Reinvestigating Who Benefits and Who Loses from Universal Childcare in Canada

Michael Kottelenberg
Queen's University

Steven F. Lehrer[1]
Queen's University and NBER

September 2011

[1]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 Michael Kottelenberg's Queen's University 2009 Masters thesis. Lehrer wishes to thank SSHRC for research support. We are responsible for all errors.

Abstract

Ongoing interest in childcare issues has been stimulated in part by a trend towards the increasing labour force participation of mothers with young children. The empirical literature evaluating the impact of childcare programs on developmental outcomes appears to reach conflicting results. On the one hand, small scale experimental studies find large benefits for disadvantaged children, whereas studies that look at programs which provide universal coverage do not find evidence of significant benefits. We argue that by examining the impact of the policy on the full distribution of the outcomes and not simply focus on mean impacts may potentially reconcile these findings. More generally, improving our understanding of who truly benefits and who loses from these policies is increasingly important as several countries worldwide are currently considering implementing a variety of universal early childhood education programs.

Specifically, we extend earlier research that utilized parametric methods to recover intent to treat estimates of access to $5 a day childcare by using a nonparametric estimators to identify causal impacts of the program itself at specific percentiles within the distribution. We find that the Quebec Family Policy significantly boost test scores for children who are most disadvantaged and located at the lower quantiles of the distribution. However, students between the 20th and 85th quantile receive significant negative impacts from child-care. As this group is the lion's share of the population, it is not surprising that the mean impacts presented in earlier work are negative in sign. Against traditional notions surrounding the implementation of universal child care, positive effects for children at the highest portion of the distribution are also found. Taken together these results indicate the importance of a distributional analysis in addition to a study of mean effects, provide a more complete picture of how childcare affects subsequent development and illustrate the trade-offs that policy-makers will be making if these policies are adopted.

Very Preliminary and Incomplete

Please Do Not Cite or Quote without Permission from the Authors

1 Introduction

Ongoing interest in child care issues has been stimulated in part by a trend towards the increasing labour participation of mothers with young children. In both Canada and the United States the employment rate of mothers with children under the age of 6 has risen substantially: in Canada from 31% in 1976 to 71% in 2008, and in the United States from 34% in 1976 to 56% in 2001.[1] This labour supply trend has been coupled with a growing focus—in both public and academic spheres—on the need for governmental participation in the provision of child care. The supply and cost of child care have increasingly become pertinent issues across North America. In 2008, CBC News reported complaints concerning the few available spaces in existing day-care facilities. The Child Care Advocacy Association of Canada suggested that, "the federal government has simply failed to meet the child care needs of Canadian families" [5]. Meanwhile, numbers taken from the Survey of Income and Program Participation demonstrate that the average American family using child care pays a whopping $6,708(US) annually for this service. As a result, among parents, politicians and policy-makers, universally provided and publicly funded child care is an issue of increasing importance.[2]

Adding support for such governments to adopt such policies is the general interpretation of the results from studies including [16] which make a strong case that investing in children at early ages will yield larger societal and individual benefits than interventions targeted at later ages. 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

[1] The statistics for Canada are derived by the author from the Canadian Labour Force Survey, while the American statistics are calculated from the Current Population Survey.

[2] Indeed, politicans have responded and this is a central element of the Liberal campaign. Some politicians have gone as far as mandating preschool in hopes of facilitating an easier transition into the first years of school.

for disadvantaged children.[3] If by providing early childhood education the government hopes to address the issues of equity in child educational outcomes or attempts to reduce other educational costs associated with lower portions of the achievement distribution, such as special education, then a greater focus must be placed on the dispersion of test scores and effects on students achieving at the lowest levels. While there is strong evidence that targeting programs is effective, current policy proposals are aimed at universal coverage and there is much less evidence on whether making such programs universal might also produce desirable effects. This pertinent issue was addressed in 2008 by Michael Baker, Jonathan Gruber, and Kevin Milligan [2] who provided an initial analysis of the impact of a unique program in North America, the Quebec Family Policy.[4] Implemented in 1997, this legislation promised universal access to five-dollar-a-day day-care across the province. The authors found a variety of negative effects for health, developmental, and behavioural measures. [5]

As such, the existing literature on the impacts of early childhood programs appear to reach conflicting results. On the one hand, small scale experimental studies find large benefits for disadvantaged children, whereas studies that look at programs which provide universal coverage do not find evidence of significant benefits. We believe that by examining the impact of the policy on the full distribution of the outcomes and not simply focus on mean impacts may potentially reconcile these findings. More generally, improving our understanding of who truly benefits and who loses from these policies is increasingly important as several countries worldwide are currently considering implementing a variety

[3] See Currie (2001) or the wide-body of evaluation research that examined either the Perry Preschool Project or the Abecedarian Project.

[4] Merrigan and Lefebvre also study the Quebec Family Policy, and its impact on maternal labour supply.

[5] This article has attracted substantial attention and while heavily critiqued by advocacy groups such as UECE and Child Care Advocacy Association of Canada, it has been praised by academics who awarded the authors the 2009 Doug Purvis Memorial Prize for a highly significant written contribution to Canadian economic policy.

of universal early childhood education programs.

Specifically, our study hopes to further dissect the Quebec Family Policy by extending Baker *et al.* (2008) (henceforth referred to as *BGM*) in three ways. First, their research evaluated the Quebec Family Policy at a time when the program was newly implemented and child care centers were trying to ramp up their services to meet the demand. The implementation of large scale social programs is rarely frictionless, and often requires time for setup, finding high quality workers and societal integration. For this reason earlier results may be capturing short-run changes that differ from the impact when child care centers are up and running in a smooth manner. Since the BGM study, there have been two new cycles of data available from the National Longitudinal Study of Children and Youth (NLSCY), enabling the present study to revisit the original results to see whether the results are robust to later time period.[6]

Second, this research extends the existing literature beyond mean effects by opening a discussion on child care's impact on the distributional characteristics of child cognitive test scores. This analysis relaxes parametric assumptions in the BGM study to identify treatment effects at specific percentiles within the distribution. In particular, we use the Athey and Imbens (2006) non-linear difference-in-difference methods to expose the heterogeneous and nuanced effects of the Quebec Family Policy that exists across child developmental scores [**?**]. In addition, we consider influence function methods proposed in Firpo *et al.* (2009) estimating the policy impact on distributional statistics such as variance, skewness, and kurtosis [12].

Third, we extend past research which focused primarily on intent to treat estimates of childcare in Canada. Using inverse propensity score reweighting procedures we attempt

[6]Note, we also considered longer run impacts of the program including the impacts of the policy at higher ages. As the common trend assumption is harder to defend in that analysis due to differing education programs across provinces (i.e. class size reductions in the early grades in Ontario) we do not report the results but are happy to make them available to any interested reader upon request.

to identify the causal impact of the program itself on both child and parent outcomes. Further, we examine whether the impacts of access to the universal childcare program vary across observable dimensions by conduction analyses by age and gender. this allows us to address two important questions: What is the appropriate age to encourage child care participation and how does child care contribute to the recent evidence on a growing gap in academic achievement between sexes?

Our main findings allow us to reconcile the conflicting evidence between the experimental studies and BGM. In particular, we find that the Quebec Family Policy significantly boost test scores for children who are most disadvantaged and located at the lower quantiles of the distribution. However, students between the 20th and 85th quantile receive significant negative impacts from child-care. As this group is the lion's share of the population, it is not surprising that the mean impacts presented in BGM are negative in sign. Against traditional notions surrounding the implementation of universal child care, positive effects for children at the highest portion of the distribution are also found. This finding is new to the child-care literature, as there are no experiments that target the most advantaged in society. Finally, our analysis of distributional statistics reveals an increase in the variance of a primary development score suggesting a movement away from homogeneous human capital investments. Taken together these results indicate the importance of a distributional analysis in addition to a study of mean effects, provide a more complete picture of how child care affects subsequent development and illustrate the trade-offs that policymakers will be making if these policies are adopted.[7]

This paper is structured as follows. Section 2 describes the context of the current study: the state of North American child care and details of the Quebec Family Policy, followed

[7] Despite traditional emphasis in the applied literature to report only mean effects of a policy, the existence of treatment effect heterogeneity in education programs is now overwhelming. A number of recent studies are reaching more definitive conclusions about the distribution of treatment effects and important applications in this regard include Bitler, Gelbach and Hoynes (2008), Djebari and Smith (2009), among others.

by an overview of the existing child care literature. An outline of the NLSCY data and the difference-in-difference empirical strategy is given in section 3. Here particular attention is placed on the necessary two parent family sample restriction, integration of newly available NLSCY data, and adjustments to the original setup. Next, in section 4, we perform the mean estimation of the policy's effect with more recent waves of the NLSCY data to determine if the results of *BGM* are robust. We find that the results in BGM are indeed robust to the inclusion of two new cycles of NLSCY data. Second, our evidence uncovers significantly greater emotional and behavioural problems for males and additionally suggests that children who attend childcare between the ages of 0-2 are at higher risk of having negative developmental outcomes. Our main results are presented in section 5, where we document the presence of heterogeneous effects across test score distributions. In particular, we find that expected benefits at the lower end of the distribution are confirmed, while children in the middle of the distribution fall behind. Against traditional notions surrounding the implementation of universal child care, positive effects for children at the highest portion of the distribution are also found. Finally, our analysis of distributional statistics reveals an increase in the variance of a primary development score suggesting a movement away from homogeneous human capital investments. Taken together these results indicate the importance of a distributional analysis in addition to a study of mean effects and also hope to provide a more complete picture of how child care affects subsequent development.[8] Finally, in the concluding section we summarize our findings, argue that policymakers should consider targeting childcare rather than developing policies that would introduce universal coverage and discuss directions for future research.

[8] Despite traditional emphasis in the applied literature to report only mean effects of a policy, the existence of treatment effect heterogeneity in education programs is now overwhelming. A number of recent studies are reaching more definitive conclusions about the distribution of treatment effects and important applications in this regard include Bitler, Gelbach and Hoynes (2008), Djebari and Smith (2009), among others.

2 Background and Previous Literature

The Quebec Family Policy, promising universal access to five-dollar-a-day day-care, is one of the most comprehensive policy measures taken by any North American government in response to child care trends and concerns.[9] In 1997, the Quebec government implemented a bold set of policies in hopes of encouraging higher birth rates, primarily by strengthening governmental support of parents. In large part, this support came in the form of a massive expansion of the child care system at a variety of age levels. Parents with children aged 0-4 would now have access to child care at a rate of $5 per day (becoming $7 per day in 2004). This program was implemented gradually; access was extended to children aged 4 in 1997, aged 3 in 1998, aged 2 in 1999 and aged 0-2 in 2000. Child care services were also increased with the introduction of full-day kindergarten and, although not officially part of the policy, more child care spaces for school aged children [34]. 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.

The Quebec Family Policy's extension of highly subsidized universally available child care to children aged 0-4 provides a unique opportunity to examine the impact of a switch to a comprehensive system of child care support. Many advocates point to the potential benefits of early childhood education to supplementing early human capital accumulation. This includes cost saving through reduction in grade retention and developmental delays. Yet, many questions remain: Does making child care more accessible, particularly for children in the earliest stages of life, also benefit development in this way? If there are

[9] In the United States particular attention has instead been placed on the development of a pre-kindergarten system: 41 of 52 states have publicly funded pre-kindergarten programs serving children to varying degrees. However, despite support for a Zero-to-Five early education model by current President Barack Obama, there remains no state in which universal child care access is provided from the earliest years of a child's life [6]. Only Oklahoma and Georgia provide true universal access for pre-kindergarten, which only targets children aged 4 [31].

mean changes who is benefiting? How does early childhood education play a role in the increasing gender gap?

Previous research into the universal subsidization or provision of child care outside of the case of Quebec has begun to address these, and other similar questions. Psychologists agree that a nurturing learning environment during the early years of a child's life is an important element in brain development, and in achieving social and cognitive potential in later life [14]. Economist James Heckman suggests that from a perspective of economic gain there is no better time to invest in human capital than in the earliest years [16]. Heckman's work stresses the importance of not only cognitive growth but also of healthy socio-emotional growth in facilitating future learning. His findings have also been supported by educational physiologists [19]. Thus, much of the subsequent child care research, discussed below, has focused on reaching out to disadvantaged children in the expectation of improving cognitive and socio-emotional outcomes by eliminating significant barriers to development in negative home environments. In applications of child care at the universal level, as in the Quebec Family Policy, a range of effects are expected. In addition to the positive gains mentioned above, negative outcomes might also be present, as children who receive strong one-on-one parental care at home may shift to potentially less effective child care with higher adult-child ratios away from home, and while in the home may be subjected to working parents affected by potentially higher levels of stress.

In past literature the presence of positive mean effects on development outcomes have been crucial to discussions surrounding the implementation of child care programs. To date, research has primarily focused on understanding these effects outside the context of universal subsidization, as many studies have attempted to quantify the mean effects of targeted and private child care programs. Within this average effect context, the conclusion has been reached that program usage is associated with an increase in emotional and behavioural problems [29, 22]. It is suggested that such effects stem from the inability of caregivers to maintain consistent, responsive, and sensitive emotional interaction, as well as from alterations in peer interactions [7, 35]. Evidence for positive cognitive gains, on the other hand, is not as consistent. For example, NICHD-ECCRN (2002a) and Magnuson *et al.* (2007) find benefits from child care, while Lefebvre and Merrigan (2002) find no change at all [27, 24, 20].

The evidence on universally implemented programs is decidedly sparse. Several studies have examined universal pre-kindergarten programs for children aged 4 in the United States, and have found significant improvement to child outcomes after program implementation [13, 17]. However, these programs remain unable to address the effects of extending full access child care to children ages 0-4. In one recent discussion paper, Havnes and Mogstad (2009) explore the impact of subsidized and universally provided child care on extended long-run outcomes, such as educational attainment and labour market participation [15]. Their study uses a difference-in-difference approach and makes use of a 1970's expansion of a universal child care program in Norway. While they find positive outcomes for the variables of interest, such long-run studies often suffer from confounding effects caused by contemporaneous policy change.

Earlier studies have also failed to address the issue of differing effects across the distribution of child developmental scores, focusing instead on mean outcomes alone. When such issues have been given attention, it is only through analysis on subsets of the overall sample. For example, studies such as Magnuson *et al.* (2007), Gormley *et al.* (2005) and Henry *et al.* (2004) discuss the bottom of the distribution by focusing on the greater positive mean effect of the treatment on children from situations of economic disadvantage, such as minorities and low income families [24, 13, 17]. While students from these sub-groups often make up a more significant proportion of students falling behind, there is still no clear indication that child care is reducing the achievement gap for the very worst of these students [33].

3 Data and Empirical Strategy

As in *BGM* , the data used to address the effects of childcare on mean and distributional outcomes is taken from the NLSCY. However, in this case cycles 6 (2004-05) and 7, (2006-07) which were not available in the previous study, will also be included. The NLSCY is a longitudinal data set compromised of a rich set of demographic and child care choice and outcome related variables. The survey's first cycle, taken in 1994-95, sampled Canadian children aged 0-11. 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.[10] In addition to a biannual follow up with the existing cohort, a new cohort of children aged 0-1 has been added at each new cycle. In each cycle approximately 2,000 children are represented at each age level allowing for the construction of a sizeable repeated cross-sectional data set.

The NLSCY is particularly useful because it contains both child developmental scores and extensive questions relating to child care usage, parental labour supply, and other demographic characteristics. This provides the opportunity to understand the effects of child care policy on a variety of childhood development and behaviour indicators. These include the revised Peabody Picture Vocabulary Test (PPVT) score for children aged 4, a standardized motor and social development score for children aged 0-3, and a series of child behavioural scores relating to hyperactivity, anxiety, physical aggression and opposition.

The natural experiment created by the implementation of the Quebec Family Policy, coupled with the repeated cross-section data of the NLSCY, makes a difference-in-difference

[10] These exclusions represent about 2% of the Canadian population.

approach an appropriate strategy for isolating a causal effect on child care use, labour supply, and development scores. The validity of a difference-in-difference method, however, relies on several key assumptions. First, as Meyer (1995) and Besley and Case (2000) highlight, it is important that natural experiments originate from a true source of exogenous change [25, 4]. They cite examples in which estimates are misleading because the direction of causation from policy change to variation in dependent variables was reversed. In the present case the change stems from the Quebec Family Policy, and in particular the implementation of the child care subsidy. Thus, this legislation must not be a response by the Quebec government to increasing maternal labour and child care trends. Although such trends were present in Quebec prior to 1997, the scope and magnitude of the Quebec child care program is such that exogeneity is plausible. Another important condition outlined by Meyer (1995) is the absence of pre-treatment effects [25]. This equally plausible assumption requires that prior knowledge of the Quebec Family Policy should not have impacted decisions by parents before the 1997 implementation of the program.

The validity of the difference-in-difference estimation also implies the conditions of common trend and common support. These conditions require a strong link between the treatment and comparison groups. In our case the similarities between these groups should be such that one could expect similar uptake in child care and maternal labour supply along two dimensions. Without policy change, uptake, or the change in usage or supply, should be the same in both the treatment and comparison groups over time, satisfying the common trend assumption. In the case of policy change for either group, the resulting uptake should be the same in both the treatment and comparison groups, thereby fulfilling the common support assumption.[11] These constraints lead to a restriction of the

[11] The common support assumption may also be weakened by differences in francophone and anglophone populations. To ensure the validity of our results this group we perform our estimation procedure on a francophone and anglophone sub-samples and find that the results are similar to the full sample. These results are available upon request.

sample to two-parent families. Such a restriction eliminates issues connected to other pre-policy subsidization, which was much higher for single parent families, and also isolates an appropriate comparison group not affected by changes in other policies during the course of the study, such as paternity leave regulations. While such a limitation reduces the reach of this study, the two-parent family remains a key focus of the universal child care debate, which is most concern with extending subsidized access to child care where it is not already available, as it is for many single-parent families.

4 Estimation of Mean Effects

To validate *BGM* results in the long-run and to show a continuity between their research and ours, this study replicates the methods used in *BGM*. The general regression technique used will separate policy effects by comparing the difference in Quebec and the rest of Canada in both pre- and post-policy period. The " difference-in-difference regression equation for the outcome of interest Y is as follows:

$$Y_{ipt} = \alpha Policy_{ipt} + \beta PROV_p + \phi YEAR_t + \lambda X_{ipt} + \varepsilon_{ipt} \tag{1}$$

where i, p, and t index individual, province, and year. The vector of covariates X, found in Table 1a, includes controls for child, parent, family, and geographic characteristics and PROV and YEAR are a series of province and time dummies. The estimated policy coefficient α captures the effect of being eligible for universally subsidized child care in Quebec on the outcome of interest. It represents the change in outcomes in Quebec pre- and post-reform relative to changes in other provinces over the same time period, where the pre-policy period, 1994-97, and the post-policy period, 2002-07, are captured by NLSCY cycles 1-2 and 5-7 respectively. Exclusions of data from 1998-2001 is made on the basis that the program was first of all not yet fully implemented during 1998-99 and rapidly expanding during 2000-01.[12]

This studies primary mean analysis deviates from *BGM* in order to accommodate for the proposed distributional analysis. To incorporate the rich set of observables provided by the NLSCY into the non-linear difference-in-difference methods (discussed in section 5) we rely on an inverse propensity scores method outlined in Hirano, Imbens, and Ridder (2003) [18].[13] Such a technique removes differences in the unconditioned distribution of scores that may arise from differences in cohorts and is accomplished by weighting the data such that observable covariates from the different sub-populations are balanced. Using the proposed inverse propensity weights, we are able to produce distributions of the test score for the treatment and comparison in the pre- and post-policy periods given a set of equivalent observables. Thus, reducing the threat that unaccounted for observables are driving the observed treatment effects.

The weights proposed by Firpo (2007) are calculated as follows,

$$\hat{\omega}(X_i, INCARE_i) = \frac{INCARE_i}{N\hat{p}(X_i)} + \frac{1 - INCARE_i}{N(1 - \hat{p}(X_i)} \tag{2}$$

where the $\hat{p}(X_i)$ is the estimated probability of a child being enrolled in childcare and $INCARE_i$ is a dummy variable for being in child care. The predicted probability is obtained through a series logit estimation, which incorporates all the *BGM* covariates and their interactions. This is done so that the chosen probability model is an approximation to a non-parametric estimation procedure, and thus is congruent with the non-parametric non-

[12] The growth in publicly funded childcare seats in Quebec was over 65,000 between 1998-2001 [21]. We also test the inclusion and exclusion of different cycles finding robust results across the different selections. 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.

[13] The application of this procedure to quantiles is discussed in Firpo (2007) [11].

linear difference-in-difference models.[14] Finally, it is important to consider that by using this weighting scheme the interpretation of estimated coefficients switch from intention-to-treat effects to treatment effect on the treated, isolating the impact of the policy on those most likely to enrol in child care.

Mean Policy Effects

In the following section, we replicate *BGM* with the newly available date, outline the mean effects of the policy as estimated using the new inverse propensity weighted method, contrast both these results with *BGM* , and finally de-construct the mean results further by presenting separate regressions for gender and age sub-samples. As this paper mainly limits itself to child and family outcomes, we focus on the child and parental behaviour indexes, child development scores, and health indicators.[15] The behavioural indexes are built from responses about the frequency of observed behaviours. [16] They include Physical Aggression, Hyperactivity and Inattention, Emotional Disorder and Anxiety and Separation Anxiety for children and Family Dysfunction, Aversive Parenting, and Mother's Depression Score for parents. The development scores consist of the Motor and Social Development (MSD) index, an index constructed from responses concerning child ability by the person most knowledgeable, and the Peabody Picture and Vocabulary Test (PPVT), a standardized language related score constructed from child responses to a test conducted by an

[14]The results presented are also verified using a logit estimation in which the covariates are not interacted.

[15]In fact, we replicate and contrast long-run estimates for all the results in *BGM* and find consistency from short to long run. Results for the maternal labour supply and child care usage variables show higher levels of uptake in both categories, an approximate increases of 33% in comparison to the estimates made by *BGM* .

[16]Presented child behavioural indexes are for children aged 2-3. Although available for children aged 4, these indexes are not composed of the same base questions. The indexes for the older children are tested and reveal a similar pattern as those for children aged 2-3.

Table 1: Independent Variables and Summary Statistics

(a) Independent Variables List

Independent Variable	Description
Province	A dummy for each province, Quebec is used as the base group.
Cycle	A dummy for each cycle, Cycle 1 is used as the base group.
Mother / Father Age	A dummy for both mother and father, each age category beginning with 16-20 and increasing by increments of 5. The last dummy created groups Mothers / Fathers aged 46 -99 together.
Mother / Father Education	A dummy for both mother and father for each level of education completed: high school drop-out, high school graduate, some post-secondary and university degree.
Mother / Father Immigration Status	A dummy for both mother and father indicating whether they have ever been landed immigrants.
Urban/Rural Area	A set of dummies indicating living setting. Categories are rural or urban setting with a population of 0-30k, 30k-100k,100k-500k, or 500k+.
Younger and Same Aged Siblings	A set of dummies indicating number of younger and same aged siblings. Categories are 0,1, or 2 or more younger and same aged siblings.
Older Siblings	A set of dummies indicating number of older siblings. Categories are 0,1, or 2 or more older siblings.
Child Age	A set of dummies indicating age of the child: 0,1,2,3 or 4.
Male Child	A dummy indicating whether the child is a male or not.

(b) Select Summary Statistics

	Average in Quebec		Average in Rest of Canada	
Covariate of Interest	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy
Living in Urban Area (500k+)	0.579	0.581	0.428	0.445
Number of Older Siblings	0.715	0.685	0.796	0.755
Number of Younger/Same Aged Siblings	0.268	0.217	0.255	0.246
Mother				
Age	30.93	31.37	31.74	32.65
Immigrant Status	0.089	0.137	0.214	0.252
High School Drop Out	0.133	0.110	0.106	0.075
University Degree	0.203	0.331	0.206	0.343
Father				
Age	33.51	34.20	34.14	35.28
Immigrant Status	0.097	0.162	0.208	0.252
High School Drop Out	0.168	0.139	0.138	0.098
University Degree	0.194	0.291	0.214	0.306
Male Child	0.509	0.514	0.509	0.514
Age of Child	2.026	1.965	1.991	2.021

N ote: The means calculated in this table are for the two-parent family sub-sample: 38648 observations. The data is split between Quebec and the rest of Canada as well as into the pre-policy period, from 1994-97, and the post policy period, from 2002-2007.

external interviewer.[17] Finally, children's health is examined using indicators for never having had a nose/throat infection, never having had an ear infection, and in "excellent health".

Using the newly available NLSCY date, we ask whether the negative child and parent effect found by *BGM* are maintained in the long run. Several reasons exist for considering this question. First, the *BGM* study looked at family and child outcomes during a time when the program was rapidly expanding to meet additional demand.[18] Second, there is reason to believe that the quality of care has improved over-time potentially mitigating negative effects.[19] Third, in long run we expect use of child care will shift towards center-based care, which is associated with better academic and language skills than other types of care arrangements [28].[20]

In terms of significance and sign both the long run estimates and the weighted results remain consistent with *BGM*. Table 2 presents three sets of policy coefficients: the exact replication of *BGM* ; the long-run estimates using the new data; and the estimates using the inverse propensity weighting procedure. In all health and behavioural measures the new estimates imply a worsening effect for children, which more consistent predicts nega-

[17] 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. It is also important to note that both developmental and behavioural measures have undergone rigourous testing.

[18] The implementation of child care for all ages in Quebec was completed in 2000, however the number of available places in subsidized daycare did not come close to satisfying demand. From 2000 to 2007 the province increased its available spaces in the program from approximately 110,000 to 200,000 [21].

[19] The Quebec government legislated in 2000 that two thirds of the staff at CPE's 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%-40% 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.

[20] The initial surge in demand for child care in Quebec will have led to growth in the provision of child care facilities. Growth in home-based care operations may have outpaced growth in institutional based facilities, due to lower start up and organizational costs. In fact, we find evidence to support this claim. Initial short-run results reported a small increase in types of home-based care. In the long-run we find a significant decline in home-based care. Graphical evidence (available upon request) clearly shows that the excess demand created by the policy change was initially met through the use of home-based care and then shifted to center-based care.

tive results than original estimates.[21] At a glance the magnitudes of the weighted results appear comparable to the un-weighted results. It is, however, important to note the distinction between the intention-to-treat effects outputted by the un-weighted procedure and the average treatment effects given by the weighted estimates. To arrive at comparable treatment effects, the intention-to-treat effects should be inflated by dividing the policy coefficient by the probability of being treated.[22] Therefore, the negative effects estimated by the weighted regressions are not as intense. Unlike the given behavioural measures, the policy effect on cognitive development, the MSD and PPVT scores, are insignificant. This stands in contrast to the intention-to-treat estimates which predict a decrease in the the MSD score and subsequently may suggest that earlier cognitive problems are more associated with an increase in maternal labour supply rather than uptake of child care. Because of the relatively strong continuity between these two estimation procedure our discussion remains brief. *BGM* provide a thorough analysis of the mean results giving us leave to dissect the driving force of these mean effects by examining sub-groups and distributional effects.

Differences by Gender and Age

Our exploration begins with the recognition that child development is complex and filled with gender specific nuance. Thus, there is an expectation that the effects of child care uptake on children will also play out differently among the sexes. The results are presented in Table 2 where confirmation of gender differences is found. These differences, however, do not occur in cognitive measures but rather in behavioural and health measures. Negative

[21] Although we discussed several reasons for negative effects to dissipate with program maturity, the inclusion of long-run data in fact shows worsening average treatment effect. This comparison is made via an average treatment effect calculated using the *BGM* short and long run estimates (Results available upon request).

[22] *BGM* hold discussion on the appropriate probability of being treated suggesting either the increase in child care 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.

Table 2: Mean Effects - Child and Family Outcomes

	Mean	BGM Method		Propensity Weighted Estimates		
	(Std. Dev.)	Original	Long Run	All	Girls	Boys
		Child Outcomes				
Hyperactivity Inattention	3.608	0.211^{α}	0.322	0.394	0.073	0.717
	(2.383)	0.208	0.117***	0.116***	0.160	0.165***
Emotional Anxiety Score	1.204	0.120	0.205	0.269	0.243	0.294
	(1.466)	0.056**	0.059***	0.065***	0.090***	0.094***
Physical Aggression Score	4.868	0.380	0.601	0.595	0.535	0.661
	(2.939)	0.085***	0.101***	0.142***	0.201***	0.201***
Separation Anxiety Score	2.626	0.098	0.164	0.302	0.069	0.535
	(1.975)	0.085	0.087*	0.090***	0.130	0.125***
MSD Score	99.185	-1.645	-1.688	-0.250	-0.379	-0.211
	(15.121)	0.461***	0.471***	0.481	0.670	0.672
PPVT Standardized Score	100.665	0.361	-0.435	-1.550	-1.349	-1.603
	(15.169)	0.752	0.765	1.046	1.520	1.470
Child in Excellent Health	0.658	-0.055	-0.049	-0.060	-0.043	-0.078
	(0.474)	0.016***	0.019**	0.014***	0.019**	0.019***
No Nose/Throat Infection	0.451	-0.140	-0.146	-0.147	-0.134	-0.160
	(0.498)	0.025***	0.016***	0.015***	0.022***	0.021***
No Ear Infection	0.524	-0.057	-0.068	-0.055	-0.035	-0.077
	(0.499)	0.020***	0.014***	0.016***	0.023	0.022***
		Family Outcomes				
Family Dysfunction Index	7.875	0.254	0.133	0.671	0.224	1.104
	(5.089)	0.173	0.167	0.144***	0.206	0.200***
Aversive Parenting	6.372	0.198	0.361	0.280	0.266	0.303
	(3.204)	0.067***	0.066***	0.080***	0.112**	0.113***
Mother's Depression Score	4.093	0.420	0.659	0.366	0.221	0.513
	(4.686)	0.119***	0.137***	0.142***	0.203	0.199***

S ignificant at the 10% level. ** Significant at the 5% level. *** Significant at the 1% level.
— Note: This table provides the estimation results for the policy dummy in Equation 1, where the *BGM* Method does not use the inverse propensity weights. The Original column reports estimation results using the cycles 1-5 whereas the Long Run column includes the newly available cycles 6,7. The estimation includes the entire sample of children aged 0-4 who have given a valid response for the corresponding variable of interest.
$^{\alpha}$ Hyperactivy and Inattention replication is different from *BGM* estimation. In cycle 4 of the NLSCY the calculation of this index was changed: two questions making up part of the index were removed and one new question was added. This difference was overcome in the replication by the merging of the existing indexes to produce one in which all questions are common.

effects are found for male children in both the separation anxiety and hyperactivity and inattention indexes, whereas results are insignificant for female children. Male children also exhibit much larger increases in their likelihood of having health related issues. In particular, the policy impact on the likelihood of not being in excellent health for male children is 7.8% for boy and only 4.3% for girls. Similarly, boys are 7.7% more likely to have had an ear infection, while girls are only 3.5% more likely.

In the absence of an effect on the female separation anxiety scores, the observed impact for male children is somewhat surprising. Benenson, Morash and Petrakos (1998) outline the existence of a higher degree of emotional closeness between girls and their mothers at the pre-school age [3]. Such a finding would suppose that female children experience more separation anxiety due to closer parental bonds. It is probable that increased hyperactivity and worse health outcomes in male children are linked closely to the nature of gendered play; interactions in child care are often gender separated, and boys tend to emphasize conflict and physical play [10]. Furthermore, both anxiety issues and hyperactivity problems can be better understood in the context of child-caregiver relations. Some research suggests that female children, not only play more cooperatively, but also receive more supervision and form closer connections to care-givers, thus providing for some of their emotional and structural needs [9, 1]. Extra supervision and relationship development for girls may be playing a significant role in driving these gendered effects.

The gender specific policy effects are also present in parenting behaviour variables. Specifically, parent's of male children experience increases to both the family dysfunction index and mother's depression score while parent's of female children do not. *BGM* discuss the possibility that some of the increase in behavioural problems might be from reporting artifacts as parents who switch into employment face higher levels of stress. We suggest that negative parental outcomes are closely linked with increase to behavioural problems and health outcomes of their children. This is congruent with psychological research which implies that increased behavioural problem can adversely effect maternal function [8]. Similarly, research has found frequent child health problems may also cause psychological distress in mothers [32]. As male children face more of these issues their parents also face increased stresses. Cumulatively, this raises concerns about lasting emotional impacts on male children, who seem to be receiving less supervision and living in more difficult home situations as a result of child care uptake. These results spark interest into the connections between emotional development, home environment, and the growing gender gap in achievement.

Just as patterns of child development across gender complicate the effects of child care, so differences in age also to play a role determining child care effects. Furthermore, an important question to be answered is whether universal child care is appropriate for children of any age. Can an optimal age to begin offering child care services be determined? Disaggregating the main sample by child age, we examine the MSD score for children aged 0-3 and find evidence that the policy effect on children varies greatly by age.[23] Newborns suffer from large negative effects at the hand of the policy, while children aged 1-2 experience negative effects but not to the same degree. On the other hand, estimates for children aged 3 suggest no change to the MSD score, similar to the Age 4 PPVT outcomes. If we are to take Heckman's argument for early human capital accumulation serious, these results suggest that child care is certainly not suited for children aged 0-2. Thus, alternative ways of encouraging development in young children should be sought [16].

[23]Table of results available upon request. point

5 Distributional Effects

5.1 Distribution Variance, Skewness, and Kurtosis

Of increasing importance to policy makers are concerns of equal opportunity for children, as more attention is being placed on enabling children to succeed regardless of socio-economic background. Investment in children at a young age is critical to future development and outcomes suggests that the issue of equal opportunity are likely best addressed in a child's earliest years. To understand child care's impact in this regard, we begin by examining the distributional statistics of developmental scores, suggesting that distribution characteristics will best reflect the heterogeneity of early human capital accumulation. Recently developed methods make it possible to identify policy effects on commonly know distributional statistics, such as variance, skewness, and kurtosis.

Computation of the policy effects on these statistics combines a straightforward transformation of the dependent variable, discussed in Firpo *et al.* (2009), and OLS regression [12]. This transformation is based on the influence function $IF(y; v(F_y))$, a means of capturing the impact of a single observation on the distribution statistic $v(F_y)$. For example, the impact of an observation y_i on the variance of the distribution F_y can be calculated using the variance influence function, $IF(y; \sigma^2) = (y_i - \mu_y)^2$, which equates to calculating the portion of total variance contributed by the single observation. In Firpo *et al.* (2009) the transformation of the independent variables is completed by recentering the influence function at the distributional statistic of interest giving it its name, the *Recentered Influence Function* (RIF).[24] By running OLS on the RIF using the same regressors as in equation 1, we are able to capture the impact of the policy on the select distribution characteristic, while controlling for other covariates. The RIF transformations for variance, skewness, and

[24] This final scaling ensures that the $E(RIF) = v(F_y)$ an important feature in some situations. In this study, the use of the influence function would be sufficient.

kurtosis are defined as follows:

$$
\begin{aligned}
RIF_{variance} &= \sigma^2 + (y_i - \mu)^2 \\
RIF_{skewness} &= \frac{\mu^3}{\sigma^3} + \frac{(y_i - \mu)^3}{\sigma^3} \\
RIF_{kurtosis} &= \frac{\mu^4}{\sigma^4} + \frac{(y_i - \mu)^4}{\sigma^4}
\end{aligned}
$$

Where μ^k is the the k^{th} moment about the mean and σ is the standard deviation.

The results of the RIF regressions for both the original and the inverse propensity weighted method are presented in table 3. Although, distributional changes are not present for the PPVT score, there are significant effects on the MSD score. Both weighted and un-weighted estimation procedures indicate significant increases in variance suggesting that there's a movement away from homogeneous child investment. Along side policy induced increase in kurtosis, a statistic which identifies the thickness of a distribution's tails, this evidence points to a growing gap between the groups of children with the highest and lowest scores. A breakdown by gender reveals that this result is primarily driven by male children. It is possible that early discrepancies in variation between the sexes, which are rooted in child care attendance, may be contributing to documented higher levels of variation in male children at later stages in their educational careers [23, 32]. To this end, we also note that although the variance effect on the PPVT score is not significant for either gender, calculated effects for male and female children have opposite signs (increases for males and decreases for females). At a similar sample size as the MSD score estimates would be significantly different from each other.[25] Further exploration of the increasing variance is

[25] Recall that the PPVT score is only recorded for children age 4, whereas the MSD score samples from a children aged 0-3.

completed in conjunction with the subsequent quantile analysis.[26]

Table 3: Mean, Variance, Skewness and Kurtosis Results I

	Mean (Std. Error)	Variance (Std. Error)	Skewness (Std. Error)	Kurtosis (Std. Error)
	Original Specification			
PPVT Score	-0.435	-3.489	-0.214	-1.339
	(1.29)	(21.015)	(0.284)	(0.931)
MSD Score	-1.688	32.781	-0.600	2.191
	(0.63)***	(11.796)***	(0.215)***	(1.066)**
	Weighted by Propensity Scores			
PPVT Score	-1.550	0.126	-0.146	-0.523
	(1.046)	(21.958)	(0.332)	(1.128)
Girls	-1.334	-19.173	0.018	-0.972
	(1.519)	(30.225)	(0.447)	(1.316)
Boys	-1.606	24.753	-0.099	0.167
	(1.469)	(31.788)	(0.484)	(1.749)
MSD Score	-0.250	30.879	-0.342	2.099
	(0.481)	(12.731)**	(0.239)	(1.192)*
Girls	-0.551	19.385	-0.397	1.810
	(0.68)	(16.847)	(0.333)	(1.674)
Boys	-0.270	44.452	-0.252	2.258
	(0.672)	(18.306)**	(0.331)	(1.605)

5.2 Identification of Quantile Effects

To more completely identify policy induced changes to the distribution of available child development scores, we present four techniques, both regression and non-parametric, designed to uncover quantile treatment effects.[27] These exciting new technique provide the opportunity to see beyond mean analysis and to uncover any underlying patterns in gains

[26] In comparison with the quantile analysis outlined below RIF regression on distributional statistics remains a resource light estimation procedure. Thus, it serves as an excellent starting point for any such distributional analysis providing clear indication of variation across distributions.

[27] In fact, we explore several other techniques as well which are not discussed for brevity sake.

and losses. Focus is first given to the unconditional quantile regressions method recently developed by Firpo *et al.* (2009), which builds off the RIF regression discussed above. This technique is employed to identify a covariate's effect on the unconditional test score distribution, which is accomplished by calculating the RIF for each quantile of interest. Firpo *et al.* (2009) define the quantile RIF as:

$$RIF(Y;q_\tau) = q_\tau + \frac{\tau - I\{Y \leq q_\tau\}}{f_Y(q_\tau)} \tag{3}$$

where $I\{\cdot\}$ is an indicator function and f_Y is the density of the marginal distribution of outcome Y. Essentially, the quantile RIF describes the impact of an outcome, Y, on the location of any given quantile. The sample counterpart of this simple transformation, defined in equation 4 and constructed of the observed τ^{th} quantile ($\widehat{q_\tau}$) and a kernel density estimator ($f_Y(\widehat{q_\tau})$), replaces the outcome variable of interest (Y_{ipt}) in the general OLS Equations 1.

$$RIF(Y;\widehat{q_\tau}) = \widehat{q_\tau} + \frac{\tau - I\{Y \leq \widehat{q_\tau}\}}{f_Y(\widehat{q_\tau})} \tag{4}$$

Such a regression produces the desired coefficient estimates of all covariates corresponding to the τ^{th} quantiles, thereby revealing the impact of an independent variable on the unconditional distribution of an outcome. This allows for an in depth look at how the implementation of such an extensive child care system impacts the patterns in gains and losses for the differing portions of the distribution.

Athey and Imbens (2006) present several non-parametric models which are designed to identify the policy effect across all quantiles and by estimating an outcomes entire counterfactual distribution. The quantile difference-in-difference is perhaps the most intuitive of these as it follows a standard difference-in-difference approach in estimating the counterfactual outcome distribution. Typically used for mean policy effects, the difference-in-difference approach compares those affected by the policy with those not affected, both before and after the intervention. Such a comparison is now made at each quantile; hence, the quantile difference-in-difference counterfactual estimation procedure relies on the assumption that between period change in a comparison group's outcome at a specific quantile adequately reflects the outcome changes which would occur in the treatment group in the absence of treatment at the same quantile. While this approach has been applied to individual quantiles in the past (Poterba, Venti, and Wise (1995) and Meyer, Viscusi, and Durbin (1995)) Athey and Imbens (2006) are first to study and apply this approach to the entire distribution of outcomes [26, 30].

The estimated counterfactual of the quantile difference-in-difference approach is calculated in the following straightforward manner. We let $F_{Y_{qt}}$ denote the CDF of the outcome of interest, y, where $q = 1$ indicates the treatment group, Quebec, (and $q = 0$ the comparison group), and $t = 1$ indicates the post-policy period (and $t = 0$ for the pre-policy period). Thus, the inverse of the CDF is given by $F^{-1}_{Y_{qt}}$. For the post-policy treatment group in the absence of treatment, the counterfactual outcome at a given percentile is defined as $F^{cf-1}_{Y_{11}}(\tau)$ and constructed using the observed data in the following equation:

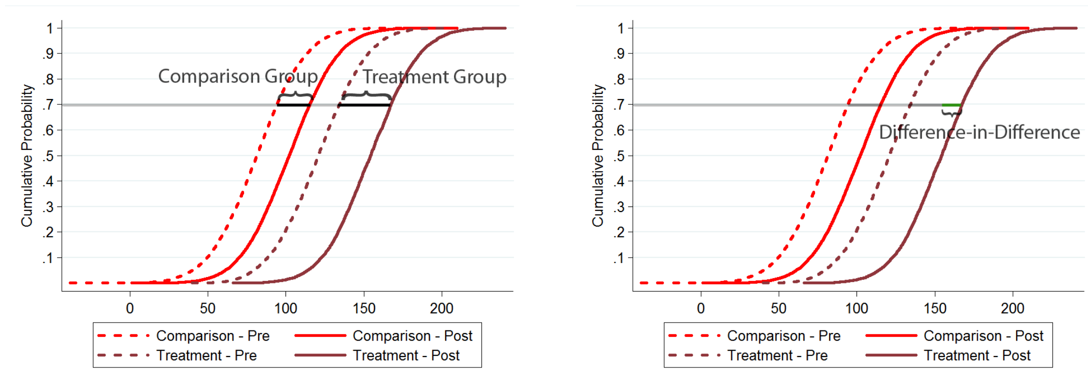
$$F^{cf-1}_{Y_{11}}(\tau) = F^{-1}_{Y_{10}}(\tau) + [F^{-1}_{Y_{01}}(\tau) - F^{-1}_{Y_{00}}(\tau)] \tag{5}$$

Thus the effect of a program at a given quantile in the distribution can be identified by the difference in the observed outcome and the estimated counterfactual one: $\Delta y(\tau) = F^{-1}_{Y_{11}}(\tau) - F^{cf-1}_{Y_{11}}(\tau)$. Graphically, this can be seen in Fig. 1 as the horizontal difference between pre- and post-treatment group CDFs subtract the horizontal distance between the the pre- and post-comparison group CDFs. It is important to note that a valid quantile difference-in-difference estimate relies on the consistent distribution of unobservables among individuals across groups and time. This disallows subpopulations to vary in unob-

Figure 1: The Quantile Difference-in-Difference Model

(a) Differences in outcome between the pre- and post-policy periods are identified for both the comparison and treatment groups at each quantile.

(b) The quantile treatment effect is the difference between these differences.



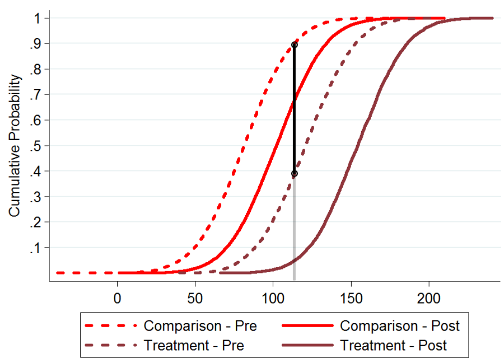
servables that make the treatment more or less effective. Unfortunately, it is not beyond belief that such differences exist between the populations in Quebec and the rest of Canada.

To address this concern Athey and Imbens (2006) also present a more generalized changes-in-change model. This estimation procedure is unique in that groups are no longer treated symmetrically as in the previous models; rather, this procedure allows for treatment to affect the treated group differently than the control group. In particular, the change-in-change model allows the distribution of unobservables across the treatment and comparison group to differ, while maintaining the assumption that individuals with similar observable and unobservable characteristics will be affected by treatment in the same way. Such an allowance can account for differences in the response to a policy. For example, the treated group may respond more positively to treatment if their distribution of unobservables is such that it is comprised of more "high return" individuals.

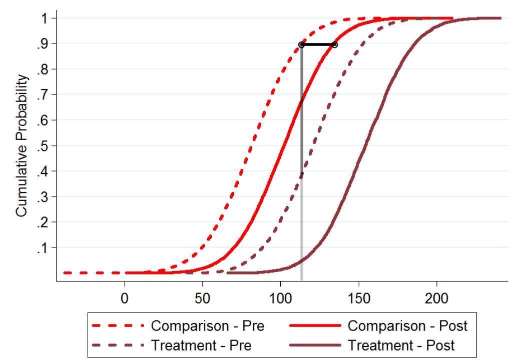
The estimation process relies on several basic assumptions. First, each period has its own mapping of observables and unobservables to outcomes in the absence of treatment.

Figure 2: The Change-in-Change Model

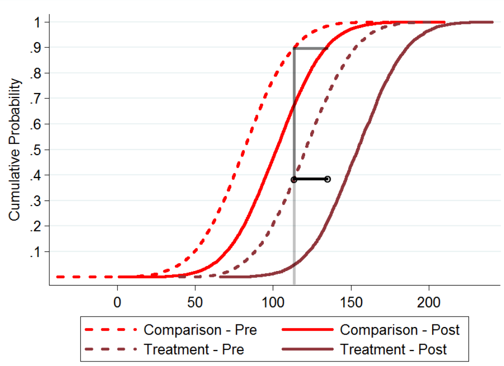
(a) Each period has its own mapping of observables and unobservables to outcomes. Therefore, individuals with similar scores have similar observables and unobservable, and thus are comparable.



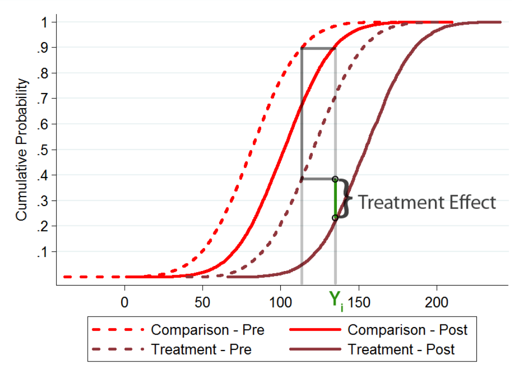
(b) The change in outcome from pre- to post-policy in the absence of treatment is identified from the comparison group individual.



(c) Given similar exposure to treatment, individuals who share observables and unobservables will experience similar change across periods. With this the counterfactual estimate is pinpointed.



(d) The difference in the CDF values of the counterfactual and observed post-policy at Y_i identify the treatment effect.



Second, each group has its own distribution of unobservables. Thus, two individuals in the same time period with the same observables and unobservables will have equal outcomes, although if these individuals are in different groups then their ranking in the distribution may be different. The counterfactual is constructed by using the distribution of unobservables of the treatment group in the first period (as understood by the observed outcomes) and mapping them into outcomes through the common second period mapping of unobservables (as outlined by second period control group outcomes). The nonparametric estimation procedure is as follows:

$$F^{cf}_{Y_{11}}(y) = F_{Y_{10}}(F^{-1}_{Y_{00}}(F_{Y_{01}}(y))) \tag{6}$$

The treatment effect, $\Delta\tau(y) = F^{cf}_{Y_{11}}(y) - F_{Y_{11}}(y)$, can be interpreted as how much more likely the treatment makes outcomes to be above a given pre-policy score. We provide the written intuition for this model accompanied by clarifying diagrams in Figure 2.

The additional heterogeneity introduced in the change-in-change model suggests that treatment effects may be different for the treated group than for the comparison group. To this end, Athey and Imbens (2006) also provide a short outline for estimating the distribution of outcomes for the comparison group had the treatment been given. This reverse change-in-change estimation operates under the same principals and assumption as the standard change-in-change estimate.

$$F^{cf}_{Y_{01}}(y) = F_{Y_{00}}(F^{-1}_{Y_{10}}(F_{Y_{11}}(y))) \tag{7}$$

In this case, the estimation relies on the distribution of outcomes for the treated group to map comparison group scores into counterfactual outcomes. In this way, the estimation produces a hypothetical treatment effect on the untreated, which can be compared to observed data to isolate the policy effect on the comparison group.

5.3 Quantile Analysis

Prior evidence on distributional statistics suggested a growing gap between the MSD score of children at the low and high ends of the distribution. In this section we explore this result by applying our quantile analysis tools. To begin, we present the un-weighted quantile analysis of each of the four discussed regression techniques in Figure 3, where calculated quantile treatment effects are plotted against pre-policy percentiles. Some care must be taken in examining these results as the calculated policy effects are interpreted differently for some analyses.

Results for the RIF regression and the quantile difference-in-difference model can be understood as the change in the location of the quantile in response to the implementation of the policy (and remains an intention to treat effect in the un-weighted results). For example, the policy estimate by the RIF regression suggests a decrease in the 10^{th} percentile by -3.8 points (corresponding to Figure 3a). The estimates from the change-in-change and reverse change-in-change models, on the other hand, are interpreted as the likelihood of a child being above the pre-policy quantile. In comparison, the estimate for the 10^{th} of the change-in-change model is -.043 (corresponding to Figure 3c), thus child MSD outcomes in the post-policy treatment group are 4.3% less likely to be above the pre-policy 10^{th} percentile.[28] While the sign of the estimates in all four methods should remain consistent, the magnitude (and subsequently the pattern) of estimates at different quantiles relative to one another are expected to vary. This is primarily due to the change in dispersion of

[28] It may also be helpful to think of the treatment effect as the change in the value of the CDF for a particular score. Thus, in the calculation of $\Delta\tau(y)$, we subtract $F^{cf}_{Y_{11}}(y)$ from $F_{Y_{11}}(y)$ rather than the $F_{Y_{11}}(y)$ from $F^{cf}_{Y_{11}}(y)$ so that positive results are considered positive change in outcomes. A lower quantile at a given score implies that more of the outcome distribution lies above.

outcomes across the distribution.[29]

An increase in variance is confirmed in the distributional results of the MSD score for each of our four estimation procedures. This results is quite striking. Strong advocates of universal child care have pointed to the extensive literature revealing positive effects of at-risk target child care programs such as Head Start. The argument is made that positive developmental gains for children in most need of the services provided through child care will counteract the losses that will occur for children taken from more effective parental care. Under this reasoning, initial negative mean results are understood to be an average of losses for highly developed children and gains for less developed children. Our estimates suggest that this may not be the case. The estimates for quantiles in the lower half of the distribution in Figure 3a are negative and mainly significant, and the largest reported effect occurs at the 3^{rd} quantile. While the estimated effects remain negative for the entire distribution their magnitude declines moving along the distribution. Although without as many significant quantile estimations, the quantile difference-in-difference estimation seen in Figure 3b mirrors this outcome. Furthermore, the change-in-change model also reflects a similar pattern, with larger negative effects in the bottom portion of the distribution relative to the top portion.

On the other hand, children in the upper portion of the distribution are seemingly unaffected by the shift to a universal child care program, as estimates are negative but mainly insignificant. This indicates that children in need of developmental support lose ground as a result of child care, and in turn casts some doubt on the ability of the Quebec

[29]Take for an example a hypothetical treatment effect such that outcome scores are improved evenly at every quantile. At the ends of the distribution, where large gaps exists between outcomes, an improvement in the outcome may be represented by the change-in-change estimate as only a small increase in the value of the CDF. On the other hand, the same improvement to outcomes in the middle of the distribution would represent a large change in the value of the CDF due to the higher density of scores. For example, at the mean there may be 5-6% change in the value of the CDF resulting from a one point score gap while at the ends of the distribution the same one point change in the score may only result in a change in the value of the CDF of 1-2%.

Family Policy to aid needy children in a similar way as targeted programs. Finally, it is interesting to note that the reverse change-in-change which represents the policy effect had it been imposed on the rest of Canada reports both larger and more consistent negative effects throughout the distribution.

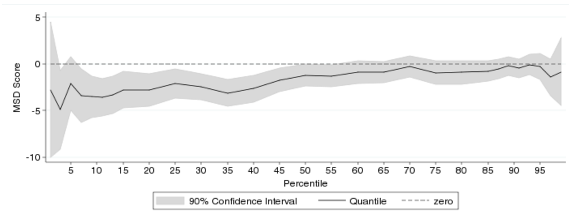
Presented in Figure 4 is the distributional analysis on the PPVT score. Although, results are mostly insignificant there is some indication of negative effects through the low to middle portions of the distribution. This is particularly evident in Figure 4d which coupled with a similar result for the MSD score suggest that a similar policy applied to the rest of Canada might have larger negative consequences.

In comparison with its un-weighted counterpart, the weighted quantile analysis presents a more expected result at the low-end of the distribution. As made clear in Figures 5 and 6 both the MSD and PPVT score estimates exhibit the anticipated positive gains in the lower portion of the distribution. This stands in contradiction to the un-weighted results and bring credence to the idea that children from less desirable home environments experience positive developmental growth due to superior care provided in childcare facilities. It is important to highlight that the RIF regressions and the quantile difference-in-difference estimations both present relatively small increases to lower end quantiles, while the change-in-change models predict estimates worthy of note. If substantial improvements occur at the lower of end the distribution, this potentially reduces future educational associated costs, in the forms of grade retention and special education. More research would be needed to verify such cost savings.

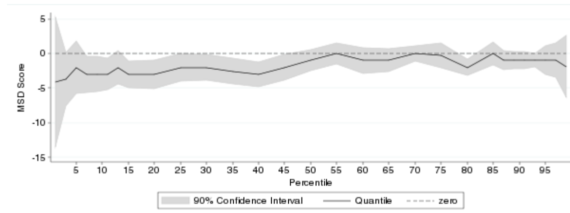
Turning to the remaining portions of the distributions, we find the MSD and PPVT scores reflect a similar pattern of gains and losses. The low end gains are followed by losses through the heart of the distribution and positive effects in the high end. These results raise concern over the treatment of the middle portion of the distribution. Why does

Figure 3: MSD - Intention-to-Treat Distributional Effects

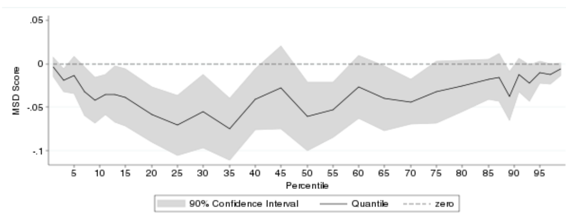
(a) RIF Regression



(b) Quantile D-in-D



(c) Change in Change



(d) Reverse Change in Change

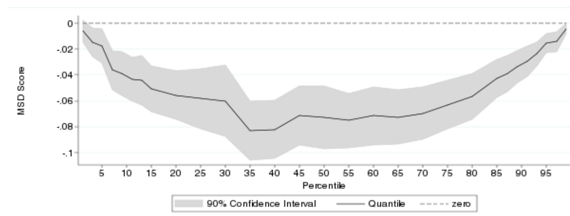
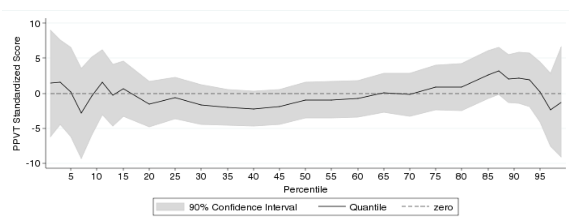
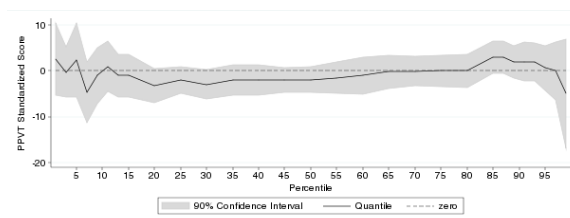


Figure 4: PPVT - Intention-to-Treat Distributional Effects

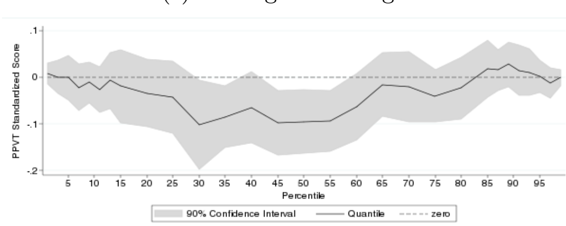
(a) RIF Regression



(b) Quantile D-in-D



(c) Change in Change



(d) Reverse Change in Change

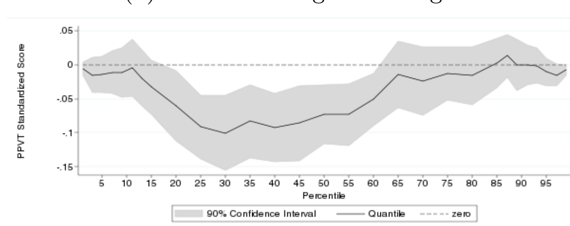
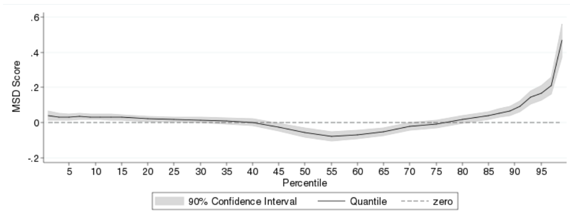
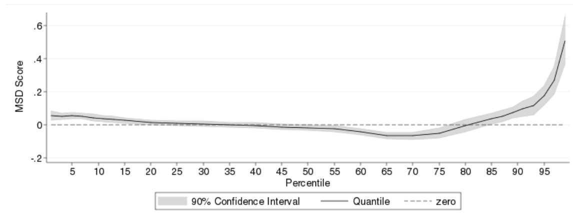


Figure 5: MSD - Average Distributional Treatment Effect on Treated

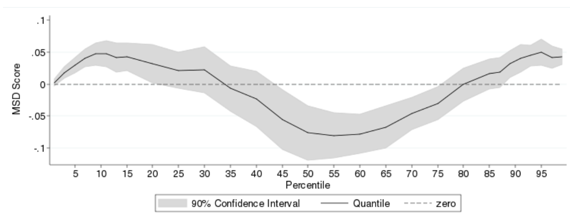
(a) RIF Regression



(b) Quantile D-in-D



(c) Change in Change



(d) Reverse Change in Change

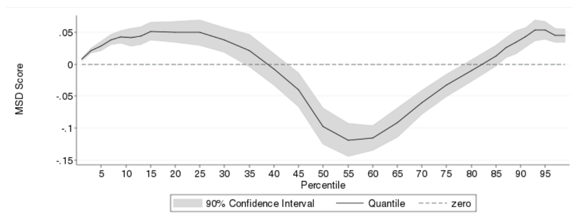
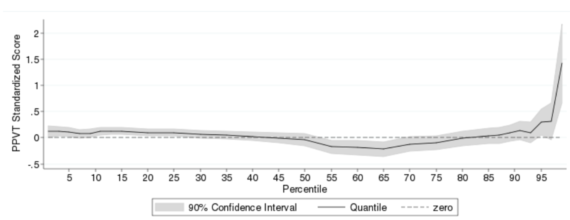
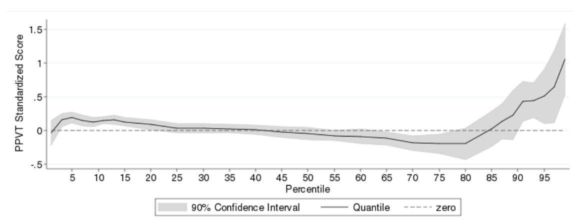


Figure 6: PPVT - Average Distributional Treatment Effect on Treated

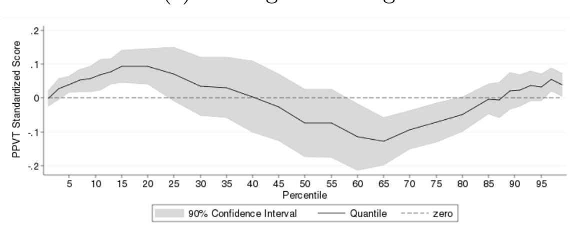
(a) RIF Regression



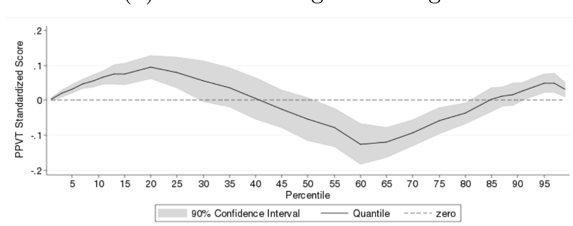
(b) Quantile D-in-D



(c) Change in Change



(d) Reverse Change in Change



childcare not appear to be effective for this group of children? Is care catering, consciously or unconsciously, to children at either ends of the distribution? Such practices would explain benefits to children in the higher portions of the distribution and would be a cause for concern for government and parent alike.

In some way this study must reconcile the differences between the weighted and un-weighted estimations. Increasing the emphasis on children most likely to attend childcare focuses attention more closely on a particular subset of the child population. Using the inverse propensity weights we examine summary statistics which suggest that the following analysis is focused on a more homogeneous group of children who tend to come from urban families with higher levels of education. On average, these children score higher than the unweighed sample for both the MSD and PPVT scores. Thus, the weighted and un-weighted populations are not readily compared and should be expected to respond differently to exposure to this new policy. Furthermore, this points to the complexity of interpreting quantile results when participation in the policy may differ across the distribution, as in the un-weighted estimation.

6 Conclusion

In this paper new evidence is brought to the intensely debated issue of publicly funded and universally provided child care by way of newly available NLSCY data and econometric techniques. First, we conducted a long-run analysis of the Quebec universal child care program. Despite significant reasons to believe that child outcomes might improve in the program's maturation, results confirm negative mean effects in both child and parental outcomes. Next, examination of the gender differences in child care related effects reveals that male children and their families are faring much worse in both behavioural and health categories. Finally, using recently developed estimation methods a variety of effects across the distribution of key development scores are discovered. In particular, positive gains are found for the tails of the distribution while large negative effects exist across the middle portion of the distribution.

Although our use of propensity weights focuses attention on those most likely to enrol in child care as a result of the policy, the array of behavioural responses to the Quebec Family Policy and the subsequent changes to child care programs in Quebec over time complicate any causal interpretation of these results. Due to shifting quality factors, changes in modes of care, increases to maternal labour supply, and changes in parental behaviour, the exact workings of a universal child care program is less than obvious. Despite this inability to differentiate between causal effects, the outcomes presented in this study still makes several contributions to child care policy debates by approaching reported results as the overall impact of the Quebec child care system.

At best, these contributions are mixed in their support for a universal child care program. The reported benefits for children of least advantage, the grounds on which universal child care is often justified, stands opposite to negative outcomes for bulk of children. Perhaps, Heckman's push for child investment and its connection to the universal provision of child care needs to be reconsider. Aligning with previous literature, these results point to targeting methods at the low end of the distribution as the most effective way to promote the well-being of children. Also, child care at the youngest ages has proven to be ineffectual. Thus, provision of care at the earliest years of life must be consider carefully. Policy-makers should weight the importance of development at the earliest stages of life appropriately and possibly consider alternative forms of investment in these children. Given current trends it does not seem likely that child care use, even at young ages, can be discouraged; this emphasizes the importance of quality of care in universal programs as the cost of cutting corners appears to be high. As our evidence suggests that program maturation and improvement of care has been ineffectual in improving child outcomes, the outlook for universal childcare is somewhat bleak.

The holistic approach of this study to the assessment of the Quebec Family Policy emphasizes the need for future research to dissect the mechanics of a universal child care program, clarifying causal relationships. Further research is also required to explore the poor health, emotional, and behavioural outcomes for male children as well as the increasing variance in outcomes. How is this evidence connected to the care provided and growing gender gap in later school achievement? Although providing an initial look at a distributional impact of child care use, the techniques used in this paper warrant further application to child care related problems. In conclusion, we suggest that the losses to children through the middle of the distribution make imperative future qualitative and quantitative research on improving outcomes for the average child in a child care settings. As a popular notion universal care child care may capture the public eye, but its implication, now and tomorrow, are far reaching and thus should be approached with evidence at the heart of the policy-making process.

7 Bibliography

References

[1] Lieselotte Ahnert, Martin Pinquart, and Michael E Lamb. Security of Childrens Relationships With Nonparental Care Providers : A Meta-Analysis. *Child Development*, 74(3):664–679, 2006.

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

[3] Joyce F Benenson, Deanna Morash, and Harriet Petrakos. Gender Differences in Emotional Closeness Between Preschool Children and Their Mothers. *Sex Roles*, 38(11-12):975–985, 1998.

[4] Timothy Besley and Anne Case. Unnatural Experiments? Estimating the Incidence of Endogenous Policies. *The Economic Journal*, 110(467):F672–F694, 2000.

[5] CBC News Your Voice. Your child-care report: Rising costs, waiting lists frustrate parents, 2009.

[6] Democratic Party. The blueprint for change: Barack Obama's plan for America. Published on behalf of Barack Obama, 2008.

[7] Peter Elfer and Katy A Dearnley. Nurseries and emotional well-being: evaluating an emotionally containing model of professional development. *Early Years*, 27(3):267–279, 2007.

[8] Frank J. Elgar, Patrick J. McGrathb, Daniel A. Waschbuschb, Sherry H. Stewart, and Lori J. Curtisc. Mutual influences on maternal depression and child adjustment problems. *Clinical Psychology Review*, 24(4):441–459, 2004.

[9] Richard A Fabes, Laura D Hanish, and Carol Lynn Martin. Children at Play: The Role of Peers in Understanding the Effects of Child Care. *Child Development*, 74(4):1039–1043, 2003.

[10] Richard A Fabes, Stephanie A Shepard, Ivanna K Guthrie, and Carol Lynn Martin. Roles of Temperamental Arousal and Gender-Segregated Play in Young Children's Social Adjustment. *Developmental Psychology*, 33(4):693–702, 1997.

[11] Sergio Firpo. Efficient Semiparametric Estimation of Quantile Treatment Effects. *Econometrica*, 75(1):259–276, 2007.

[12] Sergio Firpo, Nicole M Fortin, and Thomas Lemieux. Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973, 2009.

[13] William T Gormley Jr, Ted Gayer, Deborah Phillips, and Brittany Dawson. The effects of universal pre-k on cognitive development. *Developmental Psychology*, 41(6):872–884, 2005.

[14] Sally Grantham-mcgregor, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, and Barbara Strupp. Child development in developing countries 1: Developmental potential in the fi rst 5 years for children in. *Lancet*, 369:60–70, 2007.

[15] Tarjei Havnes and Magne Mogstad. No Child Left Behind Universal Child Care and Children s Long-Run Outcomes. 2009.

[16] James J Heckman and Dimitriy V Masterov. The Productivity Arguement for Investing in Young Children. 2007.

[17] Gary T Henry, Dana K Rickman, Bentley D Ponder, Laura W Henderson, Andrew Mashburn, and Craig S Gordon. The Georgia early childhood study: Final Report, 2004.

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

[19] Alice Sterling Honig and J. Ronald Lally. The family development research program: retrospective review. *Early Child Development and Care*, 10(1):41–62, 1982.

[20] Pierre Lefebvre and Philip Merrigan. The Effect of Childcare and on Early Education Arrangements Outcomes of Developmental. *Canadian Public Policy*, 28(2):159–186, 2002.

[21] Pierre Lefebvre and Philip Merrigan. 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, 2008.

[22] Susanna Loeb, Margaret Bridges, Daphna Bassok, Bassok Fuller, and Russel W. Rumberger. "How much is too much? The influence of preschool centers on childrens social and cognitive development," Economics of. *Economics of Education Review*, 26:52–66, 2007.

[23] David F Lohman and Joni M Lakin. Consistencies in sex differences on the Cognitive Abilities Test across countries , grades , test forms , and cohorts. *The British Psychological Society*, 79:389–407, 2009.

[24] K A Magnuson, C Ruhm, and J Waldfogel. Does prekindergarten improve school preparation and performance? *Economics of Education Review*, 26(1):33–51, 2007.

[25] Bruce D. Meyer. Natural and Quasi-Experiments in Economics. *Journal of Business & Economic Statistics*, 13(2):151–161, 2009.

[26] Bruce D. Meyer, K Viscusi, and D Durbin. Workers' Compensation and Injury Duration: Evidence from a Natural Experiment. *American Economic Review*, 85:322–340, 1995.

[27] NICHD-ECCRN. CHILD-CARE STRUCTURE -¿ PROCESS -¿ OUTCOME: Direct and Indirect Effects of Child-Care Quality on Young Children s Development. *Psychological Science*, 13(3):199–206, 2002.

[28] NICHD-ECCRN. Early child care and children' s development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, 39(1):133–164, 2002.

[29] NICHD-ECCRN. Does Amount of Time Spent in Child Care Predict Socioemotional Adjustment During the Transition to Kindergarten? *Child Development*, 74(4):976–1005, 2003.

[30] J Poterba, S Venti, and D Wise. Do 401(k) Contributions Crowd Out Other Personal Saving? *Journal of Public Economics*, 58:1–32, 1995.

[31] Pre[k]Now. State Profiles. 2009.

[32] Ellen Johnson Silver, Laurie J. Bauman, and Henry T. Ireys. Relationships of Self-Esteem and Efficacy to Psychological Distress in Mothers of Children with Chronic Physical Illness. *Health Psychology*, 14(4):333–340, 1995.

[33] Kathryn Harker Tillman, Guang Guo, and Kathleen Mullan Harris. Grade retention among immigrant children. *Social Science Research*, 35:129–156, 2006.

[34] Jocelyne Tougas. Where There s a Will, There s a Way, 2002.

[35] Elizabeth Votruba-drzal, Rebekah Levine Coley, and P Lindsay Chase-lansdale. Child Care and Low-Income Children s Development : Direct and Moderated Effects. *Child Development*, 75(1):296–312, 2004.