 and consisted of Canadian children aged zero to eleven. Biannually, a new cohort of 2,000 children aged zero to one is added to the study and data is collected from all participating cohorts. The NLSCY was also used in BGM, to which we added data from cycles six (2004-05) and seven (2006-07).

It is worth noting several additional details regarding the NLSCY sampling. This sample was restricted to Canada's ten provinces and excluded both full time members of the Canadian Armed Forces and people living on Aboriginal reserves.[12] The NLSCY contains both child developmental scores and extensive questions relating to child care usage, parental labour supply, and other demographic characteristics, providing an opportunity to study the effects of child care policy on a variety of childhood development and behaviour indicators. Child development scores include the revised Peabody Picture Vocabulary Test (PPVT) score for children aged four, a standardized motor and social development (MSD) score for children aged zero to three, and a series of child behavioural scores relating to hyperactivity, anxiety, physical aggression and opposition. The MSD scale consists of a set of 15 questions that measure dimensions of the motor, social and cognitive development of children from birth through three years of age, with questions that differ based on the child's age. Each

[12]These exclusions represent about two percent of the Canadian population.

item asks whether or not a child is able to perform a specific task.[13] The PPVT is one of the most popular standardized tests used to assess children's verbal intelligence and estimate children's scholastic aptitude. The PPVT is a short test for which an examiner says a word aloud and an examinee attempts to identify, from four options, the picture that best represents the word. Finally, in previous research (e.g. Statistics Canada, 2003; Charach et al. 2010) the child behavioural scales have been shown to reliably predict related outcomes.

Since one of our goals is to determine whether the impacts reported in BGM are transitory or permanent, we follow their sample restrictions and covariate definitions to ensure that any differences in study results are not due to choices made by the researcher. Specifically, the analyses are conducted using children from two-parent families, thereby eliminating the contaminating effects of pre-policy subsidization that generally have higher utilization rates with single-headed households. Further, this isolates an appropriate comparison group not affected by changes in other policies, such as paternity leave regulations that were altered during the data collection. While this restriction may limit the external validity of any findings, two-parent families remain a key focus of the universal childcare debate that aims to extend subsidized access to childcare to locations and individuals for which it was not previously made available. Last, observations are only included if, at the time of the interview, the child's age was four years.

Table 1 presents summary statistics on a subset of child, parent, and family variables that are used as control variables in our analyses. For each variable in this table we report the mean and standard deviation values for four samples, defined by province of residence (Quebec or the rest of Canada) and time of survey completion (pre or post the introduction of the Quebec Family Policy. As shown in Table 1, in terms of demographic characteristics, there are no striking differences between sub sample groups or group specific trends. While

[13]Two examples of MSD questions asked of parents are whether a child has ever sat up for ten minutes without assistance, or whether the child has said more than two recognizable words. This scale was developed by Dr. Gail Poe of the U.S. National Centre for Health Statistics and has also been used in collections of the National Longitudinal Survey of Youth in the United States and in recent versions of the National Child Development Survey in England.

a handful of these differences are statistically significant, we will account for these variables in the specification to minimize any bias from unobserved confounders. Across all Canadian provinces, there is a substantial increase in university completion among parents surveyed in later cohorts. Similar to BGM, Lefebvre and Merrigan (2008), and Lefebvre, Merrigan, and Roy-Desrosiers (2011), we do not believe that these statistics raise any immediate serious concerns for using the remaining Canadian provinces as a comparison group in the reduced form analyses.

Summary statistics on the outcome variables that will be examined in this study are presented in Table 2. Since the questions that underlie a number of these outcome variables are age specific, we include information on the sample size and age group of interest for each outcome in the last two columns of this table.[14] The measures presented include: behavioural indices for Physical Aggression, Hyperactivity and Inattention, Emotional Anxiety and Separation Anxiety for children and Family Dysfunction, and Mother's Depression Score for parents. To examine the child's health status we use an indicator variable based on the parent's subjective evaluation of whether their child is "excellent health", and reports indicating if the child never experienced i) a nose/throat infection, or ii) an ear infection. The last two rows of the Table 2 indicate that there were large increases in both maternal labour supply on the extensive margin and take-up of childcare following the introduction of the policy in Quebec relative to the rest Canada. With regard to the remainder of this table, there are not many substantial differences in the unconditional rates of many of the behaviours across regions in Canada prior to the policy but in the post policy periods trends in several variables, such as never experiencing a Nose or throat infection or physical aggression, differ substantially. In the next section, we will formally examine how these trends may have differed after policy implementation and if these changes were due to childcare attendance as many have hypothesized.

[14] Although available for children aged four, many of these indices are not composed of the same base questions. While not reported in the paper analyses with the indexes for the older children do indeed reveal a similar pattern as those for children aged two to three and are available upon request.

Table 1: Summary Statistics

	Quebec		Rest of Canada		Obs.
Covariates	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy	
Resides in Rural Region	0.095	0.142	0.133	0.190	38648
	(0.293)	(0.349)	(0.34)	(0.393)	
Resides in a Large City (>500K)	0.579	0.581	0.428	0.445	38648
	(0.494)	(0.493)	(0.495)	(0.497)	
Number of Older Siblings	0.715	0.685	0.796	0.755	38648
	(0.739)	(0.716)	(0.761)	(0.744)	
Number of Same Aged or	0.268	0.217	0.255	0.246	38648
Younger Siblings	(0.488)	(0.438)	(0.476)	(0.467)	
Mother's Age	30.926	31.369	31.738	32.652	38648
	(4.877)	(4.996)	(5.123)	(5.395)	
Mother's Immigrant Status	0.089	0.137	0.214	0.252	38648
	(0.285)	(0.344)	(0.41)	(0.434)	
Mother did not complete high school	0.133	0.110	0.106	0.075	38648
	(0.34)	(0.313)	(0.308)	(0.263)	
Mother is an University Graduate	0.203	0.331	0.206	0.343	38648
	(0.402)	(0.471)	(0.404)	(0.475)	
Father's Age	33.507	34.204	34.142	35.275	38648
	(5.401)	(5.817)	(5.717)	(6.17)	
Father's Immigrant Status	0.097	0.162	0.208	0.252	38648
	(0.296)	(0.368)	(0.406)	(0.434)	
Father did not complete high school	0.168	0.139	0.138	0.098	38648
	(0.374)	(0.346)	(0.345)	(0.297)	
Father is a University Graduate	0.194	0.291	0.214	0.306	38648
	(0.395)	(0.454)	(0.41)	(0.461)	
Child is Male	0.509	0.514	0.509	0.514	38648
	(0.5)	(0.5)	(0.5)	(0.5)	
Child's Age	2.026	1.965	1.991	2.021	38648
	(1.42)	(1.416)	(1.418)	(1.419)	

— Note: Each row corresponds to an independent variable and contains the mean and standard deviation (in parentheses) specific to the time and geographic region as denoted in the column header. The data is split by Quebec and the rest of Canada as well as by the pre-policy period, from 1994-97, and the post policy period, from 2002-07. The final column provides the sample size for these measurements. The NLSCY sample weights, designed to accurately reflect the make up of the Canadian population, are applied in these and all calculations throughout the paper.

Table 2: Dependent Variable Summary Statistics

Outcome Variables	Quebec		Rest of Canada		Obs.	Ages
	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy		
	Child Outcomes					
MSD Score	99.317	96.242	100.395	99.067	30472	0-3
	(15.031)	(15.28)	(15.343)	(14.81)		
PPVT Standardized Score	99.764	99.536	100.511	101.332	6585	4
	(15.139)	(15.059)	(15.277)	(15.095)		
Hyperactivity and Inattention Score	3.418	3.766	3.622	3.603	14673	2-3
	(2.573)	(2.388)	(2.43)	(2.288)		
Emotional Anxiety Score	0.967	1.395	1.080	1.311	14781	2-3
	(1.343)	(1.468)	(1.398)	(1.532)		
Physical Aggression Score	4.375	4.774	5.095	4.844	14625	2-3
	(3.041)	(2.974)	(2.961)	(2.866)		
Separation Anxiety Score	2.668	2.667	2.710	2.538	14812	2-3
	(2.029)	(1.862)	(1.991)	(1.976)		
Child in Excellent Health	0.642	0.637	0.637	0.684	38539	0-4
	(0.48)	(0.481)	(0.481)	(0.465)		
Never had a Nose /Throat Infection	0.404	0.261	0.489	0.489	31252	0-4
	(0.491)	(0.439)	(0.5)	(0.5)		
Never had an Ear Infection	0.438	0.476	0.489	0.589	31230	0-4
	(0.496)	(0.5)	(0.5)	(0.492)		
	Parent and Family Outcomes					
Mother in Excellent Health	0.406	0.399	0.388	0.398	38365	0-4
	(0.491)	(0.49)	(0.487)	(0.489)		
Father in Excellent Health	0.448	0.437	0.414	0.404	38299	0-4
	(0.497)	(0.496)	(0.493)	(0.491)		
Family Dysfunction Index	7.189	7.728	7.796	8.168	37842	0-4
	(4.979)	(5.081)	(5.156)	(5.043)		
Mother's Depression Score	4.200	4.033	4.528	3.702	34005	0-4
	(4.563)	(4.784)	(4.922)	(4.44)		
	Uptake Variables					
In Care	0.415	0.655	0.405	0.449	38166	0-4
	(0.493)	(0.475)	(0.491)	(0.497)		
Mother Working	0.530	0.669	0.591	0.621	38471	0-4
	(0.499)	(0.471)	(0.492)	(0.485)		

— Note: Each row corresponds to a variable of interest and contains the mean and standard deviation (in parentheses) specific to the time and geographic region as denoted in the column header. The final two columns provide the variable specific sample sizes and the range of ages for which the outcome is applicable. As in Table 1 the data is split between Quebec and the rest of Canada and by the pre-policy period and the post policy periods.

3 Empirical Strategy

We begin by adopting BGM's main empirical specification which estimates the effect of access to universal childcare in Quebec by comparing the difference in the evolution of a series of outcomes in Quebec relative to the rest of Canada using data from both pre- and post-policy period. The regression specification for an outcome of interest Y can be expressed as:

$$Y_{ipt} = \beta_o + \delta' Policy_{ipt} + \beta_1' PROV_p + \beta_2' YEAR_t + \beta_3' X_{ipt} + \varepsilon_{ipt} \quad (1)$$

where i, p, and t index individual, province, and year. The vector of covariates X, includes controls for child, parent, family, and geographic characteristics,[15] PROV and YEAR are respectively a series of province and time dummies. The Policy variable is an interaction between the indicator for living in Quebec after 1998, the year the Quebec Family Policy was introduced. Thus, the coefficient of interest δ provides an estimate of the average effect of being eligible to attend universally subsidized childcare in Quebec on the outcome of interest. Intuitively, δ captures any additional change in outcomes in Quebec pre- and post-reform relative to the changes that occurred in the rest of Canada over the same time period. In the statistical treatment effect literature δ is often referred to as the intent to treat (ITT) parameter.

Our work not only considers a larger number of post periods relative to BGM, but as noted in the preceding section, we dropped both the third and fourth cycle of the NLSCY data.[16] We dropped data from the fourth cycle of the NLSCY since this corresponds to the tail end of a period in which universal childcare was being introduced to children at younger ages. Figure 1 illustrates that cycle 4 corresponds to a time of substantial growth in the number of publicly funded childcare seats. Nearly 65,000 regulated childcare seat became

[15] We use the exact same set of controls as BGM a subset of which is presented in Table 1. To reduce issues related to mis-specifying the functional form of the estimating equation, each of the continuous variables reported in Table 1 including age is discretized. For example, we create a host of discrete dummy variables for the various categories measured by parental education, number of siblings and community size.

[16] Note that BGM only had cycles 1-5 available at the time of their study and thus used cycles 1-2 as the pre-policy period and cycles 4-5 as a post-policy time frame.

available between 1998-2001 (Lefebvre and Merrigan, 2008). We examined whether our full set of results were robust to the inclusion and exclusion of different cycles of the NLSCY data and we found that our results were not sensitive to the sampling scheme.[17]

Equation (1) is often referred to as a difference-in-difference estimator for which recovery of causal effects relies on the maintenance of several assumptions in the underlying data. The first assumption is the absence of anticipatory behaviour in Quebec. It is unlikely that parents in Quebec altered their childcare decisions prior to the implementation of the Quebec Family Policy, meaning this first assumption is likely met. Second, an assumption of common trend is required: in the absence of treatment, outcomes in Quebec and the rest of Canada must move along a similar path. As the various panels in Figure 2 demonstrate, the slopes of the lines in Quebec and the rest of Canada were similar (in cycles one and two) prior to the introduction of the policy, suggesting the assumption of common trend is met. The final assumption, the assumption of common support, appears plausible since ex-ante we would not expect the observed and unobserved characteristics of individuals living in Quebec to differ substantially from those living in other provinces.[18]

To move beyond estimating the impact of access to subsidized child care and to explore the effects of attending childcare, we reformulated the estimating equation. Equation 1 can be viewed as the reduced form of the following system of equations

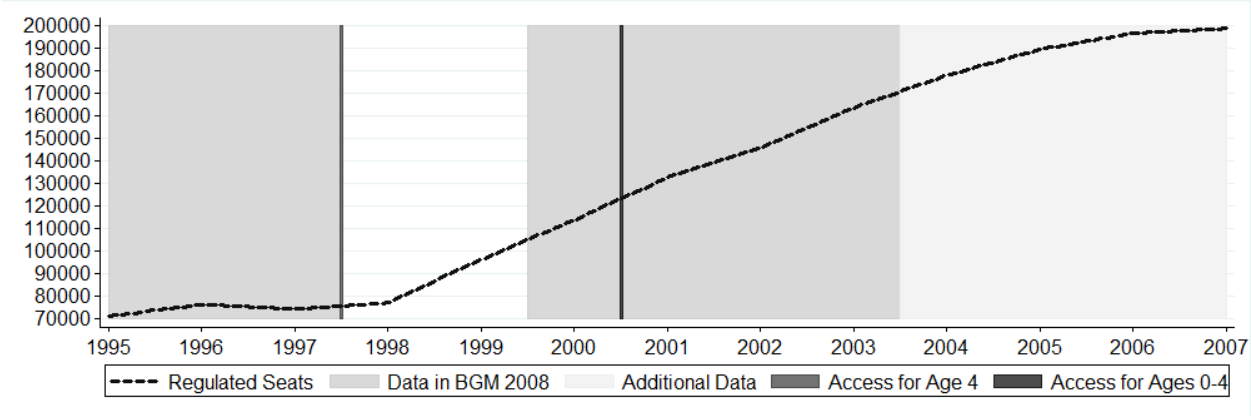
$$\begin{aligned} Y_{ipt} &= \beta_0 + \beta_1' Ccare_{ipt} + \beta_2' PROV_p + \beta_3' YEAR_t + \beta_4' X_{ipt} + \varepsilon_{ipt} \\ Ccare_{ipt} &= \gamma_0 + \gamma_1' Policy_{pt} + \gamma_2' PROV_p + \gamma_3' YEAR_t + \gamma_4' X_{ipt} + u_{ipt} \end{aligned} \tag{2}$$

where Ccare is an indicator for being in childcare. The main empirical challenge in using the

[17] The Universal childcare Benefit was introduced in 2006 providing parents $1200 annually for each child under the age of five. Removing the seventh cycle of the NLSCY, years 2006-07, does not effect our results.

[18] Despite evidence suggesting similarities between Quebec and the rest of Canada presented in Table 1 the common support assumption may also be weakened by differences in Francophone and Anglophone populations. To ensure the validity of our results we perform our estimation procedure on Francophone and Anglophone sub-samples and find that these groups are similar in their response to the policy compared to the full sample. These results are available upon request.

Figure 1: Number of Reduced Fee Childcare seats available in Quebec over Time

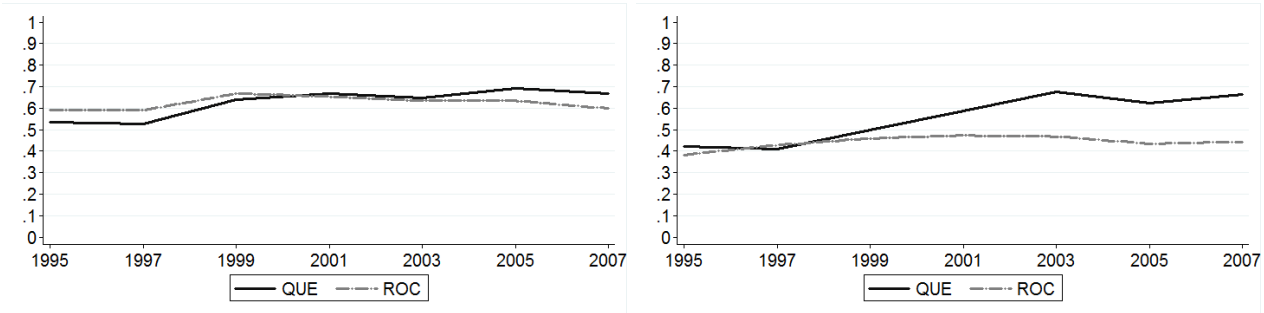


— Note: The data for this figure was taken from Lefebvre and Merrigan (2008). Some subsidies were provided both to suppliers and families reducing the cost of childcare places prior to the implementation of the Quebec Family Policy in 1997. Access was first granted to children aged 4 in September of 1997. In the three subsequent Septembers children aged three, two, and zero to one became eligible for the program. Thus, the program was fully accessible by September 2000.

Figure 2: Trends in Maternal Care and Childcare Use in Quebec and the Rest of Canada

(a) Proportion of Mothers Working

(b) Proportion of Children in Care



(c) Proportion of Children in Center Based Care

(d) Proportion of Children cared for in Another's Home

— Note: These figures present the proportions of two-parent families in each category. Figures 2c and 2d are proportions of the whole population and do not condition on whether a child is in care. Survey responses for places of care also include in one's own home, nurseries, and after school programs, the last of these two categories are less than one percent of the population.

ordinary least squares estimator to directly estimate the first equation in (2) is dealing with the endogeneity of Ccare. Placing a child in care reflects a behavioural decision made by parents and may be correlated with unobservable characteristics. To overcome the endogeneity and recover unbiased parameter estimates of $\beta_1^{'}$, we consider two alternative strategies that rely on different assumptions regarding the parents' choice to use subsidized day-care.

First, we treat the Quebec Family Policy as a natural experiment with non-compliance and use the instrumental variables estimator allowing us to recover consistent estimates of the local average treatment effect (LATE). Identification of causal effects using an instrumental variables estimator is proven in Imbens and Angrist (1994) and relies on specific conditions on how access to the policy ($Policy_{pt}$) affects childcare use ($Ccare_{ipt}$) and how access to the policy ($Policy_{pt}$) is unrelated to unobserved components of the outcome equation (ε_{ipt}). In the current study, all of these conditions appear plausible and with treatment effect heterogeneity the instrumental variables estimates of $\beta_1^{'}$ can be interpreted as the average effect of childcare for those children who only attended child-care due to the Quebec family policy.[19]

The second empirical strategy relies on an assumption that the analyst has access to all of the variables that determine the childcare attendance decision. While the selection on observables assumption that underlies this second strategy is quite strong,[20] this method arguably gets at the policy parameter of interest: the average effect of attending childcare, often referred to as the average treatment effect (ATE) in the causal inference literature. This parameter has the limitation of not making a distinction between universal childcare in Quebec and the childcare offered in the rest of Canada, and in effect treats all childcare

[19] An additional estimator that we considered was using control function methods to preserve the non-linearity of many of the outcomes variables. This involves determining the projection of the endogenous explanatory variable onto the exogenous variables and to control for endogeneity and we then add the predicted error term from this equation into the structural outcome equation. While this is asymptotically inefficient relative to MLE, it is computationally convenient. Since the control function is estimated and not the true first-stage errors, bootstrap procedures were used to account for this additional uncertainty. For space considerations, we do not report the results from the control function strategies but note that they did not differ substantially in sign, statistical significance or magnitude from the linear instrumental variables estimates.

[20] In the next section, we examine the sensitivity of our estimated treatment effects to the degree in which there is selection on unobservables.

as a homogeneous good across the sample in both time and space. Specifically we follow the procedure introduced in Hirano et al. (2003), that first estimates the second equation in (2) using the non-parametric series logit estimator.[21] Using the estimated coefficients, the predicted probabilities of receiving child-care ($\hat{p}(X_i)$) are calculated for each individual. Estimates of the average effect of attending childcare are obtained by weighted least squares estimation of the outcome equation in (2) where the estimated weight (W_i) for each individual is calculated as $W_i = \sqrt{\frac{Ccare_{ipt}}{\hat{p}(X_i)} + \frac{1-Ccare_{ipt}}{1-\hat{p}(X_i)}}$. The weights play an important role in the regression specification as they ensure that the covariates are balanced between those that attend and do not attend childcare.[22]

Prior research attempting to estimate causal impacts of Quebec family Policy has treated outcomes within and across domains for the same child as independent from one another; however, the assumption of independence across dependent variables (e.g. test scores and health metrics) may not be met, since the outcomes in multiple domains (e.g. Hyperactivity and inattention index and physical aggression index) are likely highly correlated. Making adjustments for the use of multiple outcomes has a long history in psychology (Benjamini and Yekutieli, 2001) and biostatistics (Hochberg, 1988). These techniques have also been adopted in some economic studies that examine multiple child outcomes (Kling et al., 2005; Anderson 2008; and Ding and Lehrer, 2011). Accounting for the possibility that multiple outcomes of interest correlate with one another avoids the possibility of over rejecting the null hypothesis when using univariate statistical methods. To this end we need to adjust the p-values for the multiple outcomes to reduce the likelihood of making type I errors. For each estimator and causal parameter estimated, we also report q-values that make corrections for the false discovery rate using the two-step procedure proposed by Simes (1986). Intuitively, q-

[21] The series logit estimation incorporates all of the covariates used in BGM as well as their interactions. Note that the results presented in the next section are robust to using both parametric probit and logit estimators that do not include the set of covariate interactions. As result of the inverse weighting on the propensity score, predictions that near zero and one may substantially effect results. Although not the case here differences in handling propensities at the extremes between the logit and probit models can lead to sensitivity in the estimates.

[22] This estimator has the desirable property of being doubly robust and will achieve consistent estimates as long as either the regression model or the propensity score (and thus the weights) are specified correctly.

values can be thought of as adjustments to the p-values that maintain the overall probability of making a Type I error at a fixed α (i.e. 5 percent) across the full set of outcomes.[23]

4 Results

4.1 Are the negative effects of access to universal childcare transitory or permanent?

Proponents of universal childcare often claim that the negative effects reported in BGM are transitory, and occurred as a result of the transition process that child care centres underwent in response to newly implemented policy. The rationale used in these arguments is that cycles four and five were periods during which the program was rapidly expanding to meet additional demand.[24] It may be the case that the quality of care has since improved, potentially mitigating the initial negative effects.[25] In addition, Figure 2 documents that the source of childcare provision in Quebec has shifted towards center-based care, which prior research (e.g. NICHD-ECCRN, 2004) associates with better academic and language skills relative to other types of care arrangements. The initial surge in demand for childcare in Quebec led to substantial growth in the provision of childcare facilities. Growth in home-based care operations outpaced growth in what is often termed as institutional based

[23] We also (results are available from the authors upon request) made corrections for the familywise error rate, which is more conservative than the FDR but has a greater likelihood of obtaining Type II errors.

[24] The implementation of childcare for all ages in Quebec was completed in 2000, however the number of available places in subsidized day-care did not come close to satisfying demand. Lefebvre et al. (2008) report that from 2000 to 2007 the province increased its available spaces in the program from approximately 110,000 to 200,000, as seen in Figure 1.

[25] The Quebec government legislated in 2000 that two thirds of the staff at centre de petite enfants must be trained in early childhood education (previous requirements were at one third), while at the same time wages for caregivers were scheduled to rise 35 percent to 40 percent over a four year period. The concern for quality increases culminated in legislation in August 2006, prior to the last available cycle of the NLSCY, which required two thirds of staff to have college diplomas or university degrees in early childhood education.

facilities, due to lower start up and organizational costs.[26]

Despite these shifts in provider location we do not find any evidence to support optimism for the improved effects of childcare access in the years following program introduction. The findings from which we drew these conclusions are presented in Table 3. In the first column of this table we reproduce estimates of the intent-to-treat estimates from equation (1) that were reported in BGM.[27] The second column, titled later time periods, includes results from analyses of more recent cycles of NLSCY data. The sign and statistical significance of the estimates of the effects of gaining access to universal childcare does not differ between columns for any of the child or parental outcomes. Further, the magnitude of these effects is larger for all outcomes with the exception of "Child in Excellent Health" for data from later versus earlier time periods. On average, the magnitude of the coefficient increased by approximately 43 percent relative to the estimates presented in BGM. Therefore, in contrast to described critiques, which centered around proposed mitigating effects of BGM's study, the findings from our study suggest that negative effects are not transitory and are maintained

[26]The bottom right panel of Figure 2 shows that in the period immediately after the policy was introduced (1998-2001) that there was a small increase in usage of home-based care. However, since 2001 there has been a significant decline in home-based care and corresponding growth in center based care. This suggests that the excess demand for childcare created by the policy was initially met through the use of home-based care and has since shifted to center-based care.

[27]The estimates we present differ from those presented in BGM in three ways. By first replicating the standard errors in BGM exactly, we present p-values to facilitate comparison with their q-value counterparts. Second, the estimate we present corresponding to the hyperactivity and inattention index is not consistent to BGM due to a change in how the data was collected. Specifically this index was calculated in BGM as a sum of responses to questions related to frequency of various behaviours. These questions were adjusted in cycle 4 of the NLSCY by the removal of two questions making up part of the index and one new question was added. In our work, this difference was overcome by the merging of the existing indices to produce one in which all questions are common. Finally, we do not present results for all results examined in BGM. We leave out several parental measures: consistency, hostile/ineffective parenting, aversive parenting, and satisfaction. In addition due to space constraints, we do not present results on diagnosis of asthma or having received an injury. The findings in BGM are robust to the addition of cycle six and seven data and are available upon request.

for at least ten years after program implementation.[28] Finally, while not the focus of the present paper, it is worth reporting that even after conditioning on the full set of controls, the estimates of the effect of access to childcare on maternal labour supply and childcare usage variables show higher levels of uptake in both categories when the additional cycles of data are utilized in the analyses.

In the last column of Table 3 we report q-values that correct the statistical inference procedure for multiple outcomes. Of the nine statistically significant intent to treat parameters tested in isolation, we find that all are indeed robust to multiple testing correction. This indicates that the negative and statistically significant results were not Type I errors. Taken together, the results presented in Table 3 increase our confidence that the introduction of universal childcare led to statistically significant reductions in a variety of child, parental and family measures. BGM results are indeed robust to the use of multiple testing and data from later time periods.

4.2 Are the negative effects from attending childcare?

Whether or not the negative effects found in both the current study and BGM's study can be attributed to childcare attendance remains a salient question. The BGM analyses do not directly consider childcare attendance, thereby leaving this question unanswered. To address this question we consider two identification strategies used to estimate equation (2).

Table 4 presents IV estimates of childcare that can be interpreted as a local average treatment effect, providing an estimate of the effect of child care on children in families which altered their usage as a result of the implementation of the policy. While the sign

[28] This, however, does not clarify whether negative impacts are reduced, at least to some degree, for children "in care" as the estimated report reflects the intent-to-treat effect and not actual attending childcare in some form. It is advisable to adjust intent-to-treat effects by dividing the policy coefficient by the probability of being treated. After accounting for the probability of receiving the given treatment, as calculated by the uptake in maternal labour supply and/or childcare, we find the majority of the treatment effects worsen in terms of magnitude. Of the child cognitive and behavioural outcomes only the MSD score is reduced remaining a negative effect of half of standard deviation at best. (Results available upon request). BGM hold discussion on the appropriate probability of being treated suggesting either the increase in childcare use or the increase in maternal labour supply induced by the policy. The estimates of these effects using the new NLSCY data are 0.19 and 0.11 respectively. The results above reflect the use of either of these probabilities.

Table 3: Estimates of the Causal Effect of Access to Universal Childcare

	BGM	Later Time Periods	
	(P-Value)	(P-Value)	[Q-Value]
NLSCY Data Used	1994-1997	1994-1997	
	1999-2003	2002-2007	
	Uptake Variables		
Child in Care	0.146	0.196	
	(0.000)***	(0.000)***	[0.000]***
Mother Works	0.077	0.110	
	(0.000)***	(0.000)***	[0.000]***
	Child Outcomes		
MSD Score	-1.645	-1.688	
	(0.000)***	(0.000)***	[0.000]***
PPVT Standardized Score	0.361	-0.435	
	(0.631)	(0.570)	[0.570]
Hyperactivity and Inattention Score	0.211	0.330	
	(0.118)	(0.000)***	[0.000]***
Emotional Anxiety Score	0.120	0.205	
	(0.032)**	(0.000)***	[0.001]***
Physical Aggression Score	0.380	0.601	
	(0.000)***	(0.000)***	[0.000]***
Separation Anxiety Score	0.098	0.164	
	(0.249)	(0.059)*	[0.066]*
Child in Excellent Health	-0.055	-0.049	
	(0.000)***	(0.009)***	[0.012]**
Never had a Nose/Throat Infection	-0.140	-0.146	
	(0.000)***	(0.000)***	[0.000]***
Never had an Ear Infection	-0.057	-0.068	
	(0.004)***	(0.000)***	[0.000]***
	Parent and Family Outcomes		
Mother in Excellent Health	-0.011	-0.015	
	(0.317)	(0.173)	[0.288]
Father in Excellent Health	-0.029	-0.002	
	(0.015)**	(0.912)	[0.912]
Family Dysfunction Index	0.254	0.133	
	(0.142)	(0.426)	[0.533]
Mother's Depression Score	0.420	0.659	
	(0.000)***	(0.000)***	[0.000]***

— Note: For the outcome variable in each row columns 1 (BGM sample) and 2 (later periods sample) present the estimates of the policy coefficient δ as specified in Equation (1). Each specification includes the covariates listed in Table 1 as well as province and cycle indicators. Standard errors are corrected at the province-year level then used to calculate p-values (presented in parentheses) testing the statistical significance of the corresponding estimate. The final column presents q-values which also tests the estimate in the row above for statistical significance. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

Table 4: Estimates of the Causal Effect of Attending Childcare

	IV Estimate (P-Value)	[Q-Value]	IPW Estimate (P-Value)	[Q-Value]
	Child Outcomes			
MSD Score	-7.898		1.395	
	(0.000)***	[0.000]***	(0.000)***	[0.000]***
PPVT Standardized Score	-3.133		-0.253	
	(0.531)	[0.531]	(0.536)	[0.603]
Hyperactivity and Inattention Score	1.403		0.062	
	(0.001)***	[0.001]***	(0.187)	[0.281]
Emotional Anxiety Score	0.821		-0.011	
	(0.002)***	[0.003]***	(0.704)	[0.704]
Physical Aggression Score	2.475		0.083	
	(0.000)***	[0.000]***	(0.138)	[0.248]
Separation Anxiety Score	0.662		-0.033	
	(0.066)*	[0.075]*	(0.397)	[0.510]
Child in Excellent Health	-0.233		-0.043	
	(0.037)**	[0.047]**	(0.000)***	[0.000]***
Never had a Nose/Throat Infection	-0.731		-0.081	
	(0.000)***	[0.000]***	(0.000)***	[0.000]***
Never had an Ear Infection	-0.322		-0.113	
	(0.000)***	[0.000]***	(0.000)***	[0.000]***
	Parent and Family Outcomes			
Mother in Excellent Health	-0.064		0.004	
	(0.262)	[0.437]	(0.505)	[0.505]
Father in Excellent Health	-0.005		-0.013	
	(0.957)	[0.957]	(0.030)**	[0.037]**
Family Dysfunction Index	0.700		0.193	
	(0.440)	[0.550]	(0.002)***	[0.010]***
Mother's Depression Score	3.278		-0.143	
	(0.000)***	[0.000]***	(0.012)**	[0.030]**

— Note: For the outcome variable in each row we present estimates of treatment effect of childcare, the coefficient on Ccare in Equation (2) from the instrumental variable (column 1) and inverse propensity weighting methods (column 3). Each specification includes the covariates listed in Table 1 as well as province and cycle indicators. Standard errors are corrected at the province-year level then used to calculate p-values (presented in parentheses) testing the statistical significance of the corresponding estimate. Columns 2 and 4 present q-values [also in parenthesis] based on the p-values presented in columns 1 and 3 which also test for the statistical significance of the estimate in the row above.***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

and statistical significance of these estimates mirror the intention-to-treatment parameters, the magnitude is substantially larger. These results confirm the pattern suggested in BGM. Most striking is the causal relationship between childcare use and maternal depression scores. These results suggest that mothers, whose decision to utilize childcare is tied to access to the policy, experience a type of mental stress that the average mother does not. There are many potential avenues which may cause this difference including the possibility that these are families in which i) there is a higher attachment to the role as mother, ii) mothers have less ability to handle stress in the workplace, or iii) mothers are less resilient in terms of coping with work/life balance and demands. While future research is required to understand the mechanisms at play in these results, this set of the results suggests that the hypothesized pathway of negative effects of attending childcare is driven by those most affected by the policy itself.

In column 3 of Table 4 estimates that assume selection to childcare on observed factors and use the inverse propensity score reweighting procedure described in the preceding section to recover the average causal effect of attending childcare are presented. This empirical strategy reweights the populations of users and non-users of childcare as if to restore randomization of other underlying factors, such as maternal labour, between these groups. If randomization is successfully restored then isolation of the average casual impact is easily identified by comparing outcomes between the two respective groups.

The causal estimates using the inverse propensity weighting method differ substantially from the intent-to-treat effects presented in Table 3. Estimates of the effects of attending childcare on both child cognitive and behavioural measures are statistically insignificant, with the exception of the positive effect upon MSD scores. The only child outcomes for which the IPW estimates are similar to the ITT effects are for health, indicating that going to childcare leads to a higher likelihood of not being considered in excellent health and having an ear, nose or throat infection. While there were a number of changes in estimated effects upon child outcomes, there were no major differences in terms of home environment outcomes.

Interestingly, on average, having a child in childcare led to a small but statistically significant decline in maternal depression, the sole parental and family level outcome for which the sign of the estimated effect significantly changed from Table 3 to the IPW results in Table 4.[29] Last, with only a few exception the sign and significance of the causal effects of childcare on child outcomes are similar across geographic region and time. This highlights a consistency in the average treatment effect of childcare across different environments and implies a general similarity and comparability of childcare in Quebec following the policy.

An examination of the estimates from the different empirical strategies used in the current study provides robust evidence of heterogeneous treatment effects. Children and parents from different families responded differently to both the Quebec family policy and the utilization of childcare. However, Kottelenberg (2009) did not find evidence of significant heterogeneity in the impacts of the policy across parental characteristics and household location. As such, it is challenging to identify which groups of children will benefit most from childcare policies. Yet, as evident from the contrast in the ATE and LATE estimators shown in this paper, children of parents who were on the margin of using childcare experience significantly worse outcomes relative to others who choose childcare. There are many potential scenarios that can explain who these families are.[30] For example, when making childcare decisions, parents face trade-offs between leisure versus the welfare of their child.[31] In equilibrium, the marginal rate of substitution between leisure and parental time investment into their children is equal to the price of leisure divided by the price of childcare. Some parents will change the manner in which their children receive care when the costs of childcare are lowered.

[29] In results available from the authors upon request, we repeated this exercise on samples based on geographic region and the waves of NLSCY data collection. There are no differences in the average effect of childcare on child behavioural outcomes in magnitude and statistical significance across geographic region and time. Further, since the sign of the effects of childcare on most child outcomes does not vary across geographic regions and time periods, this highlights a consistency in the average treatment effect of childcare across different environments and implies a general similarity and comparability of childcare in Quebec following the policy.

[30] There does not exist reliable statistical methods to determine who in the estimation sample changes their childcare decision due to the policy. This is a standard issue with instrumental variables estimation in that it recovers a treatment effect for a sub-population of "compliers" that cannot be identified.

[31] Assume that both parental investments and childcare are inputs into a human capital production function for the child and that these inputs have different returns.

There are multiple mechanisms by which the policy may influence childcare decisions and our parameter estimates the average effect from all of these mechanisms.

To shed light on the underlying mechanisms would require strong assumptions about the data generation process. However, these results do expand our knowledge about the consequences of the introduction of universal childcare in Quebec. Our analysis showing varied effects of publicly funded childcare complicates decisions on childcare policy. On the other hand, children from families who are most influenced by the Quebec Family Policy appear to have poorer performance on a number of outcome measures post policy. These results have implications for governmental policy; they suggest that governments should modify the design of policies that expand childcare so that even those on the margins of childcare use may experience gains. Further, additional research should be conducted to expand on the results obtained in the current study. An important limitation of this study is that the nature of NLSCY data is such that we had to treat childcare as a homogeneous good. Substantial research (e.g. Blau 1999, 2001) demonstrates that there are important underlying factors affecting individual outcomes within care that may also shed further light on why childcare works for some children and their families.

In our IV estimation, we used the introduction of the policy in Quebec to instrument for childcare attendance. To assess the suitability of our instrument we consider a simple OLS regression of the first stage regression and run an F test for the joint significance of the instrument. The results are presented in Table 5.[32] Coefficients on the instrument and exogenous regressors in both columns are reasonable in sign and magnitude. Notice that only the education of the mother, not that of the father, is significantly associated with childcare attendance in Canada. Not surprising given the high costs, families with multiple siblings are much less likely to use day-care. Finally, it is important to note that the instrument is a statistically significant predictor of childcare attendance and the F-statistic on its significance is well-above current cut-offs (i.e. Staiger and Stock, 1997) for weak instruments.

[32]This equation is identical to the propensity score equation but omits the higher-order terms and interactions as well as uses the ordinary least squares estimator in place of the instrumental variables estimator.

Last, throughout our analyses in an effort to facilitate comparisons, we used the same covariates as BGM and did not initially include maternal labour supply as an explanatory variable in equation (2). Yet, it is quite plausible that the exclusion restriction property of the instrument is violated if one does not control for maternal labour supply. The policy has direct impacts on childcare costs which may not only influence childcare attendance but also female labour force participation. These effects were reported in Table 3. As such, rather than making the heroic and implausible assumption that maternal labour supply can be treated as an exogenous explanatory variable,[33] we split the sample on the basis of whether the mother holds a job and replicated all of the analyses contained in Table 4. These results for both subsamples are reported in Table 6 are robust and in most cases do not differ in a statistically significant manner. However, the magnitude of the effect of the local average treatment effect of childcare on MSD and hyperactivity/inattention scores is larger in the sub-sample of mothers who are working. While not reported, the first stage properties of the instrument are strong in both subsamples indicating that the policy had a strong and statistically significant influence on childcare decisions for mothers who both decided to work and those who did not join the labour force. This set of results further reinforces the striking heterogeneity in the impacts of attending childcare and confidence in attributing the estimated effects to childcare itself.

4.3 Examining the sensitivity to the assumptions underlying the analyses

The assumption of selection on observables that underlies the inverse propensity weighting estimates presented in column 3 of Table 4 is potentially quite strong and concerns as to whether the presented results are driven by spurious correlations between unobserved parent

[33] While an argument could be made that this exogeneity is plausible in specifications where we are exploring the effect of childcare on child's own outcome it will clearly be of greater debate when the outcome is maternal depression. Since our existing strategy is just identified, then to treat both maternal labour supply and childcare attendance as endogenous would require another valid instrument.

Table 5: First Stage Estimates of the Childcare Attendance Equation

Covariate	Coefficient (P-Value)
Policy	0.219 (0.000)***
Mother High School Graduate	0.097 (0.000)***
Mother Some Post-Secondary	0.202 (0.000)***
Mother University Degree	0.278 (0.000)***
Father High School Graduate	0.026 (0.083)*
Father Some Post-Secondary	0.021 (0.134)
Father University Degree	-0.015 (0.349)
Mother's Immigrant Status	-0.052 (0.000)***
Father's Immigrant Status	-0.065 (0.000)***
Population < 30K	0.04 (0.000)***
Population 30-100K	0.036 (0.005)***
Population 100-500K	0.037 (0.000)***
Population > 500K	0.019 (0.113)
One Sibling Younger or of the Same Age	-0.184 (0.000)***
Two or More Siblings Younger or of the Same Age	-0.313 (0.000)***
One Sibling Older	-0.073 (0.000)***
Two or More Older Siblings	-0.188 (0.000)***
Child Gender Male	0.005 (0.532)

— Note: For each covariate column 1 presents the associated coefficient estimate from the first stage of the IV regression as described in Eq. (2). Standard errors are corrected at the province-year level then used to calculate p-values (presented in parentheses) testing the statistical significance of the corresponding estimate. Due to space restrictions we are unable to show all the covariates used including child's, mother's and father's age indicators, as well as cycle and province indicators. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

Table 6: Estimates of the Causal Effect of Attending Childcare by IV

	Mother Works [Q-Value]	Mother Does Not Work [Q-Value]
	Child Outcomes	
MSD Score	-13.057	-11.33
	[0.028]**	[0.014]**
PPVT Standardized Score	-3.929	-6.213
	[0.692]	[0.604]
Hyperactivity and Inattention Score	2.398	0.854
	[0.000]***	[0.385]
Emotional Anxiety Score	0.715	1.27
	[0.011]**	[0.091]*
Physical Aggression Score	2.966	2.596
	[0.004]***	[0.004]***
Separation Anxiety Score	0.714	0.79
	[0.273]	[0.419]
Child in Excellent Health	-0.385	-0.168
	[0.001]***	[0.599]
Never had a Nose/Throat Infection	-0.799	-0.938
	[0.000]***	[0.002]***
Never had an Ear Infection	-0.195	-0.529
	[0.273]	[0.000]***
	Parent and Family Outcomes	
Mother in Excellent Health	0.045	-0.212
	[0.613]	[0.082]*
Father in Excellent Health	0.121	-0.108
	[0.599]	[0.45]
Family Dysfunction Index	2.905	-0.34
	[0.042]**	[0.83]
Mother's Depression Score	4.727	4.84
	[0.000]***	[0.040]**

— Note: For the outcome variable in each row we present estimates of treatment effect of childcare, the coefficient on Ccare from the instrumental variable methods. The sample is split between mother's who are employed and those who are not. The specification includes the covariates listed in Table 1 as well as province and cycle indicators. Standard errors are corrected at the province-year level then used to calculate q-values (presented in parentheses) testing the statistical significance of the corresponding estimate. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

and or family characteristics and childcare use appear plausible. To explore this possibility, we use the methods developed in Altonji, Elder and Taber (2005,2008) to assess what the severity of the omitted variable bias must be to nullify the main results. In practice, this technique involves estimating a bivariate probit model,[34] and first calculating how much selection on observables effects the estimated childcare impact. Next we determine how large selection on unobservables must be to negate the documented effects.[35]

If the ratio of these two elements is greater than one our estimated treatment effect is only in danger from omitted variable bias when the unobserved variables have more explanatory power in selection than the observed ones. Thus, as this ratio grows larger so too does our confidence of the estimates presented. Since this method is designed to use binary outcome variables, we discretize the continuous test scores by assigning individuals a 0/1 score for being below or above the mean. Our results indicate that the effect of selection on unobservables must be 5.22 times as large as that from the variables we have observed to nullify the positive effect of childcare on motor and social development scores.[36]

This strengthens our confidence in the results because it suggests that the cumulative effect of missing variables must outweigh the observables by over five times in order to negate the discovered effect entirely. It is unlikely that important missing variables (i.e. earnings or innate ability) can account, in relative terms, for such a large effect. In summary, this exercise suggests that the assumption of selection on observables is plausible and that the results reported in Table 4 are not driven by unobserved confounders. However, there

[34]Formally, the bivariate probit model can be linked to equation (2) where Y is restricted to be a binary outcome. This model imposes additional assumption on the error terms

$$\begin{bmatrix} \varepsilon \\ u \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).$$

where ρ represents the correlation between the unobserved factors affecting selection of childcare and the outcome of interest.

[35]This assumes that there is equality of selection on observables and unobservables. That is, taken at random a given variable from either the observables or unobservables will in expectation affect selection similarly. This is a plausible assumption since it is much more likely that the chosen covariates account for more of selection then those not accounted for. Thus, this method establishes a lower bound for the extent of the omitted variable bias.

[36]We also discretize using the median and estimate a ratio of 4.34.

are three exceptions since the effects of childcare on the health measures did not display the same degree of robustness to missing confounders. The estimated ratio of selection on unobservables relative to selection on unobservables found for i) whether the child in excellent health, ii) child has had a nose/throat infection, and iii) child had an ear infection variables were 1.183, 0.414, and 0.773 respectively. Though arguments could be made that these ratios are close enough to one to suggest the omitted variable bias would not negate the estimated effect, these ratios are not nearly as reassuring and there is a great potential that our conclusions for the health outcomes are indeed sensitive to the plausibility of the selection on observables assumption.[37]

5 Conclusion

As a popular notion universal care childcare captures the public eye, but its implication, now and tomorrow, is far reaching and thus should be approached with evidence at the heart of the policy-making process. This paper extends earlier research that evaluates the socio-economic consequences of the Quebec family policy in several ways. First, by demonstrating that BGM's findings that access to subsidized childcare has negative impacts on individual developmental, behavioural, and health measures are robust to the inclusion of data up to ten years after the reform. Second, instrumental variable estimates suggest that children and families who choose to attend childcare as a result of the introduction of policy experience substantial declines in a variety of developmental and health outcomes. However, estimates of the average effect of attending childcare, obtained via inverse propensity score reweighting, generally show insignificant and positive effects on child development and behavioural outcomes.

Taken together, the results from this study expand our knowledge about the effective-

[37]There is no accepted threshold in the literature to determine when the degree of selection on observables is sufficient. In practice, this depends on what one would ex-ante predict is the strength of the covariates in the selection equation relative to those factors that are unaccounted for and than gauge whether these factors could bias the results.

ness of policies that subsidize childcare in Canada on developmental outcomes and reject the proposition that introducing a universal childcare program will unambiguously weaken individual and family outcomes. The large difference between the estimates of the average effect of childcare from the local average treatment effects suggest that some groups may derive more benefit from childcare than others. Understanding the sources of treatment effect heterogeneity may help us to develop an understanding of whether it would be advantageous to target child care to those who would benefit most from its provision.

An important limitation of this study is that the NLSCY treats all childcare in Canada as a homogenous good. There is a large literature on the importance of child care quality and Japel et al (2005) found that in Quebec after the policy was implemented that only 5% of programs were high quality and the majority of arrangements scored below the mid range with slightly over 10% were classified at below minimal quality. Second, the NLSCY only collects information on a narrow range of child and family outcomes and there may be other domains of importance including those measured at the community and societal level. Despite these caveats there are avenues in which one can examine if there are systematic patterns of treatment effect heterogeneity.

In this vein we have conducted two companion studies that explore treatment effect heterogeneity in the effects of access to and attending childcare. In Kottelenberg and Lehrer (2012b) we demonstrate substantial response differences in child outcomes by gender: male children and their families drive the negative effects in several key outcomes. Kottelenberg and Lehrer (2012a) move beyond simply estimating the average effect of access to and attending childcare and consider estimating the treatment effects across the full distribution of each outcome variable. Composed of negative and statistically insignificant effects, the underlying pattern in the treatment effects highlight substantial heterogeneity. The authors conclude by presenting suggestive evidence of large reductions in parental investments for these children once their children begin to attend subsidized childcare. Thus, we postulate that home inputs are important and that interventions within schools may only reinforce at

home preparation for a small fraction of the population. For the remainder, the changes in school inputs may be offset as parents substitute their investments into their children towards other activities. However, the extent and pattern of heterogeneity in parental input decisions has not been fully investigated. These results do not imply poor parenting skills, but may indicate that parents have limited knowledge of their child's human capital production function and make optimization errors when choosing inputs. Overall, as a result of the array of behavioural responses to the Quebec Family Policy, the continuing changes to childcare programs in Quebec over time, increases to maternal labour supply and changes in parental behaviour, the exact workings of universal childcare programs on child development are far from obvious. In conclusion, we suggest that further investigation using both qualitative and quantitative data is needed to improve our understanding of the pathways through which childcare contributes to the production of education, health, family and parenting outcomes.

References

Adema, William, and Peter Whiteford (2007) 'Babies and Bosses: Reconciling Work and Family Life.' Technical Report, OECD, Paris

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005) 'Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.' *Journal of Political Economy* 113(1), 151–184

—— (2008) 'Using Selection on Observed Variables to Assess Bias from Unobservables When Evaluating Swan-Ganz Catheterization.' *American Economic Review* 98(2), 345–350

Anderson, Michael L. (2008) 'Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.' *Journal of the American Statistical Association* 103(484), 1481–1495

Baker, Michael (2011) 'Innis Lecture: Universal Early Childhood Interventions: What is the Evidence Base?' *Canadian Journal of Economics* 44(4), 1069–1105

Baker, Michael, Jonathan Gruber, and Kevin Milligan (2008) 'Universal Child Care, Maternal Labor Supply, and Family Well-Being.' *Journal of Political Economy* 116(4), 709–745

Benjamini, Yoav, and Daniel Yekutieli (2001) 'The Control of the False Discovery Rate in Multiple Testing under Dependency.' *The Annals of Statistics* 29(4), 1165–1188

Blau, David M. (2001) *The Child Care Problem: An Economic Analysis* (New York: Russel Sage Foundation)

CBC News (Online) (2009) 'Daycare: The Debate Over Space available at http://www.cbc.ca/news/story/2009/02/06/f-daycare.html'

Charach, Alice, Elizabeth Lin, and Teresa To (2010) 'Evaluating the Hyperactivity/Inattention Subscale of the National Longitudinal Survey of Children and Youth.' *Statistics Canada Health Reports* 21(2), 43–50

Cleveland, Gordon (2007) 'The Benefits and Costs of Quebecs Centres de la Petite Enfance.' In 'ACFAS Conference' Trois-Rivieres

Cleveland, Gordon, Carl Corter, Janette Pelletier, Sue Colley, Jane Bertrand, and Janet Jamieson (2006) 'A Review of the State of Early Childhood Learning and development in Child Care, Kindergarten and Family Support Programs.' Technical Report

Cleveland, Gordon, Susan Colley, Martha Friendly, Donna S. Lero, and Richard Shillington (2003) 'The State of Data on Early Childhood Education and Care in Canada: National Data Project (Final Report).' Technical Report, Childcare Resource and Research Unit, Toronto

Currie, Janet (2001) 'Early Childhood Education Programns.' *The Journal of Economic Perspectives* 15(2), 213–238

Datta Gupta, Nabanita, and Marianne Simonsen (2010) 'Non-Cognitive Child Outcomes and Universal High Quality Child Care.' *Journal of Public Economics* 94(1-2), 30–43

Ding, Weili, and Steven F. Lehrer (2011) 'Experimental estimates of the impacts of class size on test scores: robustness and heterogeneity.' *Education Economics* 19(3), 229–252

Fortin, Pierre, Luc Godbout, and Suzie St-Cerny (2011) 'Economic Consequences of Quebecs Educational Childcare Policy.' In 'Early Years Economics Forum' Toronto

Havnes, Tarjei, and Magne Mogstad (2011) 'No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes.' *American Economic Journal: Economic Policy* 3(2), 97–129

Heckman, James J (2006) 'Skill Formation and the Economics of Investing in Disadvantaged Children.' *Science* 312(June), 1900–1902

—— (2011) 'Effective Child Development Strategies.' In *The PreK Debates Current Controversies and Issues,* ed. Edward Zigler, Walter S. Gilliam, and W. Steven Barnett (Baltimore: Brookes Publishing Company)

Hirano, Kieisuke, Guido W. Imbens, and Geert Ridder (2003) 'Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score.' *Econometrica* 71(4), 1161–1189

Hochberg, Yosef (1988) 'A sharper Bonferroni procedure for multiple tests of significance.' *Biometrika* 75(4), 800–802

Imbens, Guido W., and Joshua D. Angrist (1994) 'Identification and Estimation of Local Average Treatment Effects.' *Econometrica* 62(2), 467–475

Japel, Christa, Richard E Tremblay, and Sylvana Côté (2005) 'Assessing the Quality of Daycare Services Based on the Quebec Longitudinal Study of Child Development.' *IRPP Choices* 11(5), 1–42

Kling, Jeffrey R., Jeffery B. Liebman, and Lawrence F. Katz (2007) 'Experimental Analysis of Neighborhood Effects.' *Econometrica* 75(1), 83–119

Kottelenberg, Michael J. (2009) 'Long-Run Effects of the Quebec Family Policy: Revisiting Universal Child Care, Maternal Labour Supply and Family Well-Being.' *Master's Thesis, Queen's University*

Kottelenberg, Michael J., and Steven F Lehrer (2012a) 'Reinvestigating Who Benefits and Who Loses from Universal Childcare in Canada.' *Queen's University, (mimeo)*

Kottelenberg, Michael J., and Steven F. Lehrer (2012b) 'The Gender Effects of Universal Child Care in Canada: Much ado about Boys?' *Queen's University, (mimeo)*

Lefebvre, Pierre, and Phillip Merrigan (2008) 'Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada.' *Journal of Labor Economics* 26(3), 519–548

Lefebvre, Pierre, Philip Merrigan, and Matthieu Verstraete (2009) 'Dynamic labour supply effects of childcare subsidies: Evidence from a Canadian natural experiment on low-fee universal child care.' *Labour Economics* 16(5), 490–502

Lefebvre, Pierre, Phillip Merrigan, and Francis Roy-Desrosiers (2011) 'Quebec's Childcare Universal Low Fees Policy 10 Years After: Effects, Costs, and Benefits'

Mustard, Fraser (2008) 'testimony to The Standing Senate Committee on Social Affairs, Science and Technology, Thursday, February, 14, 2008, Ottawa, ON available at http://www.parl.gc.ca/Content/SEN/Committee/392/soci/45189-e.htm?comm_id=47% \&Language=E\&Ses=2'

NICHD-ECCRN (2004) 'Type of child care and childrens development at 54 months.' *Early Childhood Research Quarterly* 19(2), 203–230

Simes, R. J. (1986) 'An improved Bonferroni procedure for multiple tests of significance.' *Biometrika* 73(3), 751–754

Staiger, Douglas, and James H. Stock (1997) 'Instrumental Variables Regression with Weak Instruments.' *Econometrica* 65(3), 557–586

Statistics Canada (2003) 'Microdata User Guide: National Longitudinal Survey of Children and Youth.' Technical Report September 2002, Statistics Canada, Ottawa, ON

Tougas, Jocelyne (2002) 'Reforming Québecs early childhood care and education : The first five years.' Technical Report April, Childcare Resource and Research Unit, Toronto